BMB
Reports

Invited Mini Review

Glycoscience aids in biomarker discovery

Serenus Hua[1,2] *& Hyun Joo An*[1,2,*]

[1]Graduate School of Analytical Science and Technology, Chungnam National University, Daejeon 305-764, [2]Cancer Research Institute, Chungnam National University, Daejeon 301-721, Korea

The glycome consists of all glycans (or carbohydrates) within a biological system, and modulates a wide range of important biological activities, from protein folding to cellular communications. The mining of the glycome for disease markers represents a new paradigm for biomarker discovery; however, this effort is severely complicated by the vast complexity and structural diversity of glycans. This review summarizes recent developments in analytical technology and methodology as applied to the fields of glycomics and glycoproteomics. Mass spectrometric strategies for glycan compositional profiling are described, as are potential refinements which allow structure-specific profiling. Analytical methods that can discern protein glycosylation at a specific site of modification are also discussed in detail. Biomarker discovery applications are shown at each level of analysis, highlighting the key role that glycoscience can play in helping scientists understand disease biology. [BMB Reports 2012; 45(6): 323-330]

INTRODUCTION

Proteins are commonly decorated with long carbohydrate chains, known as glycans, during normal biosynthesis. Over 60% of all proteins (from all sources, including humans, plants, bacteria, etc) are estimated to be glycosylated (1). A large population of these glycosylated proteins may be found on the cell surface, where they are optimally poised to be the first cellular components encountered by approaching cells, pathogens, antibodies, or other molecules (2-4). The glycosylation machinery within a cell is extraordinarily complex, involving a set of over 200 competing glycosyltransferases that can each modify a nascent glycoprotein by adding specific saccharides via specific linkages (5). Up- or down-regulation of an individual transferase, thus, is amplified across the entire glycan biosynthetic pathway (and by induction all glycosylated proteins) making glycosylation an extremely sensitive indicator of intracellular conditions. Glycans and glycoproteins show great promise as a source of biomarkers for aberrant cell behavior and/or disease states such as cancer (6, 7).

The development of glycan-derived biomarkers depends heavily on the advancement of analytical technologies for elucidating the glycome. The inherent diversity and complexity of glycosylation makes glycomic analysis particularly challenging; however, mass spectrometry and associated hyphenated techniques such as LC/MS offer an elegant solution to the issues at hand. Mass spectrometry (MS) is an extremely versatile tool for probing the intricacies of complex systems such as the glycome. It provides rapid and sensitive detection of sample components and can be used as a precise tool for structural elucidation. MS has contributed significantly to recent progress towards understanding the role of the glycome in biological systems such as serum, mammalian milk, saliva, tears, etc (8-17).

Mass spectrometric techniques are often paired with online liquid chromatography (LC) in order to achieve an additional dimension (or dimensions) of separation. In relation to glycomic analysis, LC is most commonly used to separate and differentiate between the many potential structural isomers possible in glycan biosynthesis. Isomeric glycoforms of identical mass exhibit differential retention times on structure-sensitive porous graphitized carbon (PGC) (17-23) or hydrophilic interaction (HILIC) (24, 25) stationary phases and may thus be separated orthogonally to MS analysis. Once separated, fragmentation techniques such as tandem MS (MS/MS or MS^n) and/or exoglycosidase digestion may be employed to differentiate structural isomers by monosaccharide structure, connectivity, and linkage (19-24).

This review summarizes recent technological and methodological advances in the closely-related fields of glycomics and glycoproteomics. Glycomic methods for glycan compositional mass profiling as well as structure-specific chromatographic profiling are discussed in detail. More recent innovations in glycoproteomics are also discussed, with a particular emphasis on site-specific methods for analyzing glycosylation. Throughout the review, applications to glycomic and glycoproteomic biomarker discovery are highlighted.

*Corresponding author. Tel: +82-42-821-8547; Fax: +82-42-821-8541; E-mail: hjan@cnu.ac.kr
http://dx.doi.org/10.5483/BMBRep.2012.45.6.132

Received 15 June, 2012

Keywords: Biomarker, Glycomics, Glycoproteomics, Glycosylation, Mass spectrometry

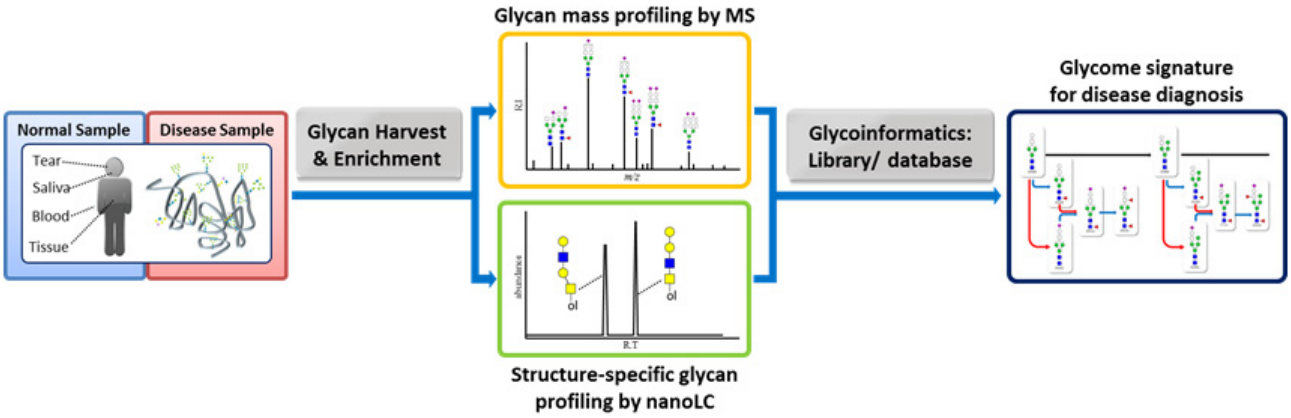

Fig. 1. Biomarker discovery platform using glycomics approach.

GLYCOMICS FOR BIOMARKER DISCOVERY: COMPOSITIONAL AND STRUCTURE-SPECIFIC APPROACHES

The search for glycan biomarkers typically begins with glycomics and other glycan-centric approaches. Glycomics analyses examine glycans which have been harvested from some biological source (commonly, glycoproteins in serum or other biological matrices). The glycan and protein moieties of a glycoprotein are separated either enzymatically (via glycanases such as peptide N-glycosidase F, a.k.a. PNGase F) or chemically (as by β-elimination). The released glycans are then enriched and analyzed.

Current methods for comprehensive analysis of glycans may be divided into two major categories- compositional mass profiling and structure-specific chromatographic profiling, both illustrated in Fig. 1. Compositional mass profiling utilizes high-resolution mass spectrometry (MS) to rapidly separate and identify glycans based on accurate mass. Glycans may be chromatographically fractionated offline and spotted for MALDI (in applications where speed is desirable) (15, 26-33); or, they may be separated online by methods such as reversed phase (RP) liquid chromatography (LC) immediately prior to introduction into the mass spectrometer via electrospray ionization (where greater sensitivity is necessary) (34-36).

Compositional mass profiling describes the glycans present in a sample in relation to the number of hexoses (such as glucose, galactose, or mannose), *N*-acetylhexosamines, deoxyhexoses (typically fucose), and sialic acids (such as *N*-acetylneuraminic acid or *N*-glycolylneuraminic acid) that each glycan is composed of. By combining the compositional information provided by MS analysis with biological knowledge of probable glycan structures, glycan structures may be inferred (37). If available, tandem MS can also be applied to further elucidate the glycan structure (2, 14, 26, 34-36, 38-42).

Glycan compositional mass profiling is an attractive option for biomarker discovery due to the simplicity of the data analysis and its potential for rapid- and high-throughput applications (when done by MALDI-MS). For example, Lebrilla and coworkers have reported extensively on the application of a MALDI- MS-based glycomics platform towards the discovery of glycan biomarkers for cancer and other diseases (7-9, 14-16, 26-29). However, compositional glycan profiling is somewhat limited by its inability to differentiate or distinguish between glycan isomers. Due to the complexity of the glycan biosynthetic pathway, a glycan composition can potentially be composed of several isomeric glycan structures with differing monosaccharide linkages or branching. Since these structures are produced by different glycosyltransferases, they often have diverging biological implications (43, 44).

To address this issue, several recent studies have focused on structure-specific chromatographic profiling of glycans (17, 18, 21, 23, 24, 45, 46). These studies utilize structure-sensitive chromatography (typically, either PGC or HILIC) to separate glycans online prior to MS analysis. Once separated, specific glycan structures may be identified based on a retention time library (21, 47) or by further fragmentation, e.g. by tandem MS (24).

The incorporation of chromatographically-derived structure specificity into existing strategies for MS-based glycan biomarker discovery promises to provide highly sensitive and specific biomarkers for many diseases. Recently, Hua *et al*. chromatographically profiled the whole serum *N*-glycome and uncovered a number of potential structure-specific biomarkers of prostate cancer (17). Meanwhile, Bäckström *et al*. examined *O*-glycan structures on MUC1, a cell-surface glycoprotein, and found differences in the glycosylation of breast, prostate, and gastric cancer samples (23).

Structure-specific glycan profiling is, at the time of writing, a relatively new technique still under development. However, the underlying analytical technology is constantly evolving, particular in the area of LC miniaturization (e.g. UHPLC and nano-LC). As the capabilities (and availability) of these instruments in-

crease, structure-specific profiling is expected to become the preferred strategy for future glycan biomarker discovery.

GLYCOPROTEOMIC APPROACHES TOWARDS DETERMINING SITE-SPECIFIC GLYCOSYLATION

Though the term "glycoproteomics" is used quite often in literature, truly integrated glycoproteomic methods are few and far between. More commonly, glycoproteins are analyzed using a variety of protein- or glycan-centric methods. A typical protein-centric approach towards glycoprotein analysis might involve enrichment for glycoproteins or glycopeptides by either lectin affinity (48-52) or hydrazide chemistry (53-56), followed by enzymatic or chemical deglycosylation of the enriched proteins or peptides. Following deglycosylation, the proteins or peptides are subjected to a standard proteomics workflow in order to identify the protein components of the sample. This approach can yield some basic information about the sites of glycosylation on a protein, but removes most if not all information about the glycan structures associated with the glycoprotein.

Conversely, glycoproteins can be analyzed using a glycan-centric (or glycomic) approach. As discussed in the previous section, this involves separation of the glycan and protein moieties of a glycoprotein, followed by analysis of the released glycans. This approach can yield extensive information about the composition and/or structure of the total glycans present in a sample (2, 14, 17, 21, 57, 58), but removes information about the glycoproteins and specific glycosylation sites from which these glycans originated.

In contrast, a truly glycoproteomic approach should integrate both glycomic and proteomic approaches to provide information about not only glycan structures but also the exact locations of these glycans on specific glycoproteins. Development of such methods is currently in their infancy; however, the general strategy involves proteolytic digestion of glycoproteins, thereby generating a mixture of glycopeptides. These glycopeptides can then be extensively interrogated by mass spectrometry for structural information about the glycan moiety as well as identity of the originating glycoprotein and site of modification (based on the peptide moiety); i.e. the glycoprotein's site-specific glycosylation.

Trypsin a commonly-used protease in proteomics workflows and, correspondingly, has been applied to determine the site-specific glycosylation of single proteins. Due to its highly specific proteolytic activity, the results of tryptic digests can easily be predicted *in silico* if the sequence of the analyte protein is known. Once tryptic peptide masses are known, it becomes a simple computational exercise to match accurate glycopeptide masses (obtained via high resolution mass spectrometry) to potential peptide/glycan combinations.

Tryptic digests have been used by many labs to determine the site-specific glycosylation of selected *N*-and *O*-glycosylated proteins, including ribonuclease B (a simple, well-characterized glycoprotein with one *N*-glycosylation site) (59); prostatic acid phosphatase (an indicator of prostate cancer) (60); bovine fetuin (a well-characterized glycoprotein with both *N*- and *O*-glycosylation) (59, 61); erythropoietin (an important biotherapeutic drug) (52, 61); horseradish peroxidase (used widely in biochemistry labs as a catalyst for oxidation) (59, 62); and haptoglobin (one of the most abundant proteins in human serum) (51, 59, 63).

However, in many cases, there are distinct disadvantages to tryptic digestion of glycoproteins. Trypsin's high substrate specificity limits the sites at which it can cleave a glycoprotein. The amino acid sequences of some glycoproteins may be fortuitously rich in arginine or lysine residues, enabling the creation of reasonably-sized glycopeptides with only a single site of glycosylation; however, other glycoprotein amino acid sequences may include extended spans without arginine or lysine residues, resulting in the creation of excessively large glycopeptides. Large glycopeptides are not only harder to detect by mass spectrometry (due to decreased ionization efficiency as well as instrumental limitations) but can also incorporate multiple sites of glycosylation, severely complicating or obfuscating site-specific analysis. This is exacerbated by the well-known phenomenon of glycoprotein trypsin resistance, which occurs when glycosylation in close proximity to a proteolytic cleavage site sterically hinders the proteolytic activity, resulting in a missed cleavage (64, 65).

One strategy to avoid such issues is to subject a glycoprotein of interest to a customized digestion scheme consisting of either sequential or simultaneous digestion by multiple specific proteases. Pompach *et al.*, for example, used endoproteinase Glu-C to cleave a problematic tryptic glycopeptide of haptoglobin with two glycosylation sites into two separate glycopeptides with one glycosylation site each, thereby enabling site-specific analysis (66). Tajiri, Yoshida, and Wada used lysylendopeptidase to create small glycopeptides of prostate specific antigen, an important biomarker for prostate cancer (67). Tajiri, Yoshida, and Wada used chymotrypsin to create *O*-glycopeptides from fibronectin after tryptic digestion failed to create any detectable *O*-glycopeptides (68). While such specifically-tailored strategies may be effective for individual glycoprotein targets, a large-scale glycoproteomic study would necessarily require a more broadly-applicable strategy for digesting glycoproteins into analyzable glycopeptides.

More recently, significant effort has gone towards development of analytical methods which utilize non-specific or broad-specificity proteases and protease cocktails to digest glycoproteins into glycopeptides and thus determine site-specific glycosylation. This represents a departure from mass spectrometry labs' traditional reliance on trypsin, inherited from the field of proteomics. Broad-specificity proteases hydrolyze peptide bonds at a number of different sites on a glycoprotein and, thus, can create glycopeptides of a roughly consistent size regardless of the amino acid sequence of the glycoprotein or steric hindrance by the glycan.

A number of broad-specificity proteases have been proposed for general use in glycoprotein digests and subsequent site-specific analysis of glycosylation. Neue *et al.* explored the utility of thermolysin in digesting *O*-glycoproteins, which often exhibit

closely-clustered glycosylation in areas rich in serine, threonine, proline, and alanine (69). Among other activities, thermolysin cleaves *N*-terminal to alanine residues, facilitating the creation of *O*-glycopeptides from within the *O*-glycosylated region with only a single glycosylation site per peptide. Proteinase K, which cleaves *C*-terminal to aliphatic, aromatic, or hydrophobic amino acids, has also been used by some groups to create small glycopeptides that are more amenable to site-specific analysis (70-73). However, perhaps the most development has gone into the use of pronase, a broad-specificity protease cocktail, as part of a generalized platform for analysis of site-specific glycosylation.

In contrast to the other proteases discussed so far, pronase is a bacterially-produced cocktail of multiple proteases with a number of different activities. As a result, pronase is remarkably robust even when dealing with markedly protease-resistant glycoproteins (74-78). Additionally, the extent of pronase digestion (and, correspondingly, the length of the resulting glycopeptides) can be modulated simply by varying the digestion time and/or the enzyme-analyte ratio (79, 80). This affords a great deal of control over the resulting digest and allows researchers to easily optimize digestions according to their experimental needs.

Due to the broad specificity of pronase, the number of *in silico* possibilities for digested glycopeptides is relatively high compared to higher-specificity proteases. In practice, however, only a few major glycopeptides are created for each glycan at each site of glycosylation. These may be easily assigned according to accurate mass using software algorithms designed to simulate non-specific cleavage of glycoproteins, such as GlycoX (81) or its in-house descendants in labs around the world. In cases where accurate mass is ambiguous or further verification is desired, tandem MS experiments can further elucidate glycopeptide composition and structure (22, 82-85).

More recently, MS-based analysis of pronase glycopeptides has been enhanced by the increasing availability of sensitive, robust, and reproducible nano-LC techniques. Online nano-LC separation of glycopeptide digests prior to MS and MS/MS detection not only increases sensitivity by reducing ion suppression, but also introduces the possibility of isomer separation and structural differentiation of isomeric protein glycoforms. Due to the structurally complex nature of glycosylation and the functional impact of small structural variations in glycans, much importance is placed on differentiating isomeric glycoforms (6, 17, 43, 86). In particular, porous graphitized carbon (PGC), a mixed-mode stationary phase popularly used in carbohydrate analysis, has been applied extensively to separation of glycan isomers (17, 18, 21, 45, 46). Lately, PGC nano-LC has been extended towards the separation of glycopeptides (22, 59, 87).

For example, Hua *et al*. recently showed that isomeric glycopeptides can be baseline resolved by chip-based PGC nano-LC (22). Glycopeptide isomer separation, combined with structural characterization by MS/MS, enabled localization of detailed *O*-glycan structures to specific glycosylation sites on trypsin-resistant *O*-glycoprotein κ-casein. Glycopeptide retention by PGC was found to be modulated by both the glycan and peptide moieties, such that the PGC retention times of glycopeptides with small peptide moieties (such as those produced by pronase or other broad-specificity proteases) were heavily affected by their glycan moieties, whereas the PGC retention times of glycopeptides with large peptide moieties (such as those produced by trypsin) were less affected by their glycan moieties. These results supported findings by Alley, Mechref, and Novotny (59) on the improved retention and separation of short tryptic glycopeptides by PGC (compared to reversed phase) and emphasize the advantages of using broad-specificity proteases to generate uniformly short glycopeptides so as to fully take advantage of the isomer-sensitive separation capabilites of PGC.

BIOMARKER APPLICATIONS OF GLYCOPROTEOMICS

Interest in glycoproteomics has spiked in recent years as the impact of glycosylation on glycoprotein structure and function has become known. Glycoproteomic methods for determining site-specific glycosylation provide unique insights into disease- related aberrations in glycosylation and have clear applications towards biomarker discovery.

Zhao *et al*. analyzed the site-specific glycosylation of α1-antitrypsin, an abundant serum glycoprotein, and were able to identify specific glycosylation patterns which varied with incidence of pancreatic cancer (43). Nakano *et al*. also studied site-specific glycosylation differences in pancreatic cancer but chose instead to focus on haptoglobin, another abundant serum protein, finding specific glycoforms which differentiated pancreatic cancer patients from both chronic pancreatitis patients and normal controls (63). Thaysen-Andersen et al. performed similar studies with tissue inhibitor of metalloproteinase-1 (TIMP-1), a known biomarker of colorectal cancer, and characterized several glycoforms of TIMP-1 that were unique to colon cancer cell lines (88).

While intriguing, these early glycoproteomic results are largely qualitative rather than quantitative, relying on stark presence/absence rather than differential levels of certain glycoforms. To find broader application and take advantage of all available data, future glycoproteomic approaches towards biomarker discovery will need to determine site-specific glycosylation in a more quantitative manner. For example, Fig. 2 shows the site-specific glycosylation on four different production batches of infliximab, a chimeric biotherapeutic, at a conserved *N*-glycosylation site in its human IgG1-derived Fc region. The high analytical reproducibility of the broad-specificity protease digests, performed in triplicate, can be seen in the miniscule error bars associated with each glycoform in the four batches. Batch-to-batch variation was tracked for 11 different glycoforms at this site of glycosylation, including two isomeric pairs.

CONCLUSION

Recent developments in MS-based glycomics and glycoproteomics have rapidly advanced the field and pushed the boundaries of glycan-derived biomarker research into new

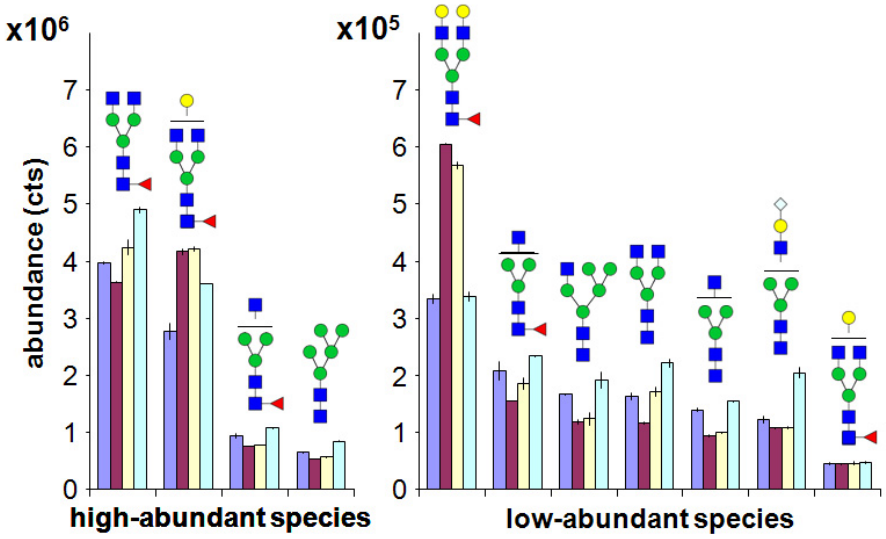

Fig. 2. Site-specific glycosylation on four different production batches of infliximab, a chimeric biotherapeutic, at a conserved *N*-glycosylation site in its human IgG1-derived Fc region. Glycan error bars denote the standard error found in three separate injections of three separate glycoprotein digests. Each color represents one of the four production batches tested.

territories. Structure-specific glycomics is now within reach and just starting to be applied to biomarker research. Meanwhile, advances in glycoproteomic analysis are beginning to yield quantitative strategies for the discovery of site-specific glycan biomarkers. These new strategies are expected to generate highly sensitive and specific biomarkers for disease from an important yet largely unexplored portion of our biology: the glycome.

Acknowledgements
We are grateful for funds provided by the Converging Research Center Program through the Ministry of Education, Science and Technology (2011K000968 for H.J.A).

REFERENCES

1. Apweiler, R., Hermjakob, H. and Sharon, N. (1999) On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database. *BBA-Gen. Subjects* **1473**, 4-8.
2. An, H. J., Gip, P., Kim, J., Wu, S., Park, K. W., McVaugh, C. T., Schaffer, D. V., Bertozzi, C. R. and Lerbilla, C. B. (2011) Extensive determination of glycan heterogeneity reveals an unusual abundance of high-mannose glycans in enriched plasma membranes of human embryonic stem cells. *Mol. Cell. Proteom.* **11**, 1-13.
3. Arndt, N. X., Tiralongo, J., Madge, P. D., von Itzstein, M. and Day, C. J. (2011) Differential carbohydrate binding and cell surface glycosylation of human cancer cell lines. *J. Cell. Biochem.* **112**, 2230-2240.
4. Li, Y.-L., Wu, G.-Z., Zeng, L., Dawe, G. S., Sun, L., Loers, G., Tilling, T., Cui, S.-S., Schachner, M. and Xiao, Z.-C. (2009) Cell surface sialylation and fucosylation are regulated by the cell recognition molecule L1 via PLCγ and cooperate to modulate embryonic stem cell survival and proliferation. *FEBS Lett.* **583**, 703-710.
5. Baum, L. G. (2002) Developing a taste for sweets. *Immunity* **16**, 5-8.
6. An, H. J., Kronewitter, S. R., de Leoz, M. L. A. and Lebrilla, C. B. (2009) Glycomics and disease markers. *Curr. Opin. Chem. Biol.* **13**, 601-607.
7. Lebrilla, C. B. and An, H. J. (2009) The prospects of glycan biomarkers for the diagnosis of diseases. *Mol. BioSyst.* **5**, 17-20.
8. An, H. J., Ninonuevo, M., Aguilan, J., Liu, H., Lebrilla, C. B., Alvarenga, L. S. and Mannis, M. J. (2005) Glycomics analyses of tear fluid for the diagnostic detection of ocular rosacea. *J. Protoeme Res.* **4**, 1981-1987.
9. Vieira, A. C., An, H. J., Ozcan, S., Kim, J.-H., Lebrilla, C. B. and Mannis, M. J. (2012) Glycomic analysis of tear and saliva in ocular rosacea patients: the search for a biomarker. *The Ocular Surface*. (In press).
10. Ninonuevo, M. R., Park, Y., Yin, H., Zhang, J., Ward, R. E., Clowers, B. H., German, J. B., Freeman, S. L., Killeen, K., Grimm, R. and Lebrilla, C. B. (2006) A strategy for annotating the human milk glycome. *J. Agric. Food Chem.* **54**, 7471-7480.
11. Barile, D., Tao, N., Lebrilla, C. B., Coisson, J.-D., Arlorio, M. and German, J. B. (2009) Permeate from cheese whey ultrafiltration is a source of milk oligosaccharides. *Int. Dairy J.* **19**, 524-530.
12. Tao, N., DePeters, E. J., Freeman, S., German, J. B., Grimm, R. and Lebrilla, C. B. (2008) Bovine milk glycome. *J. Dairy Sci.* **91**, 3768-3778.
13. LoCascio, R. G., Niñonuevo, M. R., Kronewitter, S. R., Freeman, S. L., German, J. B., Lebrilla, C. B. and Mills, D. A. (2009) A versatile and scalable strategy for glycoprofiling bifidobacterial consumption of human milk oligosaccharides. *Microb. Biotechnol.* **2**, 333-342.
14. de Leoz, M. L. A., Young, L. J. T., An, H. J., Kronewitter, S. R., Kim, J., Miyamoto, S., Borowsky, A. D., Chew, H. K. and Lebrilla, C. B. (2011) High-mannose glycans are elevated during breast cancer progression. *Mol. Cell.*

Proteom. **10**, 1-9.
15. Kronewitter, S. R., de Leoz, M. L. A., Peacock, K. S., McBride, K. R., An, H. J., Miyamoto, S., Leiserowitz, G. S. and Lebrilla, C. B. (2010) Human serum processing and analysis methods for rapid and reproducible N-glycan mass profiling. *J. Protoeme Res.* **9**, 4952-4959.
16. Barkauskas, D. A., An, H. J., Kronewitter, S. R., de Leoz, M. L., Chew, H. K., de Vere White, R. W., Leiserowitz, G. S., Miyamoto, S., Lebrilla, C. B. and Rocke, D. M. (2009) Detecting glycan cancer biomarkers in serum samples using MALDI FT-ICR mass spectrometry data. *Bioinformatics* **25**, 251-257.
17. Hua, S., An, H. J., Ozcan, S., Ro, G. S., Soares, S., DeVere-White, R. and Lebrilla, C. B. (2011) Comprehensive native glycan profiling with isomer separation and quantitation for the discovery of cancer biomarkers. *Analyst* **136**, 3663-3671.
18. Ruhaak, L. R., Miyamoto, S., Kelly, K. and Lebrilla, C. B. (2011) N-glycan profiling of dried blood spots. *Anal. Chem.* **84**, 396-402.
19. Wu, S., Tao, N., German, J. B., Grimm, R. and Lebrilla, C. B. (2010) Development of an annotated library of neutral human milk oligosaccharides. *J. Protoeme Res.* **9**, 4138-4151.
20. Wu, S., Grimm, R., German, J. B. and Lebrilla, C. B. (2010) Annotation and structural analysis of sialylated human milk oligosaccharides. *J. Protoeme Res.* **10**, 856-868.
21. Aldredge, D., An, H. J., Tang, N., Waddell, K. and Lebrilla, C. B. (2012) Annotation of a serum n-glycan library for rapid identification of structures. *J. Protoeme Res.* **11**, 1958-1968.
22. Hua, S., Nwosu, C., Strum, J., Seipert, R., An, H., Zivkovic, A., German, J. and Lebrilla, C. (2012) Site-specific protein glycosylation analysis with glycan isomer differentiation. *Anal. Bioanal. Chem.* **403**, 1291-1302.
23. Bäckström, M., Thomsson, K. A., Karlsson, H. and Hansson, G. C. (2008) Sensitive liquid chromatography-electrospray mass spectrometry allows for the analysis of the o-glycosylation of immunoprecipitated proteins from cells or tissues: application to muc1 glycosylation in cancer. *J. Protoeme Res.* **8**, 538-545.
24. Bereman, M. S., Williams, T. I. and Muddiman, D. C. (2008) Development of a nanolc ltq orbitrap mass spectrometric method for profiling glycans derived from plasma from healthy, benign tumor control and epithelial ovarian cancer patients. *Anal. Chem.* **81,** 1130-1136.
25. Bones, J., Mittermayr, S., O'Donoghue, N., Guttman, A. S. and Rudd, P. M. (2010) Ultra performance liquid chromatographic profiling of serum n-glycans for fast and efficient identification of cancer associated alterations in glycosylation. *Anal. Chem.* **82**, 10208-10215.
26. An, H. J., Miyamoto, S., Lancaster, K. S., Kirmiz, C., Li, B., Lam, K. S., Leiserowitz, G. S. and Lebrilla, C. B. (2006) Profiling of glycans in serum for the discovery of potential biomarkers for ovarian cancer. *J. Protoeme Res.* **5**, 1626-1635.
27. Kirmiz, C., Li, B., An, H. J., Clowers, B. H., Chew, H. K., Lam, K. S., Ferrige, A., Alecio, R., Borowsky, A. D. and Sulaimon, S. (2007) A serum glycomics approach to breast cancer biomarkers. *Mol. Cell. Proteom.* **6**, 43-55.
28. Leiserowitz, G. S., Lebrilla, C., Miyamoto, S., An, H. J., Duong, H., Kirmiz, C., Li, B., Liu, H. and Lam, K. S. (2008) Glycomics analysis of serum: a potential new biomarker for ovarian cancer? *Int. J. Gynecol. Cancer* **18**, 470-475.
29. de Leoz, M. L. A., An, H. J., Kronewitter, S., Kim, J., Beecroft, S., Vinall, R., Miyamoto, S., de Vere White, R., Lam, K. S. and Lebrilla, C. (2008) Glycomic approach for potential biomarkers on prostate cancer: Profiling of N-linked glycans in human sera and pRNS cell lines. *Dis. Markers* **25**, 243-258.
30. Kyselova, Z., Mechref, Y., Al Bataineh, M. M., Dobrolecki, L. E., Hickey, R. J., Vinson, J., Sweeney, C. J. and Novotny, M. V. (2007) Alterations in the serum glycome due to metastatic prostate cancer. *J. Protoeme Res.* **6**, 1822-1832.
31. Mann, B. F., Goetz, J. A., House, M. G., Schmidt, C. M. and Novotny, M. V. (2012) Glycomic and proteomic profiling of pancreatic cyst fluids identifies hyperfucosylated lactosamines on the N-linked glycans of overexpressed glycoproteins. *Mol. Cell. Proteom.* [Epub ahead of print].
32. Alley, W. R., Vasseur, J. A., Goetz, J. A., Svoboda, M., Mann, B. F., Matei, D. E., Menning, N., Hussein, A., Mechref, Y. and Novotny, M. V. (2012) N-linked glycan structures and their expressions change in the blood sera of ovarian cancer patients. *J. Protoeme Res.* **11**, 2282-2300.
33. Balog, C. I. A., Stavenhagen, K., Fung, W. L. J., Koeleman, C. A., McDonnell, L. M., Verhoeven, A., Mesker, W. E., Tollenaar, R. A. E. M., Deelder, A. M. and Wuhrer, M. (2012) N-glycosylation of colorectal cancer tissues: a liquid chromatography and mass spectrometry-based investigation. *Mol. Cell. Proteom.* [Epub ahead of print].
34. Alley, W. R., Madera, M., Mechref, Y. and Novotny, M. V. (2010) Chip-based reversed-phase liquid chromatographymass spectrometry of permethylated n-linked glycans: a potential methodology for cancer-biomarker discovery. *Anal. Chem.* **82**, 5095-5106.
35. Prater, B. D., Connelly, H. M., Qin, Q. and Cockrill, S. L. (2009) High-throughput immunoglobulin G N-glycan characterization using rapid resolution reverse-phase chromatography tandem mass spectrometry. *Anal. Biochem.* **385**, 69-79.
36. Yu, S.-Y., Chang, L.-Y., Cheng, C.-W., Chou, C.-C., Fukuda, M. and Khoo, K.-H. (2012) Priming mass spectrometry-based sulfoglycomic mapping for identification of terminal sulfated lacdiNAc glycotope. *Glycoconj. J.* 1-12. [Epub ahead of print].
37. Kronewitter, S. R., An, H. J., de Leoz, M. L., Lebrilla, C. B., Miyamoto, S. and Leiserowitz, G. S. (2009) The development of retrosynthetic glycan libraries to profile and classify the human serum N-linked glycome. *Proteomics* **9**, 2986-2994.
38. Li, B., Russell, S. C., Zhang, J., Hedrick, J. L. and Lebrilla, C. B. (2011) Structure determination by MALDI-IRMPD mass spectrometry and exoglycosidase digestions o O-linked oligosaccharides from Xenopus borealis egg jelly. *Glycobiology* **21**, 877-894.
39. An, H. J. and Lebrilla, C. B. (2011) Structure elucidation of

native N- and O-linked glycans by tandem mass spectrometry (tutorial). *Mass Spectrom. Rev.* **30**, 560-578.
40. Lancaster, K. S., An, H. J., Li, B. and Lebrilla, C. B. (2006) Interrogation of N-linked oligosaccharides using infrared multiphoton dissociation in ft-icr mass spectrometry. *Anal. Chem.* **78**, 4990-4997.
41. Ito, H., Takegawa, Y., Deguchi, K., Nagai, S., Nakagawa, H., Shinohara, Y. and Nishimura, S.-I. (2006) Direct structural assignment of neutral and sialylated N-glycans of glycopeptides using collision-induced dissociation MSn spectral matching. *Rapid Commun. Mass Sp.* **20**, 3557-3565.
42. Zhang, J., Schubothe, K., Li, B., Russell, S. and Lebrilla, C. B. (2004) Infrared multiphoton dissociation of O-linked mucin-type oligosaccharides. *Anal. Chem.* **77**, 208-214.
43. Zhao, J., Simeone, D. M., Heidt, D., Anderson, M. A. and Lubman, D. M. (2006) Comparative serum glycoproteomics using lectin selected sialic acid glycoproteins with mass spectrometric analysis: application to pancreatic cancer serum. *J. Protoeme Res.* **5**, 1792-1802.
44. De Reggi, M., Capon, C., Gharib, B., Wieruszeski, J.-M., Michel, R. and Fournet, B. (1995) The glycan moiety of human pancreatic lithostathine. *Eur. J. Biochem.* **230**, 503-510.
45. Hua, S., Lebrilla, C. and An, H. J. (2011) Application of nano-LC-based glycomics towards biomarker discovery. *Bioanalysis* **3**, 2573-2585.
46. Pabst, M., Bondili, J. S., Stadlmann, J., Mach, L. and Altmann, F. (2007) Mass + retention time = structure: a strategy for the analysis of N-glycans by carbon LC-ESI-MS and its application to fibrin N-glycans. *Anal. Chem.* **79**, 5051-5057.
47. Campbell, M. P., Royle, L., Radcliffe, C. M., Dwek, R. A. and Rudd, P. M. (2008) GlycoBase and autoGU: tools for HPLC-based glycan analysis. *Bioinformatics* **24**, 1214-1216.
48. Kreunin, P., Zhao, J., Rosser, C., Urquidi, V., Lubman, D. M. and Goodison, S. (2007) Bladder cancer associated glycoprotein signatures revealed by urinary proteomic profiling. *J. Protoeme Res.* **6**, 2631-2639.
49. Qiu, Y., Patwa, T. H., Xu, L., Shedden, K., Misek, D. E., Tuck, M., Jin, G., Ruffin, M. T., Turgeon, D. K., Synal, S., Bresalier, R., Marcon, N., Brenner, D. E. and Lubman, D. M. (2008) Plasma glycoprotein profiling for colorectal cancer biomarker identification by lectin glycoarray and lectin blot. *J. Protoeme Res.* **7**, 1693-1703.
50. Ahn, Y., Shin, P., Ji, E., Kim, H. and Yoo, J. (2012) A lectin-coupled, multiple reaction monitoring based quantitative analysis of human plasma glycoproteins by mass spectrometry. *Anal. Bioanal. Chem.* **402**, 2101-2112.
51. Miyoshi, E. and Nakano, M. (2008) Fucosylated haptoglobin is a novel marker for pancreatic cancer: detailed analyses of oligosaccharide structures. *Proteomics* **8**, 3257-3262.
52. Kurogochi, M., Amano, M., Fumoto, M., Takimoto, A., Kondo, H. and Nishimura, S.-I. (2007) Reverse glycoblotting allows rapid-enrichment glycoproteomics of biopharmaceuticals and disease-related biomarkers. *Angew. Chem. Int. Ed.* **46**, 8808-8813.
53. Zeng, X., Hood, B. L., Sun, M., Conrads, T. P., Day, R. S., Weissfeld, J. L., Siegfried, J. M. and Bigbee, W. L. (2010) Lung cancer serum biomarker discovery using glycoprotein capture and liquid chromatography mass spectrometry. *J. Protoeme Res.* **9**, 6440-6449.
54. Zhang, H., Yi, E. C., Li, X.-J., Mallick, P., Kelly-Spratt, K. S., Masselon, C. D., Camp, D. G., Smith, R. D., Kemp, C. J. and Aebersold, R. (2005) High throughput quantitative analysis of serum proteins using glycopeptide capture and liquid chromatography mass spectrometry. *Mol. Cell. Proteom.* **4**, 144-155.
55. Zhang, H., Li, X.-J., Martin, D. B. and Aebersold, R. (2003) Identification and quantification of N-linked glycoproteins using hydrazide chemistry, stable isotope labeling and mass spectrometry. *Nat. Biotech.* **21**, 660-666.
56. Zhou, Y., Aebersold, R. and Zhang, H. (2007) Isolation of N-linked glycopeptides from plasma. *Anal. Chem.* **79**, 5826-5837.
57. Tsai, H.-Y., Boonyapranai, K., Sriyam, S., Yu, C.-J., Wu, S.-W., Khoo, K.-H., Phutrakul, S. and Chen, S.-T. (2011) Glycoproteomics analysis to identify a glycoform on haptoglobin associated with lung cancer. *Proteomics* **11**, 2162-2170.
58. Dallas, D. C., Martin, W. F., Strum, J. S., Zivkovic, A. M., Smilowitz, J. T., Underwood, M. A., Affolter, M., Lebrilla, C. B. and German, J. B. (2011) N-linked glycan profiling of mature human milk by high-performance microfluidic chip liquid chromatography time-of-flight tandem mass spectrometry. *J. Agric. Food Chem.* **59**, 4255-4263.
59. Alley, W. R., Mechref, Y. and Novotny, M. V. (2009) Use of activated graphitized carbon chips for liquid chromatography/mass spectrometric and tandem mass spectrometric analysis of tryptic glycopeptides. *Rapid Commun. Mass Sp.* **23**, 495-505.
60. White, K. Y., Rodemich, L., Nyalwidhe, J. O., Comunale, M. A., Clements, M. A., Lance, R. S., Schellhammer, P. F., Mehta, A. S., Semmes, O. J. and Drake, R. R. (2009) Glycomic characterization of prostate-specific antigen and prostatic acid phosphatase in prostate cancer and benign disease seminal plasma fluids. *J. Protoeme Res.* **8**, 620-630.
61. Kuo, C.-W., Wu, I. L., Hsiao, H.-H. and Khoo, K.-H. (2012) Rapid glycopeptide enrichment and N-glycosylation site mapping strategies based on amine-functionalized magnetic nanoparticles. *Anal. Bioanal. Chem.* **402**, 2765-2776.
62. Gray, J. S. S., Yang, B. Y. and Montgomery, R. (1998) Heterogeneity of glycans at each N-glycosylation site of horseradish peroxidase. *Carbohydr. Res.* **311**, 61-69.
63. Nakano, M., Nakagawa, T., Ito, T., Kitada, T., Hijioka, T., Kasahara, A., Tajiri, M., Wada, Y., Taniguchi, N. and Miyoshi, E. (2008) Site-specific analysis of N-glycans on haptoglobin in sera of patients with pancreatic cancer: A novel approach for the development of tumor markers. *Int. J. Cancer* **122**, 2301-2309.
64. Wu, Z. L., Ethen, C., Hickey, G. E. and Jiang, W. (2009) Active 1918 pandemic flu viral neuraminidase has distinct N-glycan profile and is resistant to trypsin digestion. *Biochem. Biophys. Res. Commun.* **379**, 749-753.
65. Fujihara, J., Yasuda, T., Kunito, T., Fujii, Y., Takatsuka, H., Moritani, T. and Takeshita, H. (2008) Two N-linked glycosylation sites (Asn18 and Asn106) are both required for full enzymatic activity, thermal stability and resistance to proteolysis in mammalian deoxyribonuclease i. *Biosci. Biote-*

chnol. Biochem. **72**, 3197-3205.
66. Pompach, P., Chandler, K. B., Lan, R., Edwards, N. and Goldman, R. (2012) Semi-automated identification of n-glycopeptides by hydrophilic interaction chromatography, nano-reverse-phase lc–ms/ms and glycan database search. *J. Protoeme Res.* **11**, 1728-1740.
67. Tajiri, M., Ohyama, C. and Wada, Y. (2008) Oligosaccharide profiles of the prostate specific antigen in free and complexed forms from the prostate cancer patient serum and in seminal plasma: a glycopeptide approach. *Glycobiology* **18**, 2-8.
68. Tajiri, M., Yoshida, S. and Wada, Y. (2005) Differential analysis of site-specific glycans on plasma and cellular fibronectins: application of a hydrophilic affinity method for glycopeptide enrichment, *Glycobiology* **15**, 1332-1340.
69. Neue, K., Mormann, M., Peter-Katalinić, J. and Pohlentz, G. (2011) Elucidation of glycoprotein structures by unspecific proteolysis and direct nanoESI mass spectrometric analysis of ZIC-HILIC-enriched glycopeptides. *J. Protoeme Res.* **10**, 2248-2260.
70. Larsen, M. R., Højrup, P. and Roepstorff, P. (2005) Characterization of gel-separated glycoproteins using two-step proteolytic digestion combined with sequential microcolumns and mass spectrometry. *Mol. Cell. Proteom.* **4**, 107-119.
71. Xin, L., Zhang, H., Liu, H. and Li, Z. (2012) Equal ratio of graphite carbon to activated charcoal for enrichment of N-glycopeptides prior to matrix-assisted laser desorption/ionization time-of-flight mass spectrometric identification. *Rapid Commun. Mass Sp.* **26**, 269-274.
72. Thaysen-Andersen, M., Mysling, S. and Højrup, P. (2009) Site-specific glycoprofiling of N-Linked glycopeptides using MALDI-TOF MS: strong correlation between signal strength and glycoform quantities. *Anal. Chem.* **81**, 3933-3943.
73. Zauner, G., Koeleman, C. A. M., Deelder, A. M. and Wuhrer, M. (2010) Protein glycosylation analysis by HILIC-LC-MS of Proteinase K-generated N- and O-glycopeptides. *J. Sep. Sci.* **33**, 903-910.
74. Yu, Y. Q., Fournier, J., Gilar, M. and Gebler, J. C. (2007) Identification of N-linked glycosylation sites using glycoprotein digestion with pronase prior to MALDI tandem time-of-flight mass spectrometry. *Anal. Chem.* **79**, 1731-1738.
75. An, H. J., Froehlich, J. W. and Lebrilla, C. B. (2009) Determination of glycosylation sites and site-specific heterogeneity in glycoproteins. *Curr. Opin. Chem. Biol.* **13**, 421-426.
76. An, H. J., Peavy, T. R., Hedrick, J. L. and Lebrilla, C. B. (2003) Determination of N-glycosylation sites and site heterogeneity in glycoproteins. *Anal. Chem.* **75,** 5628-5637.
77. Li, H., Li, B., Song, H., Breydo, L., Baskakov, I. V. and Wang, L.-X. (2005) Chemoenzymatic synthesis of HIV-1 V3 glycopeptides carrying two N-glycans and effects of glycosylation on the peptide domain. *J. Org. Chem.* **70**, 9990-9996.
78. Liu, X., McNally, D. J., Nothaft, H., Szymanski, C. M., Brisson, J.-R. and Li, J. (2006) Mass spectrometry-based glycomics strategy for exploring N-linked glycosylation in eukaryotes and bacteria, *Anal. Chem.* **78,** 6081-6087.
79. Dodds, E. D., Seipert, R. R., Clowers, B. H., German, J. B. and Lebrilla, C. B. (2008) Analytical performance of immobilized pronase for glycopeptide footprinting and implications for surpassing reductionist glycoproteomics. *J. Protoeme Res.* **8,** 502-512.
80. Clowers, B. H., Dodds, E. D., Seipert, R. R. and Lebrilla, C. B. (2007) Site determination of protein glycosylation based on digestion with immobilized nonspecific proteases and fourier transform ion cyclotron resonance mass spectrometry. *J. Protoeme Res.* **6,** 4032-4040.
81. An, H. J., Tillinghast, J. S., Woodruff, D. L., Rocke, D. M. and Lebrilla, C. B. (2006) A new computer program (GlycoX) to determine simultaneously the glycosylation sites and oligosaccharide heterogeneity of glycoproteins. *J. Protoeme Res.* **5,** 2800-2808.
82. Seipert, R. R., Dodds, E. D., Clowers, B. H., Beecroft, S. M., German, J. B. and Lebrilla, C. B. (2008) Factors that influence fragmentation behavior of N-linked glycopeptide ions. *Anal. Chem.* **80,** 3684-3692.
83. Seipert, R. R., Dodds, E. D. and Lebrilla, C. B. (2008) Exploiting differential dissociation chemistries of o-linked glycopeptide ions for the localization of mucin-type protein glycosylation. *J. Protoeme Res.* **8,** 493-501.
84. Wuhrer, M., Koeleman, C. A. M., Hokke, C. H. and Deelder, A. M. (2004) Protein glycosylation analyzed by normal-phase nano-liquid chromatography-mass spectrometry of glycopeptides. *Anal. Chem.* **77**, 886-894.
85. Temporini, C., Perani, E., Calleri, E., Dolcini, L., Lubda, D., Caccialanza, G. and Massolini, G. (2006) Pronase-immobilized enzyme reactor: an approach for automation in glycoprotein analysis by LC/LC-ESI/MSn. *Anal. Chem.* **79,** 355-363.
86. Tang, Z., Varghese, R. S., Bekesova, S., Loffredo, C. A., Hamid, M. A., Kyselova, Z., Mechref, Y., Novotny, M. V., Goldman, R. and Ressom, H. W. (2009) Identification of N-glycan serum markers associated with hepatocellular carcinoma from mass spectrometry data. *J. Protoeme Res.* **9**, 104-112.
87. Froehlich, J. W., Barboza, M., Chu, C., Lerno, L. A., Clowers, B. H., Zivkovic, A. M., German, J. B. and Lebrilla, C. B. (2011) Nano-LC–MS/MS of glycopeptides produced by nonspecific proteolysis enables rapid and extensive site-specific glycosylation determination. *Anal. Chem.* **83**, 5541-5547.
88. Thaysen-Andersen, M., Thøgersen, I. B., Lademann, U., Offenberg, H., Giessing, A. M. B., Enghild, J. J., Nielsen, H. J., Brünner, N. and Højrup, P. (2008) Investigating the biomarker potential of glycoproteins using comparative glycoprofiling-application to tissue inhibitor of metalloproteinases-1, Biochimica et Biophysica Acta (BBA) - Proteins &amp. *Proteomics* **1784**, 455-463.