CrossMark

EDITORIAL

Recent advances in glycomics, glycoproteomics and allied topics

Yehia Mechref[1] · David C. Muddiman[2]

Proteins are produced by the ribosome, which translates the genetic code of messenger RNA (mRNA) into polypeptide chains. However, the number of proteins produced in a cell is significantly higher than that predicated solely by gene code translation to mRNA. Different mRNA transcripts are produced from a single gene, providing the necessary diversity of expressed proteins that exceeds what is encoded by a given genome. Proteins perform a vast array of functions, many of which are facilitated through co- and post-translational modifications (PTMs). Proteins undergo PTM to modulate many biofunctions.

Glycosylation is a common co-/post-translational modification of proteins. Approximately 70% of human proteins are glycosylated. Numerous biological functions have been associated with protein glycosylation, including cell–cell signaling, protein stability, protein localization, and immune response. The importance of protein glycosylation is readily apparent in the development and progression of human diseases such as inflammatory diseases, rheumatoid arthritis, Alzheimer's disease, Pompe disease, and cancer. These biofunctions and roles are modulated through the presence of several classes of glycan structures, including O- and N-linked glycans of proteins, glycosaminoglycans (GAGs, heparin and chondroitin sulfates), hyaluronic acids, O-linked galactose to hydroxylysines on collagen, and O-linked glucose to tyrosine on glycogen.

Our understanding of the attributes of glycans in biological functions and the development and progression of diseases has been facilitated by mass-spectrometry-based glycomics and glycoproteomics. Glycomics is focused on defining the glycan structures in a system irrespective of the sites to which such structures are attached. Glycoproteomics, on the other hand, is aimed at defining the micro- and macroheterogeneities of glycoproteins. These two fields of research are dynamic and are rapidly evolving through innovative technologies and bioinformatics by which new biological insights are being garnered.

Glycomics is an emerging research area, the overarching goal of which is to define the complex and diverse structures of glycans (the glycome) produced in a biological system in order to better assess the biological attributes of such structures. Different separation techniques, such as capillary electrophoresis and liquid chromatography, are commonly interfaced to mass spectrometry to comprehensively characterize the high complexity and structural diversity of a glycome associated with a wide range of complex biological systems. Additionally, structural characterization of glycans is achieved using several tandem mass spectrometry techniques, including collision-induced dissociation (CID), electron transfer dissociation (ETD), infrared multiphoton dissociation (IRMPD) and electron capture dissociation (ECD).

Glycoproteomics, a subset of proteomics focused on glycoproteins, seeks to identify changes in glycoprotein macroheterogeneity and microheterogeneity. Currently, several technical challenges are associated with the analysis of glycoproteins derived from biological samples. Most glycoproteins exist naturally in low abundance in biological systems.

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

✉ Yehia Mechref
   yehia.mechref@ttu.edu

✉ David C. Muddiman
   dcmuddim@ncsu.edu

[1] Department of Chemistry and Biochemistry, Texas Tech University, Lubbock, TX, USA

[2] Department of Chemistry, North Carolina State University, Raleigh, NC, USA

Meanwhile, the hydrophilic nature of glycans adversely influences the ionization efficiency of a glycoprotein. Moreover, glycoproteins commonly have more than one glycosylation site and/or other PTMs, which greatly complicates the characterization of site-specific glycosylation. The MS/MS spectra of glycopeptides contain information for both the peptide backbone and the glycan moieties. Moreover, the interpretation of MS/MS spectra requires a significant amount of data processing time and expertise; the proteomics data processing workflow is relatively robust, whereas glycoproteomics data processing is still in its infancy. These challenges are addressed though the implementation of several approaches and techniques. Enrichment methods are developed to facilitate enhanced glycoprotein/glycopeptide analysis. In order to assess the macroheterogeneity and microheterogeneity of glycoproteins, a bottom-up strategy is commonly employed. LC-MS/MS is currently the most common and powerful analytical platform for the analysis of protein glycosylation via glycopeptide analysis. Several tandem MS techniques are utilized, often sequentially, to identify the peptide backbone, glycosylation site, and diagnostic product ions indicative of a glycopeptide. Accurate and precise bioinformatics tools are necessary to automate data analysis, which facilitates the elucidation of useful biological information from complex glycoproteomics datasets.

This topical collection of articles includes three review articles and covers a broad spectrum of glycomics and glycoproteomics methods that were developed to enhance the sensitive and quantitative characterization of glycans and glycoproteins derived from biological systems. The use of reversed-phase liquid chromatography for the characterization of glycans derivatized with hydrophobic reagents is reviewed in an article by Vreeker and Wuhrer. This review highlights the advantages of RP-LC for glycan analysis and discusses the retention of glycans in relation to glycan monosaccharide composition, position, and linkage. The roles of glycans in schizophrenia, Alzheimer's disease, and cancer are reviewed in articles by Sethi and Zaia, and Kailemia et al., respectively.

Lin et al. describe the analysis of heparin by capillary electrophoresis in conjunction with negative-ion electrospray ionization mass spectrometry. The described analysis was attained using a volatile methanolic ammonium acetate electrolyte and sheath fluid. Jiang et al. illustrate an enhanced MALDI MS analysis of oligosaccharides using 2-hydrazinopyrimidine as a derivatizing reagent and MALDI co-matrix. The use of nanoporous carbon material prepared by the carbonization of a metal–organic framework or the highly selective enrichment of N-linked glycans is described by Zhang and coworkers.

Isomeric characterization of glycans is a vibrant area of glycomics. Unequivocal characterization of glycan isomers is commonly attained by MS^n or by employing separation techniques in conjunction with MS and tandem MS. Reinhold and coworkers describe the use of MS^n to determine multiple isomeric structures derived from two breast cancer cell lines, MCF-7 and MDA-MB-231. Mechref and coworkers describe the use of a porous graphitic carbon column at high temperatures in conjunction with tandem MS to characterize several permethylated glycan isomers, including linkage and antenna isomers. The use of positive and negative polarity in a gas-phase separation technique, namely ion mobility spectrometry–MS, in conjunction with metal ions to enhance glycan isomer separation, is illustrated in an article by Zheng et al. The described method was effectively employed to determine the glycotopes of the protozoan parasite *Trypanosoma cruzi*.

The development of analytical and data processing tools to enable effective characterization of the glycoprofiles of therapeutic glycoproteins is described by Planinc et al. The method employs hydrophilic interaction liquid chromatography coupled to a high-resolution mass spectrometer (HILIC-HRMS) to acquire glycan profiles which are analyzed by principal component analysis (PCA) and classified through soft independent modeling by class analogy (SIMCA). This paper described a combination of analytical and chemometric techniques that enabled the detection of small changes in glycosylation which were not detected by simply comparing extracted ion chromatograms.

Muddiman and coworkers describe the use of Skyline in conjunction with a stable isotope-labeled hydrazide tag (INLIGHT^{TM}) as a platform for the characterization and relative quantitation of the N-linked glycome in differentiating xylem of *Populus trichocarpa*. In a unique application, Mangrum et al. employed INLIGHT^{TM} with high-resolution accurate mass (HRAM) RP-LC-MS/MS to characterize heparinase-digested heparin products. The described method enabled comparative analyses of heparin, heparan sulfate, and other GAGs.

As mentioned above, sensitive glycoprotein analysis demands an enrichment step. Three articles in this topical collection describe new approaches enabling efficient glycoprotein enrichment. Glycoprotein enrichment by a diethy laminoethyl sepharose (DEAE-Sepharose) microcolumn is compared to click maltose and zwitterionic HILIC (ZIC-HILIC) by Zhu et al. The DEAE-Sepharose microcolumn enriched glycopeptides more efficiently than the other two methods. The importance of solvent composition in ZIC-HILIC is discussed by Kolarich et al. Zhang et al. described new boronate materials modulating both hydrophilic and affinity interactions to enhance the enrichment of glycopeptides. The technical reproducibility of the characterization of the O-glycosylation of mucine-type core 1 in human serum using a two-step lectin affinity enrichment is addressed in an article by Darula and coworkers. The article also addresses the roadblocks

hindering the in-depth quantitation of such *O*-glycans in biological samples.

Automated interpretation of glycoproteomics data remains an active area of research. The overarching goal of most of the efforts in this area is focused on the development of robust tools that enable automated interpretation of MS-based glycoproteomics data. Bern and coworkers describe an approach enabling automated analysis of bacterial peptidoglycan structures with Byonic software. The described approach determined several muropeptides existing at low abundances. The sequences of such peptides have not been reported previously. Additionally, GlycoPep MassList, a software tool to generate inclusion lists for glycopeptide analyses, is described and discussed by Hu et al., while Chalkley and Baker describe the use of a glycosylation site database to improve glycopeptide identification from complex biological samples.

Several papers at the end of this topical collection are dedicated to the development of cutting-edge methods in glycoproteomics. Woo et al. extended the recently developed isotope-targeted glycoproteomics (IsoTaG) approach to allow the use of either alkynyl or azido sugars to analyze the sialylated glycoproteome from PC-3 cells. The described approach enabled the characterization of 699 glycopeptides originating from 192 glycoproteins. Multiple reaction monitoring (MRM) was employed by Lebrilla and coworkers to quantify human milk glycoproteins and their glycoforms. Khatri et al. describe the need to use both proteomics and glycomics results to increase the confidence associated with the assignment of the microheterogeneity of the glycosylation sites of proteins. Sanda et al. employed data-independent LC-MS/MS analysis with soft fragmentation to monitor the changes in the glycosylation of proteins in liver cirrhosis.

The goal of this topical collection is to highlight new advances in MS-based glycomics and glycoproteomics. The articles included in this issue covered many aspects related to the challenges associated with the characterization of complex and diverse glycan biomolecules. Advances in MS-based glycomics and glycoproteomics will pave the way to a better understanding of the biological attributes of glycans. The research and review articles included in this topical collection highlight many pressing needs in the ever-vibrant and evolving research areas of MS-based glycomics and glycoproteomics.

Yehia Mechref is a professor in the Department of Chemistry and Biochemistry at Texas Tech University, Lubbock, TX, USA. He received a B.Sc. in chemistry from the American University of Beirut (Beirut, Lebanon) and a Ph.D. with an honorable mention from Oklahoma State University (Stillwater, OK, USA). His research focus is on the development of sensitive biomolecular mass spectrometry methods enabling qualitative and quantitative assessments of the roles of proteins, glycoproteins, and glycans in biological systems. Thus far, he has published 21 review articles, 14 book chapters, and 154 peer-reviewed research papers. Currently, his Scopus h-index is 46, with 5862 citations. He has received 11 US patents. He has organized and co-organized numerous symposia and conferences. He received the Barnie E. Rushing, Jr. Faculty Distinguished Research Award in 2016 and the Barnie E. Rushing, Jr. Faculty Outstanding Research Award in 2015.

David C. Muddiman is the Jacob and Betty Belin Distinguished Professor of Chemistry, and founder and director of the W.M. Keck FTMS Laboratory for Human Health Research at North Carolina State University in Raleigh, NC, USA. Prior to moving his research group to North Carolina State University, he was a professor of biochemistry and molecular biology, founder, and director of the Mayo Proteomics Research Center at the Mayo Clinic College of Medicine in Rochester, MN, USA. He is Editor of *Analytical and Biological Chemistry* and Associate Editor of the *Encyclopedia of Analytical Chemistry*, and he is on the Editorial Advisory Board of *Mass Spectrometry Reviews*, *Molecular and Cellular Proteomics*, *Rapid Communications in Mass Spectrometry*, and the *Journal of Chromatography B*. He also serves on the Advisory Board of the NIH-funded Complex Carbohydrate Research Center, University of Georgia, and the Yale/NIDA Neuroproteomics Center, Yale University. He is currently a member of the ASMS Board of Directors and the President of the United States Human Proteome Organization (US-HUPO). His group has published over 225 peer-reviewed papers and received four US patents. He is the recipient of the 2015 ACS Award in Chemical Instrumentation, the 2010 Biemann Medal from the American Society for Mass Spectrometry, the 2009 NCSU Alumni Outstanding Research Award, the 2004 ACS Arthur F. Findeis Award, the 1999 American Society for Mass Spectrometry Research Award, and the 1990–91 Safford Award, University of Pittsburgh, for Excellence in Teaching.