Molecular Omics

EDITORIAL

Check for updates

Glycomics & Glycoproteomics: From Analytics to Function

Morten Thaysen-Andersen, [iD] †[ab] Daniel Kolarich [iD] †[cd] and Nicolle H. Packer [iD] †*[abcd]

"*In analogy to the current situation in Cosmology, glycans can be considered as the "dark matter" of the biological universe: a major and critical component that has yet to be fully incorporated into the "standard model" of biology.*"

Ajit Varki and Stuart Kornfeld[1]

Outlining the topic and its importance

Glycomics and glycoproteomics define powerful technologies that help scientists to capture this biological "*dark matter*" and study all glycans and glycoproteins expressed by cells, tissues or organisms at a given time, space and condition.[2–4] While both glycomics and glycoproteomics methods have existed for several years,[5–8] key advances in the separation sciences, mass spectrometry and informatics have recently allowed for a more complete implementation and integration of glycomics and glycoproteomics in the

[a] Department of Molecular Sciences, Macquarie University, Sydney, NSW, 2109, Australia. E-mail: nicki.packer@mq.edu.au
[b] Biomolecular Discovery Research Centre, Macquarie University, Sydney, NSW, 2109, Australia
[c] Institute for Glycomics, Griffith University, Southport, QLD, 4222, Australia
[d] ARC Centre of Excellence for Nanoscale BioPhotonics, Macquarie University, Sydney, NSW, 2019, Australia
† Contributed equally.

life sciences. As exemplified by papers published in this themed issue in *Molecular Omics* "Glycomics & Glycoproteomics: From Analytics to Function", studies that concertedly apply two or more glyco-centric 'omics techniques to uncover new mechanisms in various research areas in glycobiology such as pathogen–host interactions (Cain *et al.*, DOI: 10.1039/D0MO00032A; Mule *et al.*, DOI: 10.1039/D0MO00043D; Delannoy *et al.*, DOI: 10.1039/C9MO00173E; Mthembu *et al.*, DOI: 10.1039/C9MO00175A) cancer and biomarkers (Chandler *et al.*, DOI: 10.1039/D0MO00009D; Acharya *et al.*, DOI: 10.1039/C9MO00061E) and embryogenesis (Qu *et al.*, DOI: 10.1039/D0MO00005A), are becoming common in the literature. This themed issue also illustrates that mass spectrometry-based glycoproteomics studies often use multiple orthogonal fragmentation schemes to produce complementary glycan structural information to aid glycopeptide characterisation at scale, with sensitivity, speed and accuracy (Pap *et al.*, DOI: 10.1039/C9MO00160C; Chalkley *et al.*, DOI: 10.1039/C9MO00178F). Other new technology improvements (Cordina *et al.*, DOI: 10.1039/C9MO00181F; Zhang *et al.*, DOI: 10.1039/C9MO00180H; Sethi *et al.*, DOI: 10.1039/D0MO00019A), informatics advancements (Chalkley *et al.*, DOI: 10.1039/C9MO00178F; Uh *et al.*, DOI: 10.1039/C9MO00174C; Phung *et al.*, DOI: 10.1039/C9MO00125E) and pathway analyses (del Solar *et al.*, DOI: 10.1039/D0MO00023J;

Donald *et al.*, DOI: 10.1039/C9MO00168A) also contribute to an ever-expanding analytical toolbox that now allows scientists to explore not only the structure but also the biosynthesis and function of glycans and glycoproteins.

The field has for years recognised that no single technology can provide an in-depth, often overlooked, understanding of the compositional, structural and functional heterogeneity inherently associated with protein glycosylation. Nonetheless, integration of glycomics and glycoproteomics with the many interfacing 'omics disciplines has only commenced recently. Multifaceted analytical strategies and the integration of glycomics and glycoproteomics with proteomics, transcriptomics, metabolomics and genomics are critical to dive deeper into the immensely complex and dynamic glycoproteome, the underpinning glycosylation machinery and their effects on cellular activity. Glycobiology, arguably more than other biological research areas, benefits from the synergy that can be accessed using multi-omics approaches as, for example, demonstrated in "*glycomics-assisted glycoproteomics*" studies.[9–11] In such studies, the glycomics data not only provide essential information on the fine detail of the attached glycan structures and define the boundaries of the glycome, but also enable more effective characterisation of the accompanying glycoproteomics data

and give insights into the role of glycans in protein function.

The body of research papers compiled in this themed issue and elsewhere clearly shows that glycomics and glycoproteomics have matured considerably over the past years. As the wider scientific community increasingly accesses these technologies, and the workflows and data collection, interpretation, reporting and sharing are being streamlined, we are beginning to unlock the real potential of glycomics and glycoproteomics to generate new insights into how the heterogeneous glycosylation of proteins is involved in so many biological processes.

Future directions of the field

At present, the mass spectrometric technologies, when applied to mixtures of intact glycopeptides, mostly assign the monosaccharide compositions of glycans to peptides, and only in favourable cases to discrete sites. However, more intricate structural features of glycans are often required to understand the effect of glycosylation on protein function in such areas as cellular interactions, protein cleavage, immune recognition and in biomarker discovery. Technologies that can uncover the heterogeneity of detailed structures decorating each protein site remain at this time a "holy grail" of glycobiology. Excitingly, innovative developments in "*structure-focused glycoproteomics*" employing, for example, diagnostic tandem mass spectrometry fragment ions,[12–14] sequential exoglycosidase digestions on glycopeptides,[15,16] differential liquid chromatography and ion mobility based retention,[17–21] and linkage-specific derivatisation of glycopeptides,[22] to annotate certain glycan structural features together with site information, are beginning to appear in the literature.

The other missing link for biological researchers is the means to be able to more easily interpret the volume of data obtained by the specialised mass spectrometric approaches that are being constantly expanded to cover top down, middle down and bottom up analyses of glycoproteins.[23–25] The development of glycobioinformatics tools and curated databases, together with the implementation of standardized reporting guidelines such as MIRAGE (Minimum Information Required For a Glycomics Experiment)[26–29] (https://www.beilstein-institut.de/en/projects/mirage/), to cope with these data are essential for the continued penetration of glycobiology into the many streams of life sciences. The agreed implementation of a common "glyco-language" *via* the "Symbol Nomenclature for Glycans (SNFG)"[30,31] together with the assignment of unique accession numbers to the glycan structures (https://glytoucan.org/) has started to unify and facilitate the depiction and data sharing of complex biomolecular structures across the chemical and biological sciences.

Despite this progress, the most common response of scientists to the question as to why they have not explored the glycobiology of their research question at hand has been that it is "*too hard*" – this reluctance can only be addressed by the availability of user friendly methods, software and accurate knowledge bases of glycobiological information. This information needs to be not only that acquired by the technologies of glycomics and glycoproteomics, but the seamless integration of such datasets with information delivered by well-established genomics, transcriptomics, proteomics, metabolomics and/or lipidomics technologies for an integrated and holistic understanding of cellular biology in health and disease, such as has now begun with the international glycoinformatics initiatives of GlyConnect (https://glyconnect.expasy.org/), Glygen (https://glygen.org/) and GlyCosmos (https://glycosmos.org/) under the collaborative GlySpace Alliance.[32]

Once the dynamics and functional glycan features are determined on proteins within a complex biological system it is important to translate these findings into useful practice. Coordinated efforts in glycobiotechnology and clinical glycoscience that aim to connect fundamental and applied glycobiology and to translate findings in this field into products including new therapeutic targets to treat, cure and ultimately prevent disease and new biomarkers to diagnose and prognose diseases, are ongoing.[33–35]

Not to be forgotten in the future, of course, is that glycomics does not only apply to the structure and function of proteins, but also refers to the characterisation of glycans that are found attached to lipid anchors or found as free glycans, proteoglycans, bacterial lipopolysaccharides and peptidoglycans. At this time, the technologies for the analysis of these molecules are in their infancy, and are practised by few laboratories world-wide, but must be taken into consideration as we come to realise that multi-omics analyses must be addressed across the entire breadth of molecular sciences if we are to understand the complex systems that are biology.

References

1 A. Varki, *et al.*, in *Essentials of Glycobiology*, ed. A. Varki, R. D. Cummings, J. D. Esko, P. Stanley, G. W. Hart, M. Aebi, A. G. Darvill, T. Kinoshita, N. H. Packer, J. H. Prestegard, R. L. Schnaar and P. H. Seeberger, Cold Spring Harbor Laboratory Press, Cold Spring Harbor (NY), 3rd edn, 2015–2017.

2 M. Thaysen-Andersen, N. H. Packer and B. L. Schulz, Maturing Glycoproteomics Technologies Provide Unique Structural Insights into the N-glycoproteome and Its Regulation in Health and Disease, *Mol. Cell. Proteomics*, 2016, **15**(6), 1773–1790.

3 K. Alagesan, A. Everest-Dass and D. Kolarich, Isomeric Separation and Characterisation of Glycoconjugates, *Adv. Exp. Med. Biol.*, 2018, **1104**, 77–99.

4 R. Raghunathan, *et al.*, Proteomics, Glycomics, and Glycoproteomics of Matrisome Molecules, *Mol. Cell. Proteomics*, 2019, **18**(11), 2138–2148.

5 J. Zaia, Mass spectrometry and the emerging field of glycomics, *Chem. Biol.*, 2008, **15**(9), 881–892.

6 S. J. North, *et al.*, Mass spectrometry in the analysis of N-linked and O-linked glycans, *Curr. Opin. Struct. Biol.*, 2009, **19**(5), 498–506.

7 H. Geyer and R. Geyer, Strategies for analysis of glycoprotein glycosylation, *Biochim. Biophys. Acta*, 2006, **1764**(12), 1853–1869.

8 M. Wuhrer, *et al.*, Glycoproteomics based on tandem mass spectrometry of glycopeptides, *J. Chromatogr. B: Anal. Technol. Biomed. Life Sci.*, 2007, **849**(1–2), 115–128.

9 B. L. Parker, *et al.*, Terminal Galactosylation and Sialylation Switching on Membrane Glycoproteins upon TNF-Alpha-Induced Insulin Resistance in Adipocytes, *Mol. Cell. Proteomics*, 2016, **15**(1), 141–153.

10 R. Blazev, *et al.*, Integrated glycoproteomics identifies a role of N-glycosylation and galectin-1 on myogenesis and muscle development, *Mol. Cell. Proteomics*, 2020, DOI: 10.1074/mcp. RA120.002166.

11 P. Zhao, *et al.*, Virus-Receptor Interactions of Glycosylated SARS-CoV-2 Spike and Human ACE2 Receptor, *Cell Host Microbe*, 2020, **28**(4), 586–601.e6.

12 H. Hinneburg, *et al.*, The Art of Destruction: Optimizing Collision Energies in Quadrupole-Time of Flight (Q-TOF) Instruments for Glycopeptide-Based Glycoproteomics, *J. Am. Soc. Mass Spectrom.*, 2016, **27**(3), 507–519.

13 L. Dang, *et al.*, Recognition of Bisecting N-Glycans on Intact Glycopeptides by Two Characteristic Ions in Tandem Mass Spectra, *Anal. Chem.*, 2019, **91**(9), 5478–5482.

14 C. Pett, *et al.*, Effective Assignment of alpha2,3/alpha2,6-Sialic Acid Isomers by LC-MS/MS-Based Glycoproteomics, *Angew. Chem., Int. Ed.*, 2018, **57**(30), 9320–9324.

15 D. Kolarich, *et al.*, Comprehensive glyco-proteomic analysis of human alpha1-antitrypsin and its charge isoforms, *Proteomics*, 2006, **6**(11), 3369–3380.

16 D. Kolarich, *et al.*, Glycoproteomic characterization of butyrylcholinesterase

from human plasma, *Proteomics*, 2008, **8**(2), 254–263.

17 H. Hinneburg, *et al.*, Distinguishing N-acetylneuraminic acid linkage isomers on glycopeptides by ion mobility-mass spectrometry, *Chem. Commun.*, 2016, **52**(23), 4381–4384.

18 R. S. Glaskin, *et al.*, Construction of a Database of Collision Cross Section Values for Glycopeptides, Glycans, and Peptides Determined by IM-MS, *Anal. Chem.*, 2017, **89**(8), 4452–4460.

19 A. S. Gelb, *et al.*, Composition and charge state influence on the ion-neutral collision cross sections of protonated N-linked glycopeptides: an experimental and theoretical deconstruction of coulombic repulsion vs. charge solvation effects, *Analyst*, 2019, **144**(19), 5738–5747.

20 E. S. Ji, *et al.*, Isomer separation of sialylated O- and N-linked glycopeptides using reversed-phase LC-MS/MS at high temperature, *J. Chromatogr. B: Anal. Technol. Biomed. Life Sci.*, 2019, **1110–1111**, 101–107.

21 R. Zhu, *et al.*, Isomeric Separation of N-Glycopeptides Derived from Glycoproteins by Porous Graphitic Carbon (PGC) LC-MS/MS, *Anal. Chem.*, 2020, **92**(14), 9556–9565.

22 S. Yang, *et al.*, Optimization of O-GIG for O-Glycopeptide Characterization with Sialic Acid Linkage Determination, *Anal. Chem.*, 2020, **92**(16), 10946–10951.

23 K. R. Reiding, *et al.*, Neutrophil myeloperoxidase harbors distinct site-specific peculiarities in its glycosylation, *J. Biol. Chem.*, 2019, **294**(52), 20233–20245.

24 I. Loke, *et al.*, Paucimannose-Rich N-glycosylation of Spatiotemporally Regulated Human Neutrophil Elastase Modulates Its Immune Functions, *Mol. Cell. Proteomics*, 2017, **16**(8), 1507–1527.

25 I. Loke, N. H. Packer and M. Thaysen-Andersen, Complementary LC-MS/MS-Based N-Glycan, N-Glycopeptide, and

Intact N-Glycoprotein Profiling Reveals Unconventional Asn71-Glycosylation of Human Neutrophil Cathepsin G, *Biomolecules*, 2015, **5**(3), 1832–1854.

26 D. Kolarich, *et al.*, The minimum information required for a glycomics experiment (MIRAGE) project: improving the standards for reporting mass-spectrometry-based glycoanalytic data, *Mol. Cell. Proteomics*, 2013, **12**(4), 991–995.

27 W. S. York, *et al.*, MIRAGE: the minimum information required for a glycomics experiment, *Glycobiology*, 2014, **24**(5), 402–406.

28 W. B. Struwe, *et al.*, The minimum information required for a glycomics experiment (MIRAGE) project: sample preparation guidelines for reliable reporting of glycomics datasets, *Glycobiology*, 2016, **26**(9), 907–910.

29 M. P. Campbell, *et al.*, The minimum information required for a glycomics experiment (MIRAGE) project: LC guidelines, *Glycobiology*, 2019, **29**(5), 349–354.

30 A. Varki, *et al.*, Symbol Nomenclature for Graphical Representations of Glycans, *Glycobiology*, 2015, **25**(12), 1323–1324.

31 S. Neelamegham, *et al.*, Updates to the Symbol Nomenclature for Glycans guidelines, *Glycobiology*, 2019, **29**(9), 620–624.

32 K. F. Aoki-Kinoshita, *et al.*, The GlySpace Alliance: toward a collaborative global glycoinformatics community, *Glycobiology*, 2020, **30**(2), 70–71.

33 J. Schoberer and R. Strasser, Plant glyco-biotechnology, *Semin. Cell Dev. Biol.*, 2018, **80**, 133–141.

34 L. Van Landuyt, *et al.*, Customized protein glycosylation to improve biopharmaceutical function and targeting, *Curr. Opin. Biotechnol.*, 2019, **60**, 17–28.

35 R. R. Drake, *et al.*, MALDI Mass Spectrometry Imaging of N-Linked Glycans in Tissues, *Adv. Exp. Med. Biol.*, 2018, **1104**, 59–76.