Recent advances to help cope with the complexity of MS-based glycan analysis

Pegah R. Jalili, Ph.D.
Principal Scientist, Analytical R&D
MilliporeSigma, Saint Louis, MO

Kevin Ray, Ph.D.
Senior Manager Analytical R&D
MilliporeSigma, Saint Louis, MO

In a post-translational modification reaction called glycosylation, living organisms attach carbohydrate residues known as glycans to their proteins. The resulting glycoproteins fulfill numerous functions, including as structural components of cell membranes and mediating cell interactions. Some hormones (e.g., thyrotropin, hCG) and components of the immune system (immunoglobulins, interferons) are also glycoproteins.

Of course, glycosylation also occurs in expression systems that produce biological drugs, creating a complex spectrum of products that vary significantly with regards to *in vivo* stability, solubility, activity, safety, and efficacy. Extensive characterization is therefore essential to successfully develop biological drugs. Protein glycosylation is specifically mentioned in established technical guidelines e.g., ICH Q5E and Q6B and FDA's published guidance for industry titled "Development of Therapeutic Protein Biosimilars", so regulatory authorities tend to scrutinize glycosylation data thoroughly. However, the structural complexity of N-linked and O-linked glycans can make glycoprotein analysis challenging, so it is performed at several levels: intact protein, subunit, peptide, and glycan. This article presents an update of recent advances in mass spectrometry (MS) based methods for glycan analysis at the various levels.

Intact glycoprotein analysis

Analysis of intact glycoproteins can be performed as a simple, high-throughput method to screen glycoprofiles. Detected peaks in deconvoluted spectra with specific glycan mass shifts provide information about the glycoform composition and their relative contributions. Sample preparation is minimal and the data analysis straightforward. This approach works well for low-complexity monoclonal antibodies and glycoproteins but not so much for highly complex or highly heterogeneous molecules such as erythropoietin and etanercept. Reversed phase (RP) and size exclusion chromatography (SEC) coupled to electrospray ionization mass spectrometry (ESI-MS) are usually used. RP-HPLC offers high resolution and good chromatographic peak shape, but suffers from low recovery and sample adsorption of large molecules such as antibodies, and it often has to be optimized specifically for each molecule. SEC resolution is lower but produces good quality mass spectra with short separation times, and it can be used as a generic platform. In our laboratory we routinely use SEC-ESI-MS as a high-throughput and robust approach to support clone screening and media optimization experiments.

Analysis at subunit level

Middle down or subunit analysis further simplifies the characterization of glycoproteins. For antibodies, additional detail can be achieved by treating the glycoproteins with a reducing agent to generate the light chains and heavy chains, which can be evaluated separately. Recently, subunit analysis has been greatly improved by using a *Streptococcus pyogenes* enzyme (IdeS) that cleaves IgG antibodies specifically at a unique site below the hinge region. IdeS digestion, followed by a reduction step, yields three subunits of around 25 kDa, which helps to improve mass accuracy and resolution as well as subunit-specific localization of glycosylation. Other novel enzymes that have emerged since IdeS, such as SpeB, Kgp, and IgdE, are becoming more widely used.

Glycopeptide mapping

Bottom up peptide mapping, or glycopeptide analysis, enables profiling of individual glycosylation sites and other PTMs, which is essential for the comprehensive characterization of biological proteins. Often referred to as multi-attribute methods, they yield both qualitative and quantitative information about the biologic in a single analysis. Although valuable information

can be obtained, the often overwhelming amount of data analysis required complicates this technique. Sample preparation involves denaturation, disulfide bond reduction, alkylation, and digestion of proteins into suitably sized peptides by a highly specific protease such as trypsin. A combination of multiple proteases may be necessary to generate complete glycosylation site information. Following digestion, the peptides and glycopeptides are separated and analyzed by RP-LC-MS. Detection of low-abundant glycopeptides may be challenging due to the ion suppression of co-eluting glycopeptides with high abundance of non-glycosylated peptides for complex glycoproteins. This problem can be addressed by employing HILIC chromatography as a fractionation method to selectively isolate both O- and N-glycopeptides from non-glycosylated peptides followed by RP-LC-MS/MS analysis. The emergence of improved peptide ion fragmentation methods, including collision-induced dissociation for obtaining glycan composition information, and electron-transfer dissociation to generate fragments that can confirm the peptide amino acid sequence as well as the site of glycosylation, has resulted in the accurate assignment of glycan site occupancy.

Released glycan analysis

The analysis of released glycans is the most powerful and commonly used approach for glycan composition analysis. While sample preparation is time-consuming and data analysis can be complicated, it provides valuable glycosylation profile information as well as quantification and structural information. Typically, N-linked glycans are released by a specific enzyme, such as PNGaseF, and O-linked glycans by a chemical reaction. They are purified from peptides, fluorescently labeled to increase detection sensitivity and characterized by HPLC. The elution profile of labeled glycans is compared against a labeled dextran ladder and a retention library, yielding an estimation of the number of glucose units of each species, their abundance, and a prediction of the glycan structure. Released glycan analysis has been drastically improved by introduction of new digestion methods, labeling reagents, and purification methods. Recently, we introduced the PNGase Fast kit that can cut the digestion time from around 20 hours down to 15 minutes. The labeling technique is also improved by replacing the two conventional, most commonly used fluorescent tags, 2-aminobenzamide and 2-aminobenzoic acid, with procainamide or RapidFluor-MS®. This significantly enhances the ionization efficiency for LC-ESI-MS/MS identification while still allowing quantification by fluorescence. O-glycan release methods, both reductive and non-reductive, can suffer from the stepwise removal of one sugar residue at a time from the glycans reducing end (peeling). A new O-glycan sample preparation kit recently introduced by S-BIO claims to be non-toxic while significantly reducing the peeling of the glycans and cutting the sample preparation time down to five hours.