Maximize
glycan structural information

- Monosaccharide analysis
- Glycan analysis
- Glycosylation site profiling
- Intact glycoprotein profiling
Glycosylation is one of the most important post-translational modifications of eukaryotic cell proteins. Glycan-modified proteins are involved in a wide range of biological and physiological processes including recognition and regulatory functions, cellular communication, gene expression, cellular immunity, growth, and development.

Glycan functions are often dependent on the structure of the oligosaccharide. Oligosaccharides are covalently attached to proteins primarily through two structural motifs: attached to the amide group of an asparagine, referred to as “N-linked glycans,” or attached to the hydroxyl group on serine or threonine, referred to as “O-linked glycans”. Both types of glycans are investigated as biomarkers, in order to understand changes related to complex organelle development, and as part of therapeutic protein drug development with strong indication that efficacy is affected by glycosylation. Agencies such as the United States Food and Drug Administration (U.S. FDA) and the European Medicines Agency have published documents recommending that biopharmaceutical manufacturers demonstrate satisfactory programs for understanding, measuring, and controlling glycosylation in glycoprotein-based drugs, with recent updated draft guidance from the U.S. FDA for characterization of biosimilar protein therapies. This guidance suggests that the oligosaccharide content of glycoprotein products should be understood and monitored to ensure product consistency.

Many analytical approaches have been employed in the area of glycan and glycoproteomics research, which can be generally summarized under the following categories:

- Monosaccharide analysis
- Glycan analysis
- Glycosylation site profiling
- Intact glycoprotein profiling

These analyses are performed to determine the identity and quantity of the carbohydrates present, the post-translational modification site, and the multiple glycoforms of a protein that may be present. Choosing the appropriate analysis technique depends on experimental demands and instrumental capabilities. This handbook is intended to assist the decision-making process.
Carbohydrate: A generic term used interchangeably with sugar, oligosaccharide, or glycans. This term includes monosaccharides, oligosaccharides, and polysaccharides as well as derivatives of these compounds.

CID: collision-induced dissociation

CDG: congenital disorders of glycosylation

ESI: electrospray ionization

ETD: electron transfer dissociation

FD: fluorescence detection

Glycan: A generic term for any sugar or assembly of sugars, in free form or attached to other molecules such as proteins and lipids.

GU: glucose unit

HCD: higher-energy collisional dissociation

HILIC: hydrophilic interaction liquid chromatography

HPAE-PAD: high-performance anion-exchange chromatography with pulsed amperometric detection

HPLC: high-performance liquid chromatography

HR/AM: high-resolution accurate-mass

LC/MS: liquid chromatography mass spectrometry

Monosaccharides: The simplest form of a glycan. It cannot be hydrolyzed into simpler units. Building block of glycans. Common examples include glucose, fucose, galactose, glucosamine, sialic acids, etc.

MS: mass spectrometry

MS/MS: tandem mass spectrometry

MS": multi-stage fragmentation

Oligosaccharides: A glycan polymer, consisting of two or more monosaccharides.

PGC: porous graphitic carbon

RP: reversed-phase

WAX: weak anion exchange
Monosaccharides are usually thought to be non-ionic. However, they are weak acids, and under basic conditions, pH 12-14, the carbohydrates are either partially or completely ionized and thus can be separated by anion-exchange chromatography. This chromatographic approach is well suited to base-stable polymer anion-exchange columns exemplified by the Thermo Scientific™ Dionex™ CarboPac™ columns. By combining this high-resolution anion-exchange separation with electrochemical detection, a sensitive and specific analytical method, HPAE-PAD, is possible.

Samples are acid hydrolyzed to release the monosaccharides, lyophilized, dissolved in water, and then analyzed by HPAE-PAD. This direct analysis method easily resolves, and allows quantification of, individual glycan monosaccharides, providing not only the total glycosylation of a protein, but also the amounts of specific monosaccharides. For example, this analysis provides details about fucosylation, which can impact protein function and signaling, gives an indication of the presence of O-linked glycans by monitoring changes in galactosamine content, and determines mannose content as an indication of high-mannose N-linked glycans, which have been investigated as cancer biomarkers.
Charged Monosaccharide Analysis with HPAE-PAD

Glycoprotein sialylation has been shown to be critical to bioavailability, stability, metabolism, and immunogenicity of therapeutic proteins. Although over 50 forms of sialic acid have been identified, two forms of this carbohydrate are routinely determined: N-acetyllneuraminic acid (Neu5Ac) and N-glycolylneuraminic acid (Neu5Gc). Of these, Neu5Gc is generally not found in human proteins and is potentially immunogenic. After release, sialic acids may be determined by two liquid chromatographic methods: HPAE-PAD or derivatization followed by HPLC with fluorescence detection. To analyze samples without derivatization, HPAE-PAD may be used. A rapid method has been published to determine the two most abundant sialic acids, Neu5Ac and Neu5Gc, in glycoprotein hydrolyzates. Determination of sialic acids by derivatization followed by HPLC typically uses 1,2-diamino-4,5-methylenedioxybenzene dihydrochloride (DMB) to modify the sialic acid after release from the glycoprotein to provide a fluorescent product. This HPLC method allows separation of a wider variety of modified sialic acids.

D-Mannose-6-phosphate (M-6-P) is a terminal monosaccharide of N-linked oligosaccharides that is an important intermediate in glycoprotein production and, when incorporated in the glycoprotein’s final oligosaccharide, is needed for targeting and recognition on some lysosomal proteins. Genetic glycosylation defects or malfunctions in the synthesis or processing of protein-bound oligosaccharides are known collectively as CDG. Studies of M-6-P–containing glycoproteins and M-6-P receptors are active areas for medical research with the goals of understanding and hopefully treating or controlling these and other diseases. After release from proteins, M-6-P is determined by HPAE-PAD.
Glycan Analysis

Although determining monosaccharide content in glycoproteins provides significant information, knowledge of glycan structure is valuable. Glycan analysis enables monitoring of relative quantities of a particular set of glycans. Additionally it can provide valuable information regarding changes in glycan patterns or identify aberrant glycans. Due to the complexity of a glycan structure, structural elucidation is performed after release of glycans from a protein. The preferred form of release for N-linked glycans is by enzymatic means. This can be done by treating the N-linked glycan with peptide-N-glucosidase F (PNGaseF). PNGaseF releases most glycans except those that contain α1-3 linked fucose to the reducing terminal GlcNAc. In that case, the enzyme peptide-N-glucosidase A (PNGaseA) is used. Unfortunately, there are no enzymes comparable to PNGaseF that release O-linked glycans. Instead, the preferred approach for the release of O-linked glycans is the use of chemical methods such as β-elimination. Once released, these oligosaccharides, both neutral and charged, can be analyzed by several methods to provide detailed information.

Glycan Analysis with HPAE-PAD

HPAE-PAD is a well-established glycan analysis method that separates carbohydrates with specific interactions between the hydroxyl and carboxyl groups of glycans, which are charged at high pH, and the stationary phase. This separation resolves glycans based on charge, size, composition, isomers, and linkages, and provides a profile of the overall glycosylation and individual glycans. This method is capable of separating glycans based on sialic acid linkage, providing valuable information not only about the carbohydrate sequence of a glycan, but also the subtle linkage differences that may indicate disease states. With comparison to standards, this method can provide detailed information about released neutral and charged glycans with direct detection of glycans.9-10

Workflow for glycan analysis using HPAE-PAD
Glycan Analysis with HPLC-FD

Analysis of glycans by HPLC has gained significant interest in recent years. HPLC-based glycan analysis can be categorized by the chosen detection method and the detail in the information sought. Most commonly, fluorescence-based detectors or mass spectrometers are used for detection of HPLC-separated glycans. The following section will highlight the HPLC methodology in conjunction with fluorescence detection.

Glycans do not contain fluorophores and thus must be labeled with fluorescent tags prior to analysis. Labeling is performed at the reducing end of the glycan by reductive amination, with the most common labels being 2-aminobenzamide (2-AB) or 2-aminobenzoic acid (2-AA).11-13 HPLC-based methods have been applied with fluorescent-labeled oligosaccharides for the separation of sialylated and neutral glycans from various biological samples, recombinant therapeutic proteins, growth hormones, etc. The three most commonly used column chemistries are RP, WAX, and HILIC. HILIC columns are the most widely used for glycan analysis. These columns separate sialylated and neutral glycans in one chromatographic run, providing an advantage for rapid glycan profiling. However, HILIC columns don’t separate based on charge and isomers as compared to HPAE-PAD analysis — and reduction amination methods produce some desialylation of the released oligosaccharides. A typical workflow compares the elution of 2-AB-labeled glycans against a 2-AB-labeled dextran ladder. The individual peaks from the glycan profile are then assigned a GU based on comparison to the dextran ladder. This value is correlated to glycan structures that have been previously determined. Once potential structures are identified, they are confirmed with a series of exoglycosidase digestions.

Alternatively, if a mixture of known standards is available, these standards can be analyzed under the same chromatographic conditions as the samples. The number of peaks and their area under the curve can then be used for qualitative and quantitative estimation of the unknown N-linked glycans. However, if there are completely unknown glycans present, they can be identified by MS after LC separation.

Workflow for glycan analysis using HPLC-FD
Glycan Analysis with MS

Mass spectrometry has emerged as one of the most powerful tools for the structural elucidation of glycans due to its sensitivity, small sample requirement, and ability to analyze complex glycan mixtures. In addition to the characterization of the glycan sequence, MS provides branching patterns, substituent location, and quantitative information.

Glycan identification and characterization

Due to the weak acidity or basicity of most glycans, free glycans do not ionize efficiently. Therefore, it is necessary to derivatize glycans prior to MS analysis. Derivatization methods can be classified into two categories: labeling reducing ends and protecting functional groups. Reducing-end labeling is performed by reductive amination where the most common label is 2-AB. Protection of most or all functional groups is performed by permethylation.14-17

Reductively aminated glycans are suitable for HILIC-based separation prior to MS analysis.18-20 HILIC uses both glycan size and structure in separation, thus providing the additional benefit of separating structural isomers. HILIC can also separate native glycans.19,20 The Thermo Scientific™ GlycanPac™ AXH-1 column separates labeled and unlabeled glycans using its mixed-mode surface chemistry combining WAX and HILIC retention mechanisms. The WAX functionality provides retention and selectivity for negatively charged glycans while the HILIC mode facilitates the separation of glycans of the same charge according to their polarity and size. A PGC stationary phase is also employed for glycan separation as it can separate native and reductively aminated glycans.21,22 PGC is typically used for O-linked glycans. For reductively aminated glycans, the preferred label for PGC separation is 2-aminopyridine (PA). PGC also can separate structural isomers. For MS, the benefit of reducing-end tags is that they localize the charge at the reducing end of the glycan, resulting in predominantly X,Y, and Z fragmentation ions for CID or HCD fragmentation23—greatly simplifying the structure analysis of unknowns.

Alternatively, the protection of most or all functional groups can be selected as a derivatization method. The most common of these, permethylation, improves the MS ionization response of all free glycans equally. This enables simultaneous analysis of both neutral and acidic glycans. Furthermore, permethylation increases the occurrence of cross-ring and double glycosidic cleavages during fragmentation, which is important in assessing linkage and branching information. A workflow combing permethylation and MSn has been developed. This approach enables detailed determination of glycan linkages and branching information important for structural isomer differentiation.16,17,24 Permethylated glycans can also be separated prior to analysis by MS, employing either RP or PGC.

Data analysis is a critical part of MS-based glycan analysis workflows. SimGlycan® software from PREMIER Biosoft enables interpretation of MS/MS and MSn data generated on Thermo Scientific mass spectrometers.24-26

Workflow for glycan identification and characterization using mass spectrometry
Glycan quantification

Glycan quantification has witnessed a number of method developments in recent years. One powerful and widely used approach takes advantage of reductive amination to introduce stable isotope labels onto the glycans. Similar to the 2-AB label used in glycan characterization, the 2-AA label offers many of the same benefits for LC-MS analysis. However, it is also available commercially in stable, isotopically labeled forms, 2-12C6-AA and 2-13C6-AA, and can be used for relative quantification experiments.

In stable isotopic 2-AA labeling experiments, glycans are released from proteins by enzymatic means and differentially labeled in parallel with either the 2-12C6-AA or 2-13C6-AA fluorescent tag. After labeling, samples are combined so that subsequent processing steps can be performed on the combined samples, greatly reducing sample handling variability and resulting in more accurate quantification. The equimolar, differentially labeled, paired sample is then subjected to direct infusion or LC-MS-based relative quantitative analysis.

Labeling with reductive amination tags such as 2-AA provides a number of benefits for MS analysis: 1) enhanced ionization increases efficiency in the negative ion mode, 2) labeling can be undertaken under aqueous conditions with no prior sample work-up, and 3) fragmentation of the labeled glycan occurs in a predictable manner leading to abundant ions that can be assigned unambiguously. Generally, 2-AA quantification experiments are performed by LC-MS. Because the glycans are derivatized with isotopic labels, they behave similarly within the separation stage (coeluting at the same time) but can easily be differentiated by a mass spectrometer owing to the mass differences of the two isotope labels (6 Da between the 12C6 and 13C6 isotopic pair). The relative peak intensities of multiple isotopically distinct glycans are used to determine the average change in glycan abundance in the treated sample.

The unmatched resolution and mass accuracy afforded by Thermo Scientific™ Orbitrap™ technology provide discrimination between co-eluting isobaric ions, which is essential for correct 2-AA quantification. Additionally, since quantification is dependent upon precursor ions, CID, ETD, or HCD can be used for structural elucidation.

Data analysis of 2-AA experiments requires both glycan quantification and structural elucidation.

Quantification is performed at the MS level by comparing the intensities of the 2-12C6-AA- and 2-13C6-AA-labeled precursor ions at high resolution.

Workflow for glycan quantification using mass spectrometry
Glycosylation Site Profiling with MS

Glycans can also be profiled at the peptide level, such that relative quantification of specific glycosylation sites is achieved. Sample enrichment is central to the success of this workflow by reducing overall sample complexity and facilitating sensitive, accurate analysis. The enrichment steps can be carried out at the protein level, peptide level, or both levels. The enrichment can be targeted or universal, depending on the nature of information sought. Targeted enrichment may be performed to selectively isolate a certain subset of glycopeptides or glycoproteins on the basis of specific glycan structures (e.g., the use of titanium dioxide to isolate sialylated glycopeptides).

Multiple fragmentation mechanisms are required for this workflow. A combination of HCD or CID with ETD is used. ETD provides extensive fragmentation along the peptide backbone to enable peptide sequencing while preserving glycan attachments, allowing glycosylation site localization. CID or HCD provides information about the glycan composition.

In recent years, much emphasis has been placed on developing bioinformatics tools to simplify interpretation of glycopeptide data. The recent development of a novel software tool, Byonic™ by Protein Metrics™, alleviates many of the hurdles accompanying manual interpretation. Byonic software uses data from both types of fragmentation: HCD data to identify the sugar composition and the corresponding ETD data for peptide backbone information. The final result describes the peptide sequence, glycosylation site, and glycan composition.
Intact Glycoprotein Profiling with MS

Over the past few years, recombinant monoclonal antibodies (mAbs) have gained significant importance in diagnostic and therapeutic applications. Structural features of mAbs, such as assembly of light and heavy chains with disulfide bridges, are commonly known. However, the heterogeneity of antibodies, in part due to the variation of attached sugar moieties, requires thorough characterization to verify the correctness of the overall molecule and ensure a reproducible, safe, and effective biological drug.

Intact glycoprotein profiling by mass spectrometry is a good way to ascertain expected pattern and degree of glycosylation. Due to the inherent heterogeneity and variation of the attached glycan moieties, which divides the signal from a single core protein among its multiple glycoforms, intact glycoprotein profiling is best performed on a high-resolution accurate-mass hybrid Orbitrap mass spectrometer with ESI. One of the intrinsic advantages of ESI-MS is its ability to place multiple charges on particular species. The multiple charging effect makes possible the analysis of large intact proteins using conventional mass ranges. Mass spectrometers with the highest available resolution, mass accuracy, and scan speed are preferred to isotopically resolve peaks for precise molecular mass determination and quantification.

Intact glycoprotein profiling can be performed by either direct infusion or LC/MS. For very complex samples, where the total number of isoforms might be large, LC separation is typically required prior to MS analysis.

Thermo Scientific™ Protein Deconvolution software takes full advantage of the HR/AM data produced by Orbitrap-based mass spectrometers to deconvolute the mass spectrum and identify the degree of glycosylation.

Workflow for intact glycoprotein profiling using mass spectrometry
References