The proteomic analysis of glycosylation is uniquely challenging. The numerous and varied biological roles of protein-linked glycans have fueled a tremendous demand for technologies that enable rapid, in-depth structural examination of glycosylated proteins in complex biological systems. In turn, this demand has driven many innovations in wide ranging fields of bioanalytical science. This review will summarize key developments in glycoprotein separation and enrichment, glycoprotein proteolysis strategies, glycopeptide separation and enrichment, the role of mass measurement accuracy in glycopeptide detection, glycopeptide ion dissociation methods for MS/MS, and informatic tools for glycoproteomic analysis. In aggregate, this selection of topics serves to encapsulate the present status of MS-based analytical technologies for engaging the challenges of glycoproteomic analysis.

The functional diversity and structural heterogeneity of protein-linked carbohydrates simultaneously renders glycoproteins of tremendous biological interest and unique analytical intransigence. Protein glycosylation is known to figure prominently in the activity and stability of glycoproteins [1,2]. Moreover, metabolic perturbations associated with disease are known to exert influence on the compositions and structures of glycans attached to proteins [3,4]. In light of these considerations, protein glycosylation is of special interest with respect to the overall structure and function of glycoproteins, and is also a focal point of studies in search of disease biomarkers. From an analytical perspective, elucidation of protein glycosylation in complete molecular detail is distinctly challenging on multiple levels [5]. For example, the compositional and structural heterogeneity typically associated with glycosylation (e.g., glycan compositional heterogeneity at a given glycosylation site and glycan structural heterogeneity within a given glycan composition) can easily result in dozens of glycoforms for a single glycoprotein. In addition, the branched structures of glycans and their nontemplate-driven biosynthesis allow for a multitude of structural possibilities unparalleled by other biopolymers or post-translational modifications. Thus, the study of protein glycosylation has served to catalyze numerous analytical developments and innovations.

While many analytical approaches contribute to the study of protein glycosylation, MS has become a central tool in this area [6–11]. The MS-based analysis of glycoproteins may be conducted using several different techniques (Figure 1). For example, the methods of MS-based glycomics can be used to characterize glycans that have been chemically or enzymatically released from glycoproteins without an associated characterization or identification of the protein. Alternatively, the approaches of MS-based proteomics can be applied to identify glycoproteins without regard for the sites of glycosylation or the compositions and structures of the modifying glycans. While either general approach can provide a wealth of useful information, neither is able to provide the detailed molecular characterization of protein glycosylation in a site-specific manner. In order to address protein glycosylation with this level of resolution, the characterization of glycopeptides derived from proteolytic digestion is often
**Key terms**

**Glycoform:** A specific glycosylated form of a protein that is distinguished from other such forms by the composition(s), structure(s) or attachment sites of the modifying glycans.

**Ion mobility spectrometry:** Gas-phase analog of electrophoresis in which ions traverse a cell containing a buffer gas under the influence of an electric field.

**Glycoproteomics:** The identification of protein-modifying glycans, the proteins that they modify, and the site-specific protein connectivity of individual glycan structures.

employed. Such approaches maintain a covalent connectivity between the glycan and a vestige of the polypeptide until they are able to be characterized by MS and MS/MS. These methods are truly ‘glycoproteomic’ in that they fully engage the intersection of glycomics and proteomics, both methodologically and in terms of the type of information obtained.

Accordingly, this review will focus on MS-based site-specific glycoproteomic analysis conducted at the glycopeptide level, including accurate mass MS measurements, MS/MS techniques and ion mobility spectrometry (IMS). Approaches for sample preparation (e.g., glycoconjugate separation and enrichment methods, and glycoprotein proteolysis strategies) and glycoproteomic data analysis tools will also be discussed. As an exhaustive coverage of these topics is beyond the scope of this review, a number of excellent reviews and tutorials on glycoproteomics are suggested [12–18]. Instead, we will focus in this review on recent innovations in active areas of research and development.

**Glycoprotein separation & enrichment**

**Overview**

The first step towards MS-based glycoproteomic characterization is often the isolation or enrichment of the glycosylated fraction of proteins in a complex biological mixture. A number of methods have been developed for this purpose, many of which are applicable to glycoconjugate isolation not only at the protein level, but also at the glycopeptide level (to be discussed in a succeeding section of this review) [19]. These approaches continue to be refined with respect to their efficiency and selectivity. The following section will include a brief discussion of protein-level enrichment strategies for glycoproteomics, with emphasis on recent developments in three major areas: lectin affinity methods, boronate affinity methods and hybrid affinity methods.

**Lectin affinity methods**

One of the most common approaches for glycoprotein enrichment involves the use of lectin affinity techniques [20,21]. Lectins are carbohydrate-binding proteins that each exhibit characteristic glycan-binding specificity and affinity for a given target. For example, concanavalin A (ConA) and wheat germ agglutinin (WGA) are two lectins often used for glycoprotein enrichment. While ConA primarily recognizes α-d-mannose and α-d-glucose, WGA recognizes N-acetyl-d-glucosamine and also has a high affinity towards sialic acid. Owing to the distinct carbohydrate-binding specificities of lectins, single-lectin enrichment methods are quite useful when the goal is to target only those glycoproteins that exhibit a specific type of glycosylation. Multi-lectin columns have also been developed to allow for multispecific enrichment, which is useful for capturing glycoproteins that harbor any of a targeted set of glycan motifs.

Lectins are often immobilized on agarose- or silica-based materials. In general, agarose-based columns are applicable only for offline enrichment owing to the compressible nature of the agarose columns; conversely, silica-based columns allow for online LC–MS methods, which benefit from greater analytical reproducibility and stability over time [22]. Aside from agarose- or silica-based columns, lectin functionalized magnetic microparticles have been developed as another method of lectin-based enrichment. Lectin coupled magnetic microparticles have been successfully applied to the analysis of human serum for the purposes of glycoprotein profiling and biomarker discovery [23,24]. Use of magnetic particles allows for assay automation and reduced processing time, and is thus a promising platform for high-throughput analyses.

Lubman and coworkers made use of three single-lectin affinity columns to capture sialylated glycoproteins from the serum of pancreatic cancer patients [25]. The individual lectin columns were based on *Maackia amurensis* lectin, *Sambucus nigra* agglutinin and WGA. The enriched glycoprotein fractions were then subjected to LC fractionation, and finally the target glycoproteins were characterized. This technique was reported to provide both qualitative and quantitative insight on serum glycoproteome alterations associated with pancreatic cancer. The same research group also developed a multi-lectin column (ConA and WGA) to compare serum glycoproteome profiles characteristic of pancreatic cancer and chronic pancreatitis compared with normal human sera [26]. This combination of lectins successfully identified the biomarkers unique to the two diseases. Lectin affinity methods have been coupled to selected reaction monitoring-based MS to analyze aberrant protein glycosylation as a consequence of various diseases. As one example, the lectin phytohemagglutinin-l, was utilized to target multiantennary N-linked glycan structures known
to be associated with cancer [27]. Combining this lectin-affinity approach with selected reaction monitoring-based MS allowed quantitative comparison of the glycoprofles of both normal and cancerous cells. Such emerging approaches to biomarker discovery show promise for analytical application in the future, although extensive validation of potential biomarkers (e.g., through consideration of other potentially confounding disease states and through analysis of large numbers of clinical samples) is essential and remains an active area of research.

As mentioned above, the use of multiple lectins with varying glycan-recognition specificities can be applied in concert to achieve the broad capture of glycosylated proteins from biological mixtures [28]. Novotny and coworkers compared two different strategies for applying multi-lectin affinity: a sequential multi-lectin affinity approach involving four separate lectin columns, and a parallel multi-lectin affinity technique in which the same four lectins were applied in a single enrichment step [29]. Both methods were used to identify glycosylated proteins isolated from microliter volumes of human blood serum. Their results indicated that the sequential approach to lectin affinity chromatography offered superior performance compared with the multi-lectin affinity approach, as judged by the number of identified glycoproteins following enrichment. The sequential method allowed identification of 108 glycosylated proteins, while the parallel technique identified 67 glycosylated proteins.

In another application of lectin affinity chromatography, the plant cell wall glycoproteome of *Arabidopsis thaliana* was explored by Albenne and coworkers using both single-lectin and sequential multi-lectin affinity chromatography [30]. These researchers also explored the use of boronate affinity chromatography, which will be discussed in the succeeding section. For the sequential multi-lectin study, the flow-through of the first single-lectin column was loaded into the next single-lectin column, and the process was repeated for a total of four dimensions of separation. Interestingly, in these studies, a single ConA column was found to
be the most efficient method of enrichment. Sequential enrichment by the addition of WGA, Arctocarpus integrifolia lectin and peanut agglutinin resulted in only a modest increase in glycoproteome coverage. This highlights the possibility that additional dimensions of lectin affinity enrichment may not necessarily be advantageous, particularly in light of the additional effort required for sequential multidimensional strategies.

Boronate affinity methods

Another common approach for glycoprotein enrichment involves the use of boronate affinity [31]. Compared with lectin affinity methods, boron affinity methods are inherently less specific because such enrichments are based on the formation of covalent bonds to molecules containing cis-1,2-diol (or, less frequently, cis-1,3-diol) groups. While lectins can offer specificity for a given glycan class, boronate affinity can capture a more diverse pool of glycosylated proteins.

Recent advancements of note in the field of boronate affinity include the development or improvement of various boronic acid functionalized materials. For example, multiple types of boronic acid derivatized magnetic particles have been developed [32,33]. As mentioned above, the use of such magnetic particles provides several advantages, including convenient manipulation of the particles by the application of an external magnetic field, which in turn provides the ability to automate assays and reduce sample processing times. These qualities make the use of functionalized magnetic beads or nanoparticles a promising match for high-throughput analyses. Boronate affinity materials of this general type have been shown to be useful for glycoproteome enrichment from complex samples, such as human sera. Boronate affinity enrichment of glycoproteins has also been demonstrated using novel monolithic phases [34].

As mentioned previously, enrichment of glycoproteins in the plant cell wall of A. thaliana was performed by Albenne and coworkers using two different lectin affinity approaches [30]. A third method involving boronic acid chromatography was also applied. The use of boronic acid allowed for the identification of additional glycoproteins that were not detected by lectin affinity enrichment; however, the overall number of glycoproteins enriched by this method was significantly less compared with the lectin affinity method. This application again illustrates the potential compromise between diversity and depth of the captured, enriched glycoproteome.

Multi-affinity methods

Occasionally, lectin affinity and boronate affinity have been used in tandem for enrichment purposes. For example, boronic acid–lectin affinity chromatography was reported by Guttman and coworkers in 2007 [35]. Both boronic acid chromatography and lectin affinity chromatography were employed in a single separation column. This allowed for both the specific enrichment by lectin affinity while simultaneously affording a broader coverage of enriched glycoproteins by boronate affinity. Boronic acid-decorated lectins (BAD-lectins) have also been developed by Lu et al. [36]. This material was shown to furnish enhanced overall affinity towards glycoproteins owing to a synergy between lectin-mediated and boronate-mediated carbohydrate recognition (Figure 2). The authors prepared three different BAD-lectins: ConA, Sambucus nigra agglutinin and Aleuria aurantia lectin. The lectins were then coupled to magnetic nanoparticles (BAD-lectin@MNP’s) and applied to the enrichment of glycosylated components from an artificial protein mixture. This cocktail of proteins included the glycoproteins ribonuclease B, horseradish peroxidase, and lactoferrin, in addition to the nonglycosylated proteins myoglobin and bovine serum albumin. The mixture was prepared with a total glycoprotein:total nonglycoprotein ratio of 1:50. Under these conditions, the BAD-lectin@MNP’s were found to efficiently enrich glycoproteins according to the specificity of the applied lectin, but with greater affinity than the lectin alone. The authors also made use of BAD-lectin@MNP’s for glycopeptide-level enrichment, which will be discussed in a succeeding section.

Glycoprotein proteolysis strategies

Overview

Once a glycoproteomic fraction is isolated, the glycoproteins are most commonly subjected to enzymatic proteolysis in order to generate glycopeptides that are subsequently characterized by MS analysis [37]. Because glycosylated proteins often exhibit some degree of resistance to proteolysis, a number of approaches have been developed and refined in order to obtain efficient digestion. This is of critical importance, as the sequence coverage afforded by the chosen proteolysis method affects whether relevant and informative glycopeptides can be obtained. These proteolysis strategies can be roughly categorized as either specific or nonspecific digestion methods, although alternative and combined strategies will also be discussed below.

Specific enzymatic proteolysis

Some of the commonly used proteolytic enzymes are trypsin (cleaves peptide bonds C-terminal of arginine and lysine residues), lys C (cleaves peptide bonds C-terminal of lysine residues), glu C (cleaves peptide bonds C-terminal of aspartic or glutamic acid residues) and
arg C (cleaves peptide bonds C-terminal of arginine residues). Among these, trypsin is the most widely used owing to the efficiency of digestion, the convenient size distribution of the resulting peptides, and the presence of basic amino acid residues at the C-terminus of all resulting peptides (which enhances ionization efficiency for MS analysis and directs fragmentation in MS/MS) [38]. Unfortunately, missed tryptic cleavages are common near sites of glycosylation, leading to glycopeptides with relatively lengthy amino acid chains and, occasionally, multiply glycosylated peptides. A further complication is the generation of a mixture of glycosylated and nonglycosylated peptides, in which the glycosylated fraction is typically a stoichiometrically minor component. This often necessitates subsequent enrichment of the glycosylated fraction of digestion products as discussed in a succeeding section. An advantage of specific proteolysis is the finite number of possible peptide and glycopeptide products, each of which must have some common feature (e.g., C-terminal lysine). This limits the search space in subsequent data analysis steps. Nonetheless, alternatives or adjuncts to traditional specific protease digestion are available.

Nonspecific enzymatic proteolysis

In order to address some of the challenges in specific proteolysis identified above, nonspecific proteolysis strategies have found application to glycoprotein digestion. One agent for carrying out proteolysis with relatively little specificity is proteinase K, an endopeptidase that predominantly cleaves peptide bonds at the C-terminus of aliphatic and aromatic amino acid residues. Proteinase K digestion has been successfully applied to site-specific glycosylation analysis of a range of glycoproteins, typically producing glycopeptides with relatively small peptide tags [39–42].

Another approach to nonspecific proteolysis is based on digestion with pronase – a mixture of endopeptidases and exopeptidases that is capable of hydrolyzing...
practically any peptide bond \cite{43,44}. This brings the benefit of eliminating nonglycosylated peptides from the mixture, since these are hydrolyzed to amino acids and dipeptides. By contrast, glycosylation affords some protection from peptide bond hydrolysis near the glycosylation site, presumably owing to steric interference with protease activity. At appropriately chosen digestion times and conditions, the result is a mixture of glycopeptides with various peptide ‘footprints’ about the sites of glycosylation (Figure 3). Owing to the nonspecific nature of the digestion, each glycosylation site is represented by multiple peptide footprints that can be useful for confirmatory purposes. On the balance, the complexity of the resulting glycopeptide mixture is increased compared with specific proteolysis. In addition to being carried out in homogeneous solution, pronase can be used in an immobilized form, or applied ‘in-gel’ to glycoproteins separated by electrophoresis \cite{45–49}. These approaches can be highly complementary to specific proteolysis strategies.

Other proteolysis strategies

In other developments, some effort has been directed towards increasing the speed of glycoprotein digestion. In this regard, microwave-assisted digestion strategies show particular potential. For example, Fenselau and coworkers demonstrated microwave-accelerated acid-mediated digestion of glycoproteins to generate glycopeptides in minutes \cite{50}. Microwave irradiation has also been shown to accelerate protease digestions, for example, with pronase \cite{51}.

### Glycopeptide separation & enrichment

**Overview**

Following glycoprotein-level enrichment and subsequent proteolysis, the resulting pool of peptides and glycopeptides can be further separated prior to MS analysis \cite{52,53}. This may be done in addition to protein-level enrichment, or as a standalone measure. In either case, the goal of these separations is typically either to enrich the glycosylated fraction of peptides in a complex protease digestion mixture, or to separate individual glycopeptides for analysis. Although reversed-phase LC is the most common approach for peptide-level separation in proteomics and can be applied to glycoproteomics, here we focus on those approaches with unique applicability to glycopeptide characterization.

---

**Figure 3. Nonspecific proteolysis of a theoretical glycosylated protein, and some of the potential resulting glycopeptides.** The pool of glycopeptides resulting from nonspecific proteolysis represents both the heterogeneity of glycosylation (in this case, two potential glycans at two potential sites), and the heterogeneity of the attached peptide chain resulting from digestion with pronase. Gal: Galactose; GalNAc: N-acetylgalactosamine; GlcNAc: N-acetylglucosamine; Man: Mannose; NeuAc: N-acetylneuraminic acid. Reproduced with permission from \cite{94} © Wiley Periodicals (2012).
Affinity chromatography

As discussed above, boronate affinity methods are often utilized for glycoprotein-level enrichment; in addition, this technique can be applied post-proteolysis to enrich the glycosylated fraction of analytes at the peptide level. One novel means of applying boronate affinity to capture glycopeptides following protease digestion involves the use of boronic acid functionalized mesoporous silica. This materials offers many advantages, such as high surface area, large pore volume, and a narrow distribution of pore sizes. These characteristics have been credited with enabling increased binding affinity and improvement in detection limits compared with more traditional boronate-modified phases [54,55]. As mentioned previously, the use of boronate affinity in conjunction with lectin affinity has been applied to glycopeptide-level enrichments by Lu et al. using BAD-lectin@MNPs [36]. Application of these particles to the enrichment of glycopeptides from a trypptic digest of HeLa cell membrane proteins resulted in the identification of 295 unique glycopeptides with only a 6% overlap in glycopeptides captured by all three types of MNPs. This low degree of overlap illustrates the complementary specificities of the three lectins used to prepare the BAD-lectin@MNPs.

Hydrophilic interaction LC

Hydrophilic interaction LC (HILIC) incorporates the use of a hydrophilic stationary phase and a polar organic solvent mixed in some proportion with water. The separation mechanism is believed to be very similar to a liquid–liquid solvent extraction, with analytes being partitioned between the organic fraction and the aqueous fraction of the mobile phase, the latter of which is involved in hydrating a polar stationary phase [56]. Less polar analytes elute more quickly from the column, while hydrophilic analytes interact more with the hydration layer around the stationary phase. HILIC has been performed using cellulose and agarose beads for the SPE of oligosaccharides and glycopeptides from complex mixtures [57,58]. In addition, cotton-based HILIC SPE micropipette tips have been developed for the purification and enrichment of glycopeptides [59]. These can be prepared in-house from domestic cotton pads, as demonstrated by Wührer and coworkers (Figure 4). These SPE tips were found to be suitable for enrichment of glycopeptides from trypptic digests of transferrin and immunoglobulin G. Although most commonly applied to the offline enrichment of glycopeptides, HILIC has also found application to analytical separations of glycopeptides conducted online to MS [60,61].

Graphitized carbon LC

Graphitized carbon is widely utilized as a stationary phase for the separation and enrichment of carbohydrates and glycoconjugates, including glycopeptides, often with chromatographic resolution of species arising from carbohydrate isomerism [62–64]. For glycopeptides in particular, the retention order in graphitized carbon chromatography (GCC) has been found to depend on both the peptide sequence and the carbohydrate moiety [65]. The usefulness of graphitic-based stationary phases for glycopeptide separation has also been compared with that of reversed-phase media. For example, a chip-based activated graphitized carbon stationary phase was developed and compared with a C18 (i.e., reversed-phase) chip in the analysis of glycopeptides [66]. Both stationary phases were found to exhibit adequate reproducibility of retention time, but the activated graphitized carbon chip preferentially retained hydrophilic glycopeptides, which were not effectively retained by the C18 chip. Additionally, it was noted that hydrophobic glycopeptides were better retained on the C18 chips. With the use of both media in parallel experiments, both hydrophilic and hydrophobic glycopeptides were characterized. They also found that similar mobile phase gradients could be employed for both chips, further indicating that the two media could be used together to provide more complete coverage of glycopeptides.

GCC has also been incorporated into pipette tip microcolumns, which utilize a 1:1 ratio of graphitic carbon to activated charcoal (G/A, w/w) for enrichment purposes [67]. Termed the GA-ZipTip microcolumn, this form of SPE was shown to provide efficient isolation and enrichment of glycopeptides from glycoprotein digests. The microcolumn was also found to be useful for enrichment of glycopeptides from plasma. Different ratios of these materials were tested to see the ratio effect on isolation and enrichment of glycopeptides. The ratio was found to play a key role in efficiently enriching glycopeptides from digest mixtures, as higher graphitic carbon content led to enrichment of both glycopeptides and nonglycopeptides while higher activated charcoal content led to enrichment of only a few glycopeptides. A 1:1 mixture of the two phases was found to offer the most complete isolation of glycopeptides. GCC phases have also been used in tandem with reversed-phases for two-step fractionation of glycoprotein digests. For example, Larsen et al. separated glycosylated from nonglycosylated peptides in trypsin and proteinase K digests of ovalbumin using sequential application of C18 and graphitic carbon [68].

Ion pairing chromatography

Ion pairing chromatography makes use of ions in solution to pair with and thus neutralize charged analytes. Charged sites in peptides are known to form
neutral ion pairs with oppositely charged counter ions when in solution. When ion paired, peptides become significantly more hydrophobic. Glycopeptides with charged sites will also form ion pairs, but the glycan moiety still lends greater hydrophilicity of neutral ion paired glycopeptides compared with neutral ion paired peptides. When ion pairing is used with normal-phase LC, the amount of nonglycosylated peptides in the enriched glycopeptide fraction was found to be reduced or eliminated\cite{69,70}. Ion pairing has also been applied to HILIC separations of glycopeptides\cite{71}. Similar results were shown, as ion pairing increased the difference in hydrophilicity of glycopeptides and nonglycosylated peptides, allowing for improved isolation of the glycopeptide fraction.

**Ion mobility spectrometry**

IMS is a separation technique that is performed on gas-phase ions\cite{72–74}. Although IMS can be implemented in several distinct manners, the various types of IMS experiments have some shared characteristics. In general, ions under the influence of an electric field are separated as they traverse a drift cell containing a buffer gas. The IMS drift cell is often situated between an ionization source and a mass analyzer. Ions of differing mobility travel through the drift cell with distinct velocities and/or trajectories based on their interaction with both the electric field and the drift gas, resulting in a separation that is largely based on a size-to-charge ratio. While IMS methods hold great promise for characterization of carbohydrates and glycoconjugates

---

**Key term**

**Mass defect:** Mass lost in the form of energy when a nucleus is formed; this energy is equivalent to the nuclear binding energy of a given isotope.
Glycoproteomics: recent advances in MS-based glycopeptide analysis

At this writing there are very few examples of IMS applied to intact glycopeptides. Nevertheless, these examples highlight the potential of IMS to be of great analytical utility for glycopeptide analysis. For example, high-field asymmetric wave IMS (FAIMS) has been used to separate O-linked glycopeptide isomers that have identical peptide sequences but differ in their sites of O-GlcNAc attachment \[77\]. These isomers were found to coelute in reversed-phase LC but were shown to be partially resolved in FAIMS. This separation was confirmed by post-FAIMS electron transfer dissociation (ETD), which yielded some unique fragment ions depending on the site localization of the O-GlcNAc modification. These unique fragments were convincingly shown to originate from the species partially resolved by FAIMS. These results are promising, as they show that glycosylation site can affect the gas-phase conformation of glycopeptide ions such that these can be separated or differentiated by IMS techniques.

Another IMS method, traveling wave IMS (TWIMS), has been utilized for the separation of intact glycopeptides. TWIMS has been used to separate mixtures of peptides and glycopeptides from tryptic digests of glycoproteins \[78\]. These peptides and glycopeptides had different charge states and were studied by both positive and negative electrospray ionization. This application of IMS achieved separation of multiple charge states in drift space. IMS was also shown to help reduce chemical noise and improve detection of low abundance ions. Moreover, the glycosylated fraction of peptides was found to be sorted into a different

Figure 5. Application of traveling wave ion mobility spectrometry to discriminate epimeric glycopeptides. (A) Ion mobility spectrometry arrival time distributions for the glycosylated mucin peptides PTTTPITTTTTVTPTPTTAGTQ with α-linked GalNAc (blue trace, “species 20”) and α-linked GlcNAc (red trace, “species 19”) were partially resolved by ion mobility spectrometry. (B) Arrival time distributions of epimeric monosaccharide oxonium ions produced by collision-induced dissociation of glycopeptides 19 and 20 were also partially resolved. Vertical dashed lines represent drift time determined by half the full peak width at half height. GalNAc: N-acetylgalactosamine; GlcNAc: N-acetylglucosamine.

mass/mobility space compared with nonglycosylated peptides. TWIMS has also been used to discriminate between epimeric glycans and glycopeptides. For example, Both et al. studied two mucin-type glycopeptides that differed only in the identity of the single N-acetylated hexosamine (α-linked N-acetylglactosamine vs α-linked N-acetylgalactosamine) [79]. These species could be distinguished on the basis of both partial IMS resolution of the intact glycopeptides, as well as partial resolution of the released monosaccharide oxonium ions when collision-induced dissociation (CID) was conducted prior to IMS separation (Figure 5). Overall, IMS methods have exhibited significant potential to not only coarsely segregate glycopeptides from other biomolecules, but also to discriminate among isomeric glycopeptides with subtle structural differences often not resolvable by MS alone.

**Mass measurement accuracy in glycopeptide detection**

**Overview**

MS instruments capable of mass measurement with errors on the order of one part per million (ppm) or less have become increasingly available in recent years. Accordingly, distinction of different compound classes based on high accuracy monoisotopic mass measurements is an actively developing application area. Such applications are made possible by the unique mass defects of the elements, which thus encode elemental composition information into the exact masses of molecules [80]. The use of accurate mass measurements in glycopeptide analysis will be discussed below.

**Development & applications**

The use of highly accurate mass measurements has found important application to proteomics, where constraining the tolerances of a ‘match’ to low-ppm to sub-ppm levels can eliminate many false-positive possibilities and thus improve confidence in peptide and protein identification [81]. Notably, the mass defect space occupied by a given class of molecules is limited; for example, much of the mass scale is essentially devoid of elemental compositions that could plausibly correspond to a peptide [82]. In peptide analysis for protein identification, this feature has been exploited for improving the quality of protein database searches and the reliability of protein identifications [83–86]. More recently, the use of accurate mass to infer the presence of post-translational modifications has been noted, with particular emphasis on detection of phosphopeptides based on mass alone – a distinction made possible by the mass defect of phosphorus compared with other elements encountered in peptides and their relative proportions [87,88]. The application to similar strategies to recognize glycopeptides based on accurate mass measurements has also been of interest. Relative to unmodified peptides, glycopeptides contain a greater proportion of oxygen per unit mass. Given the relatively large mass defect of oxygen as compared to the other elements found in peptides, glycopeptides are therefore enriched in mass defect when compared to peptides of equivalent mass. A general discussion of mass accuracy requirements for confident glycopeptide assignment has been provided by Desaire and Hua, while the utilization of fractional mass distributions for glycopeptide identification has been discussed by Lehmann et al. [89,90].

**Glycopeptide ion dissociation methods**

**Overview**

The detection of putative glycopeptides by MS is usually followed by MS/MS in order to confirm the monosaccharide and amino acid composition, and to ascertain the monosaccharide connectivity and polypeptide sequence. While a glycopeptide MS/MS experiment would ideally reveal all of these features, the structural information actually obtained may or may not furnish this level of detail. Indeed, the information obtained in MS/MS experiments depends greatly upon the nature of the glycopeptide precursor ion (e.g., composition, charge carrier and charge state) and the nature of the applied dissociation method (e.g., vibrational activation, electron capture/electron transfer reactions, and electronic excitation) [92]. These considerations have been detailed in the context of glycopeptide analysis elsewhere [93,94]. Here, we will briefly discuss recent developments and emerging strategies in the area of glycopeptide MS/MS analysis.

**Use of complementary MS/MS methods**

A variety of MS/MS methods are commonly applied to biomolecule analysis [95]. Most prevalent among currently available techniques are ion dissociation methods based upon either collisional vibrational activation, or electron capture/electron transfer reactions [96–98]. Since these methods lead to structurally
informative fragmentation processes through fundamentally different mechanisms, the outcomes of these MS/MS methods often yield complementary information. This complementarity has proven useful for proteomic analysis [99,100]. Likewise, in the context of glycosylated peptide analysis, this orthogonality can be exploited in order to gain highly informative MS/MS data. It is now well known that, while vibrational activation/dissociation methods such as CID and infrared multiphoton dissociation (IRMPD) most readily lead to preferential cleavage of glycosidic bonds, electron capture dissociation and ETD lead to selective scission of N-Cα bonds along the polypeptide backbone [101–104]. As a result, a number of glycopeptide analysis workflows have been devised in order to exploit multiple complementary MS/MS methods and thus gather more detailed structural information on glycopeptide ions. For example, Alley et al. demonstrated a method based on LC–MS/MS with alternating CID and ETD in order to characterize tryptic O-glycopeptides from biological mixtures [106,107]. Analysis of N-glycopeptides was carried out by Cooper and coworkers using an LC–MS/MS method in which the observation of carbohydrate oxonium ions following CID was used to trigger ETD of the same precursor ion [108]. In this way, CID and ETD were not simply alternated during the entire LC–MS/MS run; rather, ETD was only carried out if the CID spectrum yielded evidence that a given precursor ion was a glycopeptide. CID neutral loss products can also be used to trigger ETD for only those analytes likely to be glycosylated [109]. Although not widely applied to glycopeptide analysis at this writing, we note that some previously reported hybrid MS/MS approaches have significant potential to extend the usefulness of MS/MS experiments that yield complementary structural information for glycopeptides [110–115].

Use of a single MS/MS method

While the combination of two MS/MS methods can be highly informative regarding overall glycopeptide
connectivity, significant effort has also been aimed at gaining extensive glycopeptide connectivity data using a single MS/MS technique. For example, photodissociation methods for MS/MS have achieved some success in this regard [116,117]. Irradiation of trapped ions with infrared photons to bring about IRMPD is one such method [118]. Studies by Adamson and Hakansson and Bindila et al. demonstrated the ability of IRMPD to yield significant information on both the glycan and peptide moieties of N-linked and O-linked glycopeptides, respectively [119,120]. Lebriella and coworkers also found IRMPD to effectively cleave both the carbohydrate and peptide groups of N-linked and O-linked glycopeptides, with their studies focused on the precursor ion characteristics (e.g., composition, charge carrier and charge polarity) that tended to predispose a given precursor ion to one type of cleavage or the other [121,122]. Irradiation of precursor ions with ultraviolet photons to result in ultraviolet photodissociation has also found increased application to biomolecule MS/MS analysis in recent years [123,124]. The work of Zhang and Reilly and Madsen and Brodbelt has established that ultraviolet photodissociation yields concurrent fragmentation information on both the glycan and the peptide for N-linked and O-linked glycopeptides, respectively [125,126]. In addition to photodissociation methods, certain CID-based strategies are being developed with the goal of gathering both the monosaccharide connectivity and amino acid sequence of glycopeptides. Such approaches are based on the observation that the applied collision energy dictates the dominant fragment ion types appearing in a CID spectrum. For example, Kolli and Dodds studied the energy-resolved CID behavior of several protonated N-glycopeptide ions [127]. This work demonstrated that, for the analytes studied, the lowest-energy dissociation pathways involved only cleavage of glycosidic bonds. At higher collision energies, secondary and tertiary product ions were generated that ultimately yielded significant information on the peptide backbone. This observation was exploited in order to obtain composite multi-energy CID spectra that were highly informative with respect to the overall connectivity of the glycopeptides (Figure 7). Other studies have made similar use of the energy-

![Figure 7. Multi-energy collision-induced dissociation of a triply protonated tryptic glycopeptide from Erythrina cristagalli lectin. (A) The energy-modulated CID spectrum contained product ions resulting from cleavage of the glycan (labeled with blue circles) at relatively low collision energy, while increased collision energies gave rise to fragmentation of the peptide backbone (red squares). (B) The composite spectrum provided substantial coverage of the glycopeptide connectivity. Fuc: Fucose; GlcNAc: N-acetylglucosamine; Man: Mannose; Xyl: Xylose. Reproduced with permission from [127] © Royal Society of Chemistry (2014). For color images please see www.future-science.com/doi/full/10.4155/BIO.14.272.](image-url)
resolved dissociation channels of glycopeptide ions [128–131]. Thus, CID is likely to gain broader use as a means of accessing peptide sequence ions in addition to the more frequently noted glycan fragmentation products.

**Informatic tools for glycoproteomics**

**Overview**

The automated interpretation of MS-based glycoproteomics data remains a significant bottleneck in the field, with the informative tools for glycomics and glycoproteomics lagging significantly behind those of genomics and proteomics in terms of maturity, robustness, general availability, and broad applicability. Nevertheless, the development of informative capabilities to complement MS-based glycoproteomic data sets has produced a number of useful tools that continue to evolve. In general, these can be classified into two types: those that are aimed at application to MS data, and those intended to address MS/MS spectra. In the following sections, a number of software tools available to aid in interpretation of glycopeptide data will be discussed. While this discussion is not exhaustive, the interested reader is referred to Dallas et al. for further discussion of informative tools in glycoproteomics [132].

**Tools for interpretation of MS data**

A number of algorithms have been developed in order to assign potential glycopeptide compositions to experimentally observed masses. One example of this is the GlycoMod tool, which enables prediction of glycopeptide masses based on user-defined parameters for potential glycan compositions coupled with in silico specific proteolysis of a user-selected protein sequence [133]. GlycoSpectrumScan, a similar tool for assigning potential glycopeptide compositions based on MS spectra, was developed as discussed by Deshpande et al. [134]. This tool makes use of user-supplied parameters describing potential N-glycan or O-glycan compositions and the amino acid sequences of peptides obtained following protease digestion. Desaire and coworkers have developed GlycoPep DB, a tool with the ability to perform a search for plausible N-glycopeptides [135]. A distinguishing characteristic of this search tool is the use of N-glycan structure databases such that the algorithm only returns putative N-glycopeptides with compositions that have been experimentally observed. GlycoX, a tool originally developed in the MATLAB environment, has many features in common with these other utilities [136]. However, the main distinguishing characteristic of the GlycoX package is its support for glycopeptides produced by nonspecific proteolysis with pronase.

**Tools for interpretation of MS/MS data**

Mass measurements alone, even when carried out with very high accuracy, are usually insufficient to unambiguously assign the amino acid and monosaccharide composition of a glycopeptide. Therefore, MS/MS analyses are required in order to confirm composition and to probe the amino acid sequence and/or monosaccharide connectivity of potential glycopeptides. One software tool to aid in the interpretation of glycopeptide fragmentation spectra is GlycoMiner, an algorithm intended to approach MS/MS interpretation in a fashion similar to a human expert [137]. Desaire and coworkers have extended their work on GlycoPep DB, resulting in the more recently reported GlycoPep Grader, GlycoPep Detector and GlycoPep Scorer tools [138–140]. Collectively, these utilities provide support for interpretation of CID and ETD data for both N-linked and O-linked glycopeptide analytes. Goldberg et al. developed an algorithm called Cartoonist, which aided in the annotation of N-glycan MS data [141]. An evolution of this tool resulted in Peptoonist, which makes use of both MS and MS/MS data to assist in the assignment of N-glycopeptides [142]. Similarly, Chandler et al. have described a tool called GlycoPeptideSearch, which recognizes carbohydrate oxonium ions, the complete glycan loss fragment (i.e., the bare peptide) and glycosidic cleavages to match unknowns to potential peptide–glycan pairs [143]. While the majority of glycopeptide MS/MS interpretation tools discussed here handle only CID data, a program devised by Mechref and coworkers called Glycopeptide Spectra Match integrates complementary CID and ETD data in order to arrive at possible glycopeptide compositions and structures [144]. Some other examples of informative packages designed to analyze glycopeptide MS/MS data include SimGlycan, GlyDB, GlypID and the Medici Integrator suite for protein N-glycosylation analysis [145–148]. As this broad assortment of tools continues to mature, an important consideration will be the ability of these platforms to handle MS/MS spectra across a wider variety of ion dissociation methods and instrument types in a truly integrated fashion, and to further dovetail into existing proteomic data analysis tools [107]. Further improvements in software for interpretation of glycopeptide MS/MS spectra will no doubt accompany advances in theoretical prediction of glycopeptide dissociation spectra and data mining [149,150].

**Conclusion & future perspective**

As reviewed here, the field of glycoproteomics represents not only an important area of biological inquiry, but also a focal point of analytical development across multiple disciplines. No doubt, sustained innovation in this area will continue to be driven by the growing
appreciation for glycoscience and the concomitant demand for increasingly powerful glycoanalytical tools. Particularly important in this regard will be MS-based analytical methods aimed at the characterization of covalent relationships between specific glycan structures and the proteins that they modify. We anticipate this will be a particularly active area of development and ongoing maturation in the coming years. The field of MS-based glycoproteomics will surely continue to enjoy ongoing advancement contributed by glycoscientists, separation scientists, mass spectrometrists, gas phase ion chemists, informaticians, and others. Indeed, sustained progress in these areas will be essential for the continued engagement of analytical challenges in the field of glycoproteomics.

Acknowledgements
The authors would like to thank D Rathore and Y Huang for helpful input on a draft of the paper.

Financial & competing interests disclosure
Funding from the University of Nebraska–Lincoln (NE, USA), which was provided in part by the Nebraska Tobacco Settlement Biomedical Research Development Fund, is gratefully acknowledged. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

Executive summary

| Background |
| Protein glycosylation is of central importance to the structure and function of many proteins with highly varied biological roles. |
| MS-based methods for glycoproteomic analysis should ideally seek to not only identify glycans and proteins independently, but to fully characterize their site-specific, covalent relationships. |

| Glycoprotein separation & enrichment |
| Affinity-based methods such as lectin affinity and boronate affinity offer either specificity or breadth, respectively, in capturing glycosylated fractions of proteins from mixtures. |
| These methods can be combined in various ways in order to improve the depth of glycoproteome enrichment. |

| Glycoprotein proteolysis strategies |
| Both specific and nonspecific enzymatic proteolysis can contribute to the generation of glycopeptides that are informative with regard to site-specific glycosylation. |
| These proteolysis strategies can be implemented in a variety of ways (e.g., with immobilized enzymes, in protein separation gels and with microwave irradiation). |

| Glycopeptide separation & enrichment |
| The glycosylated fraction of proteolytic peptides can be enriched by affinity methods, as well as solid-phase extraction based on hydrophilic interaction LC or graphitized carbon chromatography. |
| Furthermore, hydrophilic interaction LC, graphitized carbon chromatography and ion pairing chromatography are widely used for analytical separation of glycopeptides. |
| Ion mobility methods are expected to become more important in glycopeptide analysis in the coming years. |

| Mass measurement accuracy in glycopeptide detection |
| The unique mass defects of the elements allow different classes of compounds or peptides bearing various post-translational modifications to be tentatively identified based on accurate mass measurements alone. |
| This has been applied to distinguish between peptides and glycopeptides with no prior knowledge of the biological system. |

| Glycopeptide ion dissociation methods |
| A variety of ion fragmentation methods are used in conjunction with MS/MS, some of which can provide very different and highly complementary dissociation data for glycopeptide ions. |
| Some fragmentation methods are also capable of obtaining both glycan and peptide connectivity in a single experiment, such as photodissociation methods or energy-resolved collision-induced dissociation. |

| Informatic tools for glycoproteomics |
| Software utilities to aid in the interpretation of glycoproteomic data are currently far less mature than those available for proteomics. |
| Nevertheless, a variety of informatic tools are available, each with unique strengths and limitations for processing MS and/or MS/MS data. |

| Conclusion & future perspective |
| Analytical innovation in MS-based glycoproteomics will continue to be fueled by the increasing appreciation of glycans and their biological roles. |
| Sustained progress in this regard will require the expertise of scientists across multiple disciplines. |
References

Papers of special note have been highlighted as:
• of interest; •• of considerable interest


Thorough review of glycoprotein structural characterization methods.


Review of glycoprotein-level and glycopeptide-level enrichment strategies.

24 Loo D, Jones A, Hill MM. Lectin magnetic bead array for biomarker discovery. J. Proteome Res. 9(10), 5496–5500 (2010).
30 Zhang Y, Giboulot A, Zivy M, Valot B, Jamet E, Albenne C. Combining various strategies to increase the coverage of the plant cell wall glycoproteome. Phytochemistry 72(10), 1109–1123 (2011).
Bioanalysis
Review Kolli, Schumacher & Dodds


45 Clowers BH, Dodds ED, Seipert RR, Lebrilla CB. Site determination of protein glycosylation based on digestion with immobilized nonspecific proteases and Fourier transform ion cyclotron resonance mass spectrometry. J. Proteome Res. 6(10), 4032–4040 (2007).


Review of ion mobility methods and instrumentation.


Concise review of ion mobility applications in glycoanalytical sciences.


Thorough review on high accuracy mass measurements as applied to protein and peptide analysis.

Nefedov AV, Mitra I, Brasier AR, Sadygov RG. Examining troughs in the mass distribution of all theoretically possible tryptic peptides. J. Proteome Res. 10(9), 4150–4157 (2011).


Concise review on ion dissociation methods used in MS/MS.

Review on ion dissociation methods as applied to MS/MS of glycopeptides.


**Review on ion dissociation methods as applied to MS/MS of glycopeptides.**


Concise review of informatic tools for glycoproteomic analysis.


Zhang L, Reilly JP. Extracting both peptide sequence and glycan structural information by 157 nm photodissociation of N-linked glycopeptides. *J. Proteome Res.* 8(2), 734–742 (2009).


Concise review of informative tools for glycoproteomic analysis.


Ren JM, Rejtar T, Li L, Karger BL. N-Glycan structure annotation of glycopeptides using a linearized glycan structure database (GlyDB). *J. Proteome Res.* 6(8), 3162–3173 (2007).


