Protein Science (1993), *2*, 183–196. Cambridge University Press. Printed in the USA.

Selective identification and differentiation of *N*- and *O*-linked oligosaccharides in glycoproteins by liquid chromatography–mass spectrometry

STEVEN A. CARR, MICHAEL J. HUDDLESTON, AND MARK F. BEAN

Department of Physical and Structural Chemistry, SmithKline Beecham Pharmaceuticals, King of Prussia, Pennsylvania 19406

(RECEIVED August 7, 1992; REVISED MANUSCRIPT RECEIVED October 3, 1992)

Abstract

A mass spectrometry method has been developed for selective detection of glycopeptides at the low (≤25) picomole level during chromatography of glycoprotein digests and for differentiation of *O*-linked from *N*-linked oligosaccharides. The technique involves observation of diagnostic sugar oxonium-ion fragments, particularly the $HexNAc^+$ fragment at *m/z* 204, from collisionally excited glycopeptides. Collision-induced fragmentation can be accomplished in either of two regions of a triple quadrupole mass spectrometer equipped with an atmospheric pressure, electrospray (ES) ionization source. If collisions before the first quadrupole are chosen, it is possible to enhance formation of carbohydrate-related fragment ions without distorting the distribution of peptide and glycopeptide signals by increasing the collisional excitation potential only during that portion of each scan in which the low mass carbohydrate-related ions are being detected. This procedure, requiring only a single quadrupole instrument, identifies putative glycopeptide-containing fractions in the chromatogram but suffers from a lack of specificity in the case of co-eluting peptides. Increased specificity is obtained by selectively detecting only those parent ions that fragment in Q2, the second collision region of the triple quadrupole, to produce an ion at *m/z* 204 ($HexNAc^+$). Only $(M + H)^+$ ions of glycopeptides are observed in these liquid chromatography–electrospray tandem mass spectrometry (LC-ESMS/MS) "parent-scan" spectra. *N*-linked carbohydrates are differentiated from *O*-linked by LC-ESMS/MS analysis of the digested glycoprotein prior to and after selective removal of *N*-linked carbohydrates by peptide N:glycosidase F. These methods, which constitute the first liquid chromatography–mass spectrometry (LC-MS)-based strategies for selective identification of glycopeptides in complex mixtures, facilitate location and preparative fractionation of glycopeptides for further structural characterization. In addition, these techniques may be used to assess the compositional heterogeneity at specific attachment sites, and to define the sequence context of the attachment site in proteins of known sequence. The strategy is demonstrated for bovine fetuin, a 42-kDa glycoprotein containing three *N*-linked, and at least three *O*-linked carbohydrates. Over 90% of the fetuin protein sequence was also corroborated by these LC-ESMS studies.

Keywords: glycopeptides; glycoproteins; liquid chromatography–mass spectrometry; mass spectrometry; *N*-linked oligosaccharides; *O*-linked oligosaccharides; post-translational modification; tandem mass spectrometry

The carbohydrate chains of glycoproteins have begun to be widely acknowledged as having integral roles in the functional properties of glycoproteins (Rademacher et al., 1988; Cumming, 1991). Carbohydrates linked through an asparagine side chain (Asn- or *N*-linked) have been implicated in a wide variety of functions including modulation of biological activity, recognition between cells, interaction between host and pathogen, and distribution in tissues. The profound effects that carbohydrate moieties can have on other properties of glycoproteins such as solubility, antigenicity, immunogenicity, circulatory half-life, resistance to proteolysis, and thermal stability have come under increasing scrutiny with the desire of the biotechnology industry to produce these molecules for therapeutic use in humans (Cumming, 1991; Goochee et al., 1991). Although the biological and functional roles of carbohydrate attached through the side chain of serine or threonine (*O*-linked) are less well understood, it has

Reprint requests to: Steven A. Carr, Department of Physical and Structural Chemistry, UW2940, SmithKline Beecham Pharmaceuticals, King of Prussia, Pennsylvania 19406.

been reported that *O*-glycosylation is necessary for efficient biosynthesis and secretion of certain glycoproteins such as erythropoietin (Dube et al., 1988) and for prevention of antibody-based clearance of recombinant human granulocyte–macrophage colony-stimulating factor (Gribben et al., 1990), although others have reported no apparent effect of *O*-glycosylation (Matzuk et al., 1987).

For production of glycoproteins in heterologous cells, the structure of the protein, the specific cell type employed, and the bioprocess environment all influence which glycosylation sites are utilized and the range of carbohydrate structures found at each site. Glycoprotein bioactivity or clearance can be effected by modification or removal of specific oligosaccharide structural classes from specific attachment sites (Rademacher et al., 1988; Cumming, 1991; Goochee et al., 1991). Thus the precise structures of the glycoforms, as well as their protein sequence context, must be understood in order to gain a full appreciation of their structure/function relationship.

Until recently, the carbohydrate components of glycoproteins were not generally analyzed in the context of their sequence location, principally due to the experimental difficulty in identifying and isolating glycopeptides containing specific glycosylation sites. Instead, carbohydrate has been released in toto by endoglycosidases, hydrazinolysis, or reductive elimination (the latter two methods being destructive of the protein), thereby losing the sequence context of the glycan structures. A number of techniques have recently evolved for identification and preparative fractionation at the subnanomole level of *N*-linked glycopeptides in chromatographically separated complex digests of glycoproteins. The comparative mapping methods rely on comparison of the high-performance liquid chromatography–ultraviolet (HPLC-UV) traces or total ion current traces recorded during liquid chromatography–mass spectrometry (LC-MS) of a proteolytic digest of a glycoprotein before and after treatment with endoglycosidases to identify putative *N*-linked glycopeptides (Carr & Roberts, 1986; Hemling et al., 1990; Barr et al., 1991). We have used these techniques to characterize the *N*-linked oligosaccharides in a variety of glycoproteins including tissue plasminogen activator expressed in mammalian and insect cell lines, pre-S2 hepatitis B surface antigen expressed in yeast, and in soluble CD4 (Carr et al., 1991b). Although these techniques are highly successful for glycoproteins containing relatively few attachment sites, they become increasingly less reliable as the size of the glycoprotein and the number of glycosylation sites increase. A second problem is that *O*-linked carbohydrates cannot be analyzed in this fashion, as no universal *O*-glycosidase yet exists. Another promising approach for identifying glycopeptides in glycoprotein digests involves preparative fractionation of digests by HPLC and screening of the fractions by a lectin-based dot-blot assay on polyvinylidine difluoride membrane (Hsi et al., 1991). Resin-bound lectins may also be employed to isolate glycopeptides from digests of glycoproteins (Hsi et al., 1991). A range of lectins must be employed in these experiments to ensure that all of the common structural types of carbohydrates will be detected (Haselbeck et al., 1990; Hsi et al., 1991; Kobata & Endo, 1992). Glycopeptides may also be identified by LC-MS based upon the usual presence of a heterogeneous population of oligosaccharides at a given attachment site. On reversed-phase HPLC these glycoforms partially separate, with the more heavily glycosylated glycoforms eluting first (Hemling et al., 1990). In a plot of m/z versus time, glycopeptides can be identified by the appearance of a diagonal ladder of peaks, in contrast to the peptide species that show up as vertical lines (Ling et al., 1991). Similarly, glycopeptides may be selectively identified by searching the m/z data domain for peak pairs, differing in the mass of various carbohydrate units, and noting the locations where they maximize in the chromatogram (Carr et al., 1991a). The latter studies employed liquid chromatography–electrospray mass spectrometry (LC-ESMS). Electrospray (ES), in contrast to other mass spectrometry (MS) techniques such as fast atom bombardment MS, produces multiply (as opposed to singly) charged molecular species if the analyte contains more than one possible site for proton attachment (Fenn et al., 1989).

Here we present an approach utilizing LC-ESMS and liquid chromatography–tandem MS (LC-ESMS/MS) for selective detection of both *N*- and *O*-linked glycopeptides at the low picomole level in digests of glycoproteins. The glycoform population at each site may be assessed by the relative contributions of the various molecular weight-related species to the ES mass spectra. *N*-linked glycopeptides may be distinguished from *O*-linked glycopeptides by LC-ESMS/MS analysis of the digested glycoprotein prior to, and after, selective removal of *N*-linked carbohydrates by an endoglycosidase. These techniques facilitate isolation of glycopeptides for further characterization of the carbohydrate entities. The strategy is demonstrated for bovine fetuin (Dziegielewska et al., 1990), a 42-kDa glycoprotein containing three *N*-linked and at least three *O*-linked carbohydrates.

Results

The procedure for locating and differentiating *N*- and *O*-linked carbohydrates in glycoproteins by LC-ESMS/MS is illustrated in Figure 1. The glycoprotein (>75 pmol) is cleaved enzymatically (e.g., trypsin) or chemically (e.g., cyanogen bromide), usually after reduction and alkylation, to produce a complex mixture of peptides and glycopeptides. A corresponding set of experiments may also be done on glycoprotein that has been desialylated chemically or using a neuraminidase. This helps to reduce heterogeneity of the carbohydrates and thereby improves detectability of the glycopeptides. The resulting mixture is then analyzed in three separate experiments, requiring ca. 25 pmol of digest per experiment. In the first experi-

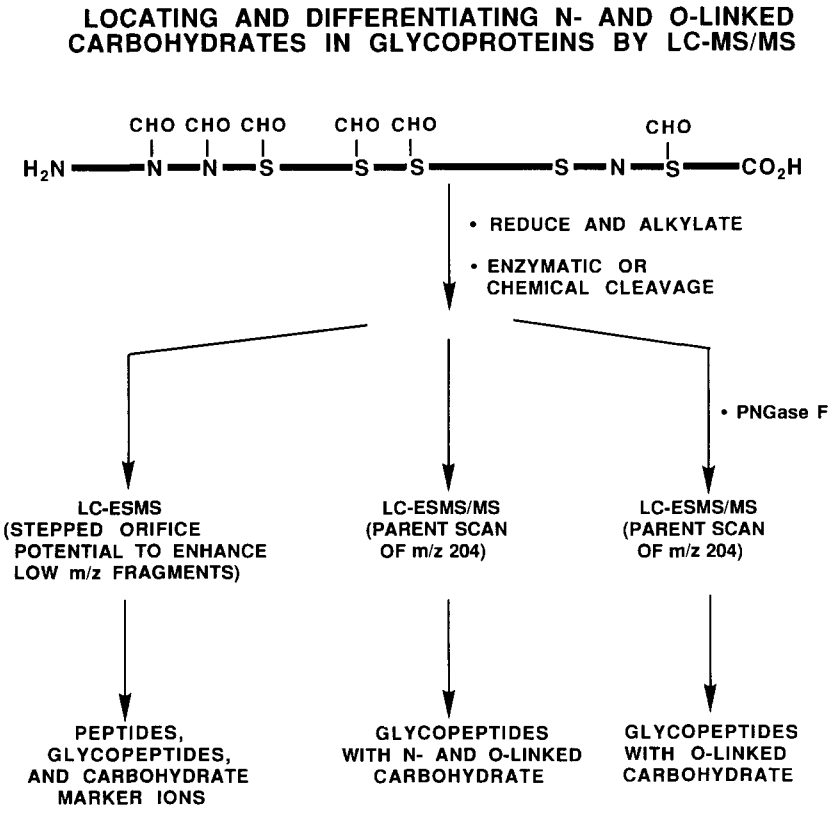

Fig. 1. Strategy for locating and differentiating *N*- and *O*-linked carbohydrates in glycoproteins by liquid chromatography–electrospray mass spectrometry (LC-ESMS) and liquid chromatography–electrospray tandem mass spectrometry (LC-ESMS/MS). Peptides are also mapped as part of the strategy.

ment, the sample is analyzed by LC-ESMS with a modified scanning protocol that enhances the production of low mass carbohydrate-specific marker ions. Data from this experiment are used to map the peptide portion of the protein and to indicate the presence of glycopeptides. In the second experiment, the sample is analyzed by LC-ESMS/MS to specifically and selectively identify glycopeptides of both *N*- and *O*-linked structural classes. In the third experiment, the sample is enzymatically digested with peptide N:glycosidase F (PNGase F) to release the *N*-linked oligosaccharides, and the sample is again analyzed by LC-ESMS/MS to identify glycopeptide signals that are due to the remaining *O*-linked carbohydrates. These experiments and the manner in which glycosylation sites and compositions of glycoforms at specific attachment sites are assessed are discussed in detail below.

Localization of glycopeptides in chromatograms by production of carbohydrate-specific marker ions during LC-MS

In the first LC-ESMS experiment, the collisional excitation potential in the region between the sampling orifice (OR) and the first mass analyzing quadrupole (Q1) of the triple quadrupole is stepped or ramped to enhance production of low mass carbohydrate-specific ions (see below) in the otherwise normal mass spectra. Without enhanced collisional excitation, the carbohydrate-specific marker ions may be weak or absent. The presence of these ions in a mass spectrum indicates that one or more of the components entering the mass spectrometer at that point in time is a glycopeptide. On the Sciex API-III triple quadrupole, the collisional excitation potential in the region prior to Q1 is controlled by the OR potential, which is sometimes referred to as the declustering potential. In the scanned or stepped OR method, the OR potential is maintained at a low, constant value (e.g., ca. 65 V) from ca. m/z 500 to m/z 2,400. This potential is sufficient for normal solvent declustering of peptide and glycopeptide molecular ions, and it is typical of the potential used on this lens for LC-ESMS. Below m/z 500, the OR potential is increased, under data system control, to high values (e.g., $\geq$120 V). Because the high collisional excitation potential is being applied prior to any mass separation, ions of all m/z are collisionally excited. However, the higher collision potential is applied only during that period of the mass analysis scan (after mass separation has occurred in Q1) when the low mass ions are being recorded. Thus, as the low mass ions are being recorded during each mass analysis scan, higher mass parent ions are being collisionally dissociated, and any low mass ($m/z < 400$) fragments produced are being detected. Conversely, as the higher mass ions are being recorded, the OR potential is relatively low, and therefore these ions are not affected. This approach is a modification of that originally proposed by Conboy (Conboy & Henion, 1991; Conboy, 1992) in

which a moderately high excitatory potential was employed across the scanned mass range. Stepping the collisional excitation potential eliminates two problems associated with the earlier approach: (1) observation of artifactual lower molecular weight glycoforms (lacking one or more nonreducing terminal sugars) that formed by collision-induced dissociations (CID), and (2) shifting of the parent ion distributions of peptides and glycopeptides to lower charge state, higher *m/z*. The latter problem is acute for glycopeptides (particularly oligomannose structures attached to relatively short peptide chains), which can have substantial mass but may exhibit low (+1 or +2) charge states because they have few available sites of protonation. Removal of charge from an already high mass glycopeptide may shift it out of the *m/z* range of the quadrupole mass analyzer.

A variety of carbohydrate-specific marker ions are observed including *m/z* 366 (Hex-HexNAc$^+$), *m/z* 292 (NeuAc$^+$), *m/z* 204 (HexNAc$^+$), *m/z* 163 (Hex$^+$), and *m/z* 147 (dHex$^+$) (Huddleston et al., 1991). Here, and throughout the text, Hex is used to refer to Man or Gal, HexNAc to GlcNAc or GalNAc, and dHex to Fuc or Xyl (or other isobaric sugars). In previous studies we determined that the fragments of *m/z* 204 and *m/z* 366 are produced by most structural types of both *N*- and *O*-linked carbohydrates. In the absence of outer-chain HexNAc or Hex-HexNAc structures they can be formed from the inner core linker structures -HexNAc-HexNAc- or -Hex-HexNAc- that are present in most *N*- and *O*-linked carbohydrates, respectively. We have found that the *m/z* 204 ion, although often yielding a somewhat lower absolute ion yield than the *m/z* 366, is, in general, a better generic indicator for *N*- and *O*-linked glycopeptides, particularly when the glycopeptide is of the oligomannose class (Huddleston et al., 1991).

The locations in the total ion current (TIC) trace (the mass spectrometer's analog of the UV absorbance trace) at which the carbohydrate-specific marker ions maximize are readily determined by plotting the reconstructed ion current trace for one or more of the marker ions. The TIC trace (*m/z* 150–2,000) and a reconstructed ion current (RIC) trace for *m/z* 204 from an LC-ESMS analysis (using the stepped orifice technique described) of 25 pmol of a trypsin/Asp-N digest of fetuin are shown in Figure 2A,B, respectively. The RIC trace of *m/z* 366 was very similar in appearance to the RIC trace of *m/z* 204 (data not shown). The summed mass spectra over the chromatographic peak eluting at ca. 30 min in Figure 2B,C illustrate the ca. twofold enhancement of the *m/z* 204 and 366 ion signal abundances in the stepped-OR experiment (Fig. 3B) relative to these ions in the "normal" LC-ESMS experiment in which the OR potential is held at a constant, relatively low value (Fig. 3A). Equally significant is the substantial reduction in the background from low *m/z* cluster ions (e.g., *m/z* 272 and *m/z* 394), further improving the signal-to-noise ratio for the marker ions in the stepped- or scanned-OR data. It is important to note that this experiment requires only a single quadrupole instrument, so that localization of glycopeptides in the chromatographic analysis is feasible without a triple quadrupole tandem mass analyzer.

While the presence of carbohydrate-specific marker ions is a strong indication that glycopeptides (either *N*- or *O*-linked) may be present, these data do not indicate which parent ions in the spectra gave rise to the carbohydrate fragment ions. For example, the LC-ESMS spectra for the glycopeptide-containing peaks eluting at ca. 30 min exhibit at least seven components (some of which show more than one charge state), and it is not possible to determine from these data which peaks are due to glycopeptides, although the component of M_r 3,218 (3^+ = 1,073.9; 2^+ = 1,610.1; Fig. 3A,B) does not correspond to a predicted tryptic or Asp-N fragment from the amino acid sequence of fetuin (Fig. 4). All of the other signals in the *m/z* range 550–1,650 are assignable to predicted peptide fragments of fetuin (Fig. 4). Previously we have used an approach for detecting glycopeptides in the LC-MS data of protein digests that involves examining the mass spectra for peaks separated by the in-chain *m/z* of carbohydrate units (Carr et al., 1990; Hemling et al., 1990). These peaks represent the parents of glycoforms of the glycopeptide having the identical peptide sequence but differing in the carbohydrate structures present, e.g., 365 (singly charged) and 182.5 (doubly charged) for glycoforms differing by Hex-HexNAc. However, when the pendant carbohydrates are relatively homogeneous, as is the case with fetuin, this procedure cannot be used. In addition, for reasons that are not presently understood, several of the glycopeptides (e.g., Ser^{253}- and Asn^{158}-containing glycopeptides; Fig. 2B) yield only weak response in the stepped or ramped OR experiments, making it possible that they would be overlooked.

Selective detection of glycopeptides

Glycopeptides may be selectively detected in the presence of peptides by scanning Q1 while monitoring for formation of specific carbohydrate marker ions (discussed above) produced by CID in Q2, the normal collision region of the triple quadrupole mass spectrometer. Although the selected marker ion is the ion detected, the spectrum corresponds to all parents that have decomposed to yield this fragment. This tandem MS experiment is referred to as a parent ion scan, and it is the second step of the strategy outlined in Figure 1. Because the goal is to detect both *N*- and *O*-linked glycopeptides, it is important to choose a marker ion that is generic to both types of carbohydrate linkages. As noted above, while both *m/z* 204 and *m/z* 366 may be employed, we have found that the *m/z* 204 signal is the most reliable in the parent ion scan mode. The total ion current trace for the LC-ESMS/MS parent

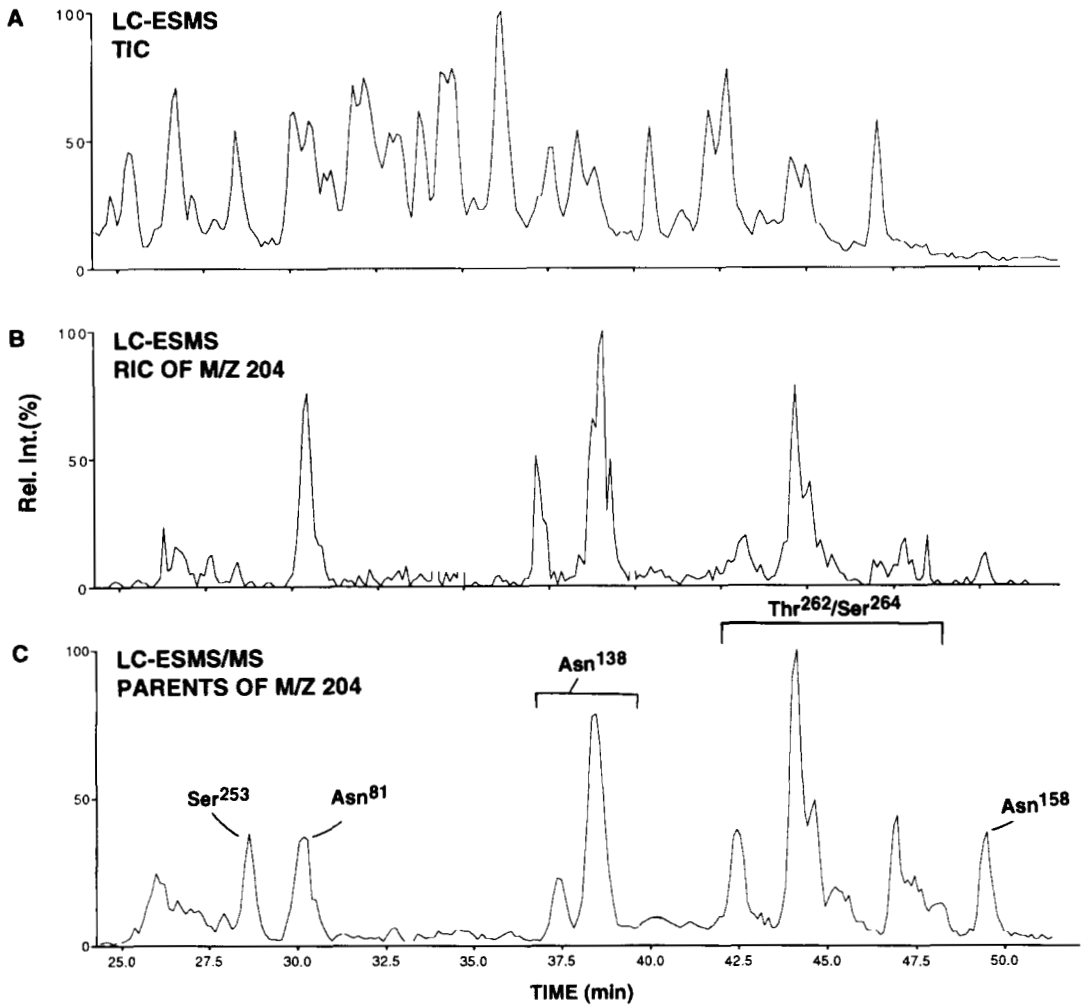

Fig. 2. A: LC-ESMS total ion current (TIC) trace with the orifice voltage ramped from 120 V at *m/z* 150 to 65 V at *m/z* 500; from *m/z* 500 to *m/z* 2,000 the orifice voltage was held constant at 65 V. **B:** Reconstructed ion chromatogram (RIC) of *m/z* 204 (HexNAc+) from the ramped orifice LC-ESMS data. **C:** LC-ESMS/MS, parent ion scan of *m/z* 204 ($HexNAc^+$), TIC trace.

ion scan experiment on a 25-pmol injection of the same tryptic/Asp-N digest of fetuin is shown in Figure 2C. More than 12 *m/z* 204-responsive peaks are readily apparent, and several regions appear to contain small amounts of poorly resolved components (e.g, 25–28 min; Fig. 2C). Examination of the mass spectra for each of these peaks indicated that the *m/z* 204 response was indeed arising via fragmentation of an *N*- or *O*-linked glycopeptide from fetuin; peptide interferences were essentially absent in these mass spectra. Glycopeptides containing *N*-linked carbohydrates were distinguished from glycopeptides containing *O*-linked carbohydrates by using the same MS technique following selective removal of *N*-linked carbohydrates with PNGase F (see Fig. 6 and Discussion, below).

The selectivity for detection of glycopeptides is evident in Figure 3C, which is the parent ion tandem mass spectrum of the components eluting at ca. 30 min. The peaks at *m/z* 1,074.0 and 951.5 correspond to the 3^+ charge state of glycopeptide Asp^{75}-Arg^{85}, in which Asn^{81} has a carbohydrate with the composition $(Hex\text{-}HexNAc)_3Hex_3\text{-}HexNAc_2$ attached (see section below on assignment of attachment site and glycoform composition). The spectra obtained for a sample that had not been previously neuraminidase treated indicated the presence of three sialic acid moieties (data not shown). Tetrasialylated carbohydrates, if present, gave signals below our detection limit (Green et al., 1988; Cumming et al., 1989). These data are consistent with the presence of triantennary complex carbohydrates at Asn^{81}, in agreement with the literature (Townsend et al., 1988; Rice et al., 1990). It is important to note that the 2^+ charge state of the glycopeptide (*m/z* 1,610.1), which gave a strong signal in the LC-ESMS data (Fig. 3A,B), is not evident in the parent ion scan data (Fig. 3C). The glycopeptide signals detected by parent ion scanning are usually shifted to higher charge state, sug-

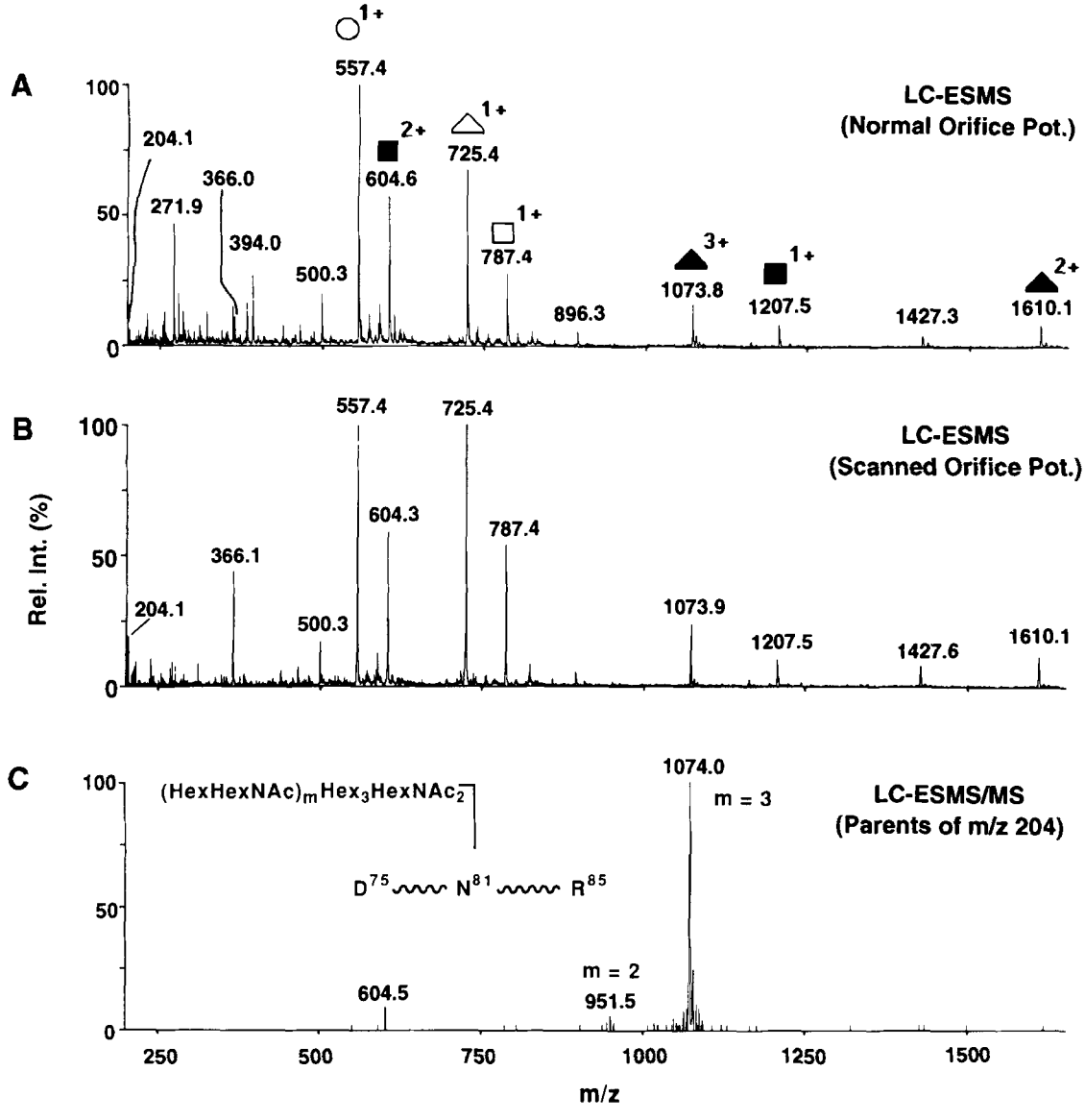

Fig. 3. Electrospray mass spectra of the glycopeptide eluting at ca. 30 min in the TIC traces shown in Figure 2. **A:** LC-ESMS using normal orifice voltage (65 V). Symbols identify a number of the parent ions present. **B:** LC-ESMS performed with the orifice voltage ramped from 120 V at *m/z* 150 to 65 V at *m/z* 500; from *m/z* 500 to *m/z* 2,000 the orifice voltage was held constant at 65 V. **C:** LC-ESMS/MS, parent ion scan of *m/z* 204 ($HexNAc^+$). The measured M_r of the Asn^{81} glycopeptide (C) is 3,219.0 (3,219.1, calculated).

gesting that the highest observed charge states of the glycopeptides more readily fragment to produce the inner core $GlcNAc^+$ ion of *m/z* 204.

The average molecular mass of the glycopeptide (or peptide) can be calculated from the spectra based on the observed masses for more than one defined charge state of the molecule (Covey et al., 1988; Fenn et al., 1989). For example, Figure 3B reveals peaks at *m/z* 1,073.9 and *m/z* 1,610.1. These can be shown to correspond to a triply and doubly protonated molecule with an M_r of $(3 \times 1{,}073.9) - 3 = 3{,}218.7$ or $(2 \times 1{,}610.1) - 2 = 3{,}218.2$. In instances where only one charge state is observed, the *m/z* differences between related ion signals (e.g., glycoforms, adducts, etc.) may be used to determine the charge state. For example, only a single charge state is apparent for the two glycopeptide peaks detected at *m/z* 1,074.0 and *m/z* 951.5 in Figure 3C. However, the *m/z* difference of 122.5 between these two signals can easily be assigned to the expected mass difference of a Hex-HexNAc moiety (365.5) if the charge state, *z*, is 3; thus the charge state is determined to be 3+ even in the absence of the 2+ or 4+ ion signals.

Two additional *N*-linked glycopeptides containing the Asn^{138} and Asn^{158} attachment sites were identified from

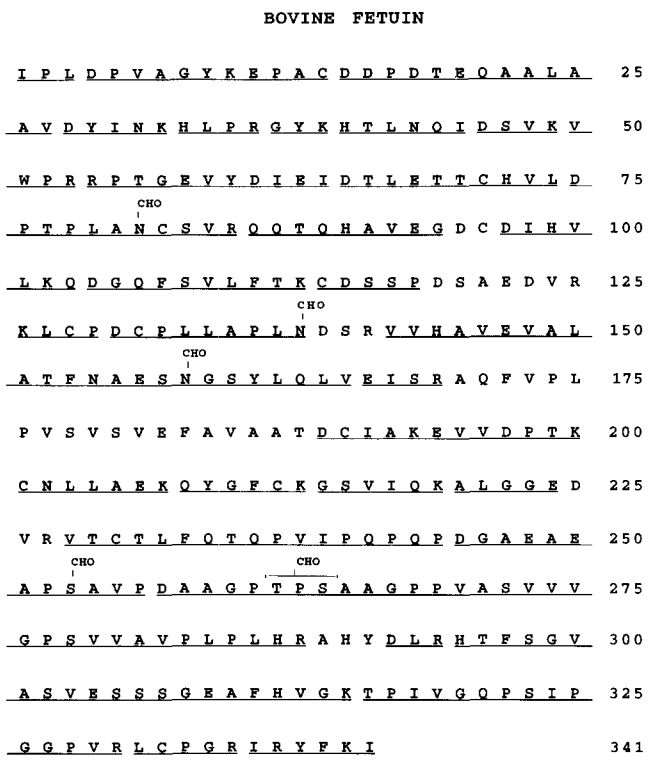

Fig. 4. The amino acid sequence for bovine fetuin, showing the glycosylation sites identified by ESMS. Underlined amino acid sequences were mapped by ESMS.

the parent ion scan data of the chromatographic peaks eluting at ca. 38 and 49 min (Fig. 2C); the corresponding ES mass spectra are shown in Figure 5B and C, respectively. Heterogeneity of the carbohydrate attached to Asn^{138} is apparent from the peaks at *m/z* 1,001.4 and *m/z* 879.6 that differ by 121.8, the in-chain *m/z* of the 3^+ charge state of Hex-HexNAc. The ES mass spectrum of the sialylated Asn^{138}-glycopeptide (obtained in a separate LC-ESMS/MS experiment on a sample of fetuin that had not been neuraminidase treated) exhibits peaks at *m/z* 1,292.6 and *m/z* 1,073.8 (Fig. 5A), which differ by 218.8, the in-chain *m/z* of the 3^+ charge state of a NeuAc-Hex-HexNAc unit (theoretical = 218.9). These data are consistent with the presence of both bi- and triantennary complex oligosaccharides at Asn^{138} (see section on site and glycoform assessment, below). In contrast, the carbohydrate attached to Asn^{158} appears to be homogeneous, consisting of only triantennary complex. Interestingly, the expected trypsin/Asp-N fragment, Val^{142} to Arg^{169}, was not observed (Fig. 4). Unexpected cleavage on the N-terminal side of Glu^{166} produced the observed glycopeptide Val^{142}-Val^{165}. The degree of sialylation at this site could not be assessed because glycopeptides containing this site were not detected in the LC-ESMS analyses prior to neuraminidase treatment. The peak eluting at ca. 37.4 min (Fig. 2C) that disappears upon PNGase F treatment (see below) yields a parent ion mass spectrum (data not shown) consistent with Pro^{129}-Asn^{138} (Fig. 4). Other *m/z* 204-responsive signals evident in the LC-ESMS/MS analysis, such as the diffuse peaks in the region 25–27.5 min and 46.7–48 min, were able to be unambiguously assigned only after their identities as *N*- or *O*-linked glycopeptides were made apparent by the experiments described below.

Differentiation of N- from O-linked oligosaccharides in glycopeptides

Glycopeptides containing *O*-linked oligosaccharides are selectively detected by LC-ESMS/MS analysis of an aliquot of the digest that has had the *N*-linked oligosaccharides released using PNGase F (Fig. 1). This is illustrated in Figure 6, which compares the total ion current traces for the LC-ESMS/MS analyses in parent ion scan mode for a trypsin/Asp-N digest of reduced and alkylated fetuin prior to (top) and after (bottom) PNGase F digestion. Many of the *m/z* 204-responsive signals are ablated after PNGase F treatment, indicating that these signals correspond to *N*-linked glycopeptides. This evidence corroborates the assignments of *N*-linked glycopeptides discussed above. The significant peaks in the total ion chromatogram that remain after PNGase F digestion correspond to *O*-linked glycopeptides. The mass spectra from the *m/z* 204 parent ion scan data of three of these *O*-linked glycopeptides are shown in Figure 7. The chromatographic peak eluting at ca. 26 min (Fig. 6B) corresponds to the glycopeptide Asp^{245}-Pro^{256} in which Ser^{253} has an *O*-linked glycopeptide with a composition of $(Hex\text{-}HexNAc)_m$ where $m = 1$ (major) or 2 (Fig. 7A). A smaller glycopeptide, Glu^{250}-Pro^{256}, containing the Ser^{253} glycosylation site elutes ca. 2 min earlier in the chromatogram (Fig. 6B). Samples that had not been neuraminidase treated exhibited molecular ions consistent with the composition $(NeuAc)_l(Hex\text{-}HexNAc)_m$ where $l = 0$, 1, or 2 for $m = 1$, and $l = 2$ for $m = 2$ (data not shown).

The chromatographic peaks eluting between 36 and 45 min are all *O*-linked glycopeptides containing the potential glycosylation sites Thr^{262}, Ser^{264}, Ser^{272}, and Ser^{278}. The parent ion scan mass spectra of the major peaks eluting at ca. 37 and 40 min are shown in Figure 7B and C, respectively. Both peaks have the identical peptide sequence Asp^{257}-Arg^{288} (Fig. 4) but differ in the number of Hex-HexNAc units attached. A smaller glycopeptide, Asp^{257}-Leu^{286}, containing these same glycosylation sites elutes at ca. 43.5 min (Fig. 6). The more heavily glycosylated forms elute earlier, as expected, from the reversed-phase separation being performed. However, the extent to which the glycoforms separate is highly variable and probably depends on the extent to which the carbohydrate interferes with specific interactions of the peptide portion of the glycopeptide with the column material. For example, the glycopeptide with an apparent carbohydrate com-

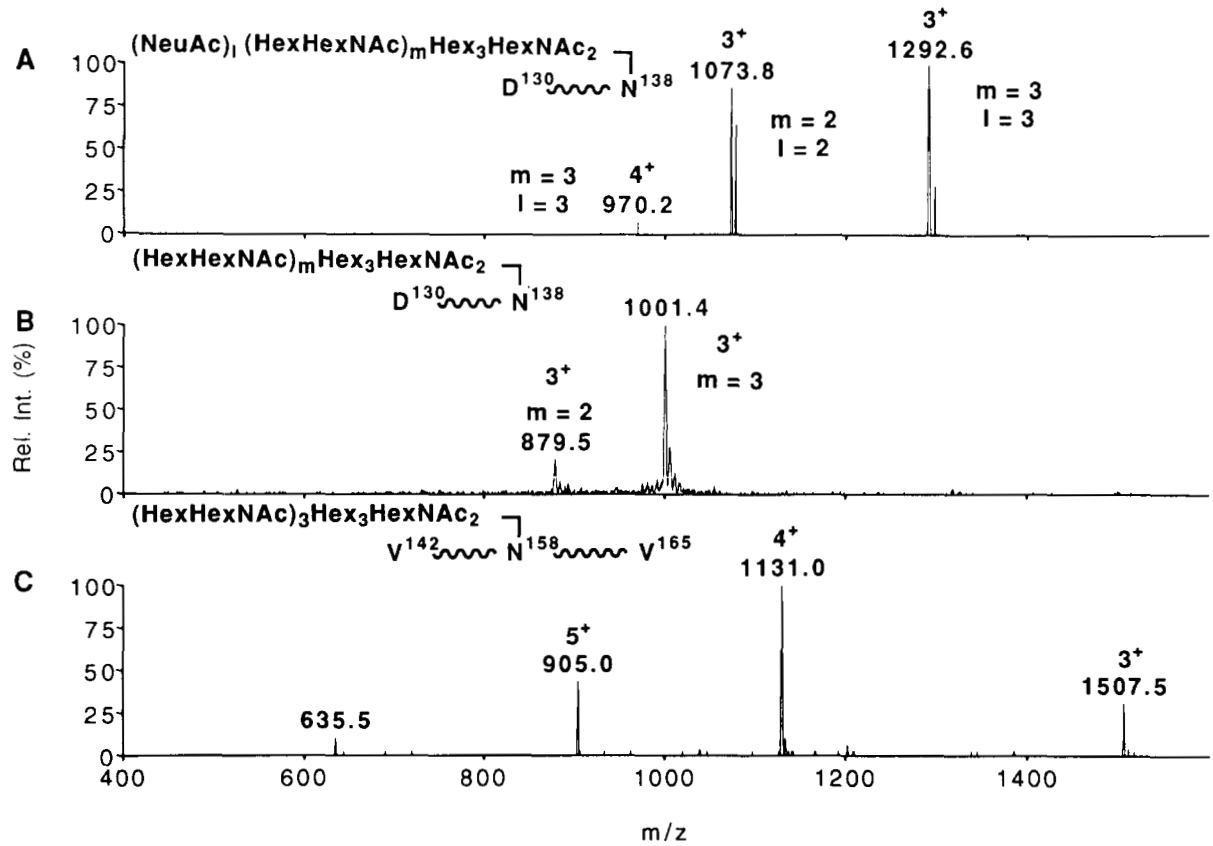

Fig. 5. LC-ESMS/MS, parents of *m/z* 204, mass spectra of *N*-linked glycopeptides. **A:** Complex carbohydrate showing bi- and trisialylated glycoforms attached to Asn^{138} (M_r meas. = 3,218.4 and 3,874.8, respectively; M_r calc. = 3,218.4 and 3,875.7, respectively). These data are from a separate LC-ESMS/MS experiment than B and C. **B:** Asn^{138}-containing glycopeptide, neuraminidase treated (peak at ca. 38 min; Fig. 2C), supports the bi- and triantennary structures (M_r meas. = 2,635.6 and 3,001.2, respectively; M_r calc. = 2,636.5 and 3,001.9, respectively). **C:** Parent ion spectrum indicates homogeneity of the glycan attached to Asn^{158} (peak at ca. 49 min; Fig. 2C). A possible triantennary complex carbohydrate is observed (M_r meas. = 4,520.0, calc. = 4,520.5).

position of (Hex-HexNAc)$_3$ (Fig. 7B) elutes ca. 2 min earlier than the glycoform with a composition of (Hex-HexNAc)$_2$ (Fig. 6). However, the glycoform with composition (Hex-HexNAc)$_1$ elutes only ca. 30 s later as a shoulder on the (Hex-HexNAc)$_2$ glycoform peak in the total ion current traces (Fig. 6). The mass spectrum of this latter peak (Fig. 7C) is shown as a composite of these two closely eluting components. In samples not treated with neuraminidase, the number of NeuAc residues was equivalent to the number of Hex-HexNAc units (data not shown). Assignment of Thr^{262} and Ser^{264} as the glycosylation sites is based primarily on literature precedent (Spiro & Bhoyroo, 1974; Nilsson et al., 1979; Edge & Spiro, 1987; Medzihradszky et al., 1990). A weak spectrum for the glycopeptide Asp^{257}-Ser^{272} with 1 and 2 Hex-HexNAc units (data not shown) was observed eluting at ca. 23.4 min (Fig. 6). Signals corresponding to this peptide with three Hex-HexNAc units attached were not observed. This suggests that either this species is present at levels that are below our detection limit, or that a fourth site of *O*-glycosylation is located at Ser^{278}. Unfortunately, peptides or glycopeptides that could be used to further localize the specific attachment sites were not observed under the cleavage conditions used. Collision-induced decomposition of all of the parent ion peaks shown in Figure 7C produced very strong product ion mass spectra on the triple quadrupole that could be used to partially sequence the peptide portion of the glycopeptide. Unfortunately the carbohydrate moieties were readily cleaved from the attachment site Ser and Thr residues, which reverted to their unmodified forms prior to production of the daughter ions detected (data not shown). Thus, it was not possible to locate the attachment sites by tandem MS of this glycopeptide.

Assignment of attachment sites and glycoform populations

Once a glycopeptide has been located and identified as either *N*- or *O*-linked using the procedures described

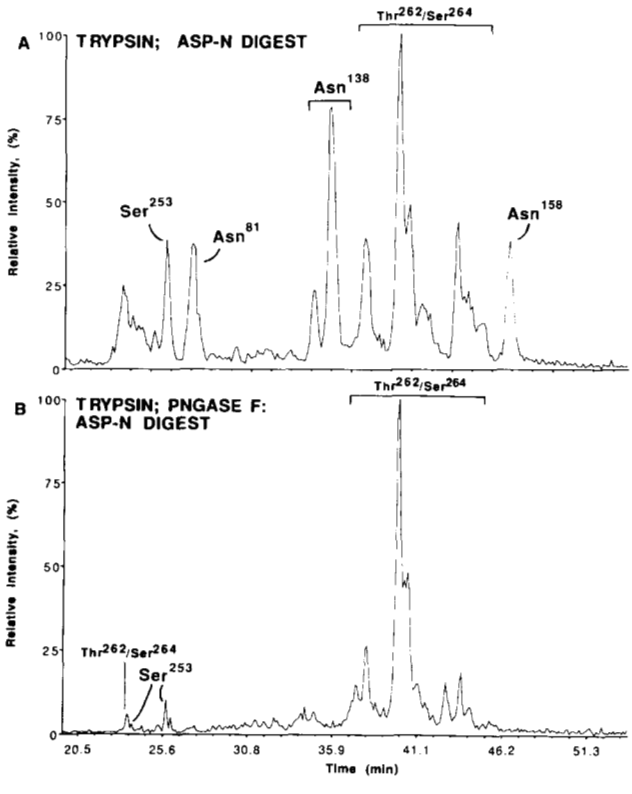

Fig. 6. Total ion current traces of a proteolytic digest of fetuin showing differences before and after treatment with PNGase F. Data are from an LC-ESMS/MS, parents of *m/z* 204, experiment. All significant peaks correspond to *O*-linked glycopeptides intact to be detected; *N*-linked glycopeptides are no longer observed. Peak labels indicate the glycosylation site(s) contained in the observed glycopeptides.

above, there are several methods for determining the specific glycosylation site used and for assessing the compositional heterogeneity of the glycoforms at that specific site. In the first method, candidate masses for the peptide containing the attachment site are derived by subtracting the in-chain masses of typical glycoprotein-associated carbohydrates from the M_r of each glycopeptide established by LC-ESMS or LC-ESMS/MS. The choice of carbohydrate masses is guided by the known range of structures typically found in glycoproteins expressed in the given cell system. In practice, the limits on the number of Hex (in-chain chemical average mass, $i = 162.14$), HexNAc ($i = 203.20$), dHex ($i = 146.14$), and NeuAc ($i = 291.26$) residues used to calculate candidate carbohydrate masses are set wider than anticipated in order to include unusual structures, if present. The protein sequence is then searched by computer for peptide sequences that match any of the candidate molecular weights for the attachment site peptides. Assignments are made based on how well the masses fit (must match to within 1 Da of expected), the presence of expected cleavage sites at the N- and C-termini of the candidate peptide (based on the proteolytic enzymes used), and the presence in the candidate peptide of either the necessary *N*- or *O*-linked attachment site residue. In the case of *N*-linked glycopeptides, the list of candidate peptides may be substantially reduced based on the requirement that the peptide contain the canonical sequence for *N*-linked glycosylation: -Asn-X-Ser/Thr- (where X = any amino acid except proline).

The *N*-linked glycopeptides discussed above will be used to illustrate this process. The M_r of the three *N*-linked glycopeptides identified in the neuraminidase-treated sample were determined to be 3,218.7 (Fig. 3), 3,001.2 (Fig. 5B), and 4,520.0 (Fig. 5C). The in-chain masses of typical *N*-linked carbohydrate moieties (e.g., 1,623.4 for the biantennary glycoform $(Hex\text{-}HexNAc)_2Hex_3HexNAc_2$ and 1,988.7 for the triantennary glycoform $(Hex\text{-}HexNAc)_3Hex_3HexNAc_2$) are subtracted from the determined M_r of the glycopeptides to obtain the list of masses for the candidate glycosylation-site peptides. For example, two of the possible candidate masses for the peptide portion of the glycopeptide of M_r 3,218.7 in Figure 3C are 1,230.0 and 1,595.3 (3,218.7 − 1,988.7 and 3,218.7 − 1,623.4, respectively). We calculated (using a computer program) that a combined proteolytic digest of fetuin using trypsin and Asp-N would produce the peptides Asp^{75}-Arg^{85}, Asp^{130}-Asn^{138}, and Val^{142}-Arg^{169}, containing potential *N*-linked glycosylation sites at Asn^{81}, Asn^{138}, and Asn^{158}, respectively (Fig. 4). The average molecular masses of these peptides are 1,230.4, 1,013.2, and 3,017.4, respectively. The match of 1,230.0 with the predicted value of 1,230.4 is very good (within 0.5 Da), indicating that the glycopeptide of M_r 3,218.7 corresponds to Asp^{75}-Arg^{85}. The carbohydrate portion, which must be attached to Asn^{81} based on the presence of only this single consensus site in the peptide, has a molecular mass consistent with that of a triantennary oligosaccharide.

Similarly, the glycopeptide of M_r 3,001.2 yields 1,012.5 and 1,377.8 for two of the possible masses for the glycosylation site peptide. Only the candidate peptide of M_r = 1,012.5 matches within 1 Da of a reasonable peptide, Asp^{130}-Asn^{138} (based on the proteases used and the presence of the consensus sequence), which has a calculated M_r = 1,013.2. The mass difference between the peptide mass and that of the glycopeptide (1,988.7) indicates that a carbohydrate with a composition consistent with that of a triantennary structure is attached to Asn^{138}. Subtraction of a wide range of in-chain carbohydrate masses from the glycopeptide with M_r = 4,520 failed to yield one of the three peptide masses predicted based on cleavage at Lys, Arg, and Asp. Therefore each of the candidate peptide masses was searched by computer against the entire sequence of fetuin without regard for the particular enzymes employed. One of the predicted masses, 2,531.3 (i.e., 4,520.0 − 1,988.7), matched the peptide Val^{142}-Val^{165}. This was deemed a likely candidate based on the fact that the peptide would form by a cleavage at a primary site on the N-terminal end (Arg^{141}) and a known secondary site for Asp-N on the C-terminal end (i.e.,

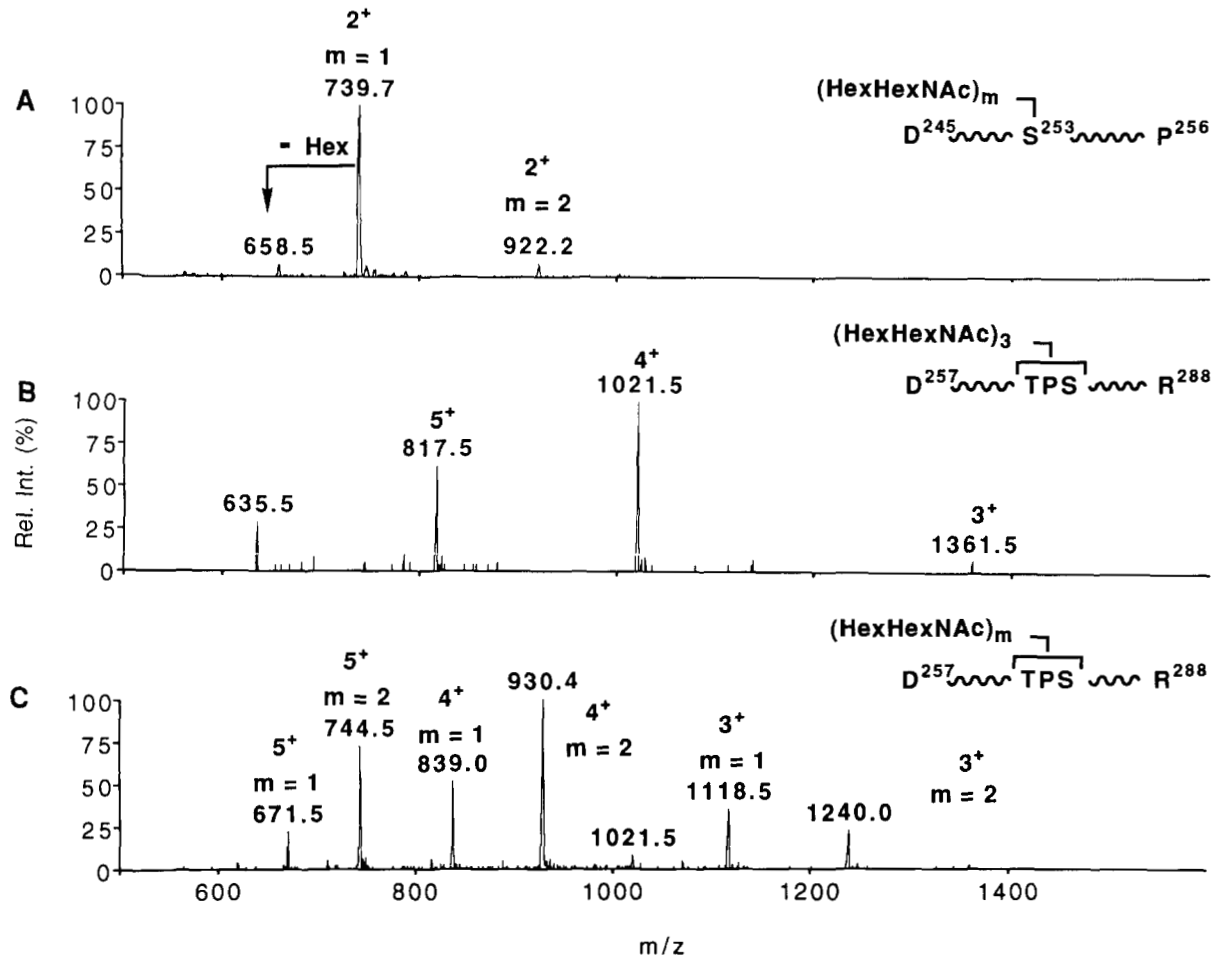

Fig. 7. Mass spectra of *O*-linked glycopeptides after digestion with PNGase F. Data are from an LC-ESMS/MS, parents of *m/z* 204, experiment (see Fig. 6). **A:** Peak at ca. 26 min corresponds to Ser^{253} glycosylation site (M_r meas. = 1,477.4, calc. = 1,478.4). Peaks at ca. 37 min (**B**) and 40 min (**C**), respectively, have the same peptide sequence Asp^{257}-Arg^{288} but differ in the number of Hex-HexNAc units ($m = 2$: M_r meas. = 3,717.6, calc. = 3,717.2; $m = 3$: M_r meas. = 3,352.5, calc. = 3,351.8).

Glu^{166}), and that it contained the third *N*-linked consensus sequence. The identity of the sequence of this glycopeptide was corroborated by automated Edman degradation (data not shown). A similar strategy is employed to identify the attachment sites and compositional heterogeneities of the *O*-linked glycopeptides. Because no consensus sequence exists for *O*-linked glycosylation, there may be ambiguity in the assignment of the specific Ser or Thr utilized if more than one of these residues is found in the same peptide (see discussion of *O*-linked glycopeptide Asp^{257}-Arg^{288}, above).

N-linked sites may also be identified by comparing the LC-ESMS-derived peptide maps obtained before and after treatment of the glycoprotein with PNGase F (Carr et al., 1990, 1991b). New M_r-related signals appearing in the mass spectra after treatment with the endoglycosidase correspond to formerly glycosylated peptides. These masses are then matched to the protein sequence, taking into account the fact that PNGase F converts the attachment site Asn to Asp, which weighs 1 Da more. The in-chain mass of the carbohydrate associated with a specific signal is determined by subtracting the M_r of the glycosylation-site peptide (attachment site as Asn) from the M_r of the glycopeptide. Possible compositions are then determined by computer using an elemental composition program in which the in-chain masses of the monosaccharides have been entered as the "elements" (Carr et al., 1990).

The MS methods described above for assessment of the glycoform populations at specific attachment sites are somewhat indirect, but they rapidly provide an informed starting point for further structural characterization of the carbohydrate. For example, after the MS techniques described here are used to locate glycopeptides in the chromatogram, the carbohydrate moieties of the isolated glycopeptides would be released by enzymatic or chemical means and the carbohydrates analyzed by other methods such as MS of the permethylated derivatives or

high-performance anion exchange chromatography (Carr et al., 1990; Barr et al., 1991).

Discussion

The LC-ESMS and LC-ESMS/MS procedures described here facilitate detection of glycopeptides in complex digests of glycoproteins, permit differentiation of glycopeptides containing *N*- vs. *O*-linked oligosaccharides, and, in the case of known protein sequences, enable sequence-specific assessment of glycoform populations without destruction of the peptide context. High selectivity and sensitivity for glycopeptide detection are obtained by enhancing the formation and detection of carbohydrate-specific marker ions by LC-ESMS. Micropreparative fractionation of the digests may be carried out simultaneously with the LC-ESMS and LC-ESMS/MS experiments to provide glycopeptides for further structural analysis. The carbohydrate moieties of glycopeptides isolated in this manner may then be released enzymatically, with PNGase F or Endo H, or chemically using hydrazinolysis (Takasaki et al., 1982) and then analyzed, for example, by high-performance anion exchange chromatography, methylation analysis, and MS (Barr et al., 1991; Carr et al., 1991b) to gain further insights into their sequences, branching patterns, and linkages. The studies described here were carried out with ca. 25 pmol of digest applied to the 320-mm internal diameter (i.d.)-packed fused silica column per experiment. Further improvement of two- to fivefold in sensitivity may be achieved by using smaller i.d. packed capillaries (e.g., 180 mm i.d.) that have recently become available commercially. Nearly 90% of the protein sequence of bovine fetuin was also corroborated in these studies, which utilized less than 100 pmol of total glycoprotein.

In the LC-ESMS analyses, we have employed a novel scanning protocol in which the collisional excitation potential of the region prior to the Q1 quadrupole is stepped or ramped with *m/z*. This technique enhances production of low *m/z* ions without loss of molecular weight information for the peptides and glycopeptides (see below) or shifting of the observed parent ion distributions to lower charge state, higher *m/z*. An added advantage of this method is that background ions from solvent clusters in the low *m/z* region of the spectrum (where they are usually most intense) are, simultaneously, reduced. The scanning protocol is easily implemented under data system control on quadrupole mass analyzers, and it is now routinely used in our laboratory for all LC-ESMS analyses of proteins and glycoproteins. In this report we have used this method for production of (positive) ions indicative of the sugar units comprising the oligosaccharide groups of the glycopeptides in order to facilitate location and preparative fractionation of glycopeptides. The ramped or stepped collisional excitation potential method could also possibly be used for detection of other modifications of interest, such as phosphates or sulfates, in either the positive or negative ion modes.

We have demonstrated the generality of using *m/z* 204 ($HexNAc^+$) as an indicator of both *N*- and *O*-linked carbohydrates. Ions of *m/z* 204 may also form by fragmentation of the peptide backbone. Such artifacts, which in practice are occasionally observed, are readily distinguished from glycopeptide-derived fragments by the absence of other carbohydrate-related ions such as *m/z* 186 and *m/z* 168, which form by water loss from *m/z* 204, and higher *m/z* ions diagnostic for carbohydrate such as *m/z* 292 and *m/z* 366 (see Results). It is also possible to use *m/z* 366 ($Hex\text{-}HexNAc^+$, or isomer) in a similar fashion to the *m/z* 204, often with excellent results. However, the yield of this fragment can be greatly reduced when the glycopeptide is of the oligomannose class, which may adversely impact on its use as a general indicator of all common oligosaccharide structural types. The presence of *N*-acetyl neuraminic acid and deoxyhexose (usually fucose) may be assessed by the RIC traces for the respective oxonium ions (*m/z* 292 and *m/z* 147, respectively).

Using the parent ion scanning protocol (LC-ESMS/MS) and monitoring for *m/z* 204 ($HexNAc^+$), we have demonstrated that the parent ions of glycopeptides are selectively detected in the presence of peptides, thereby removing interferences and unambiguously establishing which specific ions in the mass spectra are due to glycopeptides. Interestingly, the most abundant charge states of the glycopeptide-related ions in the parent ion data are shifted upward by one charge relative to the normal or stepped collisional excitation potential LC-ESMS data (e.g., from 2^+ to 3^+; see Fig. 3). This suggests that the $HexNAc^+$ fragment, which presumably derives at least in part from the sugar linked to the attachment-site Asn or Ser/Thr, is more readily formed via collisional dissociation of the higher charge state glycopeptide parents. This is consistent with the higher charge state ions taking up more energy upon collisional activation as a result of their higher translational kinetic energies. The shift to higher charge state, lower *m/z* is also analytically useful as it may facilitate observation of otherwise poorly detected high M_r glycopeptides.

The site-specific composition and heterogeneity data for the *N*-linked carbohydrates reported here are consistent with previous reports (Green et al., 1988; Townsend et al., 1988; Cumming et al., 1989; Rice et al., 1990). Compositions consistent with triantennary Asn-linked carbohydrates with the general structure $(NeuAc)_3(Hex\text{-}HexNAc)_3Hex_3HexNAc_2$ were identified at Asn^{81}, Asn^{138}, and Asn^{158}. A significant amount of glycoforms with compositions consistent with that of biantennary structures (ca. 40% based on relative peak heights in the ESMS data) were also observed at Asn^{138}. A small amount of possible biantennary complex carbohydrate may be present at Asn^{81} based on the presence of a minor peak at *m/z* 951.5 that corresponds to loss of a Hex-HexNAc

unit from the major component (Fig. 3C). The possibility that this peak is a fragment formed in the mass spectrometer cannot be ruled out without a chromatographic method for separating and quantifying these components. The collisional excitation potential used during that portion of the *m/z* scans in which the parent species of the peptide and glycopeptide parent ions are being detected is set for good sensitivity and minimal fragmentation (typically OR potential = 65 V on the Sciex). However, it is important to note that some amount of fragmentation can occur at even the lowest OR potential in ESMS, especially if the fragmentation is favorable, such as loss of lactosamine units from the outer chains of the carbohydrates (Dell, 1987). Nevertheless, biantennary complex carbohydrates have been reported to be present at all of the Asn-linked sites in fetuin (Yet et al., 1988), although the source of fetuin was different than that used in the present study. Our findings that biantennary structures are absent or present at only very low levels at Asn^{81} and Asn^{158} but are present at significant levels at Asn^{138} are consistent with the reports by Lee and coworkers (Rice et al., 1990) and Townsend et al. (1988).

The site-specific *O*-linked carbohydrate compositions and heterogeneity corroborate previous findings (Spiro & Bhoyroo, 1974; Nilsson et al., 1979; Edge & Spiro, 1987; Medzihradszky et al., 1990) but also provide new information. Specifically, we have assigned Ser^{253} unambiguously as an *O*-linked glycosylation site with the composition $(NeuAc)_l(Hex\text{-}HexNAc)_m$, where $l = 0$, 1, or 2 for $m = 1$ (major species), and $l = 2$ for $m = 2$ (minor species). A glycopeptide spanning the region containing potential *O*-linked glycosylation sites Thr^{262}, Ser^{264}, Ser^{272}, and Ser^{278} was observed with an aggregate carbohydrate composition of $(NeuAc)_l(Hex\text{-}HexNAc)_m$, where l, m = 2 (major) or 3 (minor). Thr^{262} and Ser^{264} are highlighted in the text because they have previously been shown to have carbohydrate attached (Medzihradszky et al., 1990). Our data do not rule out glycosylation at Ser^{272} or Ser^{278}. Clarifying these final attachment sites goes beyond the objectives of the present work and will probably require subdigestions sufficient to isolate each *O*-linked sugar to a unique di- or larger peptide. Tandem MS of the multiply *O*-glycosylated peptide Asp^{257}-Arg^{288} (Fig. 4) under low-energy collision conditions on the triple quadrupole resulted in elimination of the *O*-linked sugars with reversion of the attachment-site Ser or Thr to the corresponding alcohol. As a result, the sequence-specific fragment ions (which were quite intense) failed to provide any evidence for carbohydrate attachment sites. Tandem MS employing high-energy collisions on a tandem double focusing instrument may be more useful in this regard (Medzihradszky et al., 1990).

Materials and methods

Bovine fetuin from fetal calf serum (Sigma, lot 63F-9555) was reduced, carboxymethylated, and treated as previously described (Carr et al., 1990) using different combinations of the following enzymes: PNGase F (Genzyme, Cambridge, Massachusetts), TPCK-trypsin (Cooper Biomedical, Malvern, Pennsylvania), and Asp-N (Boehringer Mannheim, Indianapolis, Indiana). Samples of fetuin were also treated prior to reduction and alkylation with neuraminidase from *Vibrio cholerae* (Calbiochem, La Jolla, California) to remove sialic acid. Typically 0.1 U of enzyme was reacted with 1 nmol of glycoprotein at 37 °C overnight. Between 25 and 35 pmol of digest (as determined by quantitative amino acid analysis of the reduced and alkylated stock) were used per MS experiment. The glycopeptide models $Man_9GlcNAc_2$-Asn and Gal-GalNAc-Ser were purchased from BioCarb Chemicals (Accurate Chemical & Scientific Corp., Westbury, New York) and used without purification or desalting.

Chromatography

The gradient HPLC system is a Brownlee Labs (Santa Clara, California) Microgradient™ with dual syringe pumps, 250-μL dynamic mixer, and a Rheodyne model 8125 injector with a 20-μL replaceable sample loop. The column used in this work is an LC Packings (San Francisco, California) Fusica™ capillary column (320 μm × 15 cm) custom packed with 5-μm, 300-Å-pore size C_{18} packing from Vydac. UV detection at 215 nm was accomplished with an ABI model 785A programmable absorbance detector. The flow cell of the ABI was replaced with an LC Packings U-Z View™ capillary flow cell (35-nL volume, 7–8-mm path length). The configuration of the HPLC/MS interface uses two Rheodyne switching valves pre-injector: one to isolate the injector and column from the pump and the second to facilitate purging of the syringe pumps. A low dead volume tee (Valco) with a length of fused silica on the waste leg was used to split the flow from ca. 100 μL/min down to 4–5 μL/min to the injector and column. The capillary column was fitted directly to the column outlet on the injector. The fused silica extension tube attached to the packed fused silica column was connected to the inlet of the UV detector via a zero dead volume union (Valco). A length of 50-μm-i.d. fused silica was attached to the outlet of the UV detector and threaded through the ionspray needle on the mass spectrometer. The mobile phases used for gradient elution consisted of (A) 0.1% (v/v) trifluoroacetic acid (TFA) and (B) acetonitrile/H_2O 90:10 (v/v) containing 0.09% TFA. The gradient conditions were as follows: 0% B to 60% B linearly in 60 min; then linearly to 100% B in 10 min with no delays.

Mass spectrometry

Electrospray mass spectra were recorded on a Sciex API-III triple quadrupole mass spectrometer fitted with an articulated ionspray plenum and an atmospheric pressure ionization source (Sciex, Ontario, Canada). In LC-ESMS

(normal OR potential), the mass spectrometer was scanned over a range of *m/z* 200–2,000 at 6 s/scan in 0.33-Da steps. Orifice potential was set at 65 V. In LC-ESMS (scanned OR potential) the mass spectrometer was scanned over a range of *m/z* 150–1,950 at 6 s/scan in steps of 0.33 Da. Orifice potential was either stepped in a single jump from 140 V to 65 V at *m/z* 500 or scanned down linearly from 120 V to 65 V between *m/z* 150 and *m/z* 500; from *m/z* 500 to *m/z* 1,950 the OR was held at 65 V. In LC-ESMS/MS (parents of *m/z* 204) the first quadrupole was scanned over a range of *m/z* 500–2,000 at 6 s/scan in 0.5-Da steps. Orifice potential was set at 55 V. Argon (99.999%) was used as the collision gas with a collision gas thickness (CGT) of 7×10^{14} molecules/cm^2 (obtained using a value of ca. 700 for the CGT setting on the Sciex). The potential difference between Q0 and Q2 was ca. 43 V, and this potential, when multiplied by the charge state of the precursor ion, is approximately equal to the collision energy.

The Sciex acquisition software (Tune version 2.1) makes it possible to ramp almost any instrument voltage or setting with *m/z* in all scan modes. In order to step or ramp a parameter over only part of the mass range, the multiple-ion (MI) scan mode is used. This allows more freedom in the OR ramp or step functions during the different sections of the *m/z* scan. It is important to always set the OR value called at the scan initiation (the "state file") to the starting value of the OR ramp, which is called from the MI scan setup.

Normal MS is tuned and calibrated using a mixture of polypropylene glycols (PPG) 425, 1,000, and 2,000 (3.3×10^{-5} M, 1×10^{-4} M, and 2×10^{-4} M, respectively), in 50/50/0.1 H_2O/methanol/formic acid (v/v/v), 1 mM NH_4OAc. Normal scan ESMS and LC-ESMS spectra were recorded at instrument conditions sufficient to resolve the isomers of the PPG/NH_4^+ doubly charged ion at *m/z* 520 (85% valley definition). Tuning and calibration of parent-scan mode are also accomplished with PPG by scanning Q1 and monitoring the selected product ion (*m/z* 204) in Q3 but with the resolution set to 0 so that the entire *m/z* range is passed unfiltered. In this way, the various instrument parameters can be set up in a simulated parent-scan mode. Final optimization in true parent-scan mode is accomplished with the glycopeptide models by monitoring *m/z* 204 product ions. The analyzing quadrupole (Q1) was operated at a resolution of unit *m/z* (about 50% peak valley definition, being more resolved at low *m/z* and less resolved at high *m/z*), whereas the mass-selecting quadrupole Q3 was set to pass a 2–3-Da window around the ion of interest so as to enhance sensitivity. On the Sciex instrument, we adjusted Q1 resolution by lowering RE1 2 units and raising DM1 by 0.05 units; Q3 resolution was changed by lowering RE3 by 35–40 units while raising DM3 by 0.1. The OR potential is decreased from 65 V to 55 V to reduce formation of fragment ions prior to Q1. Other important parameters including the quadrupole offsets (related to collision energy) and the collision gas pressure are also checked at this time to maximize transmission of the 204 *m/z* ion.

Acknowledgments

The excellent technical assistance of Todd Armbruster is gratefully acknowledged. This work was supported in part by a grant from the National Institutes of Health GM-39526 to S.A. Carr. This work was presented in preliminary form at the 39th ASMS Conference on Mass Spectrometry and Allied Topics, June 1991, Nashville, Tennessee, and at the 40th ASMS Conference, June 1992, Washington, D.C.

References

Barr, J.R., Anumula, K.R., Vettese, M.B., Taylor, P.B., & Carr, S.A. (1991). Structural classification of carbohydrates in glycoproteins by mass spectrometry and high-performance anion-exchange chromatography. *Anal. Biochem. 192*, 181–192.

Carr, S.A., Armbruster, T., Hemling, M.E., Soneson, K.K., & Huddleston, M.J. (1991a). Analysis of large proteins by electrospray MS: Application to soluble complement receptor. *Proceedings of the 39th ASMS Conference on Mass Spectrometry and Allied Topics*, Nashville, Tennessee, May 19–24, pp. 483–484.

Carr, S.A., Barr, J.R., Roberts G.D., Anumula, K.R., & Taylor, P.B. (1990). Identification of attachment sites and structural classes of asparagine-linked carbohydrates in glycoproteins. *Methods Enzymol. 193*, 501–518.

Carr, S.A., Hemling, M.E., Bean, M.F., & Roberts, G.D. (1991b). Integration of mass spectrometry in analytical biotechnology. *Anal. Chem. 63*, 2802–2824.

Carr, S.A. & Roberts, G.D. (1986). Carbohydrate mapping by mass spectrometry: A novel method for identifying attachment sites of Asn-linked sugars in glycoproteins. *Anal. Biochem. 157*, 396–406.

Conboy, J.J. (1992). The determination of glycoproteins and glycopeptides by LC-MS and tandem-MS. Ph.D. Thesis, Cornell University, Ithaca, New York.

Conboy, J.J. & Henion, J.D. (1991). Ion spray LC/MS determination of glycoproteins and glycopeptides. *Proceedings of the 39th ASMS Conference on Mass Spectrometry and Allied Topics*, Nashville, Tennessee, May 19–24, pp. 1418–1419.

Covey, T.R., Bonner, R.F., & Shushan, B.I. (1988). The determination of protein, oligonucleotide and peptide molecular weights by ion-spray mass spectrometry. *Rapid Commun. Mass Spectrom. 2*, 249–256.

Cumming, D.A. (1991). Glycosylation of recombinant protein therapeutics: Control and functional implications. *Glycobiology 1*, 115–130.

Cumming, D.A., Hellerqvist, C.G., Harris-Brandts, M., Michnick, S.W., Carver, J.P., & Bendiak, B. (1989). Structures of asparagine-linked oligosaccharides of the glycoprotein fetuin having sialic acid linked to *N*-acetylglucosamine. *Biochemistry 28*, 6500–6512.

Dell, A. (1987). FAB-mass spectrometry of carbohydrates. *Adv. Carbohydr. Chem. Biochem. 45*, 19–72.

Dube, S., Fisher, J.W., & Powell, J.S. (1988). Glycosylation at specific sites of erythropoietin is essential for biosynthesis, secretion, and biological function. *J. Biol. Chem. 263*, 17516–17521.

Dziegielewska, K.M., Brown, W.M., Casey, S.-J., Christie, D.L., Foreman, R.C., Hill, R.M., & Saunders, N.R. (1990). The complete cDNA and amino acid sequence of bovine fetuin. *J. Biol. Chem. 265*, 4354–4357.

Edge, A.S.B. & Spiro, R.G. (1987). Presence of an *O*-glycosidically linked hexasaccharide in fetuin. *J. Biol. Chem. 262*, 16135–16141.

Fenn, J.B., Mann, M., Meng, C.K., Wong, S.F., & Whitehouse, C.M. (1989). Electrospray ionization for mass spectrometry of large biomolecules. *Science 246*, 64–71.

Goochee, C.F., Gramer, M.J., Andersen, D.C., Bahr, J.B., & Rasmussen, J.R. (1991). The oligosaccharides of glycoproteins: Bioprocess factors affecting oligosaccharide structure and their effect on glycoprotein properties. *Bio/Technology 9*, 1347–1355.

Green, E.D., Adelt, G., Baenziger, J.U., Wilson, S., & Van Halbeek, H.V. (1988). The asparagine-linked oligosaccharides on bovine fetuin. *J. Biol. Chem. 263*, 18253–18268.

Gribben, J.G., Devereux, S., Thomas, N.S.B., Keim, M., Jones, H.M., Goldstone, A.H., & Linch, D.C. (1990). Development of antibodies to unprotected glycosylation sites on recombinant GM-CSF. *Lancet 335*, 434–437.

Haselbeck, A., Schickaneder, E., Eltz, H.V.D., & Hosel, W. (1990). Structural characterization of glycoprotein carbohydrate chains by using digoxigenin-labeled lectins on blots. *Anal. Biochem. 191*, 25–30.

Hemling, M.E., Roberts, G.D., Johnson, W., & Carr, S.A. (1990). Analysis of proteins and glycoproteins at the picomole level by on-line coupling of microbore high-performance liquid chromatography with flow fast atom bombardment and electrospray mass spectrometry: A comparative evaluation. *Biomed. Environ. Mass Spectrom. 19*, 677–691.

Hsi, K.-L., Chen, L., Hawke, D.H., Zieske, L.R., & Yuan, P.-M. (1991). A general approach for characterizing glycosylation sites of glycoproteins. *Anal. Biochem. 198*, 238–245.

Huddleston, M.J., Bean, M.F., Barr, J.R., and Carr, S.A. (1991). Selective detection of glycopeptides in protein digests by electrospray LC-MS/MS. *Proceedings of the 39th ASMS Conference on Mass Spectrometry and Allied Topics*, Nashville, Tennessee, May 19–24, pp. 280–281.

Kobata, A. & Endo, T. (1992). Immobilized lectin columns: Useful tools for the fractionation and structural analysis of oligosaccharides. *J. Chromatogr. 597*, 111–122.

Ling, V., Guzzetta, A.W., Canova-Davis, E., Stults, J.T., Hancock, W.S., Covey, T.R., & Shushan, B.I. (1991). Characterization of the tryptic map of recombinant DNA derived tissue plasminogen activator by high-performance liquid chromatography–electrospray ionization mass spectrometry. *Anal. Chem. 63*, 2909–2915.

Matzuk, M.M., Krieger, M., Corless, C.L., & Boime, I. (1987). Effects of preventing *O*-glycosylation on the secretion of human chorionic gonadotropin in Chinese hamster ovary cells. *Proc. Natl. Acad. Sci. USA 84*, 6354–6358.

Medzihradszky, K.F., Gillece-Castro, B.L., Settineri, C.A., Townsend, R.R., Masiarz, F.R., & Burlingame, A.L. (1990). Structure determination of *O*-linked glycopeptides by tandem-MS. *Biomed. Environ. Mass Spectrom. 19*, 777–781.

Nilsson, B., Norden, N.E., & Svensson, S. (1979). Structural studies on the carbohydrate portion of fetuin. *J. Biol. Chem. 254*, 4545–4553.

Rademacher, T.W., Parekh, R.B., & Dwek, R.A. (1988). Glycobiology. *Annu. Rev. Biochem. 57*, 785–838.

Rice, K.G., Rao, N.B.N., & Lee, Y.C. (1990). Large-scale preparation and characterization of *N*-linked glycopeptides from bovine fetuin. *Anal. Biochem. 184*, 249–258.

Spiro, R.G. & Bhoyroo, V.D. (1974). Structure of the *O*-glycosidically linked carbohydrate units of fetuin. *J. Biol. Chem. 249*, 5704–5717.

Takasaki, S., Mizuochi, T., & Kobata, A. (1982). Hydrazinolysis of asparagine-linked sugar chains to produce free oligosaccharides. *Methods Enzymol. 83*, 263–268.

Townsend, R.R., Alai, M., Hardy, M.R., & Fenselau, C.C. (1988). Assessment of glycosylation-site heterogeneity using plasma desorption mass spectrometry. *Anal. Biochem. 171*, 180–191.

Yet, M.-G., Chin, C.C.Q., & Wold, F. (1988). The covalent structure of individual *N*-linked glycopeptides from ovomucoid and asialofetuin. *J. Biol. Chem. 263*, 111–117.