Glycoprotein Profiling by Electrospray Mass Spectrometry

Hui Jiang and Heather Desaire
Department of Chemistry, University of Kansas, Lawrence, Kansas, USA

Vladimir Y. Butnev and George R. Bousfield
Department of Biology, Wichita State University, Wichita, Kansas, USA

This work compares several different methods of site-specific analysis of glycoproteins using electrospray mass spectrometry. The glycoprotein, oLHα (ovine luteinizing hormone, α-sub-unit) was chosen as an appropriate example protein for these studies because of its biological relevance and extreme microheterogeneity. More than 20 unique glycoforms were detected for this glycoprotein at the Asn[56] site of oLHα. The carbohydrates present at this site affect receptor binding affinity, so understanding the great variety in the composition of these carbohydrates is important in studying ligand binding interactions. MS data was acquired on a quadrupole ion trap, a triple quadrupole, and a quadrupole time of flight mass spectrometer, and carbohydrate composition at the Asn[56] site of oLHα was determined using these instruments. Additionally, neutral loss and precursor ion scanning modes were also used to identify the glycoforms present, and these techniques were compared to the standard MS data. Of the three instruments compared in the study, the qTOF mass spectrometer achieved the lowest sample consumption, but all three instruments were useful in profiling the glycopeptide composition.   (J Am Soc Mass Spectrom 2004, 15, 750–758) © 2004 American Society for Mass Spectrometry

Structural analysis of glycoproteins is a complicated task due to the carbohydrate's complex heterogeneity. Because of the complexity of the samples and the typically small amounts of material available, mass spectrometry is one technique that could provide structural information in most cases [1]. There are two common procedures for characterization of carbohydrates on glycoproteins by mass spectrometry. The first involves release of glycans, purification, and MS analysis [2, 3]. Although removing the protein from the carbohydrates could result in simpler mass spectra, information about which carbohydrates came from which site is lost. In addition, the proton-affinity of carbohydrates is weak. Their mass spectral signal to noise ratio is low, and quite often derivatization of the samples is necessary to increase ion signal [4].

An alternate approach to analyzing the carbohydrates on glycoproteins involves digestion of the glycoprotein with endoproteases and characterizing the resulting glycopeptides [5–8]. The resultant glycopeptides may be analyzed by mass spectrometry with minimal clean-up. Peptide backbones generally have higher ionization efficiency than carbohydrates, producing more intense mass spectral signals and making accurate mass measurement easier. In addition, proteolytic fragments on either side of the amino acid carrying the carbohydrate can be used to reveal the specific site from which the carbohydrate originated [9]. When multiple glycosylation sites are present on a single glycoprotein, site-specific (glycopeptide) analysis is frequently require in order to determine which carbohydrates originated at which glycosylation site. Recently, Lebrilla and co-workers have demonstrated an effective technique for site-specific glycopeptide analysis using a nonselective enzyme and MALDI-FT MS [10]. To further develop this approach as an effective analytical technique, we compare herein various methods of site-specific glycopeptide analysis using electrospray mass spectrometry.

Several different ESI-MS strategies for detection and composition determination of glycopeptides have been previously demonstrated. For example, parent ion scanning for the characteristic glycopeptide marker ions has been carried out [1, 11]. Alternatively, direct determination of glycopeptide composition has been achieved by accurate mass measurements on a qTOF mass spectrometer for various glycoproteins [7, 8]. Quadrupole ion traps have also been used in the analysis of glycopeptides [12, 13]. The focus of this study is to profile the composition of the glycoforms present at a specific site on the protein, and compare several previously pro-

posed ESI-MS techniques in order to determine the *most effective* strategy for glycopeptide analysis on electrospray mass spectrometers. The ESI-MS techniques to be compared include composition profiling using MS data on a quadrupole ion trap, quadrupole time of flight, and triple quadrupole mass spectrometer. These studies are compared with results achieved via neutral loss scanning or precursor ion scanning experiments on a triple quadrupole mass spectrometer.

The glycoprotein chosen for these studies is luteinizing hormone (LH). This glycoprotein is one of three structurally related glycoprotein hormones involved in the regulation of reproduction. It consists of two noncovalently bound α- and β-subunits. Both the common α- and hormone-specific β-subunits are glycosylated. The glycosylation sites on both of these subunits play important roles in the structure and the function of the protein. Oligosaccharides attached to the human α-subunit at Asn^{78} are important for proper folding of the protein [14]. In addition, the type of oligosaccharides attached to a second glycosylation site (Asn^{52}) have been shown to influence receptor-binding affinity and are essential for full expression of biological activity at the cellular level [15]. (This site is numbered Asn^{56} in ovine and other mammalian α-subunits.) Glycosylation on the β-subunit has also been demonstrated to have biologically significant consequences on circulatory clearance and signal transduction [16]. Since each of the glycosylation sites have been implicated in different biological roles, it would be useful to know which carbohydrates are present at each site on the protein. By subjecting this protein to the mass spectrometric experiments herein, site-specific glycosylation information is obtained.

Experimental

Materials

The oligosaccharide internal standard for exact mass measurement, (GlcNAc-Man)$_2$Man-GlcNAc-GlcNAc (NGA2), was purchased from Prozyme (San Leandro, CA). NGA2 is a reducing carbohydrate containing a conserved trimannosyl-chitobiose core. It has a molecular weight of 1316.48650 Da.

Ovine LHα (oLHα) was obtained from purified oLH by counter-current distribution followed by Sephadex (Amersham Biosciences, Piscataway, NJ) G-100 gel filtration [17].

oLHα Glycopeptide Preparations

Samples consisting of 8 mg oLHα were dissolved in 0.126 M ammonium bicarbonate, pH 8.6, containing 1 mM CaCl$_2$, and digested for 24 h at 37 °C with proteinase K (1:10, enzyme:substrate ratio). Peptides and glycopeptides released from oLHα were separated from the protease-resistant core and protease by ultrafiltration using Amicon (Millipore, Bedford, MA) Centricon

P-10 centrifugal ultrafiltration cartridges [18]. The filtrate was evaporated to dryness in a Savant (Thermo-Savant, Holbrook, NY) SpeedVac. It was applied to a 150 mg Altech CarboGraph (Altech Associates, Inc., Deerfield, IL) graphitized charcoal cartridge, salt removed with water, and the glycopeptide-containing fraction eluted with 0.025% trifluoracetic acid in 25% acetonitrile [19]. Peptides and glycopeptides were recovered by evaporation in the SpeedVac.

Peptide Sequencing (Edman)

Samples of the proteinase K digest filtrate fractions were applied to Biobrene-coated glass fiber disks and subjected to 4 cycles of automated Edman degradation using an Applied Biosystems (Foster City, CA) model 492 Procise protein sequencer, following the manufacturer's recommended procedures.

qTOF Mass Spectrometry

Nanospray MS spectra were recorded on a Micromass Q-TOF^2 mass spectrometer. Samples were dissolved in a mixture of methanol and water (50:50 by volume) containing 0.5% acetic acid at a concentration around 1 μg/μL. Each sample (about 10 μL) was loaded onto a PicoTip (New Objective, Woburn, MA) which was mounted to the nanospray stack. A voltage of 700 to 1500 V was applied between the tip and the mass spectrometer inlet. Cone voltage was 30 V for normal MS experiments. The mass spectrometer was calibrated using cesium iodide. CID spectra were acquired at a collision energy between 25 and 40 V, with argon as the collision gas. Spectra were generally recorded between m/z 1000 and 2000 unless specified. Accurate mass measurement was done with the addition of an internal standard NGA2.

Triple Quadrupole Mass Spectrometry

Nanospray MS, parent ion scan, and neutral loss scan spectra were recorded on a Micromass Quattro Ultima mass spectrometer. Samples were dissolved in a mixture of methanol and water (50:50 by volume) containing 0.5% acetic acid at a concentration of about 3 μg/μL. Each sample (about 10 μL) was loaded to a PicoTip, as on the qTOF. A voltage of 700 to 1500 V was applied between the tip and the mass spectrometer inlet for nanospray. Cone voltage was optimized to be 100 V, at which maximum abundance of the singly charged ions was observed. For precursor ion scan and neutral loss scan, collision energy was set between 20 and 40 V. Spectra were recorded between m/z 1000 and 2000, unless specified.

Ion Trap Mass Spectrometry

Electrospray mass spectra were recorded on a Thermo Finnigan LCQ Advantage mass spectrometer (San Jose,

CA). Samples were dissolved in a mixture of methanol and water (50:50 by volume) containing 0.5% acetic acid at a concentration about 1 $\mu g/\mu L$. Each sample was injected into a mobile phase of 50% methanol in water (vol/vol) containing 0.5% acetic acid. The mobile phase entered the mass spectrometer at a flow rate of 5 $\mu L/min$ from an HPLC pump (MS Pump, Thermo, San Jose, CA). Spray voltage was maintained at approximately 4.5 kV, and N_2 was used as a nebulizing gas at a flow rate of 0.3 L/min. Spectra were recorded between m/z 1000 and 2000.

Automated Mass Assignments

Microsoft Excel macros were written in VBA (Visual Basic for Applications) to assign the glycopeptide ions in the mass spectra. The spectral data were first converted to a text file. The text file, which contained tab-delimited ASCII, was imported to Excel. The number of data points, at least several thousand, was reduced by a macro with user-selected criteria, such as relative abundance, baseline correction, and peak apex recognition. A second macro was run to display a window for information input, including mass match tolerance (in ppm), linkage (N- or O-linked), and glycopeptide combination components, which consist of peptide backbone, Hex (hexose), HexNAc (N-acetyl-hexoamine), Fuc (deoxy-hexose), sialic acid, sulfate, and sodium ion in various numbers. The macro then constituted all reasonable glycopeptide combinations. Only those combinations that were not in conflict with the mass spectrum, and matched the expected glycosylation characteristics of the glycopeptides, were considered to be reasonable. For example, for a parent ion scan of m/z 366, which is the characteristic ion for Hex-HexNAc, only those combinations that contain at least one HexNAc and one Hex were accepted as reasonable combinations. Since the glycopeptides were N-linked, the combinations were also restricted to those derived from the conserved trimannosyl-chitobiose core structure ($Man_3GlcNAc_2$). The m/z of these combinations were compared with the mass spectrum. A match was found when the m/z difference was within the chosen mass tolerance. When the process was completed, the macro generated a spreadsheet listing all the matched results showing the experimental m/z, the charge state, the matched glycopeptide combination, and the calculated m/z.

Results

Glycopeptide Isolation

After digesting the glycoprotein with proteinase K, and isolating the released peptides and glycopeptides, Edman sequencing was used to characterize the sample. This technique gave evidence that glycopeptides at the Asn[56] site had been liberated from the protein, and that the predominant peptide associated with the carbohy-
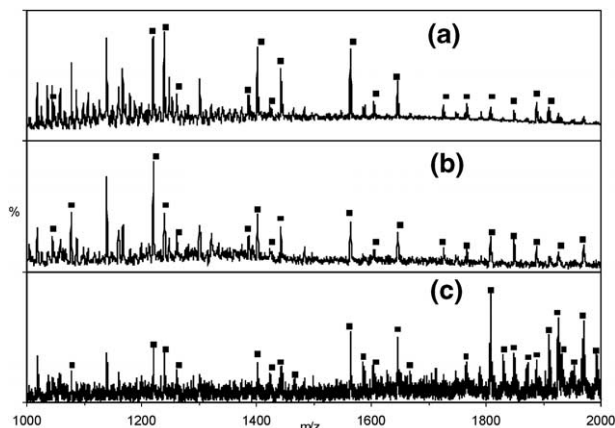

Figure 1. Mass spectra of oLHα glycopeptides on the Q-TOF (**a**), triple quadrupole (**b**), smoothed using Micromass MassLynx 3.5), and ion trap (**c**). Ions with a square are glycopeptides identified by mass match.

drate was Asn[CHO]-Ile-Thr. As expected, Asn did not show up in the first cycle of Edman sequencing because glycosylated amino acids are not detected. However, the second and third cycle of Edman sequencing showed very strong signals for Ile and Thr respectively. In the fourth cycle, the most abundant signal was more than 10 times lower than the predominant signals from cycles two or three. This suggests that the most abundant peptide present was a tripeptide, with the second and third amino-acids in the sequence being Ile-Thr. There is only one site where the dipeptide Ile-Thr exists in the protein. It is immediately following the glycosylated Asn[56]. Therefore, the evidence indicated that the glycopeptide Asn[CHO]-Ile-Thr was released in high abundance. Heterogeneity in the sequencing data indicated that other peptides were also present, as expected.

MS Analysis of Glycopeptides

The proteolytic glycopeptide digests prepared above were also analyzed by mass spectrometry. The mass spectra acquired on the qTOF, triple quadrupole and quadrupole ion trap are shown in Figure 1a, b, and c, respectively. The overall profiles of glycopeptides were obtained by MS in a range between m/z 1000 and 2000—m/z 2000 is the upper mass limit of the LCQ ion trap mass spectrometer, and the glycopeptides generated were expected to have a m/z value of greater than 1000. The spectra in Figure 1 indicate rich heterogeneity of the hormone glycopeptides.

Detection of Constituent Monosaccharides and Proteolytic Peptides

Because the digestion conditions utilized produce non-specific fragmentation of the protein, multiple peptide sequences are possible for the glycopeptides shown in Figure 1. In addition, a wide variety of carbohydrates may be present. In order to simplify the task of assign-
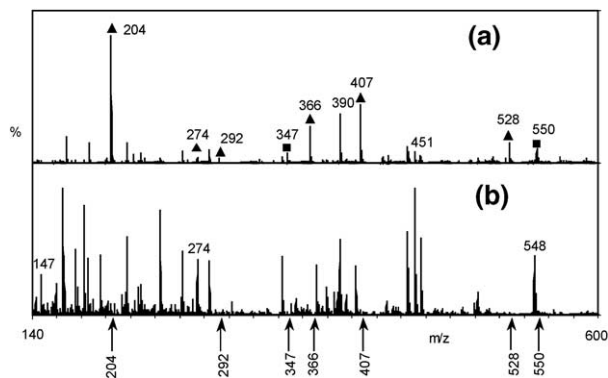
Figure 2. (**a**): CID spectrum of the oLHα glycopeptides on the Q-TOF. Glycopeptide fragment ions were generated by setting the Q-TOF in MS mode with the collision gas on and the collision energy at 30 V. Ions with a triangle indicate carbohydrate fragments. Ions with a square indicate peptide or glycopeptide fragments. (**b**): Normal mass spectrum of the oLHα glycopeptides. Only m/z 147 and 274 were present. Arrows indicate the ions observable in (**a**), but not in (**b**).

Figure 3. Generation of the characteristic ion [peptide + Glc-NAc]H$^+$ during CID. Subtracting 203 (the mass of GlcNAc) from the ion's mass gives the peptide mass, which could be used to determine the peptide composition.

ing compositions to the ions in Figure 1, additional information about the glycopeptide composition was sought out. By determining which monosaccharide components and which peptide components are present, assigning the composition of each ion is more feasible. Component analysis is determined through a mass spectrometric experiment on the qTOF. Specifically, by increasing the collision energy and adding in collision gas while the qTOF was set in MS mode, non-specific fragmentation, or CID, of all the species present is achieved, and the constituents present can be readily determined.

The spectra containing the constituents of the oLHα glycopeptides are shown in Figure 2. Ions that were formed by non-specific fragmentation include several ions attributed to mono- and oligosaccharide oxonium ions. These include m/z 204 for HexNAc (*N*-acetylhexosamine), m/z 366 for disaccharide Hex-HexNAc (Hex for hexose), m/z 407 for 2HexNAc, and m/z 528 2Hex-HexNAc. These ions confirm the presence of HexNAc and Hex in the glycopeptide sample. It should be noted that none of these ions appear at appreciable abundance in normal MS spectra obtained without fragmentation (See Figure 2b). This supports the fact that these ions are not isomeric impurities present in the samples, but fragments of the glycopeptides. High resolution mass measurements on these species were less than 20 ppm from the calculated value in every case. Other monosaccharides that have been observed in carbohydrates isolated from luteinizing hormone include sialic acid and deoxy-hexose [20, 21]. The corresponding oxonium ions (m/z 147 and 292 or its dehydrated form, m/z 274) were not detected in appreciable abundance in the CID experiment shown in Figure 2; however, that may be due to low proton affinities of these compounds.

The peptide composition can also be deduced based on the data in Figure 2. When glycopeptides undergo CID, an abundant product ion is generated by a cleav-
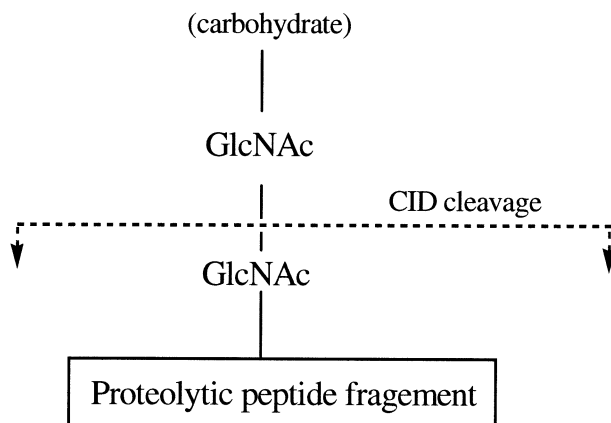
age between the first and second GlcNAc adjacent to the peptide [22] as shown in Figure 3. Such a cleavage produces ions containing the GlcNAc attached to whatever peptide was produced by proteolysis. To determine this peptide, 203 was subtracted from the masses of the ions. Then the remainders were searched for peptide compositions. A likely ion will not be smaller than m/z 336, because after 203 is subtracted from 336, the remainder is 133, which is the molecular weight of protonated Asn. (N-linked carbohydrates are always attached to Asn.) In the spectra in Figure 2a, while no ion is present at m/z 336, three unassigned ions are present at larger nominal masses: m/z 390, m/z 451, and 550. Only m/z 550 corresponds to a nominal mass that could originate from a GlcNAc-containing glycopeptide with its peptide sequence being Asn-Ile-Thr. This assignment is supported by high resolution measurements with an error of 13 ppm. The presence of peptide fragment Asn-Ile-Thr is also supported by the presence of its [M + H]$^+$ at m/z 347, and by the Edman sequencing data described above.

Automated Compositional Assignments

After determining the most likely constituents present in the glycopeptides, spectral interpretation can be facilitated by macros written in Visual Basic. The glycopeptide sequence Asn-Ile-Thr, along with all possible monosaccharide components are input into an Excel Macro, and all glycopeptide compositions that are consistent with the ions produced in the mass spectra are output. Using this approach, the composition of all the ions in Figure 1 that correspond to glycopeptides can be assigned. These assignments are displayed in Table 1.

High Resolution

On the quadrupole ion trap and triple quadrupole mass spectrometer, only low resolution data is available.

Table 1. oLHα Glycopeptides found on different mass spectrometers with different techniques

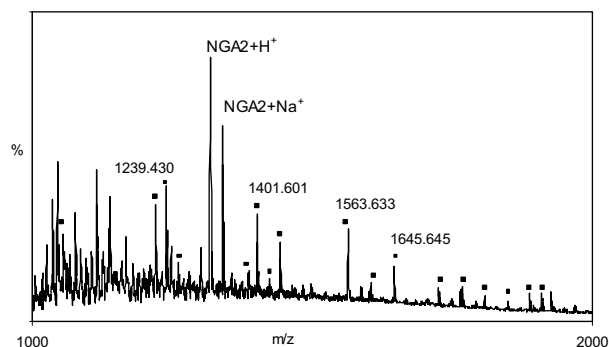| Observed m/z	Charge	Calculated m/z	Glycopeptide composition	MS			Parent ion scan			Neutral loss scan	
				qTOF	Triple Quadrupole	Ion Trap	204	163	550	80	40
1992	1+	1991.757	[NIT + 4HexNAc + 5Hex] + Na+			*					
1970	1+	1969.775	[NIT + 4HexNAc + 5Hex] + H+		*						
1954	1+	1953.780	[NIT + 4HexNAc + 4Hex + Fucose] + H+			*					
1932	1+	1931.643	[NIT + 4HexNAc + 4Hex + SO3Na] + Na+			*					
		1932.596	[NIT + 3HexNAc + 4Hex + Fucose + 2SO3] + Na+			*					*
1926	1+	1926.609	[NIT + 3HexNAc + 5Hex + 2SO3] + H+		*	*				*	
963	2+										
1910	1+	1909.661	[NIT + 4HexNAc + 4Hex + SO3] + Na+			*					*
		1909.703	[NIT + 2HexNAc + 7Hex] + Na+								
1888	1+	1887.679	[NIT + 4HexNAc + 4Hex + SO3] + H+	*	*	*				*	
944	2+										
1871	1+	1870.730	[NIT + 5HexNAc + 3Hex] + Na+	*		*					*
1848	1+	1848.748	[NIT + 5HexNAc + 3Hex] + H+	*	*	*	*				
1830	1+	1829.704	[NIT + 4HexNAc + 4Hex] + Na+		*	*					
1808	1+	1807.722	[NIT + 4HexNAc + 4Hex] + H+	*	*	*	*				
1767	1+	1766.695	[NIT + 3HexNAc + 5Hex] + H+	*	*	*			*		
1726	1+	1725.626	[NIT + 4HexNAc + 3Hex + SO3] + H+	*	*	*					
1668	1+	1667.651	[NIT + 4HexNAc + 3Hex] + Na+			*					
1646	1+	1645.669	[NIT + 4HexNAc + 3Hex] + H+	*	*	*	*	*	*		
1605	1+	1604.642	[NIT + 3HexNAc + 4Hex] + H+	*	*	*					
1586	1+	1585.598	[NIT + 2HexNAc + 5Hex] + Na+			*					
1564	1+	1563.616	[NIT + 2HexNAc + 5Hex] + H+	*	*	*	*	*	*		
1464	1+	1464.572	[NIT + 3HexNAc + 3Hex] + Na+			*					
1442	1+	1442.590	[NIT + 3HexNAc + 3Hex] + H+	*	*	*	*	*	*		
1424	1+	1423.545	[NIT + 2HexNAc + 4Hex] + Na+	*	*	*					
1402	1+	1401.563	[NIT + 2HexNAc + 4Hex] + H+	*	*	*	*	*	*		
1385	1+	1385.568	[NIT + 2HexNAc + 3Hex + Fucose] + H+	*	*	*					
1261	1+	1261.492	[NIT + 2HexNAc + 3Hex] + Na+	*	*	*					
1239	1+	1239.510	[NIT + 2HexNAc + 3Hex] + H+	*	*	*	*	*			
1220	2+	1220.452	[NIT + 6HexNAc + 4Hex + Fucose + SO3] + 2H+	*	*	*	*	*			
1077	1+	1077.457	[NIT + 2HexNAc + 2Hex] + H+		*	*	*				
1047	2+	1047.330	[NIT + 6HexNAc + 2Hex + 2SO3] + 2Na+	*					*		
		1047.352	[NIT + 4HexNAc + 5Hex + SO3] + 2Na+								
1045	2+	1044.835	[NIT + 3HexNAc + 6Hex + 2SO3] + 2H+		*						

Figure 4. Mass spectrum of the oLHα glycopeptides on the Q-TOF with accurate mass measurement. Ions with a square indicate glycopeptides identified by accurate mass match.

Figure 5. Parent ion scans on the triple quadrupole for oLHα glycopeptides. (**a**): m/z 204; (**b**): m/z 163; (**c**): m/z 550. Ions with a square indicate glycopeptides detected by parent ion scan. Spectra were smoothed using Micromass MassLynx 3.5.

Since the qTOF mass spectrometer has high resolution capabilities, high resolution data were acquired using a carbohydrate-based internal standard NGA2 (Figure 4). Each ion (singly and doubly charged) in the spectrum was matched against all possible combinations of the constituent monosaccharides using the Visual Basic macro. A match was reached when the difference between the observed m/z and the calculated m/z of a glycopeptide composition was less than 20 ppm. Therefore, all proposed glycopeptide compositions in Table 1 (which were observed on the qTOF) have met the 20 ppm criteria for mass accuracy.

Parent Ion Scan

In addition to simply obtaining the mass spectrum of the glycopeptide mixture on each mass spectrometer, precursor ion and neutral loss scanning was completed on the triple quadrupole mass spectrometer. Glycopeptides have been previously detected by precursor ion scanning of characteristic marker ions at m/z 163, 204, and 366 [11]. So these experiments were attempted herein. A series of ions were selected for detection of different constituent monosaccharides and modifications. These ions were m/z 147 for deoxy-hexose, m/z 163 for Hex, m/z 204 for HexNAc, m/z 274 for the dehydrated sialic acids, and m/z 292 for sialic acid. A glycopeptide, GlcNAc-NIT (GlcNAc is N-acetylglucosamine), at m/z 550 was also selected as a characteristic N-linked glycopeptide product ion, based on previous findings which demonstrate that the cleavage producing this ion occurs readily during collisional activation [22]. No glycopeptides were detected by precursor ion scan of m/z 147, 274, and 292. The results for the precursor ion scans of m/z 163, 204, and 550 are shown in Figure 5a, b, and c, respectively. In each spectrum, ions that can be assigned to glycopeptides are labeled with a dot, and the compositions of those ions are provided in Table 1.

Neutral Loss Scan

Another method of selective detection of glycopeptide ions is neutral loss scanning. Application of this scan
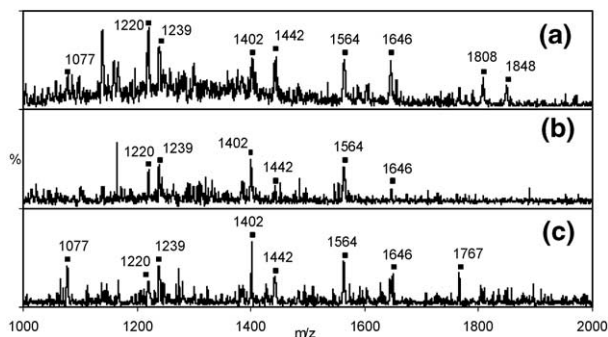
function has been reported in detection of sulfate-containing compounds by monitoring a neutral loss of 80 [23]. Since the oLHα glycoproteins have been found to be sulfated [20], it is worthwhile exploring neutral loss scan for detection of these sulfated glycoforms. Two neutral losses were selected, 80 Da for the singly charged ions and 40 Da for the doubly charged. Results are shown in Figure 6. Again, each ion that can be assigned to a glycopeptide composition is labeled with a dot, and the detected sulfated glycopeptides are also listed in Table 1.

Discussion

Method Design

While profiling glycopeptide compositions by MS has been previously reported, most examples found in the literature utilized tryptic digests or other site-specific enzymes [5, 8, 12, 13, 24]. These digestions facilitate glycopeptide analysis, because the peptide sequences produced are known a priori. However, a few examples exist where nonspecific proteases were effectively used for glycoprotein profiling [7, 10]. In the research de-
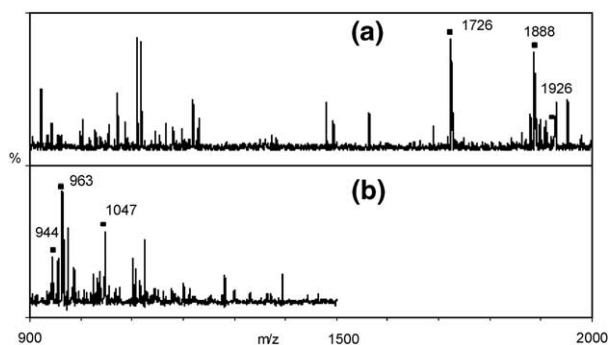

Figure 6. Neutral loss scans on the triple quadrupole mass spectrometer for detection of sulfated glycopeptides. (**a**): Neutral loss of 80 Da for singly charged ions. (**b**): Neutral loss of 40 Da for doubly charged ions. Ions with a square are glycopeptides identified by neutral loss scan. Spectra were smoothed using Micromass MassLynx 3.5.

scribed herein, a nonselective enzyme, proteinase K, was chosen. The obvious disadvantage of such an approach is that the glycopeptide sequence must be inferred from the mass spectra or other means. Additionally, non-specific proteases could produce several glycopeptides containing the same glycosylation site [7, 10]. This could complicate the mass spectra and make data interpretation more difficult. In this work, the nonspecific protease did not complicate data interpretation. The Edman degradation confirmed that one major glycopeptide was produced, and both Edman sequencing and mass spectral methods revealed the amino acid composition of the glycopeptide.

A substantial benefit from using the nonselective enzyme is that a significantly smaller peptide is present. In this case, Asn-Ile-Thr, the peptide that carries glycosylation, weighs just 346 Da. The tryptic fragment for this digestion would amass 1238 Da. So for detectors with an upper mass limit of 2000 Da, *no* glycopeptides would be detected as singly charged ions because all of the ions would weigh more than 2000 Da. (The carbohydrate portion must contain the trimannosyl core, which weighs approximately 900 Da on its own; when that is coupled to a 1238 Da peptide, the final mass is 2138 Da).

While use of a non-selective enzyme was somewhat novel in this work, so too, was the use of a custom-designed program to determine compositional assignments. A similar web-based automated assignment program GlycoMod has been used to assign carbohydrate compositions of glycopeptides in other research [12, 13, 24]. This program may be found at http://us.expasy.org/tools/glycomod/ (while the web-based program is very useful for profiling glycopeptide composition, we prefer the additional flexibility of a home-built program). Automated assignments of the glycopeptide ions not only reduced the amount of time involved in analyzing the data, it also improved the quality of the data. For example, *m/z* 1910 has two composition assignments that are within 20 ppm of the *m/z* value obtained on the qTOF mass spectrometer. One of these two possible assignments would most likely have been overlooked, if assigning the composition manually. Likewise, if only one composition is identified by the algorithm, then one can be assured that the correct composition has been determined; and other possible, isobaric compositions do not likely exist.

Glycopeptides Identified

By comparing the data in Figure 1, several important issues about glycopeptide profiling by mass spectrometry can be addressed. First, one can address the question of whether the *relative abundances* for each glycopeptide seen in the mass spectra correlate to the actual concentrations of these glycoforms present in the glycoprotein. Previously, researchers have demonstrated that the relative abundances of glycopeptides *with similar structure and sizes* can be used to infer the

amounts of each glycopeptide present [25]. This approach has been used in other works to obtain quantitative information about glycopeptides with similar carbohydrate structure [8, 10]. For the oLH sample used herein, glycopeptides with *dissimilar* structure and sizes are present. Specifically, the oLH samples contain neutral, sulfated, and sialylated carbohydrates.

Using previously published quantitative data about each of the glycoforms present on oLH, we demonstrate herein that quantitative assignments *cannot* be made for glycans with disparate functional groups. First, neutral glycoforms dominate each of the spectra in Figure 1; however previous studies on this glycoprotein indicate that the neutral glycoforms comprise only 13% of the total carbohydrate content at this site [20]. None of the mass spectrometers detected carbohydrates containing sialic acid. The sialic acid containing compounds have been previously reported to be present on 5% of the glycoforms, based on radiolabeling and fractionation by SDS-PAGE [20]. Finally, the predominant glycoform identified previously is a sulfated nonasaccharide [21]. This appears as *m/z* 1045 as a small peak in the qTOF spectrum in Figure 1a, and it is absent in the other two spectra in Figure 1. However, this ion should have been more abundant than all of the neutral glycoforms *combined*, if its ionization were not suppressed. When profiling glycoproteins by mass spectrometry, one must be aware of the strong compositional bias for detecting neutral glycoforms over sulfated and sialylated glycoforms.

While quantitative analysis of the glycopeptides is not possible with *this* MS data, each of the three mass spectrometers produced data that was qualitatively similar. (See Table 1). Most of the glycopeptide compositions identified were present in MS data from each of the three mass spectrometers. One exception includes *m/z* 1726, which appears in the qTOF and triple quadrupole data (Figure 1a and b), but not in the ion trap data (Figure 1c). By comparing the spectra in Figure 1, it is possible that the poor signal-to-noise in the ion trap spectra contributed to this ion not being readily discernable from the baseline.

One other qualitative difference among the spectra is that the ion trap data contained more sodiated glycopeptide ions than the triple quadrupole or qTOF data. Examples include *m/z* 1830, 1871, 1932, and 1992. (See Table 1). In these cases, the protonated forms of the glycopeptides were detected in the qTOF and triple quadrupole mass spectrometers, but not the sodiated adducts.

By comparing the MS data to the Parent ion scan and Neutral loss scan data in Table 1, the standard MS experiment was more successful at detecting the most glycopeptide components. The scanning experiments could not be used to identify *additional* components that were not detected in the full MS data. In addition, the signal-to-noise for the scanning experiments was quite low, and several unidentifiable peaks appeared, particularly in the neutral loss scans. While others have

proposed these methods as the best way to identify glycopeptides [22], for the samples studied herein these methods were not as effective at identifying all the glycoforms present as the standard MS data. (The scanning experiments, however, are useful because they can provide structural information about the compounds.)

In summary, each of the three mass spectrometers utilized herein provided approximately the same amount of qualitative information about the glycopeptide composition. Several neutral glycopeptides were identified in each of the three mass spectrometers. These carbohydrates have not been identified by previous methods, because the neutral carbohydrate content at this site is too complex for separation and characterization by other methods [20]. In terms of assigning the *m/z* values obtained from the mass spectrometer to a glycopeptide composition, the qTOF data is perhaps most useful because of its increased degree of mass accuracy. However, useful structural assignments could be made on data originating from all three mass spectrometers. In terms of sample consumption, the qTOF (with nanospray source) consumed approximately 0.05 μg of sample, while the triple quadrupole MS (with nanospray source) used 0.5 μg of sample, and the quadrupole ion trap required 10 μg of sample, for the data in Figure 1a, b, and c, respectively. A clear sensitivity advantage is given to the qTOF mass spectrometer.

Conclusion

Several different mass spectrometric methods were compared in order to define the most effective strategy for glycopeptide analysis. In these studies, standard MS data revealed more glycoforms present than precursor ion scanning or neutral loss scanning experiments. Of the three instruments compared in the study, the qTOF mass spectrometer required the smallest sample consumption, but all three instruments were useful in profiling the glycopeptide composition.

This work also demonstrated that nonselective enzymes may produce useful glycopeptide fragments. Nonselective enzymes were preferred to trypsin in *this* study because one major glycopeptide was produced, and this peptide had a lower molecular weight than the expected tryptic fragment. However, these fortuitous results may not be typical for other glycoproteins.

Finally, this is the first example of site-specific carbohydrate profiling of luteinizing hormone that describes the neutral glycoforms present. These studies revealed that the carbohydrate composition at the Asn[56] site of LH is quite heterogeneous—with more than 20 unique glycoforms detected. It is known that the carbohydrates present at this site effect receptor binding affinity, so understanding the great variety in the composition of these carbohydrates is important in studying ligand binding interactions.

References

1. Harvey, D. J. Identification of Protein-Bound Carbohydrates by Mass Spectrometry. *Proteomics* **2001,** *1,* 311–328.
2. Rouse, J. C.; Vath, J. E. On-the-Probe Sample Cleanup Strategies for Glycoprotein-Released Carbohydrates Prior to Matrix-Assisted Laser Desorption-Ionization Time-of-Flight Spectrometry. *Anal. Biochem.* **1996,** *238,* 82–92.
3. Scanlan, C. N.; Pantophlet, R.; Wormald, M. R.; Saphire, E. O.; Stanfield, R.; Wilson, I. A.; Katinger, H.; Dwek, R. A.; Rudd, P. M.; Burton, D. R. The Broadly Neutralizing Anti-Human Immunodeficiency Virus Type 1 Antibody 2G12 Recognizes a Cluster of $\alpha\rightarrow2$ Mannose Residues on the Outer Face of gp120. *J. Virol.* **2002,** *76,* 7306–7321.
4. Harvey, D. J. N-(2-Diethylamino)Ethyl-4-Aminobenzamide Derivative for High Sensitivity Mass Spectrometric Detection and Structure Determination of N-Linked Carbohydrates. *Rapid Commun. Mass Spectrom.* **2000,** *14,* 862–871.
5. Mechref, Y.; Novotny, M. V. Structural Investigations of Glycoconjugates at High Sensitivity. *Chem. Rev.* **2002,** *102,* 321–369.
6. Huddleston, M. J.; Bean, M. F.; Carr, S. A. Collisional Fragmentation of Glycopeptides by Electrospray Ionization LC/MS and LC/MS/MS: Methods for Selective Detection of Glycopeptides in Protein Digests. *Anal. Chem.* **1993,** *65,* 877–884.
7. Juhasz, P.; Martin, S. A. The Utility of Nonspecific Proteases in the Characterization of Glycoproteins by High-Resolution Time-of-Flight Mass Spectrometry. *Int. J. Mass Spectrom.* **1997,** *169/170,* 217–230.
8. Zhu, X.; Borchers, C.; Bienstock, R. J.; Tomer, K. B. Mass Spectrometric Characterization of Glycosylation Pattern of HIV-gp120 Expressed in CHO Cells. *Biochem.* **2000,** *39,* 11194–11204.
9. Dell, A.; Morris, H. R. Glycoprotein Structure Determination by Mass Spectrometry. *Science* **2001,** *291,* 2351–2356.
10. An, H. J.; Peavy, T. R.; Hedrick, J. L.; Lebrilla, C. B. Determination of *N*-Glycosylation Sites and Site Heterogeneity in Glycoproteins. *Anal. Chem.* **2003,** *75,* 5628–5673.
11. Carr, S. A.; Huddleston, M. J.; Bean, M. F. Selective Identification and Differentiation of N- and O-Linked Oligosaccharides in Glycoproteins by Liquid Chromatography-Mass Spectrometry. *Protein Sci.* **1993,** *2,* 183–196.
12. Ma, S.; Hill, K. E.; Burk, R. F.; Caprioli, R. M. Mass Spectrometric Identification of N- and O-Glycosylation Sites of Full-Length Rat Selenoprotein P and Determination of Selenide-Sulfide and Disulfide Linkages in the Shortest Isoform. *Biochem.* **2003,** *42,* 9703–9711.
13. Zhen, Y.; Caprioli, R. M.; Staros, J. V. Characterization of Glycosylation Sites of the Epidermal Growth Factor Receptor. *Biochem.* **2003,** *42,* 5478–5492.
14. Matzuk, M. M.; Boime, I. Site-Specific Mutagenesis Defines the Intracellular Role of the Asparagine-Linked Oligosaccharides of Chrionic Gonadotropin β-Subunit. *J. Biol. Chem.* **1988,** *263,* 17106–17111.
15. Butnev, V. Y.; Gotschall, R. R.; Butnev, V. Y.; Baker, V. L.; Moore, W. T.; Bousfield, G. R. Hormone-Specific Inhibitory Influence of α-Subunit Asn[56] Oligosaccharide on in Vitro Subunit Association and FSH Receptor Binding of Equine Gonadotropins. *Biol. Repro.* **1998,** *58,* 458–469.

16. Matsuk, M. M.; Keene, J. L.; Boime, I. Site Specificity of the Chorionic Gonadotropin N-Linked Oligosaccharides in Signal Transduction. *J. Biol. Chem.* **1989,** *264,* 2409–2414.

17. Liu, W.-K.; Ascoli, M.; Ward, D. N. Ovine Lutropin Subunit Isolation: Comparison of Salt Precipitation and Countercurrent Distribution Procedures. *J. Biol. Chem.* **1977,** *252,* 5274–5279.

18. Bousfield, G. R.; Baker, V. L.; Gotschall, R. R.; Butnev, V. Y. Carbohydrate Analysis of Glycoprotein Hormones. *Methods* **2000,** *21,* 15–39.

19. Packer, N. H.; Lawson, M. A.; Jardine, D. R.; Redmond, J. W. A General Approach to Desalting Oligosaccharides Released from Glycoproteins. *Glycoconj. J.* **1998,** *15,* 737–747.

20. Weisshaar, G.; Hiyama, J.; Renwick, A. G. C. Site-Specific N-Glycosylation of Ovine Lutropin. Structural Analysis of One- and Two-Dimensional Proton NMR Spectroscopy. *Eur. J. Biochem.* **1990,** *192(741),* 751.

21. Thotakura, N. R.; Blithe, D. L. Glycoprotein Hormones: Glycobiology of Gonadotrophins, Thyrotrophin, and Free α-Subunit. *Glycobiology* **1995,** *5,* 3–10.

22. Ritchie, M. A.; Gill, A. C.; Deery, M. J.; Lilley, K. Precursor Ion Scanning for Detection and Structural Characterization of Heterogeneous Glycopeptide Mixtures. *J. Am. Soc. Mass Spectrom.* **2002,** *13,* 1065–1077.

23. Yagami, T.; Kitagawa, K.; Futaki, S. Analysis of Sulfated Tyrosine-Containing Peptides by Liquid Secondary-Ion Mass Spectrometry with Constant Neutral-Loss (80 amu) scanning. *Anal. Sci.* **1995,** *11,* 1025–1028.

24. Hakansson, K.; Emmett, M. R.; Marshall, A. G.; Davidsson, P.; Nilsson, C. L. Structural Analysis of 2D-Gel-Separated Glycoproteins from Human Cerebrospinal Fluid by Tandem High-Resolution Mass Spectrometry. *J. Proteome Res.* **2003,** *2(6),* 581–588.

25. Liu, T.; Li, J. D.; Zeng, R.; Shao, X. X.; Wang, K. Y.; Xia, Q. C. Capillary Electrophoresis-Electrospray Mass Spectrometry for the Characterization of High-Mannose-Type N-Glycosylation and Differential Oxidation in Glycoproteins by Charge Reversal and Protease/Glycosidase Digestion. *Anal. Chem.* **2001,** *73,* 5875–5885.