research articles Journal of proteome research

Site Determination of Protein Glycosylation Based on Digestion with Immobilized Nonspecific Proteases and Fourier Transform Ion Cyclotron Resonance Mass Spectrometry

Brian H. Clowers,[†] Eric D. Dodds,[†] Richard R. Seipert,[†] and Carlito B. Lebrilla*[,†,‡]

Department of Chemistry and School of Medicine, University of California, Davis, Davis, California 95616

Received May 25, 2007

An improved method for site-specific characterization of protein glycosylation has been devised using nonspecific digestion with immobilized pronase combined with Fourier transform ion cyclotron resonance mass spectrometry (FTICR−MS). This procedure was demonstrated using ribonuclease B (RNase B) and κ-casein (κ-csn) as representative N-linked and O-linked glycoproteins, respectively. Immobilization of the pronase enzymes facilitated their removal from the glycopeptide preparations, and was found to prevent enzyme autolysis while leaving the proteolytic activities of pronase intact. Increased digestion efficiency, simplified sample preparation, and reduced sample complexity were consequently realized. To supplement this technique, a refined glycopeptide search algorithm was developed to aid in the accurate mass based assignment of N-linked and O-linked glycopeptides derived from nonspecific proteolysis. Monitoring the progress of glycoprotein digestion over time allowed detailed tracking of successive amino acid cleavages about the sites of glycan attachment, and provided a more complete protein glycosylation profile than any single representative time point. This information was further complemented by tandem MS experiments with infrared multiphoton dissociation (IRMPD), allowing confirmation of glycopeptide composition. Overall, the combination of immobilized pronase digestion, time course sampling, FTICR−MS, and IRMPD was shown to furnish an efficient and robust approach for the rapid and sensitive profiling of protein glycosylation.

Keywords: protein glycosylation • N-glycosylation • O-glycosylation • electrospray • FTICR−MS • IRMPD • immobilized enzyme digestion

Introduction

Affecting roughly 50% of the eukaryotic proteome, protein glycosylation is one of the most common forms of post-translational modification.[1] Not surprisingly, the range of structural complexity and frequency of protein glycosylation are in accordance with their resulting biological roles,[2−4] making their characterization of imminent importance. Despite the ubiquity of glycosylation, analysis of these modifications remains a formidable analytical challenge due to the range of glycan structures combined with the variation in attachment site.[5]

Although other forms of protein glycosylation have been reported in the literature,[4,6] the two primary types of protein glycosylation are those found attached to asparagine residues, termed N-linked, and those connected to serine or threonine residues, referred to as O-linked. Though O-linked carbohydrates may be found at any given serine or threonine residue, N-linked glycosylation is sensitive to the two amino acids following the asparagine residue in the protein sequence. The known consensus sequences for N-linked glycosylation are NXS, NXT, and the less common NXC, where X is any amino acid other than proline.[4,6] Building upon a common pentasaccharide core, N-linked glycans can be classified into three general types: high-mannose, complex, and hybrid.[5] While these three classes give rise to a respectable array of glycan structures, those based on the eight known O-linked glycan cores encompass a wider structural range.[5]

Knowledge of the possible glycan core structures and of the amino acid residues they may modify is useful; however, these considerations alone do not fully describe the diversity of protein-modifying glycans. In addition to glycan class, core type, and attachment site, the overall compositions of the glycans and their heterogeneity at a particular attachment site must be determined to thoroughly characterize protein glycosylation.

A sizable collection of analytical techniques are often necessary to probe glycan attachment site, composition and structure.[7−10] As it affords the benefits of selectivity, sensitivity, and specificity, mass spectrometry has been increasingly used to address the complexity of protein glycosylation.[10−12] Combined with mass spectrometry, enzymatic and chemical methods of glycan release provide an effective means to probe the range of glycan structures present, but often fail to elucidate

* To whom correspondence should be addressed. Tel, (530) 752-6364; Fax, (530) 754-5609; E-mail, cblebrilla@ucdavis.edu.
[†] Department of Chemistry.
[‡] School of Medicine.

Published on Web 09/07/2007
10.1021/pr070317z CCC: $37.00

the site of glycosylation and the heterogeneity of glycosylation at a specific site.[10,13−15] To determine glycosylation site, methods designed for mapping of other post-translational modifications have been employed, though in many cases determining glycan composition using such methods remains elusive.[13,16,17] Current proteomic strategies for exploring protein glycosylation often use trypsin-based digestion methods followed by liquid chromatography prior to analysis using mass spectrometry.[14,15,18,19] Trypsin based methods offer the advantage of being directly portable from standard protein analysis approaches. Unfortunately, analysis of the glycosylated peptide contingent in the resulting mixture suffers from the charge suppression of glycans in the presence of tryptic peptides, and is further complicated by increased chemical heterogeneity induced by reduction and alkylation steps. These issues aside, the activity of trypsin is often hindered by post-translational modifications,[20] further minimizing the effectiveness of trypsin-based glycoprotein analysis techniques.

An alternative method involves the use of nonspecific protease digestion in order to circumvent many of the challenges associated with traditional tryptic methods.[21−23] Possessing an extensive range of activities, nonspecific proteases are capable of breaking down virtually all proteins into their individual amino acid components. The proteolytic capability of nonspecific proteases often extends to both native and denatured proteins, thus minimizing or eliminating the requirement of reduction and alkylation steps inherent to most tryptic digestion methods. One appealing agent for nonspecific digestion of glycoproteins is pronase, which consists of various serine and metallo-endopeptidases along with exopeptidases capable of cleaving amino acid residues from either the C- or N-terminus.[24] Because hydrolysis of amide bonds adjacent to a site of glycosylation is sterically hindered by the glycan itself, the use of pronase with appropriately controlled digestion time provides the means to preserve the site of glycosylation and produce glycopeptides of varying lengths.[25] The primary disadvantages of this method stem from the potential for protease−protease digestion, which has the unfortunate twofold effect of complicating the sample with protease-derived peptides as well as reducing the activity of the protease mixture overall.[21,22] Extended digestion periods of up to 48 h are required for some glycoproteins due to a reduction in pronase activity. These limitations, however, may be greatly mitigated by immobilizing the pronase enzymes, thus allowing production of comparatively clean glycopeptide preparations with reduced digestion times.

Despite the significant influence of protein glycosylation on biological function, carbohydrates as post-translational modifiers remain most often neglected by current protein analysis approaches. This is primarily due to the formidable analytical challenges associated with the characterization of protein glycosylation. Recently, Temporini et al. demonstrated a novel on-line method that uses an immobilized pronase reactor to generate N-linked glycopeptides.[26] Their technique also detailed the use of porous graphitized carbon enrichment and normal phase liquid chromatography following enzyme digestion, with assignment of glycopeptides based upon tandem mass spectrometry experiments and measurement of average m/z values.

Here, we describe a key refinement of an earlier technique to probe glycan attachment site and composition based on digestion of target glycoproteins with immobilized nonspecific proteases. This approach utilizes the highly accurate m/z measurements afforded by Fourier transform ion cyclotron resonance mass spectrometry and temporal sampling of the glycoprotein digestion. This procedure maximizes throughput while minimizing chemical interference and makes use of reagents and conditions directly compatible with mass spectrometry ionization sources. In addition to being applicable to both N- and O-linked glycoproteins, this technique is demonstrated to allow control of the resulting glycopeptide sequence length and virtually eliminate protease autolysis products.

Experimental

Chemicals and Materials. Bovine ribonuclease B (RNase B; SwissProt accession number P61823), bovine κ-casein (κ-csn; SwissProt accession number P02668), pronase E, ammonium acetate, sodium azide, sodium phosphate, hydrochloric acid, formic acid, and ethanolamine were obtained from Sigma-Aldrich (St. Louis, MO). Cyanogen bromide activated sepharose 4B beads (average particle diameter, 90 μm) were acquired from GE Healthcare (Piscataway, NJ). Solutions for electrospray were comprised of a 50:50 (vol/vol) mixture of acetonitrile and deionized water with 0.1% formic acid.

Pronase Immobilization. To reconstitute the lyophilized medium and remove manufacture-related contaminants, 150 mg portions of CNBr activated sepharose 4B beads were suspended in 1 mL of 1 mM hydrochloric acid in a 1.5 mL screw cap micro-centrifuge tube. These suspensions were gently agitated for approximately 15 min prior to centrifugation and subsequent removal of the supernatant. Each bead preparation was then rinsed with ~800 μL of 100 mM total phosphate buffer (pH 7.4), centrifuged, and removed of supernatant. This wash step was repeated for a total of four cycles. Following the final wash cycle, each batch of beads was treated with 500 μL of the 100 mM phosphate buffer, plus 100 μL of 10 mg/mL Pronase E dissolved in deionized water. These mixtures were gently agitated for 24 h at room temperature to ensure quantitative covalent coupling of proteases to the beads via reaction of primary amino groups with the CNBr activated sepharose support.[27] After centrifugation and supernatant removal, four washings were performed using ~800 μL of the phosphate buffer to ensure removal of any excess enzyme. After removing the supernatant from the last wash cycle, 500 μL of 1 M aqueous ethanolamine (pH 9.0) was added to block all remaining active sites on the beads. This reaction was allowed to proceed with mild agitation for 2 h at room temperature. To remove the residual ethanolamine, another series of four wash cycles was performed; however, these washes were done using ~800 μL of 100 mM NH_4OAc buffer (pH 7.2). If glycoprotein digestion was not to be carried out immediately, the pronase coupled beads were suspended in 100 mM NH_4OAc containing 0.05% (w/v) NaN_3 and stored at 4 °C. It should be noted that prior to protease coupling and active site blocking, reagents containing primary amino groups (e.g., tris buffered solutions) may interfere with successful enzyme immobilization. Before and after each use, the immobilized pronase beads were washed thoroughly using the appropriate buffer for enzyme digestion (100 mM NH_4OAc) or storage (100 mM NH_4OAc with 0.05% NaN_3).

Glycoprotein Digestion. Stock glycoprotein solutions were prepared at a concentration of approximately 50 μM in the ammonium acetate buffer. Typically, 100 μL of glycoprotein stock solution and 200 μL ammonium acetate buffer were combined with an aliquot of pronase coupled beads that had been prepared on the scale described above. Combined with gentle agitation, the digestion was allowed to proceed at 37

°C. Glycoprotein digestion mixtures were centrifuged and the supernatants subsampled at various time points ranging from 90 min to 24 h to monitor the progress of the reaction. Great care was taken to avoid transferring any pronase-coupled beads along with the supernatant. Glycopeptide solutions were then stored at −20 °C prior to analysis.

All glycopeptide supernatants were diluted into an electrospray-compatible solvent system for direct MS analysis. All digests were diluted at least 10-fold to a final composition of 50% aqueous acetonitrile with 0.1% formic acid (by volume). When desired, the immobilized pronase suspension was reused following another round of washing to remove any residual materials from the previous digestion.

Instrumentation and Analysis. Nanoelectrospray was performed using a Picoview nanoESI stage (New Objective, Woburn, MA) held at 1800−2400 V with respect to the sample cone interface. Ions passed through the 390 micron aperture of the atmospheric pressure interface (Z-spray ion source, Waters, UK) into the hexapole accumulation region of a 9.4 T Fourier transform ion cyclotron resonance mass spectrometer (IonSpec QFT, Irvine, CA). Ions were externally accumulated in the hexapole for up to 4 s prior to injection to the ICR cell by an RF-only quadrupole ion guide. When appropriate, candidate m/z values for tandem MS were selected within the ICR cell using stored-waveform inverse Fourier transform (SWIFT) isolation prior to infrared multiphoton dissociation (IRMPD) using a 10.6 μm CO_2 laser (Parallax Laser Inc., Waltham, MA). To ensure the isolated ion cloud received maximum exposure during each IR laser pulse, the laser beam radius was enhanced to 0.5 cm using a beam expander (Synrad Laser, Mukilteo, WA). The expanded beam was then transmitted into the ICR cell through a BaF_2 window (Bicron Corp., Newbury, OH). The degree of fragmentation was optimized by varying the IRMPD laser pulse between 500 and 2000 ms. Glycopeptide spectra were externally calibrated using the IonSpec Omega data station, and each spectrum was default thresholded and deisotoped using the IonSpec PeakHunter software. Monoisotopic peak lists were exported in text file format for further processing.

Glycopeptide Search Routine. The user interface and glycopeptide search algorithm was designed and written using IGOR Pro 5 (Wavemetrics, Inc., Lake Oswego, OR). Though written for use with IGOR Pro, the corresponding routines are available upon request and may be easily adapted to other programming languages. To identify candidate glycopeptide compositions, the program requires the user to load both a FASTA protein sequence file and an experimental monoisotopic peak list. The *in silico* digestion scheme identifies all potential glycosylation sites for either N- or O-linked glycosylation based on the FASTA sequence. All serine and threonine residues are recognized as potential O-linked glycosylation sites, whereas asparagine residues occurring as part of the consensus sequences NXT, NXS, and NXC (where X represents any amino acid aside from proline) are identified as potential sites for N-linked glycosylation. Because nonspecific protease digestion produces successively smaller peptide sequences as digestion time increases, the *in silico* digestion generates a list of peptides containing the possible glycosylation site(s) with lengths ranging from a user specified maximum down to a single amino acid residue. After generating a list of candidate peptide moieties, the user is prompted to provide search constraints outlining the number and type of monosaccharide residues to be coupled with the *in silico* peptide list. Once the candidate peptides have been generated and carbohydrate search parameters have been defined, the user is then prompted to specify a mass tolerance constraint prior to searching the experimental peak list for possible matches. If a combination of candidate peptides and possible glycans possesses a calculated m/z ratio that matches an experimental m/z ratio within the tolerance provided by the user, the peptide sequence and possible glycan composition is entered into a report table for further review. While this routine bears some similarities to the GlycoX software tool previously reported by this laboratory,[23] the IGOR Pro version dispenses with the in-house spectral processing tools in favor of directly searching thresholded, deisotoped peak lists produced by commercial software (in this case, IonSpec PeakHunter). Moreover, this newer routine features a convenient graphical user interface, as opposed to the MATLAB command line driven operation of GlycoX.

Because a large number of candidate peptides and glycans may be generated for a relatively simple system, it is imperative that the user provide experimentally valid constraints and mass tolerances on the search algorithm to minimize the unavoidable occurrence of false positive matches. It should be noted that, in all circumstances, this approach to glycopeptide analysis only provides a list of potential glycopeptides, with the user being ultimately responsible for the correct assignment of glycopeptides. Furthermore, the use of complementary information related to the glycoprotein structure, independently determined glycan compositions, or putative sites of glycosylation can greatly minimize the number of false positives.

Results and Discussion

Analysis of Bovine RNase B. Bovine RNase B (P61823, EC 3.1.27.5) is a well-characterized glycoprotein containing a single site of glycosylation at ^{60}N. This site is known to possess a heterogeneous population of high-mannose type oligosaccharides containing between 5 and 9 mannose residues.[9] The array of *N*-glycan structures found on RNase B make this protein an ideal candidate to illustrate the ability of immobilized nonspecific proteases for rapid site specific determination of glycosylation.

An attractive feature of immobilized nonspecific protease digestion for glycan site determination is the ability to observe the degradation of the peptide moiety as a function of time. Figure 1 illustrates the transition of RNase B glycopeptide sequence length over the course of a 24 h digestion cycle. As expected, the relative levels of the observed m/z ratios shift toward lower mass ions as digestion time increases. Overall, the most intense peaks observed after 1.5 h of pronase digestion of RNase B were found between m/z 1600 and 2500, whereas after 6 h of digestion the range of glycopeptides occupied the m/z range between 1400 and 2200. To highlight the shift in mass range as a function of digestion time, the peak intensities above m/z 2200 have been expanded 4-fold in Figure 1. Also worthy of note is the similarity between the spectra obtained for 6, 12, and 24 h, indicating completion of digestion after ~6 h. Importantly, no new species were observed to appear after prolonged incubation, indicating that no detectable degree of protease autolysis occurred.

Using the peak abundance information found at each sampled time point and the sequential degradation mechanism of immobilized pronase, the cleavage of amino acid residues from the glycopeptide may be tracked. For example, the peak located at m/z 1968.8115 after 1.5 h of digestion corresponds to an *N*-glycan consisting of 2 *N*-acetylhexosamine (HexNAc)

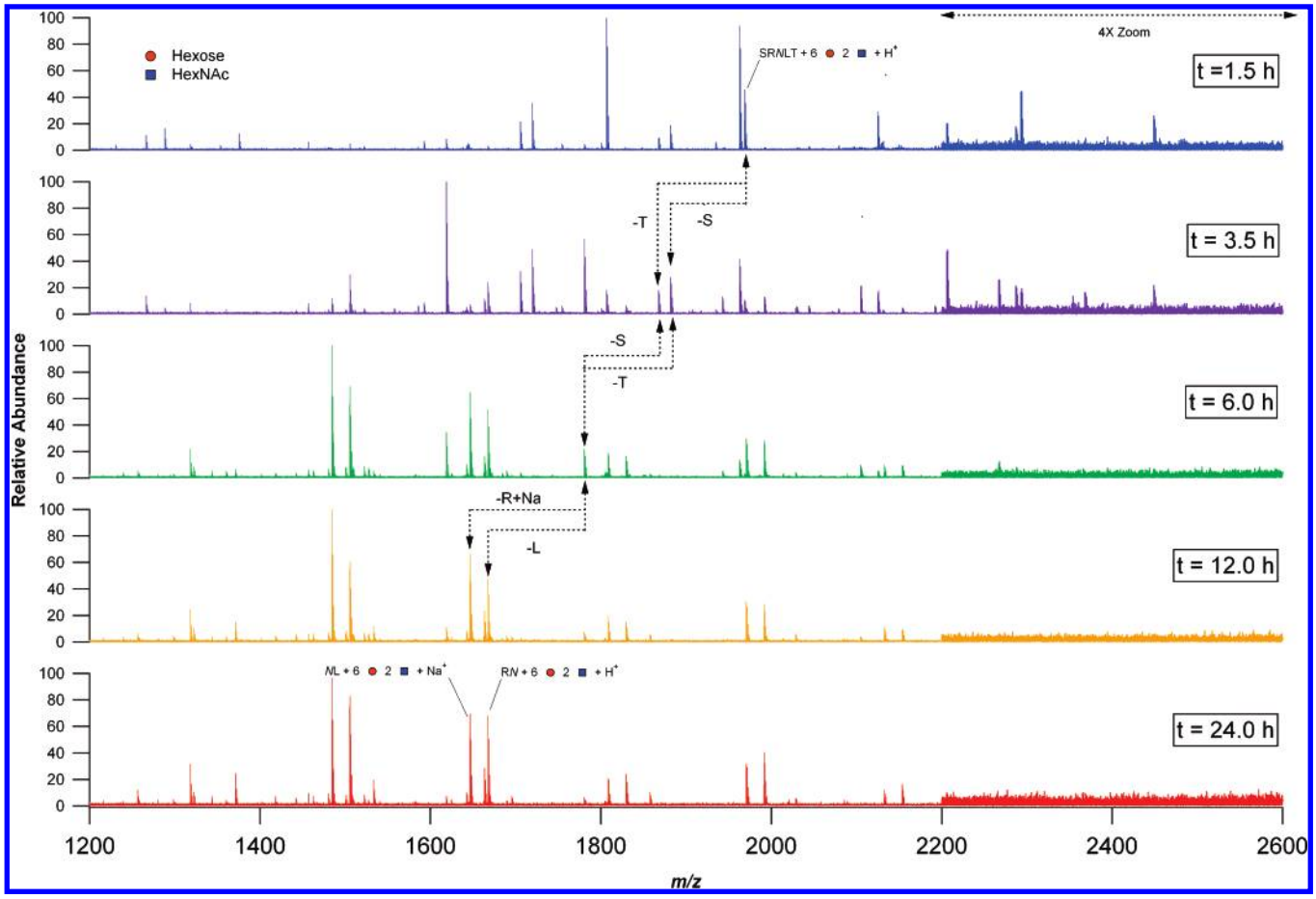

Figure 1. Positive mode nanoESI FTICR mass spectra of RNase B N-linked glycopeptides resulting from time course digestion using bead-immobilized pronase. As the majority of species observed were singly charged, these traces illustrate the glycopeptides distribution shift from larger species (higher *m*/*z*) to smaller glycopeptides (lower *m*/*z*). Combining the information obtained at each time point allows a comprehensive picture of glycosylation site location to be developed.

and 6 hexose (Hex) residues attached to the asparagine residue of the peptide SRNLT (5 ppm mass error). The successive loss of amino acids from both the N and C termini is traced in Figure 1 to the dipeptides NL and RN, respectively, which become major products after 6 h of digestion. Fully annotated glycopeptide spectra corresponding to 1.5 and 24 h of digestion are provided in Figure 2.

Careful examination of the glycopeptides found at each time point revealed a trend linking ionization behavior to the changing amino acid composition of the observed glycopeptides. Specifically, whether the glycopeptide observed was protonated or sodiated was dependent on the presence or absence of a basic amino acid residue. As seen in Figure 2, all RNase B derived glycopeptides containing an arginine residue were observed as protonated ions, whereas those lacking an arginine residue were observed primarily as sodium adducts. This observation is directly related to the relative gas phase affinities of the different glycopeptide species for protons or sodium ions. Thus, although the gas-phase proton affinity of arginine containing glycopeptides is very high, the production of the sodiated species evidently becomes favored in the absence of a highly basic amino acid residue.

In addition to illustrating the range of glycopeptide lengths observed over the course of digestion and the corresponding ionization behavior of the glycopeptide products, Figure 2 emphasizes two significant advantages of using immobilized pronase as a means to identify sites of glycosylation. The first of these is the increase in enzymatic activity (requiring only 6 h to fully digest ribonuclease B) relative to glycoprotein digestion using free pronase (which has typically been performed over the course of 24−48 h).[21,22,28] The second advantage afforded by this method is the apparent absence of pronase autolysis products, thus simplifying spectral interpretation. As evident from the assignments made in Figure 2, the vast majority of peaks observed using this method were confidently attributed to the target glycoprotein.

Tables 1 and 2 summarize the observed peak assignments made at 1.5 and 24 h and provide the observed glycopeptide masses, compositions, ion types, and mass errors for each assigned glycopeptide. These assignments were made by constraining the search algorithm to identify possible protonated or sodiated high mannose glycopeptides found within 25 ppm of the theoretical value. The root-mean-square errors for the 1.5 and 24 h glycopeptide assignments were 3 and 8 ppm, respectively. It should be noted that relaxing the *m*/*z* search tolerances and the compositional glycan constraints drastically increases the number of false positives. Although the information provided by time course measurements of glycoprotein digestion greatly improves the confidence in assigning glycosylation position, these data are directly complemented by tandem mass spectrometry experiments.

The eight amino acid peptide that includes the glycosylation site of RNase B at ^{60}N is ^{57}KSRNLTKD^{64}. Given the digestion mechanism of pronase, the possibility exists that either of the

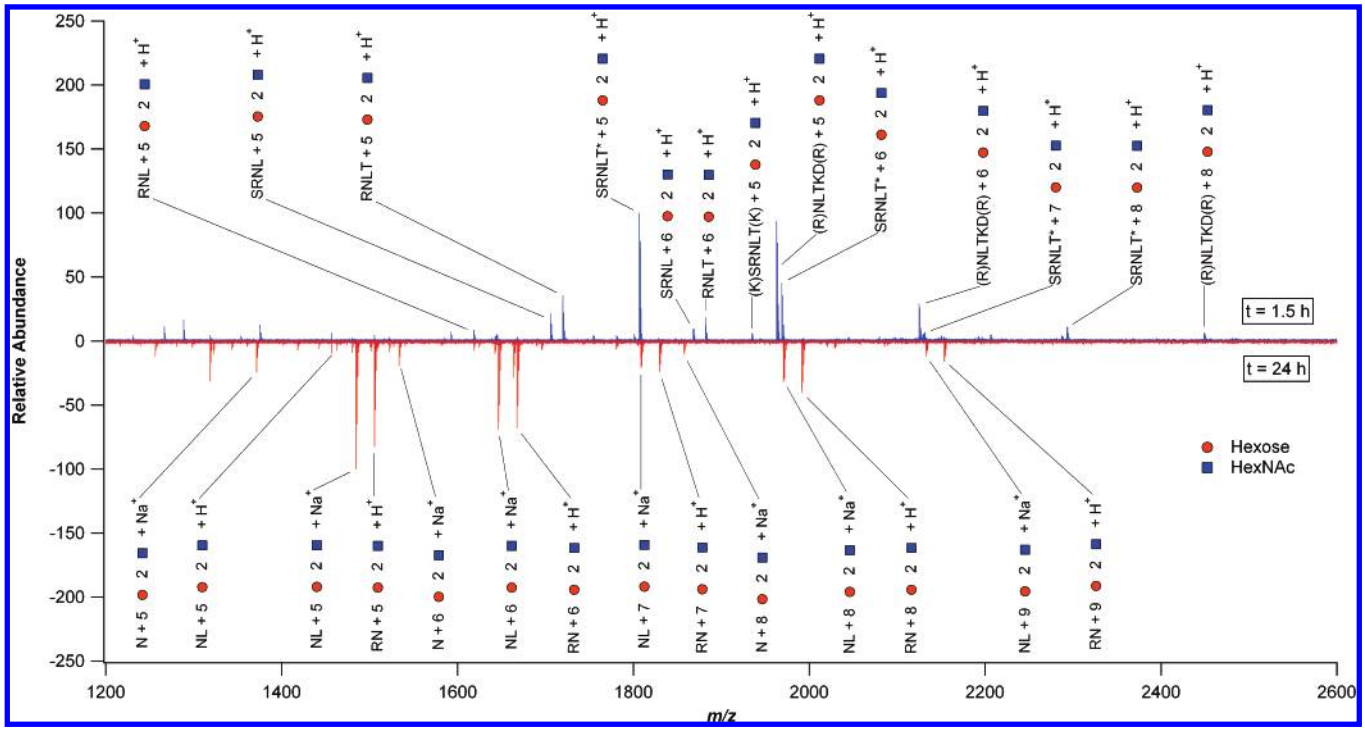

Figure 2. Using the high mass accuracy provided by FTICR combined with a straightforward *in silico* digestion and glycan identification algorithm, these two spectra illustrate the comprehensive peak assignment possible across a broad range of immobilized pronase digestion times. The upper trace corresponds to the glycopeptides distribution observed after 1.5 h of digestion and the lower trace after 24 h.

Table 1. Glycopeptides Derived from RNase B Following 1.5 h Incubation with Bead-Immobilized Pronase

GP *m*/*z*	peptide[a]	site	glycan	ion type	Δ*m*/*z* (Da)[b]
1618.6719	^{59}RNL^{61}	^{60}N	$HexNAc_2$ Hex_5	$[M+H]^+$	0.0032
1705.7055	^{58}SRNL^{61}	^{60}N	$HexNAc_2$ Hex_5	$[M+H]^+$	0.0047
1719.7229	^{59}RNLT^{62}	^{60}N	$HexNAc_2$ Hex_5	$[M+H]^+$	0.0064
1806.7591	^{58}SRNLT^{62}	^{60}N	$HexNAc_2$ Hex_5	$[M+H]^+$	0.0106
1867.7533	^{58}SRNL^{61}	^{60}N	$HexNAc_2$ Hex_6	$[M+H]^+$	0.0003
1881.7714	^{59}RNLT^{62}	^{60}N	$HexNAc_2$ Hex_6	$[M+H]^+$	0.0022
1934.8446	^{57}(K)SRNLT(K)^{63}	^{60}N	$HexNAc_2$ Hex_5	$[M+H]^+$	0.0012
1962.8507	^{59}(R)NLTKD(R)^{65}	^{60}N	$HexNAc_2$ Hex_5	$[M+H]^+$	0.0124
1968.8115	^{58}SRNLT^{62}	^{60}N	$HexNAc_2$ Hex_6	$[M+H]^+$	0.0102
2124.8946	^{59}(R)NLTKD(R)^{65}	^{60}N	$HexNAc_2$ Hex_6	$[M+H]^+$	0.0035
2130.8544	^{58}SRNLT^{62}	^{60}N	$HexNAc_2$ Hex_7	$[M+H]^+$	0.0003
2292.9023	^{58}SRNLT^{62}	^{60}N	$HexNAc_2$ Hex_8	$[M+H]^+$	0.0046
2448.9948	^{59}(R)NLTKD(R)^{65}	^{60}N	$HexNAc_2$ Hex_8	$[M+H]^+$	0.0019

[a] Parentheses are used to indicate ambiguous assignment of terminating amino acid residues. [b] All mass errors are absolute values (i.e., represented as non-negative). The root-mean-square mass error of the assignments was 3 ppm (externally calibrated).

Table 2. Glycopeptides Derived from RNase B Following 24 h Incubation with Bead-Immobilized Pronase

GP *m*/*z*	peptide	site	glycan	ion type	Δ*m*/*z* (Da)[a]
1371.4574	^{60}N	^{60}N	$HexNAc_2$ Hex_5	$[M+Na]^+$	0.0081
1462.5591	^{60}NL^{61}	^{60}N	$HexNAc_2$ Hex_5	$[M+H]^+$	0.0085
1484.5431	^{60}NL^{61}	^{60}N	$HexNAc_2$ Hex_5	$[M+Na]^+$	0.0065
1505.5766	^{59}RN^{60}	^{60}N	$HexNAc_2$ Hex_5	$[M+H]^+$	0.0081
1533.5076	^{60}N	^{60}N	$HexNAc_2$ Hex_6	$[M+Na]^+$	0.0108
1646.5937	^{60}NL^{61}	^{60}N	$HexNAc_2$ Hex_6	$[M+Na]^+$	0.0087
1667.6270	^{59}RN^{60}	^{60}N	$HexNAc_2$ Hex_6	$[M+H]^+$	0.0105
1808.6415	^{60}NL^{61}	^{60}N	$HexNAc_2$ Hex_7	$[M+Na]^+$	0.0137
1829.6744	^{59}RN^{60}	^{60}N	$HexNAc_2$ Hex_7	$[M+H]^+$	0.0159
1857.6044	^{60}N	^{60}N	$HexNAc_2$ Hex_8	$[M+Na]^+$	0.0196
1970.6924	^{60}NL^{61}	^{60}N	$HexNAc_2$ Hex_8	$[M+Na]^+$	0.0156
1991.7259	^{59}RN^{60}	^{60}N	$HexNAc_2$ Hex_8	$[M+H]^+$	0.0172
2132.7363	^{60}NL^{61}	^{60}N	$HexNAc_2$ Hex_9	$[M+Na]^+$	0.0245
2153.7712	^{59}RN^{60}	^{60}N	$HexNAc_2$ Hex_9	$[M+H]^+$	0.0247

[a] All mass errors are absolute values (i.e., represented as non-negative). The root-mean-square mass error of the assignments was 8 ppm (externally calibrated).

two peptides ^{58}SRNLT^{62} ($[M + H]^+$: 590.3256) and ^{59}NLTKD^{64} ($[M + H]^+$: 590.3144) may be formed. These two protonated peptides differ in mass by 0.0112 Da; consequently, ^{58}SRNLT^{62} + Man-5 and ^{59}NLTKD^{64} + Man-5 differ in mass by only 6 ppm. While FTICR−MS is in principle capable of resolving and distinguishing these two possibilities, signals potentially matching multiple feasible glycopeptides within a few parts per million were interrogated by tandem MS using IRMPD to remove any ambiguities in glycopeptide compositional assignment. IRMPD of the putative Man-5 glycopeptide at *m*/*z* 1806.7591 is shown in Figure 3. After losing one *N*-acetylhexosamine residue and 5 hexose residues, the remaining glycopeptide dissociated into smaller fragments detailing the composition and sequence of the peptide backbone. The observed fragmentation behavior was in agreement with a recent report by Adamson and Håkansson detailing IRMPD of high-mannose type glycopeptides.[29] Using this information, the glycopeptide ^{58}SRNLT^{62} + 2 HexNAc + 5 Hex was confirmed as the correct glycopeptide assignment.

A similar conflicting assignment involving sulfated glycopeptides produced by another nonspecific protease, proteinase K, was recently reported. In that study, tandem mass spectrometry was also required to resolve an initially ambiguous assignment made on the basis of accurate mass.[30] As reports in the literature and the above example illustrate, the occurrence of false positives and overlapping possibilities must be considered when assigning glycopeptides derived from nonspecific proteolysis, even when mass measurements are made with a high degree of accuracy.[23,30,31] Although tandem mass spectrometry was able to resolve the ambiguities in both cases, the utility of supporting information in addition to accurate mass measurement should not be overlooked. To guard against

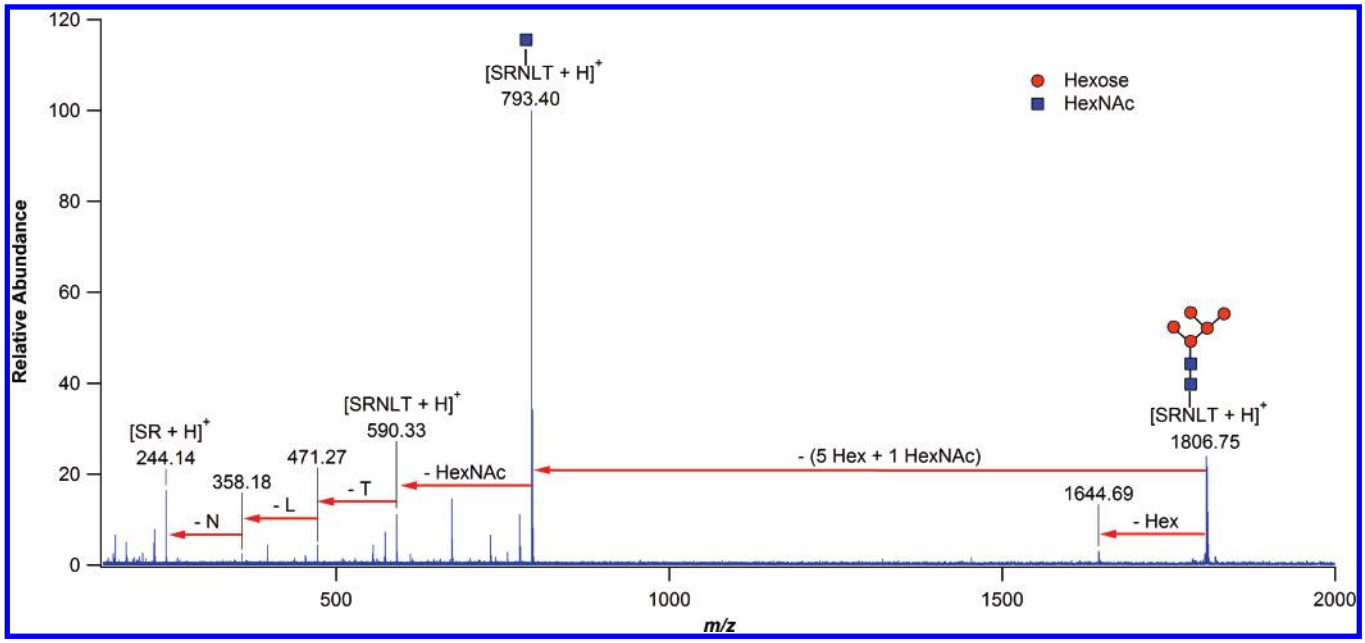

Figure 3. Infrared multiphoton dissociation of the peptide ^{58}SRNLT^{62} connected to Man-5. These data distinguish this glycopeptide assignment from the other peptide possibility of ^{59}NLTKD^{64}.

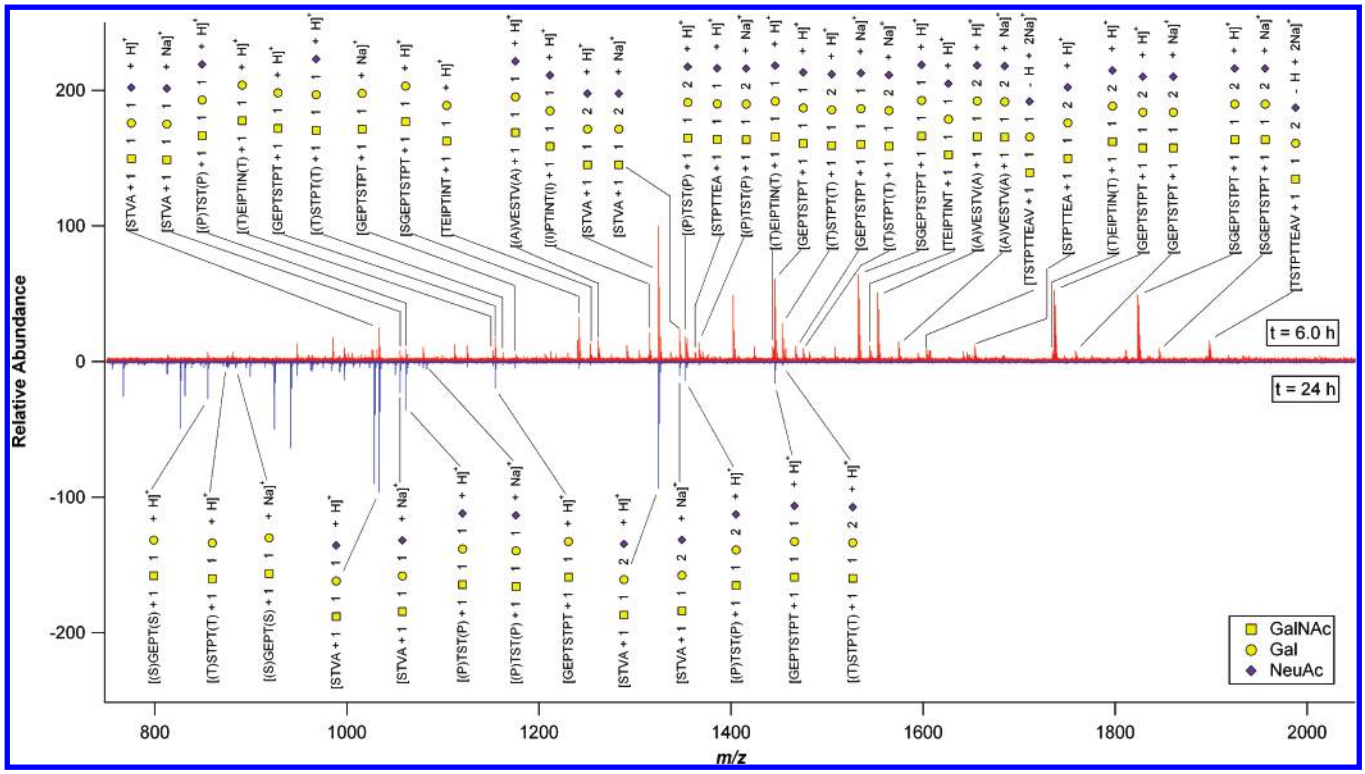

Figure 4. Positive mode nanoESI FTICR mass spectra of κ-csn O-linked glycopeptides following 6 and 24 h of incubation with bead-immobilized pronase.

false identification, assignments were only made for candidate glycopeptides for which multiple glycoforms of a given peptide tag were observed and/or multiple peptide tags corresponding to a specific site of glycosylation were observed containing a range of glycan heterogeneity. Typically, the observation of a single glycopeptide candidate is insufficient to assign a site of glycosylation.

Analysis of Bovine κ-Csn. Bovine κ-csn served as a second model glycoprotein for further evaluating the effectiveness of bead-immobilized pronase for glycosylation site analysis. As one of the major protein constituents of bovine milk, κ-csn contributes to the stabilization of casein micelles which serve as vehicles for the suspension of insoluble calcium phosphate.[32] In its natural state, κ-csn is found as a monomer as well as an intermolecularly disulfide bridged multimer consisting of up to ten or more copies of the protein.[33] According to the current Swiss Prot annotation, the C-terminal domain of the protein is known to harbor several sites of post-translational modifica-

Table 3. Glycopeptides Derived from κ-csn Following 6.0 h Incubation with Bead-Immobilized Pronase

GP *m*/*z*	peptide[a]	site	glycan	ion type	Δ*m*/*z* (Da)[b]
1033.4260	^{162}STVA^{165}	^{163}T	$HexNAc_1$ Hex_1 $NeuAc_1$	$[M+H]^+$	0.0047
1055.4051	^{162}STVA^{165}	^{163}T	$HexNAc_1$ Hex_1 $NeuAc_1$	$[M+Na]^+$	0.0075
1061.4223	^{151}(P)TST(P)^{155} or ^{153}STPT^{156}	^{152}T or ^{154}T	$HexNAc_1$ Hex_1 $NeuAc_1$	$[M+H]^+$	0.0033
1152.5454	^{138}(T)EIPTIN(T)^{145}	^{142}T	$HexNAc_1$ Hex_1	$[M+H]^+$	0.0064
1154.4877	^{149}GEPTSTPT^{156}	^{152}T or ^{154}T	$HexNAc_1$ Hex_1	$[M+H]^+$	0.0071
1162.4689	^{152}(T)STPT(T)^{158} or ^{159}(AV)EST(VA)^{165}	^{152}T, ^{154}T, ^{157}T, or ^{163}T	$HexNAc_1$ Hex_1 $NeuAc_1$	$[M+H]^+$	0.0044
1176.4690	^{149}GEPTSTPT^{156}	^{152}T or ^{154}T	$HexNAc_1$ Hex_1	$[M+Na]^+$	0.0076
1241.5207	^{148}SGEPTSTPT^{156}	^{152}T or ^{154}T	$HexNAc_1$ Hex_1	$[M+H]^+$	0.0061
1253.5911	^{138}TEIPTINT^{145}	^{142}T	$HexNAc_1$ Hex_1	$[M+H]^+$	0.0084
1261.5388	^{159}(A)VESTV(A)^{165}	^{163}T	$HexNAc_1$ Hex_1 $NeuAc_1$	$[M+H]^+$	0.0029
1314.5947	^{140}(I)PTINT(I)^{146}	^{142}T	$HexNAc_1$ Hex_1 $NeuAc_1$	$[M+H]^+$	0.0099
1324.5203	^{162}STVA^{165}	^{163}T	$HexNAc_1$ Hex_1 $NeuAc_2$	$[M+H]^+$	0.0058
1346.5044	^{162}STVA^{165}	^{163}T	$HexNAc_1$ Hex_1 $NeuAc_2$	$[M+Na]^+$	0.0037
1352.5156	^{151}(P)TST(P)^{155} or ^{153}STPT^{156}	^{152}T or ^{154}T	$HexNAc_1$ Hex_1 $NeuAc_2$	$[M+H]^+$	0.0054
1362.5443	^{153}STPTTEA^{159}	^{154}T or ^{157}T	$HexNAc_1$ Hex_1 $NeuAc_1$	$[M+H]^+$	0.0088
1374.5005	^{151}(P)TST(P)^{155} or ^{153}STPT^{156}	^{152}T or ^{154}T	$HexNAc_1$ Hex_1 $NeuAc_2$	$[M+Na]^+$	0.0024
1443.6408	^{138}(T)EIPTIN(T)^{145}	^{142}T	$HexNAc_1$ Hex_1 $NeuAc_1$	$[M+H]^+$	0.0064
1445.5837	^{149}GEPTSTPT^{156}	^{152}T or ^{154}T	$HexNAc_1$ Hex_1 $NeuAc_1$	$[M+H]^+$	0.0064
1453.5611	^{152}(T)STPT(T)^{158} or ^{159}(AV)EST(VA)^{165}	^{152}T, ^{154}T, ^{157}T, or ^{163}T	$HexNAc_1$ Hex_1 $NeuAc_2$	$[M+H]^+$	0.0076
1467.5588	^{149}GEPTSTPT^{156}	^{152}T or ^{154}T	$HexNAc_1$ Hex_1 $NeuAc_1$	$[M+Na]^+$	0.0132
1475.5436	^{152}(T)STPT(T)^{158} or ^{159}(AV)EST(VA)^{165}	^{152}T, ^{154}T, ^{157}T, or ^{163}T	$HexNAc_1$ Hex_1 $NeuAc_2$	$[M+Na]^+$	0.0070
1532.6118	^{148}SGEPTSTPT^{156}	^{152}T or ^{154}T	$HexNAc_1$ Hex_1 $NeuAc_1$	$[M+H]^+$	0.0104
1544.6870	^{138}TEIPTINT^{145}	^{142}T	$HexNAc_1$ Hex_1 $NeuAc_1$	$[M+H]^+$	0.0079
1552.6283	^{159}(A)VESTV(A)^{165}	^{163}T	$HexNAc_1$ Hex_1 $NeuAc_2$	$[M+H]^+$	0.0088
1574.6067	^{159}(A)VESTV(A)^{165}	^{163}T	$HexNAc_1$ Hex_1 $NeuAc_2$	$[M+Na]^+$	0.0123
1606.6436	^{152}TSTPTTEAV^{160} or ^{157}(T)EAVESTVA(T)^{166}	^{152}T, ^{154}T, ^{157}T, or ^{163}T	$HexNAc_1$ Hex_1 $NeuAc_1$	$[M-H+2Na]^+$	0.0107
1653.6309	^{153}STPTTEA^{159}	^{154}T or ^{157}T	$HexNAc_1$ Hex_1 $NeuAc_2$	$[M+H]^+$	0.0175
1734.7314	^{138}(T)EIPTIN(T)^{145}	^{142}T	$HexNAc_1$ Hex_1 $NeuAc_2$	$[M+H]^+$	0.0113
1736.6727	^{149}GEPTSTPT^{156}	^{152}T or ^{154}T	$HexNAc_1$ Hex_1 $NeuAc_2$	$[M+H]^+$	0.0128
1758.6511	^{149}GEPTSTPT^{156}	^{152}T or ^{154}T	$HexNAc_1$ Hex_1 $NeuAc_2$	$[M+Na]^+$	0.0163
1823.7001	^{148}SGEPTSTPT^{156}	^{152}T or ^{154}T	$HexNAc_1$ Hex_1 $NeuAc_2$	$[M+H]^+$	0.0175
1845.6809	^{148}SGEPTSTPT^{156}	^{152}T or ^{154}T	$HexNAc_1$ Hex_1 $NeuAc_2$	$[M+Na]^+$	0.0185
1897.7326	^{152}TSTPTTEAV^{160} or ^{157}(T)EAVESTVA(T)^{166}	^{152}T, ^{154}T, ^{157}T, or ^{163}T	$HexNAc_1$ Hex_1 $NeuAc_2$	$[M-H+2Na]^+$	0.0044

[a] Parentheses are used to indicate ambiguous assignment of terminating amino acid residues. [b] All mass errors are absolute values (i.e., represented as non-negative). The root-mean-square mass error of the assignments was 6 ppm (externally calibrated).

Table 4. Glycopeptides Derived from Bovine κ-csn Following 24 h Incubation with Bead-Immobilized Pronase

GP *m*/*z*	peptide[a]	site	glycan	ion type	Δ*m*/*z* (Da)[b]
855.3487	^{148}(S)GEPT(S)^{153}	^{152}T	$HexNAc_1$ Hex_1	$[M+H]^+$	0.0021
871.3774	^{152}(T)STPT(T)^{158} or ^{159}(AV)EST(VA)^{165}	^{152}T, ^{154}T, ^{157}T, or ^{163}T	$HexNAc_1$ Hex_1	$[M+H]^+$	0.0005
877.3298	^{148}(S)GEPT(S)^{153}	^{152}T	$HexNAc_1$ Hex_1	$[M+Na]^+$	0.0013
1033.4259	^{162}STVA^{165}	^{163}T	$HexNAc_1$ Hex_1 $NeuAc_1$	$[M+H]^+$	0.0048
1055.4096	^{162}STVA^{165}	^{163}T	$HexNAc_1$ Hex_1 $NeuAc_1$	$[M+Na]^+$	0.0030
1061.4225	^{151}(P)TST(P)^{155} or ^{153}STPT^{156}	^{152}T or ^{154}T	$HexNAc_1$ Hex_1 $NeuAc_1$	$[M+H]^+$	0.0031
1083.4078	^{151}(P)TST(P)^{155} or ^{153}STPT^{156}	^{152}T or ^{154}T	$HexNAc_1$ Hex_1 $NeuAc_1$	$[M+Na]^+$	0.0003
1154.4911	^{149}GEPTSTPT^{156}	^{152}T or ^{154}T	$HexNAc_1$ Hex_1	$[M+H]^+$	0.0036
1324.5162	^{162}STVA^{165}	^{163}T	$HexNAc_1$ Hex_1 $NeuAc_2$	$[M+H]^+$	0.0099
1346.5079	^{162}STVA^{165}	^{163}T	$HexNAc_1$ Hex_1 $NeuAc_2$	$[M+Na]^+$	0.0001
1352.5176	^{151}(P)TST(P)^{155} or ^{153}STPT^{156}	^{152}T or ^{154}T	$HexNAc_1$ Hex_1 $NeuAc_2$	$[M+H]^+$	0.0034
1445.5897	^{149}GEPTSTPT^{156}	^{152}T or ^{154}T	$HexNAc_1$ Hex_1 $NeuAc_1$	$[M+H]^+$	0.0005
1453.5672	^{152}(T)STPT(T)^{158} or ^{159}(AV)EST(VA)^{165}	^{152}T, ^{154}T, ^{157}T, or ^{163}T	$HexNAc_1$ Hex_1 $NeuAc_2$	$[M+H]^+$	0.0015

[a] Parentheses are used to indicate ambiguous assignment of terminating amino acid residues. [b] All mass errors are absolute values (i.e., represented as non-negative). The root-mean-square mass error of the assignments was 3 ppm (externally calibrated).

tion, including two potential sites phosphorylation (^{148}S and ^{170}S) and six potential sites of glycosylation (^{142}T, ^{152}T, ^{154}T, ^{157}T, ^{163}T, and ^{186}T). The carbohydrate modifications consist of monosialylated and disialylated core 1 type O-linked glycans.[34,35] Recently, analysis of bovine κ-csn heterogeneity by two-dimensional gel electrophoresis revealed ten distinct forms of the protein, which were accounted for as differentially phosphorylated and heterogeneously glycosylated copies of two major allelic variants.[36] Subsequent studies of individual glycoforms demonstrated glycan occupancy at three of the six potential sites but did not rule out low-level glycosylation of the other potential sites.[37] Due to these many levels of complexity, κ-casein represents an exceptional challenge in the context of glycoproteomic profiling.

The 6 and 24 h bead-immobilized pronase digests of κ-csn are presented in Figure 4. As with the N-linked glycopeptides of RNase B, the O-linked glycopeptides of κ-csn were shifted to lower masses corresponding to smaller peptide moieties as the digestion was allowed to progress over time. Products of the 6 h incubation were observed over a range of approximately *m*/*z* 1000−1900; by comparison, the 24 h digest produced glycopeptides ranging from about *m*/*z* 850 to 1450.

Complete listings of the glycopeptide assignments for the 6 and 24 h digests are presented in Tables 3 and 4, respectively. The 33 glycopeptide signals observed in the 6 h digest contained between four and nine amino acid residues, with five potential glycosylation sites represented. The 13 glycopeptides assigned in the 24 h digest were mapped to only four of the

potential glycosylation sites and carried predominantly tetra- and pentapeptide tags, with the exception that two of the assigned glycopeptides contained eight amino acid residues. At both durations of incubation, all known O-linked glycan compositions were observed. The comparison of the two digestion times reiterates the value of tracking the progression of nonspecific proteolysis; although the 6 h digestion products were larger, more numerous, and more prone to ambiguous assignment, this time point provided clear evidence for glycosylation at ^{142}T; conversely, the products of the 24 h incubation provided no evidence of glycosylation at this position. Regardless of digestion time, no indication of glycosylation at ^{186}T was found.

While the amino acid and monosaccharide compositions for most of the O-linked glycopeptides could be assigned based on accurate mass measurement, the exact peptide sequence or precise site of glycosylation remained ambiguous in several cases. For example, the glycopeptide observed at 1453.5672 was assigned as the glycan $HexNAc_1$ Hex_1 $NeuAc_2$ attached to a peptide with a sequence of TSTPT, STPTT, AVEST, or ESTVA. This single assignment serves to implicate four potential sites of glycosylation: ^{152}T, ^{154}T, ^{157}T, and ^{163}T. The four candidate amino acid sequences correspond to two possible amino acid compositions, but only involve a single elemental composition. Therefore, all four potential glycopeptides are equally plausible and cannot be distinguished by single-stage MS alone. This example illustrates the need for particular caution in assigning site-specificity to O-linked glycosylation. Since the sites of O-linked glycosylation are not restricted to a particular protein consensus sequence, glycopeptide assignment at the level of elemental composition may still be insufficient for unambiguous establishment of the modification sites. As with N-linked glycan composition and site identification, only candidate peptides and O-glycans displaying multiple peptide and/or glycan tags were assigned. Despite the limitations inherent to the assignment of O-linked glycosylation sites, these results demonstrate that bead-immobilized pronase can produce informative O-linked glycopeptides of tunable length from a complex, heterogeneously glycosylated protein.

Conclusions

Identification and characterization of O- and N-linked glycosylation sites was achieved using a three tiered approaching involving nonspecific glycoprotein digestion with immobilized pronase, high mass accuracy mass spectrometry experiments of time specific glycopeptide preparations, and data screening against a list of possible glycosylation sites and glycan compositions. The significant advantages of this approach lie in the inherent efficiency of nonspecific glycoprotein digestion and the pronounced reduction in chemical noise previously introduced by enzyme autolysis. These benefits are evident in that for each time course measurement the vast majority of observed peaks could be assigned to glycopeptides.

These experiments demonstrated that even with a high degree of mass accuracy, the full complexity of glycopeptides could not be elucidated using a single stage of mass spectrometry alone. Rather, the ambiguities that existed could be minimized or entirely eliminated using tandem mass spectrometry and digestion time course measurements to track the regression of the glycopeptide length. Due to the ease of implementation, this method affords researchers a rapid and effective means to complement the information provided by other tools for glycoprotein analysis.

Acknowledgment. We would like to thank Dr. Suzanne Miyamoto for the generous loan of the incubation apparatus used for glycoprotein digestion. Additional appreciation is extended to Dr. Jerry L. Hedrick for his encouragement and helpful discussions. Funding for this work was provided by the National Institutes of Health (GM 49077), the California Dairy Research Foundation (06 LEC-01-NH), and Dairy Management Incorporated.

References

(1) Walsh, G.; Jefferis, R. Post-translational modifications in the context of therapeutic proteins. *Nat. Biotechnol.* **2006**, *24*(10), 1241−52.
(2) Dwek, R. A. Glycobiology: Toward understanding the function of sugars. *Chem. Rev.* **1996**, *96*(2), 683−720.
(3) Varki, A. Biological Roles of Oligosaccharides - All of the Theories Are Correct. *Glycobiology* **1993**, *3*(2), 97−130.
(4) Spiro, R. G. Protein glycosylation: nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds. *Glycobiology* **2002**, *12*(4), 43r−56r.
(5) Varki, A. *Essentials of glycobiology*; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, 1999; pp xvii, 653.
(6) Hofsteenge, J.; Muller, D. R.; de Beer, T.; Loffler, A.; Richter, W. J.; Vliegenthart, J. F. New type of linkage between a carbohydrate and a protein: C-glycosylation of a specific tryptophan residue in human RNase Us. *Biochemistry* **1994**, *33*(46), 13524−13530.
(7) Lo-Guidice, J. M.; Lhermitte, M. HPLC of oligosaccharides in glycobiology. *Biomed. Chromatogr.* **1996**, *10*(6), 290−296.
(8) Rohrer, J. S.; Cooper, G. A.; Townsend, R. R. Identification, Quantification, and Characterization of Glycopeptides in Reversed-Phase Hplc Separations of Glycoprotein Proteolytic Digests. *Anal. Biochem.* **1993**, *212*(1), 7−16.
(9) Fu, D.; Chen, L.; O'Neill, R. A. A detailed structural characterization of ribonuclease B oligosaccharides by 1H NMR spectroscopy and mass spectrometry. *Carbohydr. Res.* **1994**, *261*(2), 173−186.
(10) Harvey, D. J. Structural determination of N-linked glycans by matrix-assisted laser desorption/ionization and electrospray ionization mass spectrometry. *Proteomics* **2005**, *5*(7), 1774−1786.
(11) Zaia, J. Mass spectrometry of oligosaccharides. *Mass Spectrom. Rev.* **2004**, *23*(3), 161−227.
(12) Park, Y.; Lebrilla, C. B. Application of Fourier transform ion cyclotron resonance mass spectrometry to oligosaccharides. *Mass Spectrom. Rev.* **2005**, *24*(2), 232−264.
(13) Morelle, W.; Canis, K.; Chirat, F.; Faid, V.; Michalski, J. C. The use of mass spectrometry for the proteomic analysis of glycosylation. *Proteomics* **2006**, *6*(14), 3993−4015.
(14) Wuhrer, M.; Koeleman, C. A.; Deelder, A. M.; Hokke, C. H. Normal-phase nanoscale liquid chromatography-mass spectrometry of underivatized oligosaccharides at low-femtomole sensitivity. *Anal. Chem.* **2004**, *76*(3), 833−838.
(15) Schulz, B. L.; Packer, N. H.; Karlsson, N. G. Small-scale analysis of O-linked oligosaccharides from glycoproteins and mucins separated by gel electrophoresis. *Anal. Chem.* **2002**, *74*(23), 6088−6097.
(16) Angel, P. M.; Lim, J. M.; Wells, L.; Bergmann, C.; Orlando, R. A potential pitfall in 18O-based N-linked glycosylation site mapping. *Rapid Commun. Mass Spectrom.* **2007**, *21*(5), 674−682.
(17) Wuhrer, M.; Hokke, C. H.; Deelder, A. M. Glycopeptide analysis by matrix-assisted laser desorption/ionization tandem time-of-flight mass spectrometry reveals novel features of horseradish peroxidase glycosylation. *Rapid Commun. Mass Spectrom.* **2004**, *18*(15), 1741−1748.
(18) Bunkenborg, J.; Pilch, B. J.; Podtelejnikov, A. V.; Wisniewski, J. R. Screening for N-glycosylated proteins by liquid chromatography mass spectrometry. *Proteomics* **2004**, *4*(2), 454−465.
(19) Kawasaki, N.; Itoh, S.; Ohta, M.; Hayakawa, T. Microanalysis of N-linked oligosaccharides in a glycoprotein by capillary liquid chromatography/mass spectrometry and liquid chromatography/tandem mass spectrometry. *Anal. Biochem.* **2003**, *316*(1), 15−22.
(20) Wang, P.; Li, G.; Granados, R. R. Identification of two new peritrophic membrane proteins from larval Trichoplusia ni: structural characteristics and their functions in the protease rich insect gut. *Insect Biochem. Mol. Biol.* **2004**, *34*(3), 215−227.
(21) 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*, 217−230.

(22) 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*(20), 5628−5637.
(23) An, H. J.; Tillinghast, J. S.; Woodruff, D. L.; Rocke, D. M.; Lebrilla, C. B. A new computer program (GlycoX) to determine simultaneously the glycosylation sites and oligosaccharide heterogeneity of glycoproteins. *J. Proteome Res.* **2006**, *5*(10), 2800−2808.
(24) Narahashi, Y. Proteolytic Enzymes. *Methods Enzymol.* **1970**, *19*, 651.
(25) Spiro, R. G. Characterization of carbohydrate units of glycoproteins. *Methods Enzymol.* **1966**, *8*, 26−52.
(26) Temporini, C.; Perani, E.; Calleri, E.; Dolcini, L.; Lubda, D.; Caccialanza, G.; Massolini, G. Pronase-immobilized enzyme reactor: An approach for automation in glycoprotein analysis by LC/LC-ESI/MSn. *Anal. Chem.* **2007**, *79*(1), 355−363.
(27) Axen, R.; Porath, J.; Ernback, S. Chemical coupling of peptides and proteins to polysaccharides by means of cyanogen halides. *Nature* **1967**, *214*(5095), 1302−1304.
(28) Yu, Y. Q.; Fournier, J.; Gilar, M.; Gebler, J. C. Identification of N-linked glycosylation sites using glycoprotein digestion with pronase prior to MALDI tandem time-of-flight mass spectrometry. *Anal. Chem.* **2007**, *79*(4), 1731−1738.
(29) Adamson, J. T.; Hakansson, K. Infrared multiphoton dissociation and electron capture dissociation of high-mannose type glycopeptides. *J. Proteome Res.* **2006**, *5*(3), 493−501.
(30) Irungu, J.; Dalpathado, D. S.; Go, E. P.; Jiang, H.; Ha, H. V.; Bousfield, G. R.; Desaire, H. Method for characterizing sulfated glycoproteins in a glycosylation site-specific fashion, using ion pairing and tandem mass spectrometry. *Anal. Chem.* **2006**, *78*(4), 1181−1190.
(31) Jiang, H.; Desaire, H.; Butnev, V. Y.; Bousfield, G. R. Glycoprotein profiling by electrospray mass spectrometry. *J. Am. Soc. Mass Spectrom.* **2004**, *15*(5), 750−758.
(32) Swaisgood, H. E. Review and update of casein chemistry. *J. Dairy Sci.* **1993**, *76*(10), 3054−3061.
(33) Rasmussen, L. K.; Hojrup, P.; Petersen, T. E. The multimeric structure and disulfide-bonding pattern of bovine kappa-casein. *Eur. J. Biochem.* **1992**, *207*(1), 215−222.
(34) Minikiewicz, P.; Slangen, C. J.; Lagerwerf, F. M.; Haverkamp, J.; Rollema, H. S.; Visser, S. Reversed-phase high-performance liquid chromatographic separation of bovine kappa-casein macropeptide and characterization of isolated fractions. *J. Chromatogr. A* **1996**, *743*(1), 123−135.
(35) Pisano, A.; Packer, N. H.; Redmond, J. W.; Williams, K. L.; Gooley, A. A. Characterization of O-linked glycosylation motifs in the glycopeptide domain of bovine kappa-casein. *Glycobiology* **1994**, *4*(6), 837−844.
(36) Holland, J. W.; Deeth, H. C.; Alewood, P. F. Proteomic analysis of kappa-casein micro-heterogeneity. *Proteomics* **2004**, *4*(3), 743−752.
(37) Holland, J. W.; Deeth, H. C.; Alewood, P. F. Analysis of O-glycosylation site occupancy in bovine kappa-casein glycoforms separated by two-dimensional gel electrophoresis. *Proteomics* **2005**, *5*(4), 990−1002.

PR070317Z