Chemoenzymatic Synthesis and MS^n Characterization of a Homogenous Glycoprotein

*John. E. Schiel, Mark S. Lowenthal, Karen W. Phinney
National Institute of Standards and Technology
Analytical Chemistry Division
100 Bureau Drive, Stop 8392
Gaithersburg, MD 20899

Abstract

Glycoproteins exist as heterogeneous mixtures of glycoforms due to glycosylation at multiple sites with various glycans. The glycosylation pattern has significant implications for clinical diagnostics and biopharmaceutical development because it affects stability, toxicity, and activity. Complete glycoprotein characterization involves numerous multidisciplinary techniques, including many invaluable mass spectrometry-based approaches. A well defined glycoprotein standard containing a single glycan of known composition, linkage, and steroechemistry would be of great value for the comparison and evaluation of glycoprotein analysis techniques.

The current project describes the rapid one-pot chemoenzymatic rearrangement of bovine ribonuclease B (RNase B) to contain a single glycosylation site with a known trisaccharide glycan. The procedure involved enzymatic removal of native glycans from the protein, leaving only the penultimate N-acetylglucosamine residue attached. A reactive disaccharide oxazoline derivative was then synthesized and stereospecifically added to the deglycosylated RNase through Endo-β-N-Acetylglucosaminidase M catalyzed chemoenzymatic transglycosylation. Oxazoline formation conditions were optimized using mass spectrometry, and the product verified based on its collision induced dissociation (CID) mass spectrum. Enzymatic removal of native glycans as well as formation of the desired homogeneous product was also monitored using mass spectrometry. LC-MS^n using four sequential rounds of CID was used to verify that the original glycosylation site had in fact been reorganized to contain the new glycan. The techniques described herein are not limited to this analyte or glycan, and should be amenable to the synthesis of numerous homogeneous glycoconjugates with judicious choice of enzyme/substrate combinations. The combined use of chemoenzymatic synthesis and mass spectrometry-based characterization show promise for the development of homogeneous glycoprotein reference materials. These materials may prove useful for increasing our understanding of currently used and innovative glycoprotein analysis techniques.

Keywords: glycoprotein, chemoenzymatic synthesis, mass spectrometry, oxazoline, glycosylation

* Corresponding Author: John E. Schiel, john.schiel@nist.gov, 301-975-3137

INTRODUCTION

Proteins found in biological systems consist of an amino acid backbone that is often modified at numerous sites after translation. These post translational modifications can play a significant role in the protein's stability, toxicity, and activity. One common PTM is glycosylation. This modification has recently been of interest in clinical diagnostics and the development of biopharmaceuticals because of the impact changes in glycosylation patterns can have on a protein's form and function. Aberrant glycosylation patterns have been shown to correlate with numerous disease states and have become useful for clinical diagnostics [1,2]. The glycosylation pattern has also become of great interest in developing biopharmaceuticals as it can have a large effect on the safety, stability, and efficacy of a drug [3-5].

Numerous methods have been developed for qualitative glycoprotein analysis that focus on characterizing the glycosylation sites, patterns of glycosylation, and microheterogeneity of glycans present at each site [6-10]. High performance liquid chromatography (LC) and mass spectrometry (MS) have become invaluable tools in the analysis of protein glycosylation. Matrix assisted laser desorption ionization (MALDI) and electrospray ionization (ESI) have been used with time-of-flight (TOF) mass analyzers for monitoring the intact mass of a glycoprotein and assessing completion of enzymatic glycan removal [8,11]. Glycans are often removed from a protein using PNGase F, Endo H, or Endo H_f (a fusion protein of Endo H and maltose binding protein). The released glycans can be labeled at the reducing terminus with a fluorescent dye, permethylated, or analyzed directly after enzymatic release using LC and/or MS [8,10]. Collision induced dissociation (CID) with various mass analyzers can be used to fragment glycosidic bonds, with higher energy or multiple rounds of dissociation being used to obtain

cross ring cleavages and additional information on glycan branching [10]. Finally, analysis of glycopeptides after trypsin or pronase digestions has also become common using MS experiments [6,7,9]. LC-MS/MS of glycopeptides can yield information on the glycan composition (CID) or amino acid sequence and therefore glycosylation site (ETD) [7,12]. Multiple rounds of CID can also be used in LC-MS^n experiments to gain information on glycan composition (MS^2) and peptide sequence (MS^3 and MS^4) [7]. Proteomic comparison of tryptic glycopeptides after PNGase release in ^{16}O and ^{18}O labeled water can also be used to identify glycosylation sites, however pre-treatment conditions must be carefully controlled to reduce false positive results [13]. Each of the methods described above can provide various degrees of information regarding protein glycosylation (e.g. composition, branching information, relative quantification, and/or site identification). Numerous factors can result in method or interlaboratory variability, the extent of which has recently become increasingly studied [14-16]. A more detailed understanding of glycan-dependent effects such as enzymatic efficiency, labeling efficiency, sample loss during glycoconjugate purification/enrichment, isomeric resolution using different LC methods, and fragmentation efficiency and pattern for each of the available techniques is necessary to increase the accuracy and precision of glycoprotein characterization. This reproducibility and method comparability is crucial to ensuring accuracy of clinical diagnostic tests, consistency of biopharmaceutical products, and will play a major role in the need for assessment of biosimilar/biobetter drug applications.

Detailed comparability studies on naturally occurring glycoproteins are difficult due to inherent glycan heterogeneity and lack of an established true identity. A glycoprotein containing a known, synthetically added glycan at a single glycosylation site would circumvent these pitfalls and serve as a reference material for all available methods. One recent methodology for

obtaining a homogeneous glycosylation structure has been the chemoenzymatic rearrangement of glycoconjugates using Endo-β-N-Acetylglucosaminidases [17-28]. The protein of interest is first de-glycosylated with the enzyme Endo H (or related enzyme), which cleaves the entire glycan except the penultimate GlcNAc residue. The remaining GlcNAc can then act as a handle for a second enzyme (e.g. Endo A for high mannose type glycans or Endo M for complex, hybrid, or high mannose) to act as a glycosyltransferase to transfer the homogeneous glycan of interest to the appropriate site. Transglycosylation using these enzymes is stereospecific for the desired β-(1→4)-glycosidic linkage [27]. Therefore, if a glycan of known composition, linkage, and stereochemistry is used, the product will be glycosylated with only the glycan of interest with no microheterogeneity. This method was recently shown to be capable of glycosylation remodeling of RNase B to contain homogenous glycosylation of either the pentasaccharide core or an un-natural heptasaccharide product in good yield [22,27]. In this case, good yields were obtained because: (1) oxazoline-modified glycans were used as activated derivatives to be added to the GlcNac peptide and (2) hydrolysis of the product by the enzyme (which shows both transferase and endoglycosidase activity) was not observed because Endo A acts to only hydrolyze high mannose type glycans [22,27]. In an effort to prepare an enzyme with low glycosidase activity (hydrolysis) and high transferase activity, multiple point mutants of the Endo M enzyme were prepared. The N175A mutant showed very high transferase activity and no glycosidase activity allowing the transfer of a high-mannose type glycan of natural configuration to a glycopeptide in a 60% yield [29]. Endoglycosidase mutants were also recently used to prepare homogeneously glycosylated forms of RNase B [19]. Natural high mannose and complex type homogeneous glycans were formed using mutant enzymes of Endo M and Endo A which displayed solely transglycosylation activity when oxazoline activated glycans were used as substrates [19].

This methodology has been shown to be very efficient at glycoprotein and glycopeptide remodeling, however, currently there is only one commercially available enzyme (Endo M) for this type of reaction. In addition, previous chemoenzymatic synthesis procedures have relied on relatively difficult glycan synthesis to form oxazoline derivatives requiring multiple rounds of protection, de-protection, and purification prior to reaction with the deglycosylated protein. Recently, a relatively simple aqueous phase synthesis of an oxazoline intermediate has been developed [30]. In a recent publication, a complex carbohydrate oxazoline was synthesized in this manner, purified, and effectively used in the glycan rearrangement of RNase B using Endo A and a mutant form of Endo M [20].

In the current report, the potential of aqueous oxazoline synthesis and its use without further purification was explored with Endo M catalyzed transglycosylation to form a homogeneous glycoprotein as described in Fig. 1. The mannose-glucosamine (ManGlcNAc) oxazoline derivative was formed through a nucleophilic substitution catalyzed by 1-chloro-1,3-dimethylimidazolinium chloride (DMC). Intramolecular rearrangement of the reacted β-anomer resulted in the necessary oxazoline intermediate. The glycan release was performed using Endo H_f, therefore allowing simple filtration to purify the GlcNAc RNase (GRNase). The oxazoline reaction product was directly mixed with GRNase and Endo M without further purification to form the desired homogeneous glycoprotein. A set of mass spectrometry related techniques were developed to monitor oxazoline synthesis for yield optimization. In addition, LC-MS^n methods were developed capable of glycan identity and site classification, an initial characterization required to verify the glycan homogeneity of an engineered glycoprotein.

EXPERIMENTAL

Reagents. The recombinant Endo-β-D-acetylglucosaminidase H_f was purchased from New England Biolabs (Ipswitch, MA). The Endo-β-D-acetylglucosaminidase M was obtained from TCI Chemicals (Portland, OR). The 2-acetamido-2-deoxy-4-O-(β-D-mannopyranosyl)-D-glucose (ManGlcNAc) was from Toronto Research Chemicals (North York, Ontario). Ribonuclease B, 2-chloro-1,3-dimethylimidazoline chloride (DMC), and triethylamine (Et_3N) were purchased from Sigma (St. Louis, MO). Filtration was performed using Amicon Ultra centrifugal filtration devices (3 and 30 KDa MWCO) from Millipore (Billerica, MA). Burdick and Jackson LC/MS grade water and acetonitrile were purchased from VWR (West Chester, PA) and used for mass spectrometry solution preparation. Tryptic digestion was performed using Waters Rapigest Surfactant (Milford, MA) and Promega Trypsin Gold (Madison, WI). All other reagents used were of reagent grade or higher quality.

Oxazoline Synthesis. The procedure followed was only slightly modified from the literature [30]. Briefly, ManGlcNAc was dissolved in H_20 at a concentration of 4 mg/mL and cooled to 0^oC. Triethylamine and DMC were added in a 3- and 9-fold molar excess versus ManGlcNAc, respectively. This mixture was allowed to react for one hour at 0^oC and used in the transglycosylation as described below without any further purification. A 15 µL aliquot of this sample was mixed with 200 µL of 1 mmol/L NaOH in 50/50 (v/v %) methanol/water for mass spectrometric analysis.

Endo H_f Glycan Release. Briefly, 20 mg RNase B was dissolved in 6.3 mL of 50 mmol/L sodium citrate (pH 5.5) and 700 µL of Endo H_f solution as provided by the manufacturer. This mixture was allowed to react overnight at 37^oC.

Purification of the resulting GRNase was achieved using the Amicon Ultra centrifugal devices. The mixture was first spun in an Eppendorf 5404 Centrifuge (Hauppauge, NY) through a 30 KDa MWCO filter (3000 x g) to separate the Endo H_f from the GRNase and released glycans. An additional 9 mL of water was flowed through the filter (3000 x g) to optimize product collection. The flow through was combined and placed into a 3KDa filter and solvent exchanged into LC/MS grade water with multiple spins (4000 x g) and water additions. The final filtrate was brought to 7 mL with LC/MS grade water and determined to have a total protein content of 0.88 mg/mL based on UV-Vis absorbance at 280 nm collected on an Agilent 8452 UV-Vis spectrophotometer (Santa Clara, CA).

Chemoenzymatic Transglycosylation. Three individual samples were prepared in order to monitor the kinetics of the reaction. An aliquot of the purified GRNase was mixed with the oxazoline reaction product at a ratio of 150 µL oxazoline reaction product per 1 mg GRNase (as determined by A280). This solution was then adjusted to pH 6.0 and 100 mmol/L potassium phosphate using a 10X potassium phosphate buffer (1 mol/L). Finally, 40 mU of Endo M / mg of GRNase were added and the mixture was shaken for 1, 3, or 20 hours. At each desired time point, one of the samples was immediately taken through the 30/3 KDa MWCO filtration procedure to separate the Endo M from the RNase products, thereby effectively quenching the reaction. The final solvent exchange was performed using LC/MS grade 0.1 % (v/v) formic acid in water to a final volume of 200 µL. An additional 200 µL of 0.1 % (v/v) formic acid in acetonitrile was added to this mixture to prepare a solution of 400 µL of reaction product in 50/50 (v/v %) acetonitrile/water with 0.1 % (v/v) formic acid for direct MS infusion.

A second sample was also prepared in the same manner as described above, except the 30 KDa MWCO filtration step was not performed. The 3 KDa MWCO solvent exchange was used

to prepare the reaction product in pure water. This sample was digested using trypsin and analyzed by LC-MS/MS as described below.

Tryptic Digestion. Tryptic digestion was performed by dissolving 0.1 mg of protein in 100 µL of 50 mmol/L ammonium bicarbonate and 100 µL of 0.2 % (v/v) Rapigest and boiling for 2 minutes. DTT was added to a final concentration of 5 mmol/L and incubated at 60^{o}C for 30 minutes. Iodoacetamide was added to a final concentration of 15 mmol/L and reacted in the dark for 30 minutes at room temperature. The digestion was performed at 37^{o}C for 16 hours with a $CaCl_2$ concentration of 1 mmol/L and a 1:10 (mol/mol) trypsin to protein ratio. After incubation, HCl was added to the sample to a final concentration of 100 mmol/L. This was incubated for an additional 45 minutes at 37^{o}C followed by centrifugation to remove the acid hydrolyzed Rapigest surfactant. The acidic sample was neutralized with an appropriate quantity of 500 mmol/L ammonium bicarbonate buffer. This sample was then diluted 10-fold prior to injection onto the LC-MS^{n} system.

Mass Spectrometry. Mass spectra of the intact glycoproteins were collected by direct infusion (5 µL /min) on an ESI-LTQ XL mass spectrometer operated in positive ion mode. As described above, mass spectra of the intact glycoproteins were performed using 50/50 (v/v %) acetonitrile/water with 0.1 % (v/v) formic acid. Carbohydrate and oxazoline mass spectra were performed in 1 mmol/L NaOH dissolved in 50/50 (v/v %) methanol/water.

LC-MS^{n} of the tryptic digests were performed on the ESI-LTQ XL mass spectrometer using an Agilent 1100 HPLC system for sample introduction. Injections corresponding to 20, 100, or 200 pmol of digested protein were made onto a 15 cm x 2.1 mm i.d. (d_P = 3 µm) Discovery Bio C18 column at 0.2 mL/min. A binary gradient of water with 0.1 % (v/v) formic acid (solvent A) and ACN with 0.1 % (v/v) formic acid (solvent B) was applied as 0 min 98/2

(%A/%B), 2 min 98/2, 5 min 90/10, 35 min 50/50, 37 min 5/95, 39 min 5/95, 41 min 98/2, 60 min 98/2. The LTQ XL was set to collect data dependent MS^4 data using an inclusion mass list for MS^2 parent selection, followed by two rounds of additional CID using the base peak from the previous scan. The initial round of CID was performed using a collision energy of 25, activation Q of 0.25, and an activation time of 30 ms. The second and third rounds of CID (for MS^3 and MS^4) were performed in the same manner, only with a collision energy of 35. Dynamic exclusion was enabled and set to exclude a selected precursor ion for 5 seconds after a three repeat counts were recorded. The ion optics for LC-MS/MS were tuned using angiotensin infusion.

DISCUSSION

Material Selection. *Endo-β-D-Acetylglucosaminidase M.* A variety of chemoenzymatic transglycosylations of various acceptors including asparagine analogs, peptides, and proteins have been accomplished using Acetylglucosaminidase enzymes [17-28]. Transglycosylation has been performed on RNase B using Endo A,[19,22,23] mutants of Endo A and Endo M,[18-20] and on an IgG Fc fragment using Endo A [28]. The recombinant mutant forms of Endo A and Endo M have point mutations that decrease their hydrolytic activity, and therefore increase their transglycosylation activity. These enzymes would be desirable in the current experiment, however, they are not available commercially. The only currently available enzyme is Endo M, therefore, this enzyme was chosen for the initial proof of principle studies.

ManGlcNAc Oligosaccharide. Endo M is capable of transglycosylation of a GlcNAc-asparagine receptor with a range of oxazoline substrates. ManGlcNAc was shown to be added with a 98% yield, however, significant hydrolysis (“off”) activity begins to limit yield when larger glycan structures are used [24,25]. In order to obtain a proof of principle for the aqueous

synthesis, the disaccharide (ManGlcNac) was chosen to be the initial glycan added to RNase B. This small glycan is expected to give the highest yield because Endo M does not have "off" hydrolytic activity for this substrate [25].

Native RNase B. Bovine Ribonuclease B (RNase B) has commonly been used in past experiments as an acceptor for chemoenzymatic transglycosylation [20,22,23,27]. It is a relatively small protein (124 amino acids) that contains a single N-linked glycosylation site [31,32]. Five high mannose type glycoforms exist (Man5 – Man9) and have been well characterized with mass spectrometric methods.

A spectrum of intact, native RNase B is shown in Fig. 2(a). Numerous charge states corresponding to the addition of 7 to 13 protons are apparent in the spectrum. In addition to the charge state envelope, the glycoform heterogeneity is clearly observed for each charge state, and labeled at M5 through M9 for the various high mannose type glycoforms.

The spectrum of the intact glycoprotein gives some information on the relative difference in mass of the various glycoforms, however, for an unknown protein many additional experiments are necessary to characterize the glycosylation. One commonly used method is to perform tryptic digestion of the protein and characterize the glycosylated peptides using LC-MS^n analysis.

The known amino acid sequence of RNase B was used to predict tryptic cleavage sites. This information along with the masses of Man5 – Man9 glycans known to be associated with RNase B allowed estimation of the m/z expected for RNase B glycopeptides. The expected glycopeptides ionized as doubly protonated ions and were found to elute between 4.1-6.2 minutes. Fig. 2(b) is the average MS^1 mass spectrum collected for the RNase B sample eluting

over this time period. All of the known glycoforms are identified in the spectrum corresponding to glycosylation at Asn34 and the predicted tryptic peptide sequence NLTK.

Once again, although Fig. 2(b) depicts the glycan structure as being known, the mass alone does not ensure the glycan structure is in fact as labeled due to the isobaric nature of many monosaccharides. For this reason, CID of the glycopeptide peaks is necessary in order to ensure correct assignment. Fig. 2(c) is a representative CID spectrum of the m/z = 927.35 Da parent glycopeptide ion. CID fragmentation of a glycopeptide typically results in cleavage of the glycan at the glycosidic bonds, with Y- and B-type fragments(according to the nomenclature of Domon and Costello)[33] dominating the spectrum [9]. Fig. 2(c) shows a series of 2^+ Y-ion peaks, 1^+ Y-ion peaks, and a single 1^+ B-type ion at m/z = 1176.16. Additional B/Y dual glycosidic cleavage peaks are also visible in the spectrum, including B_4/Y_4 at 1014.2 Da, $B_4/Y_{3\alpha}$ at 852.15, and B_4/Y_3 at 690.14 Da. The combined fragmentation spectrum is characteristic for the indicated glycopeptide, and allows confirmation of the glycan identity.

Additional rounds of CID were also taken on the glycopeptide. The MS^3 spectrum was a data dependent scan of the base peak from the MS^2 spectrum, which was the peptide-GlcNAc ion. Fragmentation of this species resulted in loss of the GlcNAc residue, so the MS^3 spectrum predominantly contained the peptide ion only. This ion was selected for the MS^4, in which the peptide fragmented to give amino acid sequence information as depicted in Fig. 2(d). A combination of m/z from the MS^1, glycan fragmentation in MS^2, and amino acid sequencing in MS^4 gives complete glycan type and glycosylation site verification of the glycopeptides.

Oxazoline Synthesis and Mass Spectral Characterization. The optimized reaction conditions for synthesis of the ManGlcNAc oxazoline are listed in the Experimental Section. Mass spectrometry with direct infusion was used to characterize the products of this reaction.

This analysis method allowed for assessment of product under a given set of conditions very quickly, ensuring oxazoline product was formed prior to mixture with GRNase. Initial conditions tested for MS infusion were a dilution of the sample in an acetonitrile/water solvent containing 0.1 % (v/v) formic acid. It was found that the weak acid was sufficient to reverse the reaction of oxazoline formed back to the starting material.

Alternative conditions for MS using NaOH in methanol/water were found to be optimal for product characterization. Addition of sodium has been shown in previous reports to increase the ionization efficiency of glycans, as well as increase the amount of structural information attainable due to increased occurrence of cross ring cleavages [10,34]. A characteristic spectrum obtained for the oxazoline synthesis reaction product is shown in Fig. 3(a). The labeling of these peaks refers to the theoretical masses listed in Fig. 1. Due to the complexity of the reaction mixture, adduct peaks corresponding to adduct of Et_3NH^+ and Cl^- are also observed. Additional verification of these assignments was performed using CID of each of the associated peaks. Fig. 3(b) is the CID spectrum obtained for the sodium adduct of pure ManGlcNAc. The predominant ion formed is from a neutral water loss, which has been shown to primarily result from the cleavage of the reducing terminal C-OH bond [35]. The proposed mechanism for this fragmentation involves abstraction of a nearby proton by the reducing oxygen, followed by water loss [35]. Labeling studies with sugars containing C2 hydroxyl groups indicate that the proton originated from the C2 hydroxyl, which is not present in a GlcNAc. However, it is likely the mechanism for water loss is similar to that depicted in Fig. 1 for the oxazoline synthesis. Abstraction of the amine proton and ring formation would be an energetically favored fragmentation, explaining the large water loss.

All four glycosidic cleavages, and a few cross ring cleavages were also observed. Some of these ions could result from numerous cross ring losses, however, labeling studies have shown that these types of fragmentations typically occur from cleavage of the reducing sugar [35]. The proposed mechanism involves opening of the reducing terminus to the aldehyde form, followed by a reverse aldol type reaction through the formation of a favorable six-membered ring intermediate [35]. The $^{0,2}A_2$ ion can form via this mechanism, and the $^{2,4}A_2$ fragment can form by two successive reverse aldol type rearrangements. These ions are typically the main cross ring fragments observed for sodiated oligosaccharides with a free reducing terminus [10]. The ion at 346 Da could also be the result of a cross ring fragmentation ($^{0,4}X_0$, $^{0,4}X_1$, $^{1,3}X_1$, or $^{2,4}X_1$), but the direct reverse aldol type reaction is not accessible for these cleavages, and therefore explains its low relative intensity.

Fig. 3(c) is the CID mass spectra collected for the sodium adduct of the oxazoline product. The most identifiable difference compared to the ManGlcNAc is that the fragments requiring a free reducing terminus are no longer present in high abundance. The water loss is much less intense and the $^{0,2}A_2$ peak is no longer present indicating the reducing terminus is no longer free, but instead modified as the oxazoline ring. Some peaks corresponding to cross ring fragmentation are still present, and likely result from a less energetically favorable mechanism. The $^{3,5}A_2$ fragment and proposed three-bond fragment at m/z = 287 are indicated in Fig. 3(c). Both of these fragments are in agreement with oxazoline formation. The ion at 328 could be a number of isobaric fragments ($^{0,4}X_0$, $^{0,4}X_1$, $^{1,3}X_1$, or $^{2,4}X_1$), and does not give any additional structural identification.

As with ManGlcNAc, all four glycosidic cleavages are seen. The B_1 and C_1 ion in both structures occurs at the same m/z because the non-reducing terminus is identical. However, in

the oxazoline product the Y and Z ions are shifted to a lower mass, a result of the oxazoline ring formed at the reducing terminus. The relative intensity of the Y ion is also much larger in the oxazoline structure, which is typical of a sugar modified at the reducing terminus. The oxazoline can no longer undergo the fragmentation involving the free reducing terminus, and the Y fragmentation pathway becomes the most energetically favored.

Spectra were also collected for CID of the Et_3NHCl adduct peaks listed in Fig. 3(a). The first round of CID on these species resulted in loss of the Et_3NHCl adduct. A second round of CID caused an identical fragmentation pattern to that of the starting material (m/z = 526.95) or the oxazoline product (m/z = 524.92), verifying the identity of these species.

GRNase Preparation. The next step in the synthetic procedure was to prepare the glycoprotein to be a sufficient acceptor for the transglycosylation reaction. Endo H_f was used to prepare deglycosylated RNase B, resulting in a glycoprotein containing only a single GlcNAc residue at the original Asn 34 glycosylation site (GRNase). This reaction was performed as described in the experimental section and GRNase was purified from the Endo H_f enzyme using membrane filtration. Fig. 4(a) is a representative mass spectrum of the GRNase reaction product after treatment with Endo H_f. The heterogeneity due to the presence of various glycoforms is no longer present. Instead, a charge state envelope corresponding to a homogeneous population of GRNase is present. The theoretical mass of GRNase was calculated to be 13885.4 based on the amino acid sequence of the RNase apoprotein, the reduction in mass due to the four disulfide bonds in the protein, and the presence of a single GlcNAc residue at Asn 34. Each of the peaks observed in Fig. 4(a) are within 0.1 to 0.5 Da of the theoretical m/z. This indicates that Endo H_f was efficient at deglycosylating RNase B, and this product should now be sufficient for use in the transglycosylation reaction.

One item to consider here is the total yield of GRNase. Absorbance measurements indicate that approximately 30% of the original RNase B used was actually recovered as GRNase, despite the fact that the reaction goes to completion. Although the Endo H_f allows for simple membrane purification (MW = 70 KDa), there is significant loss of GRNase due to adsorption on the 30 KDa membrane when the GRNase is flowed through. However, there was little to no loss observed on the 3 KDa MWCO filter. In future experiments, this loss can be avoided using a different purification method, such as the use of bead-immobilized Endo H or Endo H_f (e.g. onto sepharose, paramagnetic beads, etc.), thereby allowing for simple purification but reducing adsorptive losses significantly.

Homogeneous MG_2RNase. A one-pot transglycosylation approach was performed using the GRNase and the oxazoline reaction product directly. These materials were mixed with Endo M at the desired pH, and the reaction was allowed to proceed at room temperature. The reaction was monitored using infusion of the products into the mass spectrometer as described in the experimental section. Fig. 4(b) is a representative spectrum obtained for the MG_2RNAse reaction product. Over time, peaks corresponding to the calculated m/z of the trisaccharaide-containing product begin to form, demonstrating the reaction progress.

Cumulative areas of the MG_2RNAse vs. GRNase peaks were used to determine relative yield over time. It was found that typical enzyme kinetics were followed, with approximately 8%, 13%, and 19% relative yield after 1, 3, and 20 hours respectively. The leveling of the relative yield at 19% after 20 hours indicate that Endo M does not show “off” activity for this glycan, as has been observed in the past [25]. However, the enzyme activity was quite low and kinetics were relatively slow compared to previous report using this same enzyme/substrate combination. For example, Rising et. al found that Endo M was capable of transferring purified

ManGlcNAc oxazoline to an amino acid acceptor with a 98% yield in 3 hours [25]. The slower kinetics observed using GRNase as the acceptor are in part due to inherently slower kinetics involved with using a protein vs. an amino acid GlcNAc as the acceptor. In addition however, it is likely that the relatively large ionic strength of the one-pot approach (due to DMC and ET_3NH^+) results in a lower enzyme activity and reaction rate. Purification of the oxazoline product prior to transglycosylation would likely result in a much more quantitative yield. This would add an additional step to the method, but a significant increase in yield and reaction rate may offset any loss in sample and/or time.

Although MS^1 data indicated the desired MG_2RNase product based on m/z, the site specificity of the transglycosylation should be verified. A sample of the transglycosylation reaction product was digested using trypsin, and subjected to LC-MS^n analysis for this purpose. Our results for native RNase B indicate that under the digestion conditions used, a glycopeptide corresponding to $ManGlcNAc_2$-NLTK should be observed. The $[M+2H]^{2+}$ ion corresponding to this glycopeptide (m/z = 522.2 Da) eluted at a retention time of 8.1 minutes. MS^n fragmentation of this ion (Fig. 5) showed very similar behavior to that observed for native RNase B (Fig. 2). Fig. 5(a) depicts the MS^2 of the parent ion. Oxonium marker ions for GlcNAc and ManGlcNAc are present, indicating the ion is a glycopeptide. Characteristic losses of a mannose (1+ and 2+ series), followed by a GlcNAc (1+ series) are also observed, verifying the glycan identity.

The MS^3 of the base peak (GlcNAc-peptide) from MS^2 showed fragmentation typical of glycopeptides, forming predominantly singly charged peptide ion with the entire glycan removed. A subsequent round of CID on this peptide (MS^4) is shown in Fig. 5(b). This spectrum is nearly identical to that shown in Fig. 2(d). The y- and b- ion series allow for positive

identification of the transglycosylation site, and verifies the product of the one-pot transglycosylation to be the RNase apoprotein with a $ManGlcNAc_2$ glycan attached at Asn-34.

The MG_2RNAse produced is a non-native glycan, in that N-linked glycans of this structure do not occur in nature. The production of a non-native glycan was used in this case due to the limited commercial availability of Endoglycosidase enzymes capable of these types of reactions. Many naturally occurring enzymes (including the commercially available Endo M) have significant "off" activity for natural glycans. However, recent trends have shown that simple point mutations of naturally occurring enzymes can be made to result in enzymes with solely "on" activity for natural glycans [18-20]. Wild type Endo A, initially thought to have little transglycosylation activity for complex type glycans, was found to produce a high product yield when performed under higher substrate and enzyme concentrations [20]. In fact, Ribonuclease B was recently remodeled in a one-pot fashion using purified oxazoline to form Ribonuclease C [20]. Future use of appropriate enzyme/substrate combinations show great promise for the rapid production of homogeneous glycoproteins in a one-pot fashion. It seems from the current study, and comparison to previous literature results, that purification of the oxazoline prior to transglycosylation will result in increased yield when the DMC/Base oxazoline synthesis method is utilized. Despite the additional purification step added for oxazoline purification, the aqueous phase synthesis still represents an increase in throughput compared to previous synthetic strategies. Future use of immobilized enzymes may also reduce any protein purification protocols necessary, resulting in very rapid routes to pure, homogeneous glycoproteins.

As an increasing number of homogeneous glycoproteins are prepared, it is crucial the structure and product identity are well characterized. The current study describes methodology useful for rapid product verification through LC-MS^n studies to verify glycan identity, and

glycosylation site of the product protein. Additional characterization using PNGase release, permethylation, 2-AB labeling, etc. can be carried out as well on homogeneous glycoproteins amenable to these methods to ensure product identity. Well characterized homogeneous glycoproteins could then serve as reference materials for the harmonization of glycoprotein analysis techniques (e.g. mass spectrometry methods) through inter-laboratory studies. In addition, the availability of such materials would allow for detailed assessment of the rate, efficiency, precision, and glycan specific effects of commonly used sample pre-treatment steps such as PNGase release, pronase digestion, derivatization, etc. A more detailed understanding of these principles would lead to more accurate and consistent glycoprotein characterization, aiding in the assurance of biopharmaceutical quality and identification of glycan related biomarkers for clinical diagnostics.

CONCLUSIONS

A rapid "one-pot" chemoenzymatic approach for the engineering of protein glycosylation was described. Ribonuclease B was effectively deglycosylated using Endo H_f, leaving the intact protein with a single GlcNAc residue attached at the original glycosylation site. A one-step sugar oxazoline synthesis was then performed on ManGlcNAc to form the necessary reactive intermediate. Mass spectrometry was used to identify the formation of the desired oxazoline sugar and was shown to be an effective tool for reaction monitoring.

The deglycosylated RNase B was mixed with the oxazoline reaction product and Endo M in order to enzymatically attach the desired sugar, forming a homogenous $ManGlcNAc_2$-Ribonuclease glycoform. $LC\text{-}MS^4$ methodology was developed using a tryptic digest of the product for verification of glycosylation site and identity. A glycopeptide containing the specific $ManGlcNAc_2$ attached at the desired Asn 34 was observed. The first stage of CID was used to

sequence and verify the glycan composition, while a second and third stage of CID allowed sequencing of the tryptic peptide confirming the glycosylation site.

The synthetic techniques described herein can be applied to a wide range of glycan/glycoprotein systems with the proper enzyme/substrate combination. This method allows the use of a glycan substrate of known composition, branching, and stereochemistry to ensure a product containing a homogenous, well defined glycan. Continued development and availability of glycosyltransferase enzymes should decrease substrate specificity requirements, allowing synthesis of homogeneous glycoproteins containing specific glycans commonly found in nature. The combined use of rapid chemoenzymatic glycan rearrangement and MS-based product verification has the potential to produce well-characterized homogeneous glycoproteins. Such materials could serve as standards useful for the detailed assessment and harmonization of glycoprotein characterization techniques. This increased understanding is necessary to ensure the accuracy and precision in clinical glycoprotein measurements, compare glycosylation of different biopharmaceutical lots, and establish biosimilar product identity.

DISCLAIMER

Commercial equipment, instruments, and materials are identified throughout this paper to adequately specify the experimental procedure. Such identification does not imply recommendation or endorsement by NIST nor does it imply that the equipment, instruments, or materials are necessarily the best available for the purpose.

Reference List

[1] An HJ, Kronewitter SR, de Leoz MLA, Lebrilla CB, Glycomics and disease markers. Current Opinion in Chemical Biology 13, 601-607 (2009)

[2] Packer NH, von der Lieth CW, Aoki-Kinoshita KF, Lebrilla CB, Paulson JC, Raman R, Rudd P, Sasisekharan R, Taniguchi N, York WS, Frontiers in glycomics: Bioinformatics and biomarkers in disease - An NIH White Paper prepared from discussions by the focus groups at a workshop on the NIH campus, Bethesda MD (September 11-13,2006). Proteomics 8, 8-20 (2008)

[3] Higgens E, Carbohydrate analysis throughout the development of a protein therapeutic. Glycoconjugate J. 27, 211-225 (2010)

[4] Kawasaki N, Itoh S, Hashii N, Takakura D, Qin Y, Huang XY, Yamaguchi T, The Significance of Glycosylation Analysis in Development of Biopharmaceuticals. Biological & Pharmaceutical Bulletin 32, 796-800 (2009)

[5] van Berkel PHC, Gerritsen J, Perdok G, Valbjorn J, Vink T, van de Winkel JGJ, Parren PWHI, N-Linked Glycosylation is an Important Parameter for Optimal Selection of Cell Lines Producing Biopharmaceutical Human IgG. Biotechnology Progress 25, 244-251 (2009)

[6] An HJ, Froehlich JW, Lebrilla CB, Determination of glycosylation sites and site-specific heterogeneity in glycoproteins. Current Opinion in Chemical Biology 13, 421-426 (2009)

[7] Dalpathado DS, Desaire H, Glycopeptide analysis by mass spectrometry. Analyst 133, 731-738 (2008)

[8] Huhn C, Selman MHJ, Ruhaak LR, Deelder AM, Wuhrer M, IgG glycosylation analysis. Proteomics 9, 882-913 (2009)

[9] Wuhrer M, Catalina MI, Deelder AM, Hokke CH, Glycoproteomics based on tandem mass spectrometry of glycopeptides. J.Chromatogr.B. 849, 115-128 (2007)

[10] Zaia J, Mass spectrometry of oligosaccharides. Mass Spectrom.Rev. 23, 161-227 (2004)

[11] Prater BD, Connelly HM, Qin Q, Cockrill SL, High-throughput immunoglobulin G N-glycan characterization using rapid resolution reverse-phase chromatography tandem mass spectrometry. Analytical Biochemistry 385, 69-79 (2009)

[12] Alley WR, Mechref Y, Novotny MV, Characterization of glycopeptides by combining collision-induced dissociation and electrom transfer dissociation mass spectrometry data. Rapid Comm.Mass Spectrom. 23, 161-170 (2009)

[13] Angel PM, Lim JM, Wells L, Bergmann C, Orlando R, A potential pitfall in O-18-based N-linked glycosylation site mapping. Rapid Comm.Mass Spectrom. 21, 674-682 (2007)

[14] Wada Y, Azadi P, Costello CE, Dell A, Dwek RA, Geyer H, Geyer R, Kakehi K, Karlsson NG, Kato K, Kawasaki N, Khoo KH, Kim S, Kondo A, Lattova E, Mechref Y, Miyoshi E, Nakamura K, Narimatsu H, Novotny MV, Packer NH, Perreault H, Peter-Katalinic J, Pohlentz G, Reinhold VN, Rudd PM, Suzuki A, Taniguchi N, Comparison of the methods for profiling glycoprotein glycans - HUPO Human Disease Glycomics/Proteome Initiative multi-institutional study. Glycobiology 17, 411-422 (2007)

[15] Wagner-Rousset E, Bednarczyk A, Bussat MC, Colas O, Corvaia N, Schaeffer C, Van Dorsselaer A, Beck A, The way forward, enhanced characterization of therapeutic antibody glycosylation: Comparison of three level mass spectrometry-based strategies. J.Chromatogr.B. 872, 23-37 (2008)

[16] Wuhrer M, Koeleman CAM, Hokke CH, Deelder AM, Mass spectrometry of proton adducts of fucosylated N-glycans: fucose transfer between antennae gives rise to misleading fragments. Rapid Comm.Mass Spectrom. 20, 1747-1754 (2006)

[17] Fujita K, Yamamoto K, A remodeling system for the oligosaccharide chains on glycoproteins with microbial endo-beta-N-acetylglucosaminidases. Biochimica Et Biophysica Acta-General Subjects 1760, 1631-1635 (2006)

[18] Heidecke CD, Ling ZL, Bruce NC, Moir JWB, Parsons TB, Fairbanks AJ, Enhanced glycosylation with mutants of endohexosaminidase A (Endo A). Chembiochem 9, 2045-2051 (2008)

[19] Huang W, Li C, Li B, Umekawa M, Yamamoto K, Zhang X, Wang LX, Glycosynthases Enable a Highly Efficient Chemoenzymatic Synthesis of N-Glycoproteins Carrying Intact Natural N-Glycans. J.Am.Chem.Soc. 131, 2214-2223 (2009)

[20] Huang W, Yang Q, Umekawa M, Yamamoto K, Wang LX, Arthrobacter Endo-B-N-Acetylglucosaminidase Shows Transglycosylation Activity on Complex-Type N-glycan Oxazolines: One-Pot Conversion of Ribonuclease B to Sialylated Ribonuclease C. Chembiochem 11, 1350-1355 (2010)

[21] Li B, Zeng Y, Hauser S, Song HJ, Wang LX, Highly efficient endoglycosidase-catalyzed synthesis of glycopeptides using oligosaccharide oxazolines as donor substrates. J.Am.Chem.Soc. 127, 9692-9693 (2005)

[22] Li B, Song HJ, Hauser S, Wang LX, A highly efficient chemoenzymatic approach toward glycoprotein synthesis. Org.Let. 8, 3081-3084 (2006)

[23] Ochiai H, Huang W, Wang LX, Expeditious chemoenzymatic synthesis of homogeneous N-glycoproteins carrying defined oligosaccharide ligands. J.Am.Chem.Soc. 130, 13790-13803 (2008)

[24] Rising TWDE, Heidecke CD, Moir JWB, Ling ZL, Fairbanks AJ, Endohexosaminidase-catalysed glycosylation with oxazoline donors: Fine tuning of catalytic efficiency and reversibility. Chemistry-A European Journal 14, 6444-6464 (2008)

[25] Rising TWDF, Claridge TDW, Moir JWB, Fairbanks AJ, Endohexosaminidase M: Exploring and exploiting enzyme substrate specificity. Chembiochem 7, 1177-1180 (2006)

[26] Rising TWDF, Claridge TDW, Davies N, Gamblin DP, Moir JWB, Fairbanks AJ, Synthesis of N-glycan oxazolines: donors for endohexosaminidase catalysed glycosylation. Carb.Res. 341, 1574-1596 (2006)

[27] Wang LX, Chemoenzymatic synthesis of glycopeptides and glycoproteins through endoglycosidase-catalyzed transglycosylation. Carb.Res. 343, 1509-1522 (2008)

[28] Wei YD, Li CS, Huang W, Li B, Strome S, Wang LX, Glycoengineering of human IgG1-Fc through combined yeast expression and in vitro chemoenzymatic glycosylation. Biochem. 47, 10294-10304 (2008)

[29] Umekawa M, Huang W, Li B, Fujita K, Ashida H, Wang LX, Yamamoto K, Mutants of Mucor hiemalis endo-beta-N-acetylglucosaminidase show enhanced transglycosylation and glycosynthase-like activities. J.Biol.Chem. 283, 4469-4479 (2008)

[30] Noguchi M, Tanaka T, Gyakushi H, Kobayashi A, Shoda S, Efficient Synthesis of Sugar Oxazolines from Unprotected N-Acetyl-2-amino Sugars by Using Chloroformamidinium Reagent in Water. J.Org.Chem. 74, 2210-2212 (2009)

[31] Smyth DG, Stein WH, Moore S, Sequence of Amino Acid Residues in Bovine Pancreatic Ribonuclease - Revisions and Confirmations. J.Biol.Chem. 238, 227-& (1963)

[32] Witte K, Sears P, Martin R, Wong CH, Enzymatic glycoprotein synthesis: Preparation of ribonuclease glycoforms via enzymatic glycopeptide condensation and glycosylation. J.Am.Chem.Soc. 119, 2114-2118 (1997)

[33] Domon B, Costello CE, A Systematic Nomenclature for Carbohydrate Fragmentations in Fab-Ms Ms Spectra of Glycoconjugates. Glycoconjugate J. 5, 397-409 (1988)

[34] Harvey DJ, Collision-induced fragmentation of underivatized N-linked carbohydrates ionized by electrospray. Journal of Mass Spectrometry 35, 1178-1190 (2000)

[35] Hofmeister GE, Zhou Z, Leary JA, Linkage Position Determination in Lithium-Cationized Disaccharides - Tandem Mass-Spectrometry and Semiempirical Calculations. J.Am.Chem.Soc. 113, 5964-5970 (1991)

Figure Captions

Fig. 1. Global scheme for "one-pot" chemoenzymatic synthesis of a homogeneous glycoprotein.

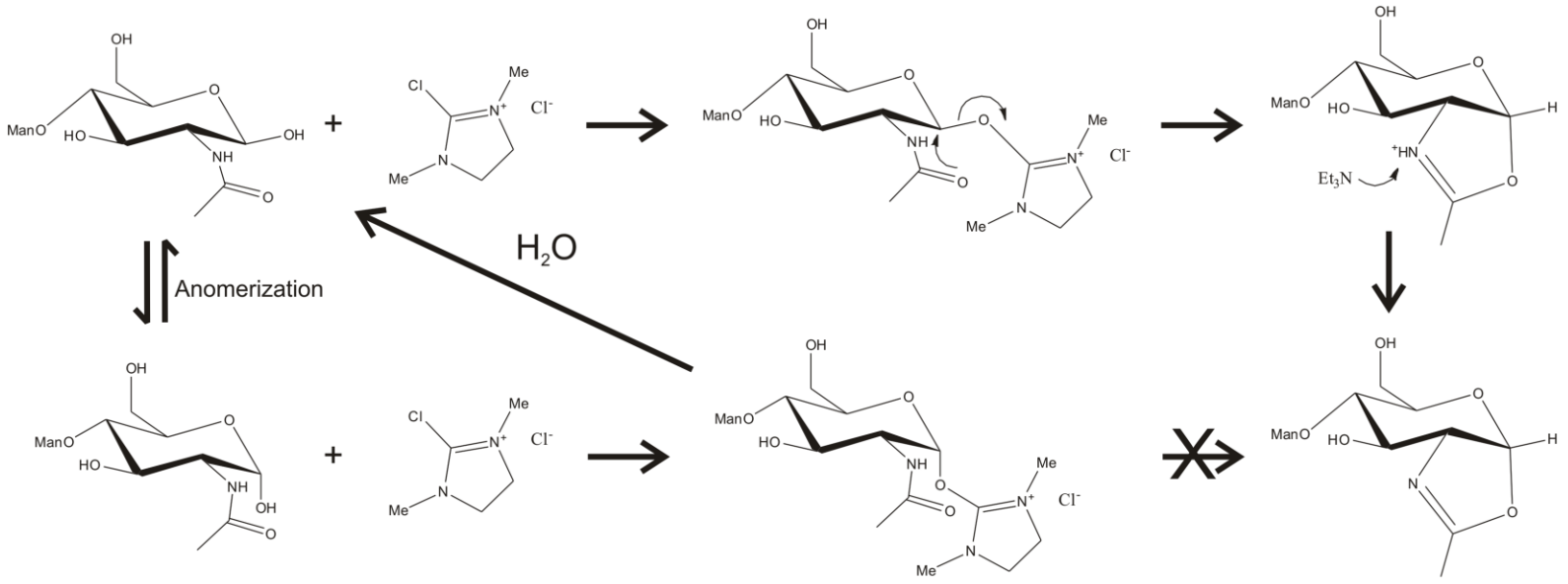
Fig. 2. Representative (a) direct infusion mass spectra of native RNAse B and LC-MS^n data collected for a tryptic digest of native RNase B. The average MS^1 spectrum (b) over the elution period of 4.1 to 6.2 minutes where all five glycopeptides eluted. Representative (c) MS^2 and (d) MS^4 spectra obtained for the Man6 glycopeptide demonstrating characteristic glycan fragmentation in MS^2 and peptide fragmentation in MS^4 for complete characterization of the glycopeptide identity.

Fig. 3. Representative (a) MS^1 spectra for the monitoring of oxazoline synthesis progression and MS^2 spectra of (b) native ManGlcNAc and (c) oxazoline product.

Fig. 4. Infusion mass spectra collected for intact protein after (a) Endo H_f treatment to form GRNase and after (b) Endo M catalyzed transglycosylation to form MG_2RNase.

Fig. 5. LC-MS^n spectra collected for tryptic digest of MG_2RNase reaction product. The mass spectra collected are (a) MS^2 of the glycopeptides demonstrating glycan fragmentation and (b) MS^4 resulting in peptide fragmentation for complete glycopeptides characterization and verification of product formation.

Figure 1

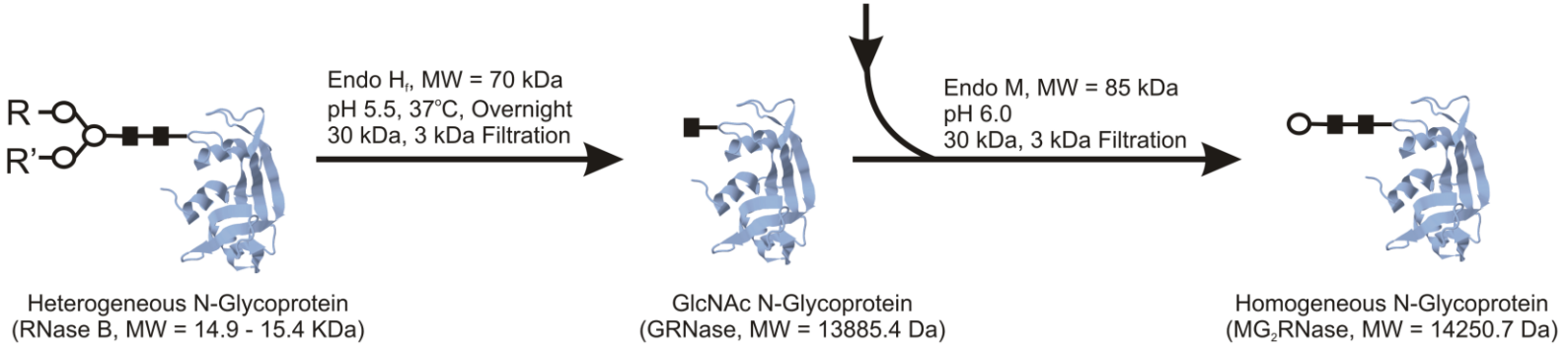


Man GlcNAc (a)	
Ion	m/z
$[a+H]^+$	384.1
$[a+Na]^+$	406.1
$[a+Na+Et_3NH+Cl]$	543.2

α–anomer product ($c+Cl^-$)	
Ion	m/z
$[c]^+$	480.2

Man GlcNAc oxazoline (b)	
Ion	m/z
$[b+H]^+$	366.1
$[b+Na]^+$	388.1
$[b+Na+Et_3NH+Cl]^+$	525.2

Oxazoline Reaction Products



Heterogeneous N-Glycoprotein
(RNase B, MW = 14.9 - 15.4 KDa)

GlcNAc N-Glycoprotein
(GRNase, MW = 13885.4 Da)

Homogeneous N-Glycoprotein
(MG_2RNase, MW = 14250.7 Da)

Figure 2

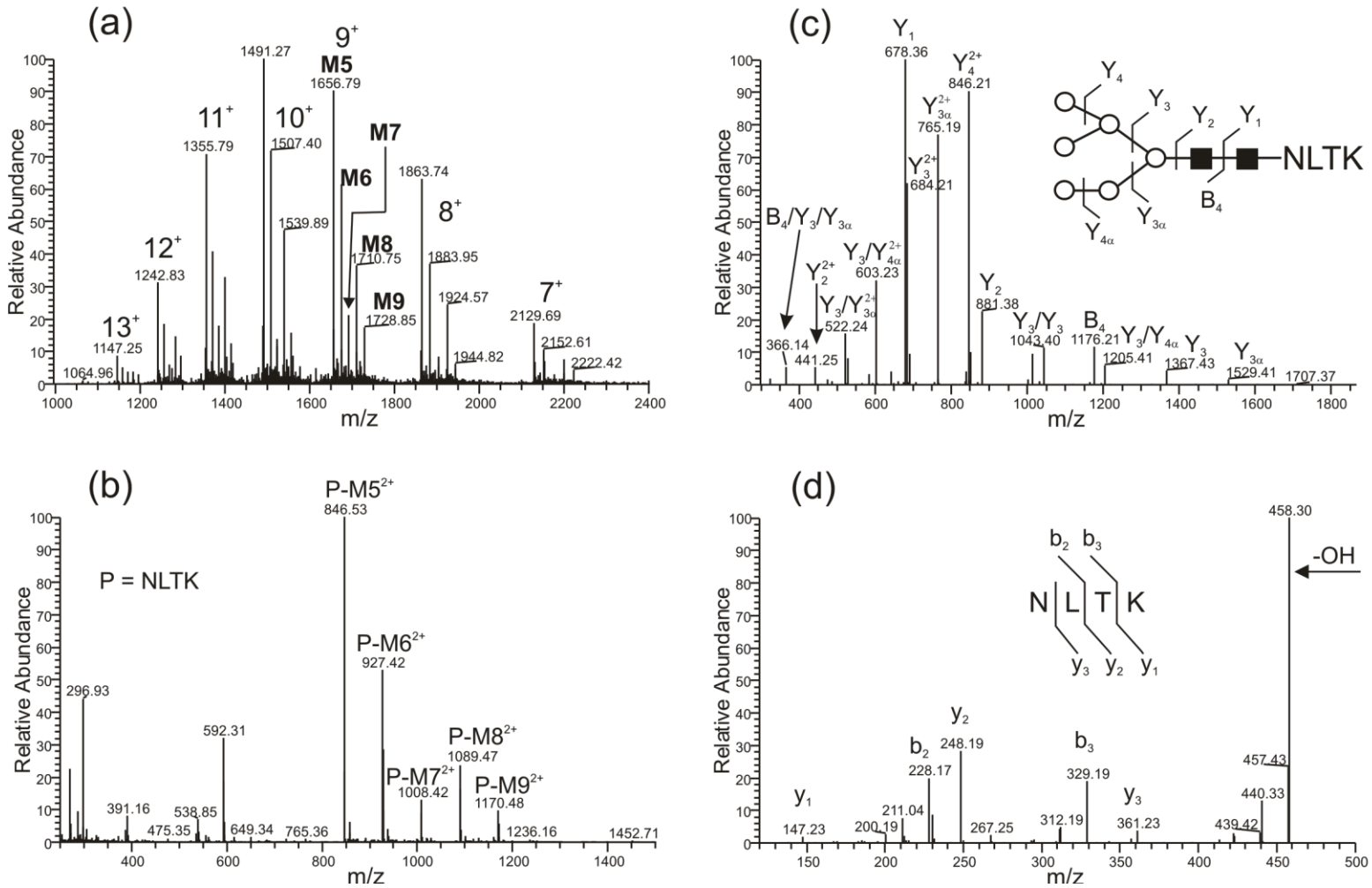

(a)
Relative Abundance
m/z
9+
M5
1656.79
1491.27
11+
1355.79
10+
1507.40
M7
M6
1863.74
8+
1539.89
12+
1242.83
M8
1710.75
1883.95
M9
1728.85
1924.57
7+
2129.69
13+
1147.25
2152.61
1064.96
1944.82
2222.42
(b)
P = NLTK
P-M5 2+
846.53
P-M6 2+
927.42
296.93
592.31
P-M8 2+
1089.47
P-M7 2+
1008.42
P-M9 2+
1170.48
391.16
538.85
475.35
649.34
765.36
1236.16
1452.71
(c)
Y1
678.36
Y4 2+
846.21
Y3α 2+
765.19
Y3 2+
684.21
B4/Y3/Y3α
Y3/Y4α 2+
603.23
Y2 2+
Y3/Y3α 2+
522.24
Y2
881.38
Y3/Y3
1043.40
B4
1176.21
Y3/Y4α
Y3
Y3α
366.14
441.25
1205.41
1367.43
1529.41
1707.37
Y4
Y3
Y2
Y1
NLTK
B4
Y3α
Y4α
(d)
458.30
-OH
b2 b3
N L T K
y3 y2 y1
y2
b2
248.19
228.17
b3
329.19
457.43
y1
y3
211.04
440.33
147.23
200.19
267.25
312.19
361.23
439.42

Figure 3

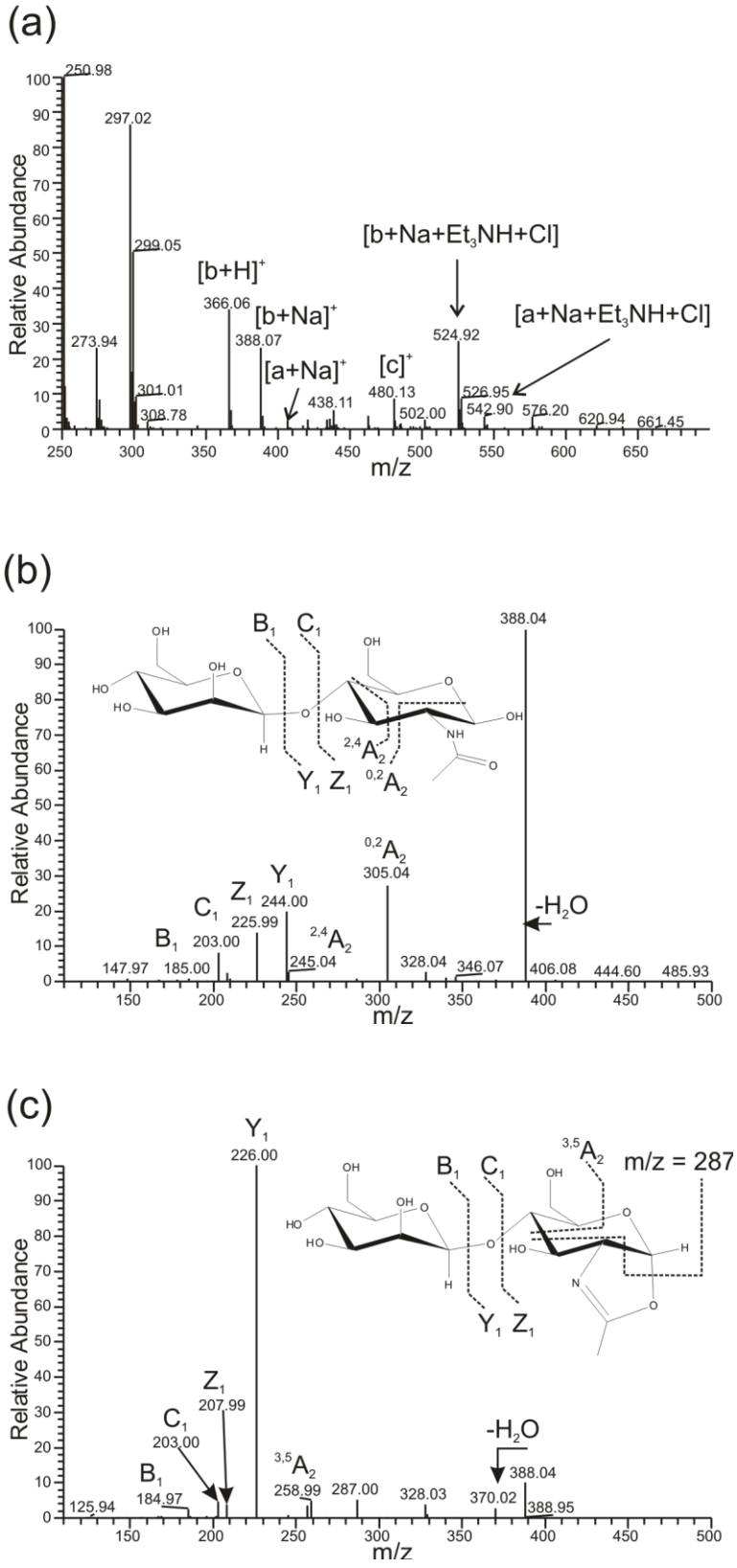

(a)
Relative Abundance
m/z
250.98
297.02
299.05
273.94
301.01
308.78
366.06
388.07
438.11
480.13
502.00
524.92
526.95
542.90
576.20
620.94
661.45
[b+H]+
[b+Na]+
[a+Na]+
[c]+
[b+Na+Et3NH+Cl]
[a+Na+Et3NH+Cl]
(b)
Relative Abundance
m/z
B1
C1
Y1
Z1
0,2A2
2,4A2
388.04
305.04
244.00
225.99
203.00
245.04
147.97
185.00
328.04
346.07
406.08
444.60
485.93
-H2O
(c)
Relative Abundance
m/z
Y1
226.00
Z1
207.99
C1
203.00
B1
184.97
125.94
3,5A2
258.99
287.00
328.03
370.02
388.95
388.04
-H2O
m/z = 287

Figure 4

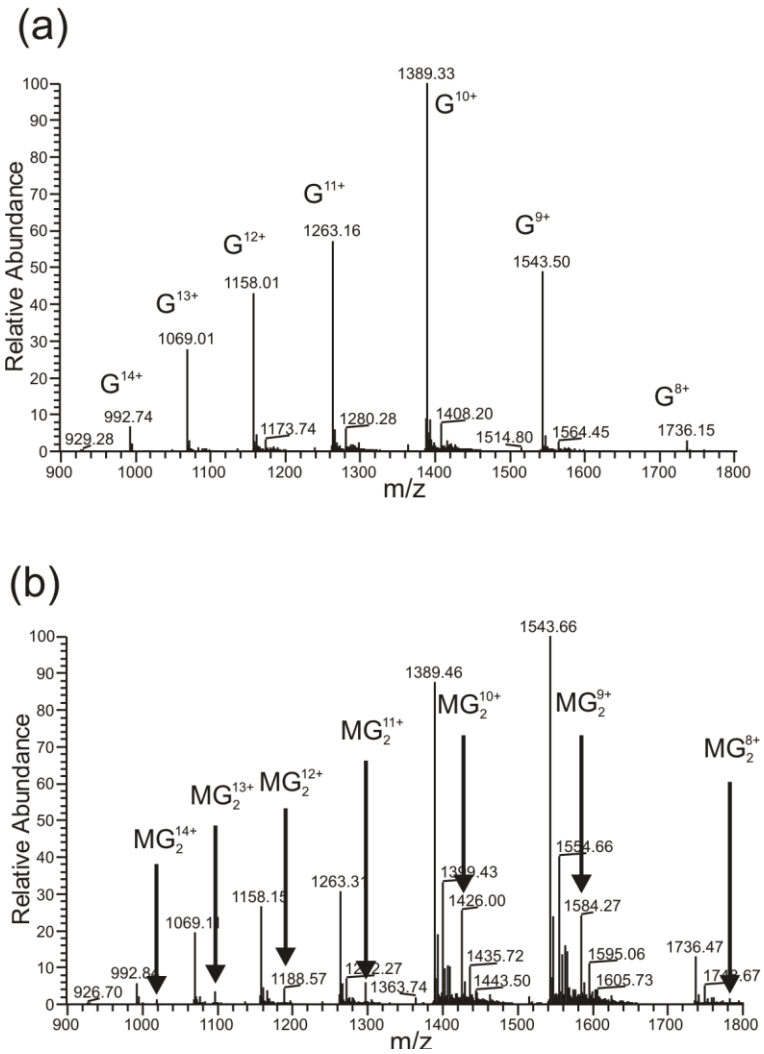
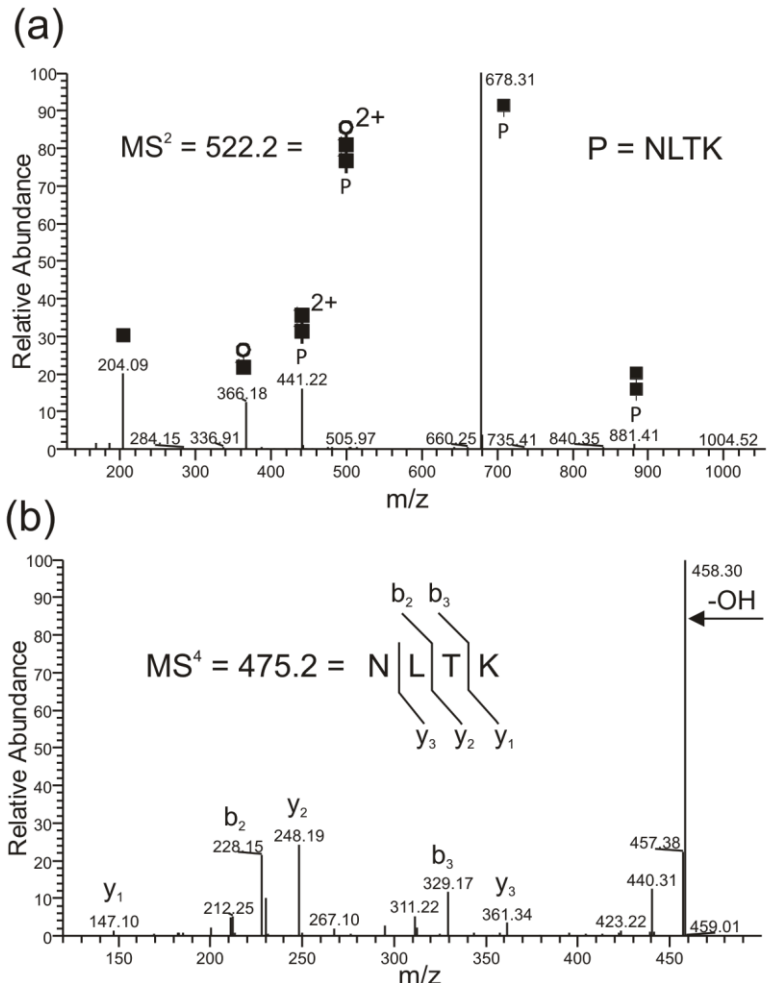

(a)
Relative Abundance
m/z
G14+
G13+
G12+
G11+
G10+
G9+
G8+
992.74
1069.01
1158.01
1263.16
1389.33
1543.50
1736.15
929.28
1173.74
1280.28
1408.20
1514.80
1564.45
(b)
Relative Abundance
m/z
MG2 14+
MG2 13+
MG2 12+
MG2 11+
MG2 10+
MG2 9+
MG2 8+
1543.66
1389.46
1554.66
1584.27
1426.00
1399.43
1263.31
1158.15
1069.11
992.8
926.70
1188.57
1435.72
1443.50
1595.06
1605.73
1736.47

Figure 5


(a)
MS² = 522.2 =
2+
P
P = NLTK
678.31
204.09
366.18
441.22
881.41
Relative Abundance
m/z
(b)
MS⁴ = 475.2 =
N L T K
b₂ b₃
y₃ y₂ y₁
458.30
-OH
b₂
y₂
b₃
y₃
y₁
228.15
248.19
329.17
361.34
147.10
212.25
267.10
311.22
423.22
457.38
440.31
459.01
Relative Abundance
m/z