Anal Bioanal Chem (2012) 403:2279–2289
DOI 10.1007/s00216-012-5749-5

ORIGINAL PAPER

LC-MS/MS biopharmaceutical glycoanalysis: identification of desirable reference material characteristics

John E. Schiel · Jennifer Au · Hua-Jun He ·
Karen W. Phinney

Received: 11 November 2011 / Revised: 10 January 2012 / Accepted: 13 January 2012 / Published online: 22 February 2012
© Springer-Verlag (outside the USA) 2012

Abstract Glycosylation, the enzymatic addition of carbohydrates to a protein, is one of the most abundant post-translational modifications found in nature. There is variability in the number, location, and identity of glycans attached. As a result, a glycoprotein consists of a number of glycoforms with different combinations of glycans, potentially resulting in different stability, toxicity, and activity. This is especially important in the biopharmaceutical industry where product consistency and safety are vital. Glycoprotein analysis involves numerous mass spectrometry based techniques, each of which provides various aspects of characterization. The current paper describes two commonly used analytical techniques for glycoprotein characterization. In one experiment, nonspecific proteolysis is combined with a two-tiered mass spectrometry approach (MALDI-TOF and LC-MS/MS) to gain glycosylation site and glycan identity. In a second approach, glycans were enzymatically released, labeled with a fluorescent dye, and analyzed using LC-Fluorescence-MS/MS to give glycan identification and relative quantification. The type and degree of information yielded by each method is assessed in an effort to identify desired reference material characteristics for improving biopharmaceutical glycoanalysis.

Published in the special issue *Young Investigators in Analytical and Bioanalytical Science* with guest editors S. Daunert, J. Bettmer, T. Hasegawa, Q. Wang and Y. Wei.

Electronic supplementary material The online version of this article (doi:10.1007/s00216-012-5749-5) contains supplementary material, which is available to authorized users.

J. E. Schiel (✉) · J. Au · H.-J. He · K. W. Phinney
National Institute of Standards and Technology,
Analytical Chemistry Division,
100 Bureau Dr., Stop 8392,
Gaithersburg, MD 20899, USA
e-mail: john.schiel@nist.gov

John Schiel
has been a Research Chemist at the National Institute of Standards and Technology since July 2009. He was the recipient of the *Bioanalysis* Young Investigator Award and an ACS Division of Analytical Chemistry Fellowship. John's research is directed toward improving the reliability of mass spectrometry based glycoprotein characterization through reference material development.

Keywords Biopharmaceutical · Glycoprotein · Glycan · Mass spectrometry · Standards · Reference material

Introduction

Proteins are complex structures composed of a sequence of amino acids folded into a three-dimensional structure, and in many cases, may have multiple associated subunits. This complexity is often increased by the addition of post-translational modifications (PTM) such as the addition of phosphate or carbohydrate (glycan) chains to specific sites. One common PTM is the addition of a glycan to the side chain of an asparagine residue in the consensus sequence N-X-S/T, where X can be any amino acid other than proline. These N-linked glycans can be greatly variable in structure and a given protein may contain one of many glycans at the same site [1]. Therefore, glycoproteins exist as a mixture of glycoforms as opposed to a single, well-defined structure.

The pharmaceutical industry has developed a number of structurally complex glycoproteins into drugs [2–4]. Although protein sequence is directed by DNA, glycosylation can be variable and change with growth conditions, between batches, and when different cell lines are used [3, 4]. Glycosylation can affect the stability, toxicity, and activity of a given drug, and therefore must be monitored [2–4]. Analysis of biopharmaceutical glycosylation is often performed through the enzymatic release of all N-linked glycans from a protein population, followed by fluorescent labeling and high-performance liquid chromatography (HPLC) or electrophoretic analysis [2, 5–8]. Mass spectrometry can also be used (with or without fluorescent labels) as a detection method, and glycan sequencing can be performed using liquid chromatography tandem mass spectrometry (LC-MS/MS) strategies [8–10].

Protease digestions have also proven useful for glycoanalysis. LC-MS^n (multiple sequential stages of fragmentation) of a trypsin digest can give both glycosylation site and glycan identity, however, the complexity of this mixture can lead to ion suppression of the glycopeptides and glycopeptide enrichment can often be difficult [9, 11, 12]. An alternative method is the use of a protease mixture (e.g., pronase) capable of cleaving all peptide bonds [11–23]. The result is a mixture of primarily amino acids; however, steric hindrance around the glycosylation site results in a peptide tag of 1 to 6 amino acids with the glycan attached at its original site. This digestion procedure allows for simple and efficient purification of the glycopeptides prior to mass spectrometry, thereby reducing ionization suppression [14, 16, 17]. The resulting sample can be analyzed using matrix assisted laser desorption ionization mass spectrometry (MALDI) and/or LC-MS^n [12, 14–18, 21–23]. Software has recently been developed capable of matching accurate mass against a theoretical library of glycan compositions and the known protein sequence to automate compositional matching [13, 15].

This manuscript describes pronase digestion and analysis of rituximab (a representative biopharmaceutical) glycosylation with MALDI-time of flight mass spectrometry (MALDI-TOF) and LC-MS/MS as well as 2-aminobenzamide (2-AB) fluorescence labeling and LC-Fluorescence-MS/MS. The combined methods allow for identification of glycosylation site, glycan identity, and relative glycan quantification. Specific attention is devoted to how well-defined reference materials could benefit these and other glycoanalytical strategies.

Experimental

Reagents Pronase, CNBr-activated sepharose 4B, 2,5-dihydroxybenzoic acid (DHB), ethanolamine, PNGase F, 2-aminobenzamide, GlycoProfile cartridges, and RNase B were obtained from Sigma (St. Louis, MO). Rituximab was obtained from RxUSA (Port Washington, NY). Purified 2-AB glycan standards (G0, Man5, G0F, G1F, G2F) and 2-AB human IgG standard kit were purchased from Prozyme (Hayward, CA). Burdick and Jackson LC-MS grade water and acetonitrile (ACN) with and without 0.1% (*v*/*v*) formic acid were obtained from VWR (West Chester, PA). Carbograph Extract Clean columns were purchased from Grace (Deerfield, IL). All other chemicals used were reagent grade or better.

Glycoprotein digestion using pronase Pronase was immobilized onto CNBr-activated sepharose using a previously published procedure [16]. The overall ratio of reagents used was 1mg pronase/150 mg CNBr activated beads/1 mg of glycoprotein and was allowed to digest at 37 °C for 20 h. An additional digestion step (23 hours at 37 °C) was performed on rituximab by adding 1 mg dissolved pronase to the supernatant from the immobilized digestion. Glycopeptides were purified into three fractions (10%, 20%, and 40% ACN (*v*/*v*) using graphitized carbon SPE according to a previously published procedure [12].

Glycopeptide MALDI-TOF Each pronase SPE fraction was reconstituted in approximately 200 µL of water per 1 mg of glycoprotein digested. Matrix was prepared to contain 0.01 g DHB in 200 µL of 50%/50% (*v*/*v*) ACN/water. Spotting solutions were prepared by mixing equal volumes of matrix and resuspended sample to result in a final spotting solution with approximately 25 µg DHB/µL in 25% (*v*/*v*) ACN. MALDI-TOF spectra were collected on an ABI Voyager 4700 mass spectrometer operated in the positive reflectron mode. The laser intensity was set to 4,500 to 5,000. The spectra were externally calibrated based on the 4700 calibration mixture from Applied Biosystems (Foster City, CA).

Graphitized carbon LC-MS/MS Following MALDI-TOF analysis, each fraction that yielded peaks in MALDI-TOF spectra was combined. This sample was diluted with 0.1% (*v*/*v*) formic acid in water to prepare a solution containing approximately 5 pmol glycopeptide/µL and 60 µL was injected in the LC-MS/MS experiment.

LC-MS/MS was performed on a Thermo ESI-LTQ-XL mass spectrometer using an Agilent 1100 HPLC system for sample introduction. Injections were made onto a Thermo Hypercarb column (100×2.1 mm i.d., 5 µm particle diameter) at 0.3 mL/min. The LC run consisted of a binary gradient of 0.1% (*v*/*v*) formic acid in water (solvent A) and 0.1% (*v*/*v*) formic acid in acetonitrile (solvent B). The gradient used was (%A/%B, *v*/*v*) 97.0%/3.0% at 0 min, 97.0%/3.0% at 4.6 min, 83.1%/16.9% at 36.5 min, 56.9%/41.3% at 54.7 min, 10.0%/90.0% at 63.8 min, 10.0%/90.0% at

82.0 min, 97.0%/3.0% at 91.2 min, and 97.0%/3.0% at 127.0 min. The LTQ-XL was set to collect data in a data dependent MS^1/MS^2 mode. CID spectra were collected using an isolation width of 2 Da, collision energy of 25, an activation Q of 0.25, and an activation time of 30 ms. Dynamic exclusion was enabled and set to exclude a selected precursor ion for 5 s after 5 consecutive MS^2 had been collected for that ion.

The ion optics of the mass spectrometer was tuned by infusing an RNase B pronase digest. The infused solution was prepared to contain approximately 30 pmol glycopeptides/μL dissolved in 0.1% (*v*/*v*) formic acid in 25%/75% (*v*/*v*) mixture of ACN/water. This mixture was infused at 5 μL/min with a 0.3 mL/min LC makeup flow of 0.1% (*v*/*v*) formic acid in 25%/75% (*v*/*v*) ACN/water. The ion optics were autotuned to the 753.3 Da peak, later shown to be the RN-Man5 glycopeptide. The transfer capillary temperature was set to 275 °C.

2-AB labeling and purification Glycans were enzymatically released from purified rituximab using PNGase F in 20 mmol/L ammonium bicarbonate buffer (pH 7.8) for 2 h at 37 °C. MALDI-TOF analysis of the intact protein before and after glycan release was used to ensure complete glycan release. The sample was analyzed as an equal volume mixture with 20 μg sinapinic acid/μL in 50/50 (*v*/*v*) ACN/water containing 0.1% (*v*/*v*) TFA matrix. Released glycans were separated by protein precipitation in 75% ethanol and the supernatant was dried in a vacuum centrifuge. Glycans were reconstituted in 10 μL of 30%/70% (*v*/*v*) acetic acid/DMSO containing 60 μg/μL $NaCNBH_4$ and 50 μg/μL 2-AB (65 °C for 3 h). Labeled glycans were purified using a Glycoprofile glycan clean-up cartridge according to the manufacturer's instructions.

2-AB glycan LC-fluorescence-MS/MS The dried 2-AB glycans were reconstituted by adding 125 μL of water per 250 μg of rituximab starting material. Each of the purified glycan standards was reconstituted with water to a concentration of 1 pmol/μL. LC-F-MS/MS analysis was performed on 75, 20, and 80 μL injections of the rituximab, standards, and human IgG 2-AB labeled glycans, respectively.

Online LC-fluorescence-MS/MS was performed using the system described previously with a Linear Brand fluorescence detector online between the LC and MS. A slightly modified procedure from the literature was used [8]. The samples were injected onto a Zorbax Stable-Bond Rapid resolution HT C18 column (50×2.1 mm i.d., particle diameter=1.8 μm) at 0.333 mL/min and 50 °C. A binary gradient was used for the analytical separation (Solvent A=0.1% (*v*/*v*) acetic acid in water, solvent B=5%/95% (*v*/*v*) acetonitrile (ACN)/water with 0.1% acetic acid (*v*/*v*). The analytical gradient consisted of (%A/%B, *v*/*v*) 70.0%/30.0% at 0 min, 70.0%/30.0% at 2 min, 51.9%/48.1% at 35 min, 41.8%/58.2% at 42 min, and 41.8%/58.2% at 50 min. A third solvent (solvent C) consisting of 80%/20% (*v*/*v*) ACN/water with 0.1% (*v*/*v*) acetic acid was used to rinse the column for 10 min, followed by equilibration for 20 min under the original starting conditions prior to the next sample injection.

Fluorescence signal ($\lambda_{excitation}$=330 nm and $\lambda_{emission}$=420 nm) was recorded using Xcalibur software and an ss420x A/D converter from Agilent. The mass spectrometer was set to collect in a data dependent MS^1/MS^2 mode. MS^2 CID spectra were collected using an isolation width of 3 Da, collision energy of 35, an activation Q of 0.25, and an activation time of 30 ms. Dynamic exclusion was enabled and set to exclude a selected precursor ion for 5 seconds after 10 consecutive MS^2 had been collected for that ion.

The ion optics of the mass spectrometer were tuned to a 1-pmol/μL (water) solution of 2-AB G0 glycan obtained from Prozyme. This solution was infused into the mass spectrometer at 30 μL/min with an LC makeup flow of 70%/30% (solvent A/solvent B, *v*/*v*) at 0.3 mL/min using a mixing tee. The ion optics were autotuned to the $[M+2H]^{2+}$ ion observed at 719.2 Da. The transfer capillary was set to 250 °C.

Data analysis See Electronic Supplementary Material for data analysis procedures using SimGlycan, Glycopeptide Finder, and Xcalibur Software [24].

Results and discussion

Rituximab pronase digestion

Rituximab MALDI-TOF Glycopeptide peaks were present in the 20 and 40% fractions and ionized as $[M+Na]^+$ and $[M+2Na-H]^+$ adducts (Fig. 1). The 40% fraction contained only peaks that were also in the 20% fraction. A total of 22 potential glycopeptide ions were identified in the spectra, corresponding to 16 potential glycopeptides. Nearly every major peak in the MALDI spectrum resulted in a potential match except for the two at *m*/*z*=1,668.57 Da and *m*/*z*=1,830.63 Da. These peaks were initially thought to be due to water loss during the ionization process, however additional data collected during LC-MS/MS described below indicates they are modified glycopeptide.

Rituximab LC-MS/MS The overall designation of each chromatographic peak after MS^2 analysis is provided in Fig. 2 (a, b), as well as Table 1. Figure 2 (c, d) shows the representative XIC chromatograms corresponding to the *m*/*z* for N- and NS-glycopeptides with the G1F glycan. Both of these *m*/*z* co-elute with one of the major peaks in the base peak chromatogram, and seem to show a "double peak"

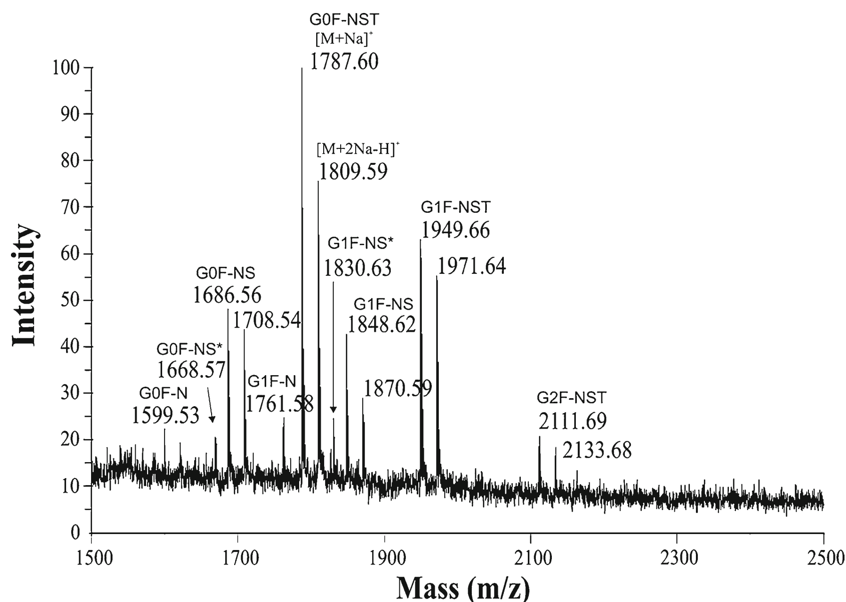


Fig. 1 Representative MALDI-TOF spectrum of rituximab pronase digest (20% fraction). The "*" represents glycopeptide verified later to be −18 Da product of indicated species. See text for details

elution behavior. The double peak elution behavior represents resolution of the NS-G1F isomers with galactose attached to either the 1,3 or 1,6 antennae. Graphitized carbon LC-MS has recently been shown to be capable of resolving a wide variety of pronase-derived glycopeptide isomers [25, 26]. Ma et al. performed exoglycosidase digestion and CE-LIF mobility shift assays to determine that the dominant G1F isomer present in rituximab contains galactose on the 1,6 arm [7]. This indicates the earlier eluting peak corresponds to the glycopeptide isomer with a 1,6 galactosylated antennae.

Although the G1F glycopeptide isomers were resolved, the attached peptide had little effect on the elution behavior in this case. This is depicted in Fig. 2 (c, d), where the XIC chromatograms show the N- and NS- glycopeptides of the G1F glycan co-elute. In fact, for the G0F, G1F, and G2F glycans, XIC chromatograms showed the peptide portion (N, NS, or NST) had little effect on retention. The peptide does play a significant role in elution of other species such as the G0-NS and G0-NST glycopeptides. This reiterates the fact that graphitized carbon can have complex retention behavior affected by glycan, peptide, or both.

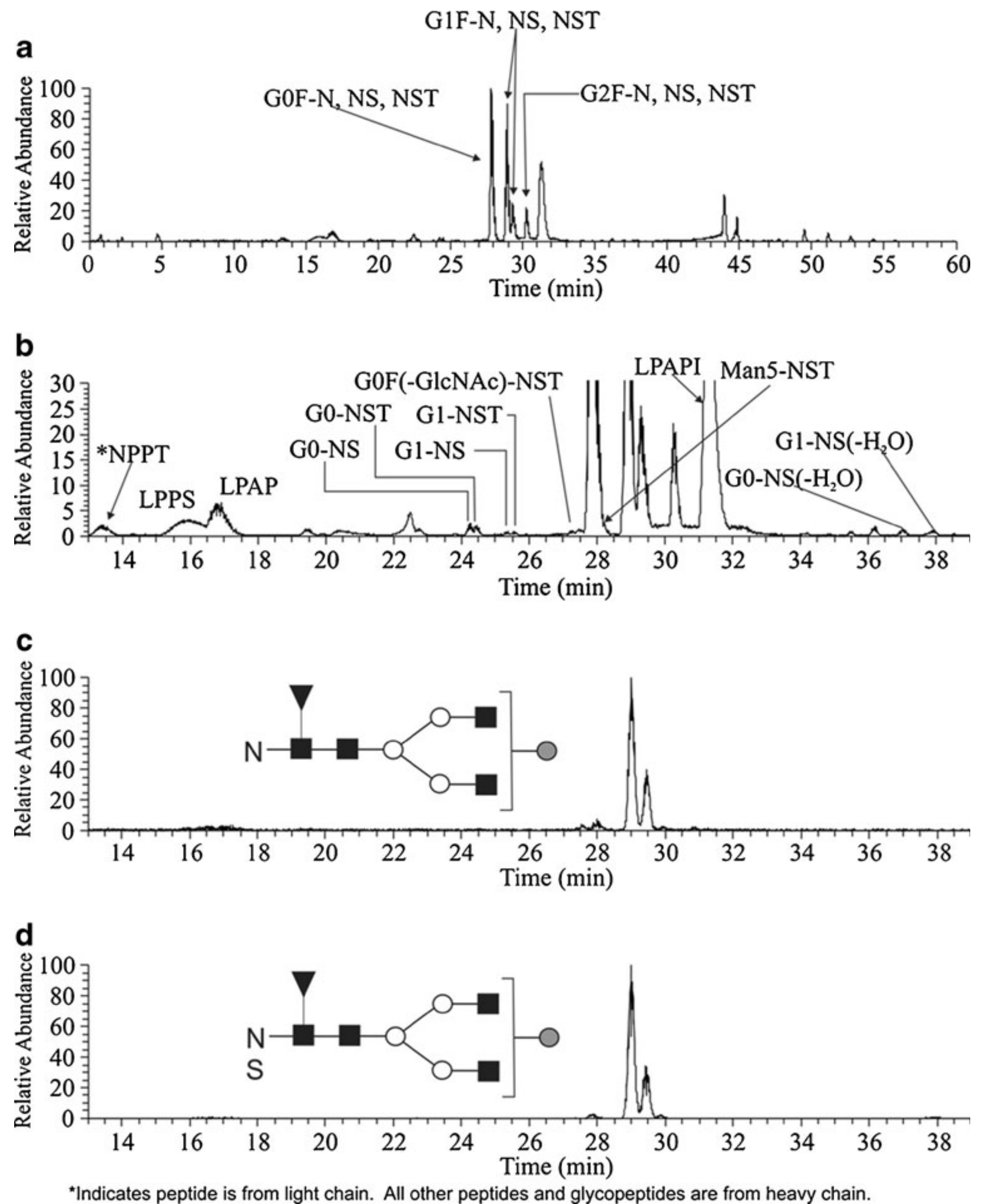
The averaged MS^2 spectra for each of the NS-G1F isomers are shown in Fig. 3 (a, b). In both cases, the characteristic oxonium ion is present at *m*/*z*=366 Da and a Y-ion series is present indicative of a NS-G1F glycopeptide. It should be noted that the loss of a terminal galactose residue (e.g., −162 Da from 1^+ parent ion) is not observed. This fragmentation has been reported previously for 2-AB labeled G1F glycan [27]. Preferential cleavage of the GlcNAc-Man bond likely occurs through attack of the NAc carbonyl oxygen on its own reducing terminus to form a 5-membered oxazoline ring, explaining this bond's lability. An interesting observation in the MS^2 spectra for each of these peaks was a difference in the relative intensity of the $Y_{4\alpha}$ and Y_4 fragments. Under the conditions shown, all G1F glycopeptides (N, NS, and NST) showed a similar intensity profile of the $Y_{4\alpha}$ and Y_4 fragments and seem to offer an additional artifact for isomer identification. Similar CID experiments and fragment relative intensities have been used in a number of studies to differentiate galactose attached to the 1,3 or 1,6 linked branch [28–31].

The XIC chromatograms prepared from the MALDI peak list and MS^2 spectra were used to verify a total of seven glycopeptides in the MALDI spectrum. These are listed in Table 1 and account for every peak in the MALDI spectrum except for two (*m*/*z*=1,668.57 Da and *m*/*z*=1,830.63 Da). In addition, a number of chromatographic peaks remained unassigned at this point. For this reason, the search for low abundance species was performed as described in the Electronic Supplementary Material. An artifact seen quite often was a base peak greater than 500 Da (M_h) that would fragment solely into an ion exactly at *m*/*z*= $[(M_h-1)/2+1]^+$. Over the same retention time, there was typically an ion in the MS^1 at this smaller *m*/*z*. CID spectra of this smaller *m*/*z* revealed that these were peptides partially ionizing as dimers and are listed in Fig. 2(b).

The final low abundance peaks were shown to have MS^2 spectra with glycosidic fragmentation. To ensure that no false positive glycopeptides were reported, a low abundance glycopeptide had to have either (1) a different retention time than a high abundance glycopeptide or (2) for co-eluting glycopeptides the low abundance glycan could not have an *m*/*z* corresponding to a terminal monosaccharide residue loss from the highly abundant species. A total of eight additional glycopeptides were identified and are listed in Table 1 as only observed in the LC-MS/MS method.

There were two additional species that eluted between 37 to 39 min and had an MS^2 profile indicative of glycan

Fig. 2 Chromatograms obtained for LC-MS/MS of rituximab including **(a)** base peak chromatogram (m/z=500 to 2,000 Da), **(b)** expanded base peak chromatogram from 13 to 40 min, and **(c)** and **(d)** extracted ion chromatograms for m/z=870.4 Da and m/z=913.8 Da



fragmentation. The mass of these species was calculated to be exactly 18 Da less than the NS-G0F and NS-G1F species, but with a much different retention time. Manual interpretation of the MS^2 spectra was consistent with G0F and G1F fragmentation. It seems that these peaks are due to a degradation product of NS-G0F and NS-G1F and a rearrangement specific to the NS-peptide was causing their formation. The MALDI-TOF spectra also indicate the presence of these two species, as their masses correspond to $[M+Na]^+$ ions that would form at the previously unidentified m/z of 1,668.57 and 1,830.63 Da.

Pronase digestion results in a mixture whereby glycopeptides can be easily purified from sample matrix, remaining amino acids, and most small peptides as opposed to a specific protease where many more peptides are expected to interfere. MALDI-TOF was only able to identify the three major glycans found in rituximab due to ion suppression effects; however, these glycans account for the vast majority of the glycans present on this protein. The high mass accuracy of the MALDI data was useful to limit the number of potential glycopeptides prior to LC-MS/MS analysis. Graphitized carbon LC-MS/MS of the pronase digest was used to further characterize rituximab glycosylation site and glycan identity in this experiment. Seven glycans were identified on rituximab, each of which were observed as NST-glycopeptides, an amino acid sequence only present on the rituximab heavy chain starting at N297. The data have therefore confirmed the glycosylation site at N297 to be glycosylated with each of the glycans listed in Table 1. Pronase digestion and combined MALDI-TOF and LC-MS/

Table 1 Summary of glycopeptides observed in MALDI-TOF and LC-MS/MS of rituximab pronase digest

Glycan structure	Glycan name	Peptide (X) observed MALDI-TOF [a](m/z =[M+Na]$^+$)	Peptide (X) observed LC-MS/MS (m/z = [M+2H]$^{2+}$) Verified with CID
	G0F	N (1599.53) [b,c]NS (1686.56) [b]NST (1787.60)	N (789.5) [c]NS (832.9) NST (883.9)
	G1F	N (1761.58) [b,c]NS (1848.62) [b]NST (1949.66)	N (870.4) [c]NS (913.8) NST (964.5)
	G2F	[b]NST (2111.69)	N (951.4) [c]NS (994.9) NST (1045.4)
	G1		NS (840.8) NST (891.3)
	G0		NS (759.8) NST (810.4)
	Man 5		NST (769.3)
	G0F-GlcNAc		NST (781.8)

[a] The *m*/*z* listed is for the most abundant adduct observed for the ion. In the MALDI-TOF spectrum, glycopeptides ionized predominantly as singly charged Na^+ ions

[b] Ions corresponding to the $[M+2Na-H]^+$ (22 Da higher than *m*/*z* listed) were also identified in the MALDI-TOF spectrum for these species

[c] Peaks corresponding to a water loss, thought to be due to cyclization of the peptide backbone were also observed for these species in the respective experiment

MS analysis therefore have served as a useful tool in the characterization of biopharmaceutical glycosylation site and glycan type.

Rituximab 2-AB glycan LC-F-MS/MS

2-AB glycan standards Conditions for analysis of 2-AB labeled glycans were initially optimized using pre-labeled glycans and a human IgG glycan mixture purchased from a commercial source (Electronic Supplementary Material). These materials were found to be sufficient for optimizing data analysis and chromatographic conditions, however, residual contaminants corresponding to isomers or similar glycans were often found. Therefore these materials were not deemed useful as absolute quantitative standards.

Rituximab 2-AB glycans The fluorescence profile of rituximab 2-AB glycans is shown in Fig. 4. The Man5, G0, G0F, G1F, and G2F glycans were identified based on matching parent *m*/*z* and MS^2 spectra with the glycan standards as well as a positive match in SimGlycan. Each of these five species also had HK adduct MS^2 spectra that matched with those collected for the standards. An expanded view of the fluorescence signal in the range of 0 to 30 min reveals low abundance detail in the chromatogram (Fig. 4(b)). Mass spectral content of these peaks revealed there were *m*/*z*

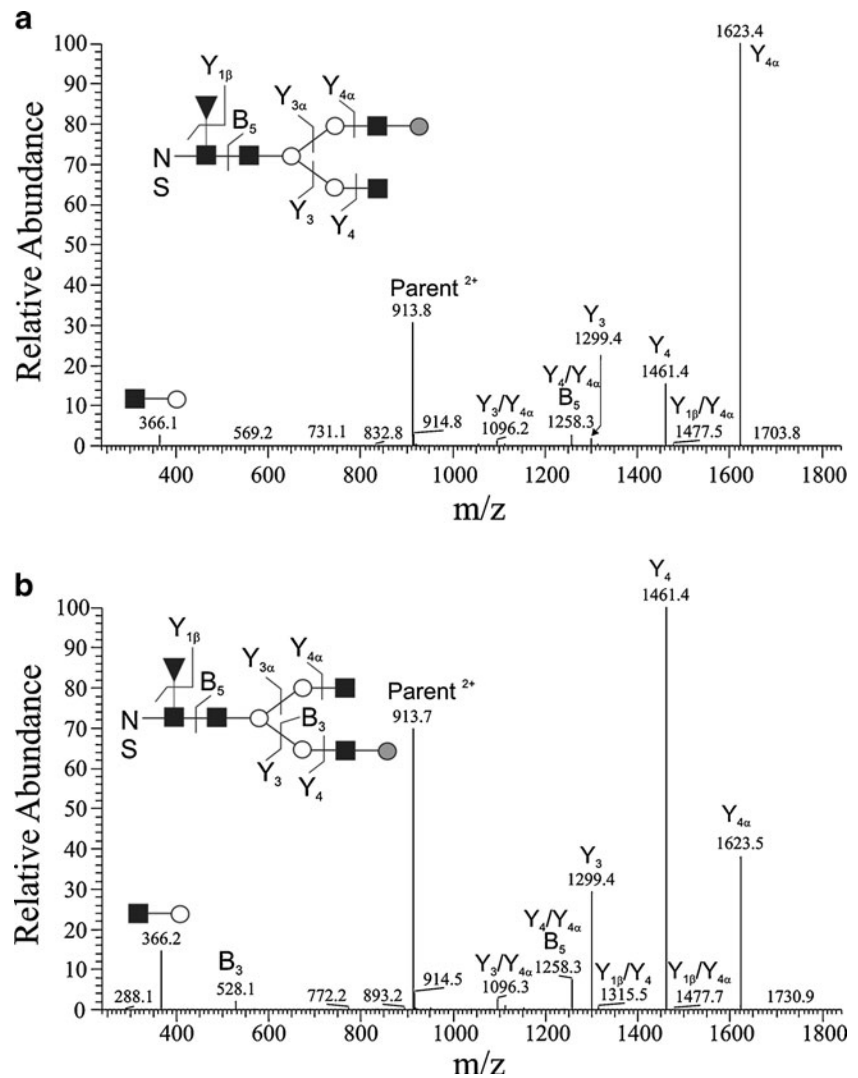


Fig. 3 CID spectra of G1F-NS isomers from rituximab resolved during LC-MS/MS and eluting at **(a)** 29.0 (galactose on 1,6 antennae) and **(b)** 29.5 min (galactose on 1,3 antennae)

corresponding to 2-AB labeled glycans. The Man8, Man7, Man 6 (at 11.1 min), Man5/G1, and G0FBi were matched via MS^1 only (no MS^2 spectra were triggered). Despite this, the parent *m*/*z* as well as retention behavior consistent with that reported previously using this method allows a reasonably certain identification [8]. All other glycans were identified based on parent *m*/*z* and MS^2 matching in SimGlycan. All of these species fit well with the predicted fragmentation of the $[M+2H]^{2+}$ parent ion, giving a Y-ion series indicative of the assigned structure, as well as some B-type fragmentation.

Figure 4(a) shows only a single G1F peak at 36.5 min, indicating the G1F positional isomers galactosylated on different antennae, was not resolved as they were in the pronase method. However, glycan species that seem to be low abundance isomers of the forms listed in Electronic Supplementary Material Table S1 were identified in the 2-AB method (peaks listed with an “i” in Fig. 4(b)). These peaks often co-eluted with species designated in Fig. 4(a), but were of larger *m*/*z* and could not have been due to in-source fragmentation. In addition, they had significantly different retention times from those of the more abundant form. The isomer peaks eluting at approximately 18 min were also seen in the standards injections, but those at 26.3 min were not. In both cases, MS^2 spectra were collected for the isomer peaks that looked identical or nearly identical to the MS^2 spectra noted for the higher abundance forms. In the current case, as well in previous reports using this method, the mass spectra are not sufficiently different to absolutely identify linkage of these isomer species [8]. Additional experiments such as permethylation analysis, exoglycosidase digestion, or NMR would be required for complete branching analysis. Despite this, the combined use of retention and MS data revealed the presence of glycan isomers. The current data also indicates that although retention time can often be matched to indicate elution of a specific glycan, the MS and MS^2 spectra increase reliability of identifying all glycans contributing to a given peak.

Relatively good resolution of most species was achieved; however, co-elution in some cases (e.g., the 26.3 min peak) makes designation of glycan specific peaks impossible. As

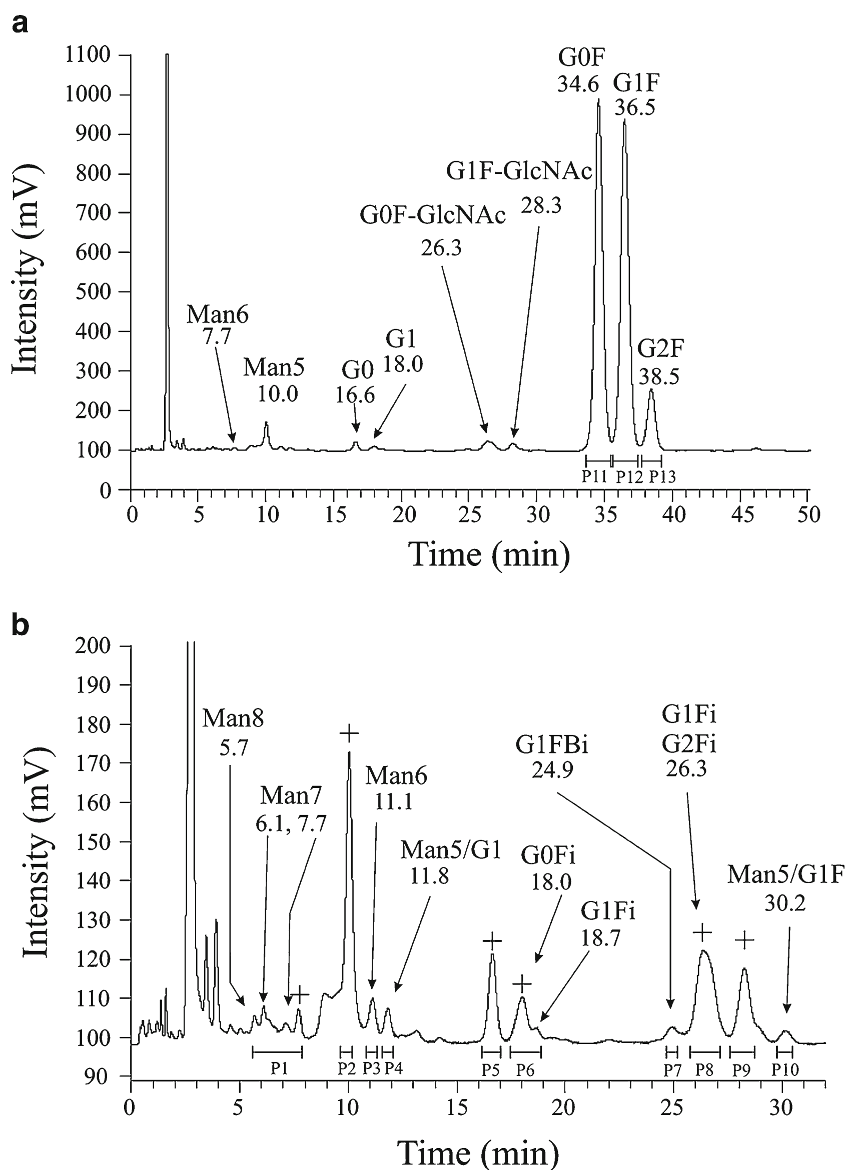


Fig. 4 Fluorescence chromatogram recorded for 2-AB labeled rituximab glycans over **(a)** the entire analytical gradient and **(b)** expanded region from 1 to 30 min. Designations in part **(a)** were verified with MS^2 data. See text for designations made in part **(b)**. Symbols: "+" peak was designated with high abundance glycan in part **(a)**, "i" indicates that these species are thought to be isomers of the forms depicted in Electronic Supplementary Material Table S1, "Px" refers to the peak number designation in Table 2

an alternative, unresolved species were lumped into representative "peaks" as depicted by the Px designations in Fig. 4(a). Each of these peak areas was then integrated and the quantitative results can be seen in Table 2 for triplicate analysis. The quantitative results show that although overlapping peaks were observed, the designations allowed for relatively good injection reproducibility for the same preparation and serves as a method for relative comparison of glycosylation profiles as may be necessary in comparative studies.

Potential for reference material development

Currently, qualitative and quantitative lot-to-lot variability is controlled through comparability testing. Each manufacturer uses a representative lot (reference standard) that is validated for analytical consistency and stability in order to establish the comparability of future lots. This has been a successful practice in ensuring the safety and efficacy of a given biopharmaceutical product, however it does not allow method-specific or glycan-specific variables to be assessed between manufacturers using slightly different techniques. Current comparability testing and the use of in-house reference standards would be greatly supplemented by widely available and commonly used reference materials that could be used as common benchmarks for assessing biopharmaceutical glycosylation. Although the current report used only a small cross section of available glycoanalytical techniques, many desirable features of a glycoanalytical reference material are apparent from the data as described below.

Applicability and availability MALDI-TOF and LC-MS/MS of a pronase digest was shown in this work to be capable of identifying glycosylation site, while LC-F-MS/

Table 2 Summary of rituximab 2-AB glycans observed in LC-MS/MS

Peak	Glycan	Percent abundance (±1 S.D.)	Observed in Ref. 9	Observed in Ref. 6
P1	Man8	.86 (±0.12)	—	Y
	Man7		—	Y
	Man6		—	Y
P2	Man5	2.05 (±0.17)	Y	Y
P3	Man6	0.49 (±0.13)	—	Y
P4	Man5/G1	0.28 (±0.10)	—	—
P5	G0	0.73 (±0.04)	Y	Y
P6	G1	0.67 (±0.06)	Y	—
	G0Fi		—	—
	G1Fi		—	—
P7	G1FBi	0.24 (±0.03)	—	—
P8	G0F-G1cNAc	1.71 (±0.06)	Y	Y
	G1Fi		—	—
	G2Fi		—	—
P9	G1F-G1cNAc	1.05 (±0.02)	—	—
P10	Man5/G1F	0.21 (±0.05)	—	—
P11	G0F	42.44 (±0.23)	Y	Y
P12	G1F	41.65 (±0.43)	Y	Y
P13	G2F	7.60 (±0.06)	Y	Y
—	Man9	—	—	Y
—	Sialylated Forms	—	—	Y

MS of released and fluorescently labeled glycans yielded relative quantification. Additional methods such as permethylation and MS^n could also be used to gain branching information. Due to the multi-method nature of glycoanalysis, the most representative reference material would be an intact glycoprotein capable of undergoing both proteolytic and glycan release methods.

Both the 2-AB and pronase methods produced information on glycan identity (Tables 1 and 2); however a combination of the methods was shown to give the most complete description of all glycan species. The total quantity of 2-AB glycan injected was approximately sixfold higher than for the pronase method, but the MS signal intensity used to identify glycans was higher for pronase-derived glycopeptides due to their improved ionization efficiency (e.g., NST-G0F glycopeptide gave approximately an order of magnitude more intense MS signal than the 2-AB labeled G0F glycan). It is therefore somewhat surprising that LC-F-MS/MS of 2-AB labeled rituximab glycans identified certain low abundance species that were not observed in the pronase method. A number of method or glycan-dependent factors other than injection quantity could have contributed to differences in glycans observed (e.g., enzyme efficiency, glycan degradation, sample loss, etc.). Each of these factors highlights the need for a widely available reference material that would allow a benchmark comparison of different methods and aid in identifying sources of variability.

Tables 1 and 2 compare the qualitative identifications made in the current report with those reported in the literature using other methodology including LC-MS/MS of tryptic glycopeptides [9] and negative ion MALDI analysis of fraction collected 2-AA glycans [6]. Differences in glycans identified are potentially due to slightly different sample pretreatment procedures, different ionization (MALDI vs. ESI methods), or different detection modes (negative vs. positive). For example, various sialylated forms were observed by Kamoda that were not observed in the current sample [6]. Negative mode mass spectrometry is a convenient method to increase detection of sialylated forms; however, analysis using the rapid resolution 2-AB method of human IgG (Electronic Supplementary Material) was sufficient to observe the sialylated G1F and G2F forms. Therefore, they should have been detected in the 2-AB method had they been present in the rituximab sample. The fact that these forms were not detected in the current experiment could have been due to any number of experimental factors as described above or simply lot-to-lot variation of the particular samples analyzed in the literature vs. those assessed in the current report. A well characterized, widely available material with known glycosylation would allow longitudinal comparisons because lot-to-lot variability would no longer be a factor. This would provide a means for longitudinal quality control, aid in qualification of new reference standard lots, and be useful in establishing method comparability for submission of biosimilars.

Potential reference material composition The glycans identified on rituximab represent some of the glycans that are commonly found on biopharmaceuticals, although the glycan identities found on a particular product are dependent on a number of factors including expression system and conditions, drug identity, etc. A glycoprotein reference material will not be able to cover all of the expected glycans present in biopharmaceuticals as a whole. However, IgG-based therapeutics represents the largest fraction of recombinant therapeutics, and therefore recombinant IgG-based materials containing biopharmaceutically relevant glycosylation would be immediately applicable to the largest number of QC applications. An attempt is typically made to humanize glycosylation, therefore the optimum reference material would contain the broadest coverage of human IgG glycans possible. This type of reference material would serve as a realistic sample for quality control of established methods and assessment of novel methodology. Method qualification based on the reference material would infer the method's usefulness for characterization and comparability of a new therapeutic. This would provide a common data set by which regulators could assess submissions from different manufacturers and dataset for different products.

In addition to containing common glycans, the inevitable presence of isomers must also be considered during reference material development. The rituximab MALDI compositional matching and SimGlycan MS^2 spectral interpretation in the current report used theoretical glycans based on previous biosynthetic knowledge to infer linkage as well as prepare theoretical fragmentation patterns for comparison, practices common in glycoanalysis. In the case of pronase-derived G1F positional isomers, fragment ion intensities were useful for identifying antennae positional isomers (Fig. 3), but CID fragmentation alone was not sufficient to differentiate isomers in all cases observed in the 2-AB study. The MS/MS data and biosynthetic filter results alone would have incorrectly designated many isomers as high abundance forms listed in Electronic Supplementary Material Table S1. Retention data were used in the current study to identify these as isomers, and additional methodology (permethylation analysis, NMR, etc.) would be useful to completely designate their branching pattern. As analytical technologies improve, the presence of isomers in glycoproteins is increasingly being reported [24, 26, 32], and the isomeric content of a well-characterized reference material should be available to the extent possible with current technology.

Chemoenzymatic synthesis has been used to add a wide variety of glycans to protein acceptors [33, 34]. The ability to tailor this glycosylation in vitro to contain specific glycans (e.g., if a given glycan is deemed to be a critical quality attribute) shows great promise for detailed analysis of glycan-specific and method-dependent variables. The synthetic nature of homogeneous glycoprotein synthesis could potentially be used to prepare reference materials with specific glycan isomers [35]. Preparation of materials with isomeric purity would be useful to establish retention behavior and/or fragmentation behavior for these species allowing side-by-side comparisons to unknown products. A greater understanding of these variables would allow manufacturers to better choose analytical methods for their own products when a specific product or process related glycan variant is considered to be important.

Quantitative properties In the current study, relative glycan quantification was achieved through enzymatic release of glycans, labeling of the reducing terminus, and measurement of the resultant fluorescent signals (Table 2). Each of these reactions was monitored to have gone to completion by MALDI of the protein before and after deglycosylation and assessing the LC-F-MS/MS mass spectra for unlabeled glycans. In both cases, the reactions were shown to be quantitative to within the detection limits of the method and were therefore useful for relative quantification. However, it is still assumed that sample loss is consistent from one preparation to the next, and that losses of each glycan type are equivalent.

MS spectral counting was not assessed for quantitative capabilities in the current study due to the inherent nature of ion trap instruments. Although quantification in trap instruments is possible when scan trapping times and ion accumulation settings are considered, other instrumentation (e.g., QTOF) is more readily suited for this type of quantification and promising results have been reported for released glycans (native and derivatized), specific proteolytic peptides, and nonspecific proteolytic peptides [26, 36–38]. Such methods have been suggested to offer the advantage of reduced sample pre-treatment steps; however, each may also suffer from incomplete proteolysis, ion suppression effects, and differences in species-specific ionization efficiencies. A number of MS-based relative quantification schemes have also recently been developed using isotopically labeled derivatizing agents or via incorporation during expression [39–41]. To date, absolute quantification of a given glycan has not been achieved and represents a difficult challenge due to the inherent complexity of carbohydrates. However, even a small cohort of reference materials of known glycan concentration would be useful for establishing a common benchmark for relative glycoanalytical schemes to be assessed. The ability to set acceptance criteria for a glycan-related critical quality attribute in an absolute fashion would allow for establishment of much more stringent quality control limits and lead to higher confidence of safety and efficacy. Establishing absolute quantification of heterogeneous glycoprotein reference materials such as recombinant IgG will require development of novel methodology. As an alternative, the semi-synthetic nature of chemoenzymatically synthesized materials may allow a level of quantification through purity validation of

a homogeneously glycosylated protein with one glycosylation site.

Conclusions

LC-MS/MS of a pronase digest was shown to be capable of identifying glycosylation site and glycan identity while LC-F-MS/MS of released and fluorescently labeled glycans yielded relative quantification and glycan identity. Glycoanalysis of a given analyte requires a multi-method approach to gain the most complete characterization, many of which can be complimentary and confirmatory. Assessment of the methods described herein indicates development of a widely available, well-characterized reference material with quantitative properties would greatly supplement current comparability testing. Recombinant IgG containing biopharmaceutically pertinent glycans and homogeneous glycoproteins synthesized using chemoenzymatic methods both show great promise as potential materials to supplement in-house reference standards. Glycoanalytical reference material development represents a future milestone in the advancement of glycoanalysis measurement science that would greatly benefit biopharmaceutical development.

Acknowledgment The authors would like to thank Dr. Carlito Lebrilla for technical guidance regarding pronase digestion and Dr. Illarion Turko for assistance performing MALDI-MS.

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

1. Stanley P, Schachter H, Taniguchi N (2010) N-glycans. In: Varki A, Cummings RD, Esko JD, Freeze HH, Stanley P, Bertozzi CR, Hart GW, Etzler ME (ed) Essentials of glycobiology, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Springs Harbor
2. Higgens E (2010) Glycoconjug J 27:211–225
3. Kawasaki N, Itoh S, Hashii N, Takakura D, Qin Y, Huang XY, Yamaguchi T (2009) Biol Pharm Bull 32:796–800
4. van Berkel PHC, Gerritsen J, Perdok G, Valbjorn J, Vink T, van de Winkel JGJ, Parren PWHI (2009) Biotech Prog 25:244–251
5. Kamoda S, Nomura C, Kinoshita M, Nishiura S, Ishikawa R, Kakehi K, Kawasaki N, Hayakawa T (2004) J Chromatogr A 1050:211–216
6. Kamoda S, Ishikawa R, Kakehi K (2006) J Chromatogr A 1133:332–339
7. Ma S, Nashabeh W (1999) Anal Chem 71:5185–5192
8. Prater BD, Connelly HM, Qin Q, Cockrill SL (2009) Anal Biochem 385:69–79
9. Stadlmann J, Pabst M, Kolarich D, Kunert R, Altmann F (2008) Proteomics 8:2858–2871
10. Wuhrer M, Deelder AM, Hokke CH (2005) J Chromatogr B 825:124–133
11. Huhn C, Selman MHJ, Ruhaak LR, Deelder AM, Wuhrer M (2009) Proteomics 9:882–913
12. An HJ, Peavy TR, Hedrick JL, Lebrilla CB (2003) Anal Chem 75:5628–5637
13. An HJ, Tillinghast JS, Woodruff DL, Rocke DM, Lebrilla CB (2006) J Proteome Res 5:2800–2808
14. An HJ, Froehlich JW, Lebrilla CB (2009) Curr Opin Chem Biol 13:421–426
15. Clowers BH, Dodds ED, Seipert RR, Lebrilla CB (2007) J Proteome Res 6:4032–4040
16. Dodds ED, Seipert RR, Clowers BH, German JB, Lebrilla CB (2009) J Proteome Res 8:502–512
17. Yu YQ, Fournier J, Gilar M, Gebler JC (2007) Anal Chem 79:1731–1738
18. Juhasz P, Martin SA (1997) Int J Mass Spectrom 169:217–230
19. Liu X, McNally DJ, Nothaft H, Szymanski CM, Brisson JR, Li JJ (2006) Anal Chem 78:6081–6087
20. Liu X, Chan K, Chu IK, Li JJ (2008) Carb Res 343:2870–2877
21. Temporini C, Perani E, Calleri E, Dolcini L, Lubda D, Caccialanza G, Massolini G (2007) Anal Chem 79:355–363
22. Wuhrer M, Koeleman CAM, Hokke CH, Deelder AM (2005) Anal Chem 77:886–894
23. Zauner G, Koeleman CAM, Deelder AM, Wuhrer M (2010) J Sep Sci 33:903–910
24. Nwosu CC, Seipert RR, Strum JS, Hua SS, Zivkovic AM, German BJ, Lebrilla CB (2011) J Proteome Res 10:2612–2624
25. Nwosu CC, Seipert RR, Strum JS, Hua SS, An HJ, Zivkovic AM, German BJ, Lebrilla CB (2011) J Proteome Res 10:2612–2624
26. Hua S, Nwosu CC, Strum JS, Seipert RR, An HJ, Zivkovic AM, German BJ, Lebrilla CB (2011) Anal Bioanal Chem. doi:10.1007/s00216-011-5109-x
27. Chen XY, Flynn GC (2007) Anal Biochem 370:147–161
28. Flynn GC, Chen XY, Liu YD, Shah B, Zhang ZQ (2010) Mol Immunol 47:2074–2082
29. Ito H, Takegawa Y, Deguchi K, Nagai S, Nakagawa H, Shinohara Y, Nishimura SI (2006) Rapid Commun Mass Spectrom 20:3557–3565
30. Takegawa Y, Deguchi K, Ito S, Yoshioka S, Sano A, Yoshinari K, Kobayashi K, Nakagawa H, Monde K, Nishimura SI (2004) Anal Chem 76:7294–7303
31. Takegawa Y, Ito S, Yoshioka S, Deguchi K, Nakagawa H, Monde K, Nishimura SI (2004) Rapid Commun Mass Spectrom 18:385–391
32. Reinhold V, Ashline DJ, Zhang H (2010) Unraveling the structural details of the glycoproteome by ion trap mass spectrometry. In: Practical aspects trapped ion mass spectrometry. Taylor and Francis, Boca Raton
33. Huang W, Li C, Li B, Umekawa M, Yamamoto K, Zhang X, Wang LX (2009) J Am Chem Soc 131:2214–2223
34. Huang W, Yang QA, Umekawa M, Yamamoto K, Wang LX (2010) Chembiochem 11:1350–1355
35. Schiel JE, Lowenthal MS, Phinney KW (2011) J Mass Spectrom 46:649–657
36. de Leoz ML, Young LJ, An HJ, Kronewitter SR, Kim J, Miyamoto S, Borowsky AD, Chew HK, Lebrilla CB (2011) High-mannose glycans are elevated during breast cancer progression. Mol Cell Proteom 10(1). doi:10.1074/mcp.M110.002717
37. Ivancic MM, Gadgil HS, Halsall HB, Treuheit MJ (2010) Anal Biochem 400:25–32
38. Rebecchi KR, Wenke JL, Go EP, Desaire H (2009) J Am Soc Mass Spectrom 20:1048–1059
39. Atwood JA, Cheng L, Alvarez-Manilla G, Warren NL, York WS, Orlando R (2008) J Proteome Res 7:367–374
40. Orlando R, Lim JM, Atwood JA, Angel PM, Fang M, Aoki K, Alvarez-Manilla G, Moremen KW, York WS, Tiemeyer M, Pierce M, Dalton S, Wells L (2009) J Proteome Res 8:3816–3823
41. Zhang P, Zhang Y, Xue XD, Wang CJ, Wang ZF, Huang LJ (2011) Anal Biochem 418:1–9