Available online at www.sciencedirect.com
ScienceDirect
Carbohydrate Research 343 (2008) 758–768
Carbohydrate RESEARCH

Comparative RP-HPLC for rapid identification of glycopeptides and application in off-line LC-MALDI-MS analysis

Yoshimi Kanie, Akiko Enomoto, Satoshi Goto and Osamu Kanie*

Mitsubishi Kagaku Institute of Life Sciences (MITILS), 11 Minamiooya, Machida-shi, Tokyo 194-8511, Japan

Received 24 September 2007; received in revised form 8 December 2007; accepted 12 December 2007
Available online 23 December 2007

Abstract—Despite the increasing attention being paid to the functions of glycoproteins, their structural analysis is still difficult and hinders functional investigations. Structural analysis of post-translationally modified proteins is thought to be achieved using methods frequently utilized in proteomics research; however, the same methods cannot be used for glycosylated proteins. One of the difficulties associated with the physiochemical properties of glycopeptides and peptides is that the detection of the former is considerably more difficult, because of the existence of glycoforms that increase molecular weight and reduces quantities of individual species. Thus, difficulties are often faced in finding glycopeptide(s) by using MS when analyzing peaks (or fractions) obtained after proteolytic digestion and HPLC. One simple yet difficult solution to this problem would be to develop a purification method that provides better resolution. Our intention has been to address this issue by using a combination of conventional methods. We found that a method consisting of a combination of rough fractionation using a reverse-phase cartridge column under acidic conditions and comparative RP-HPLC, where the two chromatograms obtained using phosphate and borate buffers under basic conditions were compared, is effective for MS-based structural analysis. The applicability of the method in glycoprotein analysis was examined using various samples including ribonuclease B (RNase B), IgG1, ovalbumin (OVA), and asialo fetuin (ASF). The results suggest that the method is useful in the analysis of glycoproteins.

Keywords: Borate esters; Glycoprotein; Glycopeptides; RP-HPLC; Alkaline pH; Tandem mass spectrometry

1. Introduction

In mammalian species, one of the most abundant post-translational modifications (PTM), amongst various PTMs such as phosphorylation, glycosylation, acylation, and alkylation, of secretory and membrane-anchored proteins is glycosylation. The glycan structures of glycoproteins are involved in physiochemical properties such as solubility, resistance against proteolysis, and structural stability. Moreover, glycosylation is considered to play important roles in biological processes, including protein conformation, molecular recognition, and cellular interaction.[1–3] Changes in glycan structures and populations induce protein function disorders. Cancer metastasis,[4] chronic rheumatoid arthritis,[5] and carbohydrate deficient glycoprotein syndrome[6] are now known to be related to glycan structural changes. As awareness of the importance of the glycan moiety of glycoproteins increases, determination of the entire structure of a glycoprotein is necessary to obtain a better understanding of the relationship between structure and function.

Mass spectrometry (MS)-based structural analysis is often advantageous because of the low sample consumption and fewer sample handling steps. The analysis of glycoproteins and glycopeptides have been often carried out after cleavage of glycans despite the increasing number of handling steps. The reason for this is the difficulty associated with the physiochemical properties of glycopeptides and peptides and that the former ionizes rather poorly. A direct analysis of glycopeptide is preferable if possible.The method generally proceeds as follows: (1)

* Corresponding author. Tel.: +81 42 724 6238; fax: +81 42 724 6317; e-mail: kokanee@mitils.jp

0008-6215/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.carres.2007.12.008

tryptic digestion of a protein, (2) separation of peptides and glycopeptides, (3) sequencing of peptide fragments and identification of a protein by MS, and (4) sequencing of glycan structures and determination of the sites of glycosylation. Tryptic digestion, which produces polypeptides with a basic residue at the C-terminus, is advantageous for MS measurements in positive mode. In the mass spectrometric analysis of peptides, the observation of signals are affected by the nature of individual peptides. On the other hand, the analysis of glycopeptides becomes significantly more difficult because of the existence of glycoforms that increase molecular weight and reduce quantities of individual species. When both the peptides and glycopeptides are present in a given fraction after proteolytic digestion, it is usual that the signals of the glycopeptides are difficult to detect. Therefore, the separation of tryptic fragments into as many peaks as possible should be the main focus.

Several methods have been proposed for the isolation of glycopeptides in peptide mixtures to address the above issues[7,8] based on 'selectively adsorbing' carbohydrates using porous graphitized carbon,[9,10] cellulose columns,[11] and lectin columns.[12–17] We found that these methods are not sufficient for glycoproteomic investigation after tryptic digestion of glycoproteins. Porous graphitized carbon has an affinity for oligosaccharides and can be used to isolate glycopeptides with relatively short amino acid sequences; however, its application to tryptic glycopeptides is limited. Cellulose columns have hydrogen bonding properties that are insufficient for the complete separation of peptides and glycopeptides. Because affinity chromatography using lectins depends on the substrate specificity of each lectin, it is not suitable for glycopeptides with an unknown variety of glycan forms.

RP-HPLC has been conventionally used to separate peptides obtained from enzymatic digestion of proteins, including glycoproteins. Traditionally, protein digests are separated under acidic conditions, which is suitable in combination with on-line MS analysis.[18] More recently, the separation of tryptic digests under basic conditions was achieved, which showed a different elution profile compared with the one obtained under acidic conditions.[19] However, the basis of the separation is the hydrophobic interaction between alkyl chains in the stationary phase and the side chains of peptides. A dominating factor that influences retention time is the sequence of a peptide,[20] which is not effective for the separation of peptides and glycopeptides. Regardless of the difficulty, an LC-based separation method has to be employed before MS analysis. LC-ESI-MS, Nano-LC-ESI-MS, or LC-MALDI-MS has thus been frequently used to analyze peaks one-by-one.[18,21]

The polyol structures of carbohydrates in glycopeptides can reversibly form anionic borate esters in the presence of tetrahydroxyborate ion, $B(OH_4^-)$, at alkaline pH (i.e., pH 8–10), which might be effectively used to distinguish glycopeptides from peptides. The use of this property in GE,[22] CE,[23] and liquid chromatography[24–26] has been reported for the separation in individual oligosaccharides released from glycoproteins and peptides with glycan variations. The expected principle of elution by using basic borate buffer is shown in Figure 1. The negative charge of the complexes has a relatively

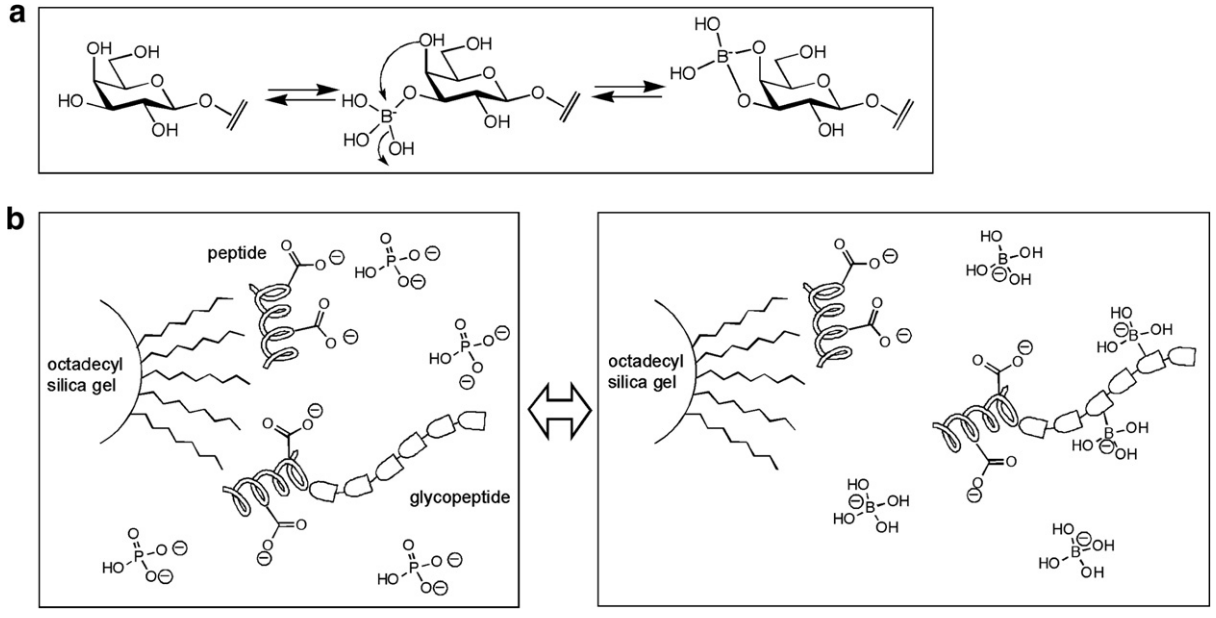

Figure 1. Principle of the elution system under basic conditions with borate buffer. (a) The polyol functional groups in carbohydrates reversibly form anionic borate esters with the tetrahydroxyborate ion, $B(OH_4^-)$, at alkaline pH. (b) Glycopeptides become more hydrophilic due to the negative charge of the borate ester, which affects retention time of glycopeptides.

Table 1. List of amino acid sequences and theoretical masses of tryptic peptides with potential glycosylation sites for model glycoproteins

Sequence	Average mass (theoretical)	Type of N-linked glycan	Glycoprotein
NLTK	475.5	High-mannose	RNase B
EEQYNSTYR	1190.2	Complex	IgG1
YNLTSVLMAMGITDVFSSSANLSGISSAESLK	3295.7	High-mannose, Hybrid	OVA
LC[a]PDC[a]PLLAPLNDSR	1742.0	Complex	ASF
VVHAVEVALATFNAESNGSYLQLVEISR	3018.3	Complex	ASF
RPTGEVYDIEIDTLETTC[a]HVLDPTPLANC[a]SVR	3674.0	Complex	ASF

Underscored amino acids correspond to N-glycosylation sites.
[a] Carbamidomethyl cysteine.

weaker adsorption to the stationary phase due to the hydrophilic nature of the compound. Although, the technique has been used to separate individual glycopeptides from mixtures of glycopeptides,[27] it is difficult to separate glycopeptides without any overlap with other peptides from tryptic digests.

Potential sites of N-glycosylation are readily identified with the consensus sequence, Asn-X-Ser/Thr where X can be any amino acid except proline. Table 1 shows the amino acid sequence of tryptic peptides with glycosylation site(s) and structures of the oligosaccharides of model glycoproteins. Some of the glycoproteins such as RNase B, IgG1, and OVA possess a single site of N-glycosylation while ASF possesses three sites of N-glycosylation. The peptide fragments obtained after tryptic digestion have different peptide length, chemical and physical properties. It should be noted that these proteins contain a variety of N-glycan structures such as high-mannose, hybrid, and complex-type oligosaccharides.[27–32] For the separation technique to analyze such glycopeptides to be effective, it is necessary to provide individual glycan-containing fractions from a proteolytic mixture consisting of a variety of compound.

Here, we report a separation method that relies on 'comparative RP-HPLC', where the retention times of individual peaks observed in the chromatogram using borate buffer and phosphate buffer at alkaline pH are compared to distinguish glycopeptides from a tryptic digest. Prior to the comparative RP-HPLC, the samples are fractionated using acidic conditions on a RP cartridge column (Fig. 2). We succeeded in rapidly identifying glycopeptides in several model glycoproteins. This method further enabled MS/MS analyses of the glycopeptides obtained.

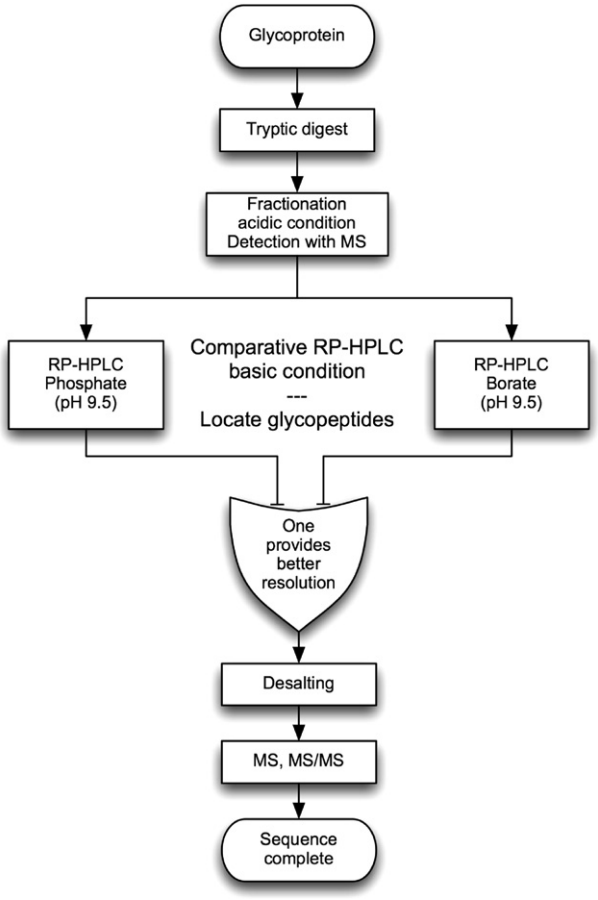

Figure 2. Flow chart of glycoprotein analysis. Comparison of the chromatograms obtained by RP-HPLC using phosphate buffer and borate buffer provides information regarding the retention times of glycopeptides.

2. Materials and methods

2.1. Chemicals and materials

Sequencing grade modified trypsin was purchased from Promega (Madison, WI, USA). PNGase F [EC 3.2.2.18] (proteomics grade), ribonuclease B (RNase B) from bovine pancreas, ovalbumin (OVA) from chicken egg white, and asialo fetuin (ASF) from fetal calf serum were purchased from Sigma–Aldrich (St. Louis, MO, USA). Human IgG1 from myeloma was obtained from Cosmo Bio (Tokyo, Japan). RapiGest™ SF, Sep-Pak C18 cartridge column, and 2,5-dihydroxybenzoic acid (DHB) were purchased from Waters (Milford, MA, USA). Dithiothreitol (DTT) and iodoacetamide were purchased from Nacalai Chemical (Kyoto, Japan). Trifluoroacetic acid (TFA) and formic acid were purchased from Wako Pure Chemical (Osaka, Japan). ZipTips (C18) were obtained from Millipore (Billerica,

MA, USA). All solvents were HPLC grade or LC/MS grade.

2.2. Sep-Pak C18 cartridge column

The fractionation of tryptic digests of glycopeptides using a Sep-Pak C18 cartridge column was achieved by stepwise elution with 10–40% acetonitrile containing 0.1% formic acid and 0.02% TFA for every 10%. The fractions were collected in 4 mL after conditioning the cartridge with 2% acetonitrile containing 0.1% formic acid and 0.02% TFA (4 mL) followed by washing with 80% acetonitrile containing 0.1% formic acid and 0.02% TFA (10 mL).

2.3. HPLC

All chromatographic separations were performed on a Waters 600E multisolvent delivery system equipped with a U6K injector and 486 multiwavelength UV detector (injection volume: 50–200 μL). The data were analyzed using Smart Chrom software (KYA TECH, Japan).

For the separation of tryptic digests of IgG1 and OVA under acidic conditions, a 4.6 mm i.d. × 150 mm, 5 μm Waters XBridge C8 column (Waters, St. Louis, MO, USA) was operated at room temperature, with a mobile phase of water containing 0.1% formic acid and 0.02% TFA (A) and 70% acetonitrile containing 0.1% formic acid and 0.02% TFA (B). The proportion of B was programmed to linearly increase from 0% to 100% over 45 min beginning 5 min after injection.

For the separation of tryptic digests of IgG1 under alkaline conditions, the column was operated at room temperature, using a mobile phase of 2% acetonitrile in 50 mM phosphate buffer at pH 9.5 or 50 mM borate buffer at pH 9.5 (C) and 30% acetonitrile in 50 mM phosphate buffer at pH 9.5 or 50 mM borate buffer at pH 9.5 (D). The proportion of D was programmed to linearly increase from 0% to 100% over 28 min beginning 5 min after injection.

For the evaluation of retention time in alkaline conditions, the column was operated at room temperature, using a mobile phase of 2% acetonitrile in 50 mM phosphate buffer with a pH from 7.0 to 11.0 or 50 mM borate buffer with a pH from 7.0 to 11.0 (E) and 30% acetonitrile in 50 mM phosphate buffer with a pH from 7.0 to 11.0 or 50 mM borate buffer with a pH from 7.0 to 11.0 (F). The proportion of F was programmed to linearly increase from 0% to 100% over 28 min beginning 2 min after injection.

The proportion of D was programmed to linearly increase from 0% to 100% over 28 min beginning 2 min after injection for the separation of tryptic digests of RNase B. The column was operated using a mobile phase of 10% acetonitrile in 40 mM phosphate buffer at pH 9.5 or 40 mM borate buffer at pH 9.5 (G) and 30% acetonitrile in 40 mM phosphate buffer at pH 9.5 or 40 mM borate buffer at pH 9.5 (H) for the separation of other tryptic digests. The proportion of H was programmed to linearly increase from 25% to 100% over 30 min after injection for 40% elution of Sep-Pak C18 cartridge column or from 0% to 80% in 30 min after injection for 20% elution of Sep-Pak C18 cartridge column.

All analyses were performed at a flow rate of 1.0 mL/min and detected at 215 nm.

2.4. Mass spectrometry

MALDI-TOF MS spectra were acquired using a Voyager mass spectrometer (Applied Biosystems, Foster City, CA). An acceleration voltage of 20 kV and a nitrogen laser at 337 nm were used. Mass spectra were acquired in positive linear mode. MS/MS measurements were performed using an Ultraflex TOF/TOF mass spectrometer equipped with a reflector (Bruker Daltonics GmbH, Bremen, Germany). An acceleration voltage of 25 kV and nitrogen laser at 337 nm were used. Metastable ions generated by laser-induced decomposition of the selected precursor ions were analyzed without any additional collision gas. Precursor ions were accelerated to 9.5 kV and selected in a timed ion gate. The fragments were further accelerated by 26 kV in the LIFT cell, and their m/z values were analyzed after the ion reflector passage. Mass spectra were acquired in positive linear mode. DHB (10 mg/mL) in 10% methanol was used as the matrix.

2.5. General method of tryptic digestion of glycoprotein

The glycoprotein (1 mg) was dissolved in 500 μL of 50 mM ammonium bicarbonate (pH 8.0) containing 0.1% RapiGest, and incubated with 2.5 μL of 1 M DTT for 30 min at 60 °C. The sample solution was allowed to cool to room temperature, 15.5 μL of 500 mM iodoacetamide was added, and the solution stood for 30 min in the dark at room temperature. Then, modified trypsin was added to a final volume of 1:50 (w/w) of glycoprotein and $CaCl_2$ solution was added in a final concentration of 1 mM. The tryptic digestion proceeded at 37 °C overnight. The reaction was terminated by the addition of 500 μL of 1% TFA.

3. Results and discussion

3.1. HPLC of glycopeptides from IgG1 in acidic and alkaline conditions using phosphate and borate buffer

To examine the effect of the RP-HPLC conditions, we first carried out separation of the tryptic digest of IgG1 (Fig. 3). Although many peptide fragments were

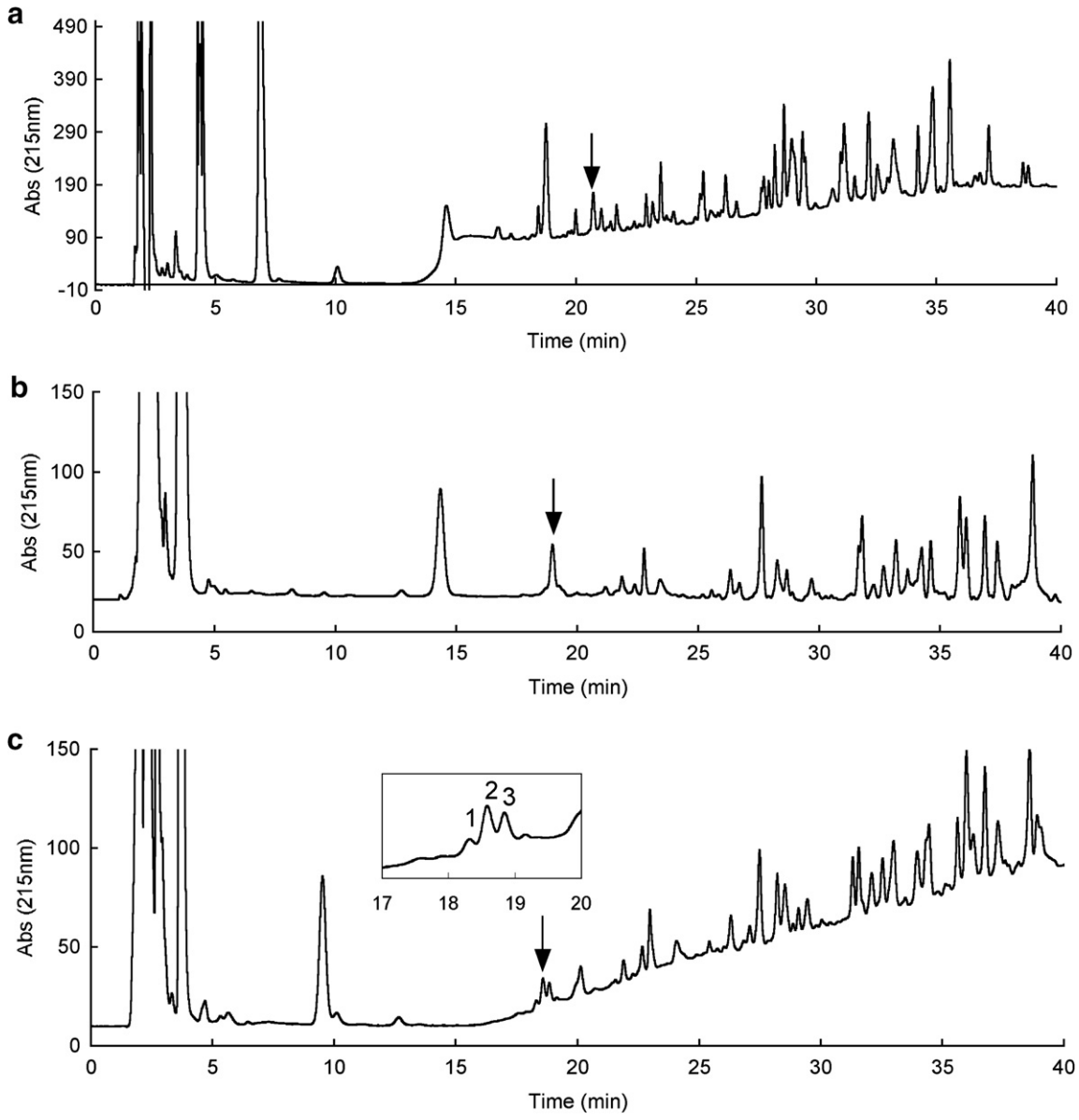

Figure 3. Reverse-phase chromatograms of tryptic peptides from human IgG1. (a) Chromatogram obtained under acidic conditions; (b) chromatogram obtained with 50 mM phosphate buffer at pH 9.5; (c) chromatogram obtained with 50 mM borate buffer at pH 9.5. Insert indicates the expansion of the region of glycopeptides. Details of elution conditions are described in Section 2.

observed after tryptic digestion of the 146 kDa glycoprotein, it was relatively easy to detect glycopeptides using MS because of the hydrophilic properties of the backbone peptide and arginine at the C-terminus with a theoretical mass of 1190.2 Da. Under RP-HPLC in acidic conditions, all glycopeptides in Table 2 were eluted all together as shown by the arrow in the chromatogram (Fig. 3a). RP-HPLC under alkaline conditions was examined next (Fig. 3b and c). Under these conditions, a mobile phase containing 3% acetonitrile was required to elute the glycopeptides. Figure 3b and c shows the separations of tryptic digests with phosphate and borate buffer at pH 9.5 using a linear gradient with an acetonitrile concentration from 2% to 30%. The glycopeptides appeared at 19 min in phosphate and at 18–19 min as three separate peaks in borate buffer (indicated by arrows), respectively, without any overlapping peptides. The mass spectra were measured for fractions containing each peak in Figure 3c after desalination using ZipTip C18, and the structures of glycopeptides in three fractions were determined. The observed major signals for the three peaks in the chromatogram (Fig. 3c) were assigned to be m/z 2958.1 (peak 1), 2796.0 (peak 2), and 2633.9 (peak 3) (Table 2). The observed glycan structures at peaks 1–3 contained at least two hexoses and three *N*-acetylhexosamines (Table 2). According to the synthetic pathway, two *N*-acetylglucosamines (GlcNAc) and three mannoses (Man) are present at the core

Table 2. List of identified glycopeptides from human IgG1

Fraction (#)	Observed *m/z*	Theoretical glycan mass	Identified glycan structure
1	2958.1[a]	1769.6	$Hex_5HexNAc_4Fuc$
2	2796.0[a]	1607.4	$Hex_4HexNAc_4Fuc$
	2649.1	1461.3	$Hex_4HexNAc_4$
	2592.5	1404.2	$Hex_4HexNAc_3Fuc$
	2430.2	1242.1	$Hex_3HexNAc_3Fuc$
3	3161.9	1972.7	$Hex_5HexNAc_5Fuc$
	2633.9[a]	1445.3	$Hex_3HexNAc_4Fuc$
	2487.1	1299.1	$Hex_3HexNAc_4$
	2267.8	1079.9	$Hex_2HexNAc_3Fuc$

[a] Dominant signal in the fraction.

of glycan structures. Therefore, extra HexNAc and Hex are considered to be GlcNAc attached to the core Man_3 and galactose (Gal) further added to the GlcNAc. The difference in the retention time was considered due to the ability to form borate esters between polyols and $B(OH_4^-)$ in alkaline pH. Furthermore, the observed main signals for each peak differs from the number of Gal units only, which indicates that the resolution of the chromatography is quite high. Most of the reported neutral glycans[31] were identified in this direct analysis of tryptic digest.

3.2. Fractioning tryptic digests by Sep-Pak C18 in acidic conditions

Despite the success in the isolation of glycopeptides from simple peptides in the tryptic digest of IgG1, it is usually time-consuming to compare the retention times of all detected peaks from one chromatogram to another. Therefore, fractionation on a C18 Sep-Pak, which is expected to be carried out prior to HPLC separation, was examined next. For this purpose, the use of acidic buffer was considered to be preferable because the remaining silanol on the stationary phase is protonated and the effect of silanol on the sample is suppressed, resulting in a superb recovery of analyte.

Glycopeptides from RNase B were obtained as a mixture of peptides in the 10% acetonitrile fraction by Sep-Pak C18 in acidic conditions followed by detection using MALDI-TOF-MS. A pair of RP-HPLC chromatograms (50 mM phosphate buffer at pH 9.5 using a linear gradient of 2–30% acetonitrile) for tryptic digest of RNase B, are shown in Figure 4. As is clear from these chromatograms, fractionation on a C18 Sep-Pak dramatically reduced the number of observed peaks; thus,

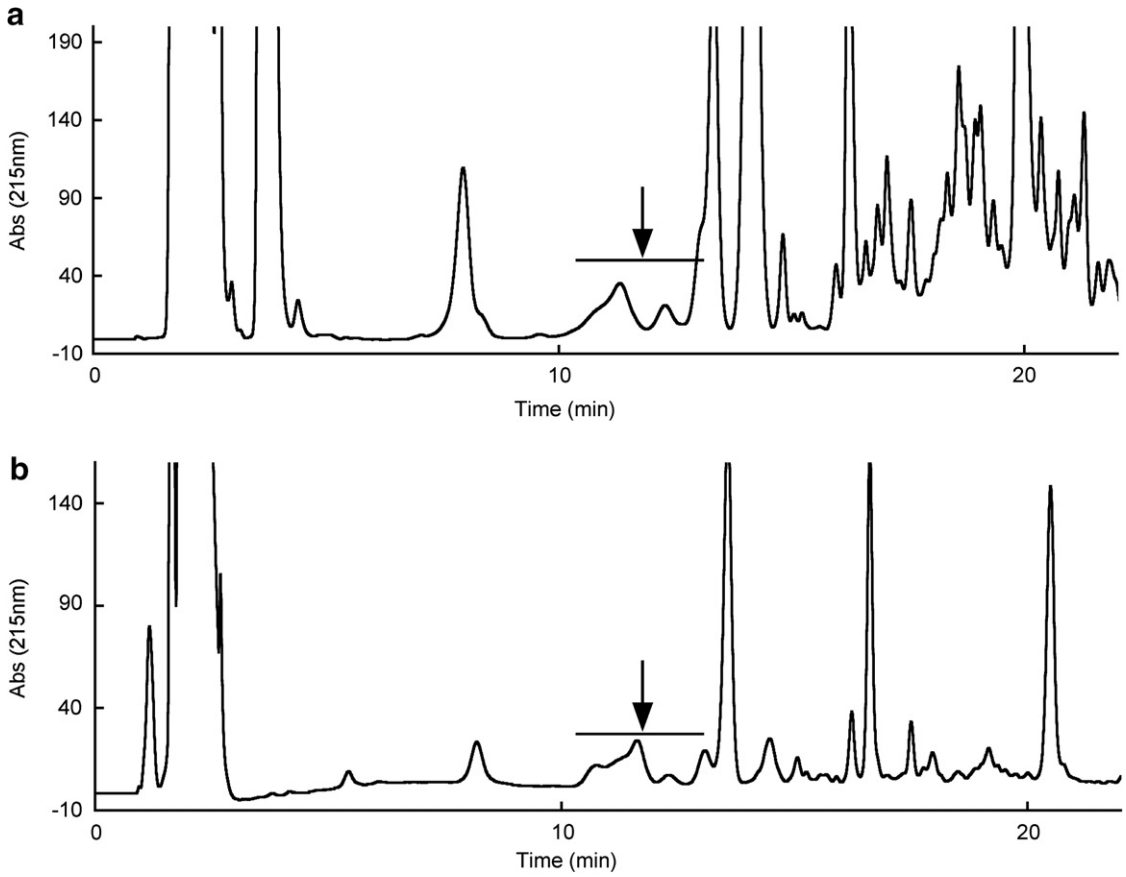

Figure 4. Reverse-phase chromatograms of tryptic peptides from RNase B. (a) Chromatogram obtained for the tryptic digest at pH 9.5 using 50 mM phosphate buffer without any treatment; (b) chromatogram obtained at pH 9.5 using 50 mM phosphate buffer after fractionation using Sep-Pak C18.

the step is considered to be effective before carrying out further HPLC.

3.3. Effect of mobile phase pH on retention time of glycopeptides

We examined the dependency of the retention time of the glycopeptide with *m*/*z* 1692.8 obtained from tryptic digest of RNase B, a 15 kDa glycoprotein, using borate and phosphate in the range of pH 7.0–11.0. Figure 5 shows the schematic structure of the glycopeptide and a graph of the changes of an average retention time at three injections at each pH. The retention times of both buffers were the same at pH 7.0, but the retention time of borate was shortened at higher pH compared to phosphate. This result agrees with borate ester formation. In addition, the longer retention times observed at higher pH for both buffers are considered to be due to the total charge of peptides. (Lys ε-amino group, $pK_a = 10.5$) To distinguish peaks associated with glycopeptides and simple peptides, we wanted to use a buffer that provided the largest difference in retention time. The data suggested that the use of buffer with a pH of 9 should be effective.

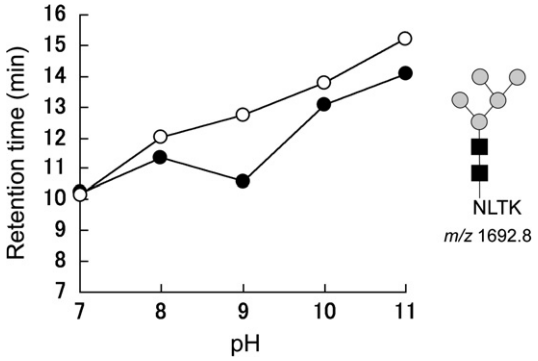

Figure 5. Schematic structure of a glycopeptide with *m*/*z* 1692.8 and retention time with buffers of different pH where the effects of phosphate and borate buffers were compared. Structure: gray circle, Man; black square, GlcNAc. Graph: open circle, phosphate; closed circle, borate.

3.4. Comparative RP-HPLC (phosphate buffer vs borate buffer) of tryptic digests of OVA and ASF

We further examined two other model glycoproteins, OVA, and ASF, to demonstrate the eligibility and the general applicability of the above method. First, comparative RP-HPLC of OVA was examined. OVA is a 43 kDa glycoprotein containing high-mannose and hybrid-type oligosaccharides.[32] The backbone peptides of glycopeptide fragments, with a theoretical average mass of 3295.7 Da and lysine at the C-terminus obtained from tryptic digestion, are relatively hydrophobic in nature. Figure 6a indicates a chromatogram of RP-HPLC under acidic conditions. The MS analysis of the fraction shown in Figure 6a provided the spectrum in Figure 6b where signals of glycopeptides are confirmed together with those of peptides. The 40% acetonitrile fraction of Sep-Pak C18 fraction was analyzed under basic conditions with both buffers (Fig. 7a and b) using an acetonitrile gradient from 15% to 30%. Here, we used buffers with pH 9.5 instead of pH 9.0 as it was suggested above because of the ease of preparations. The elution patterns of the main peaks were comparable to each other. The

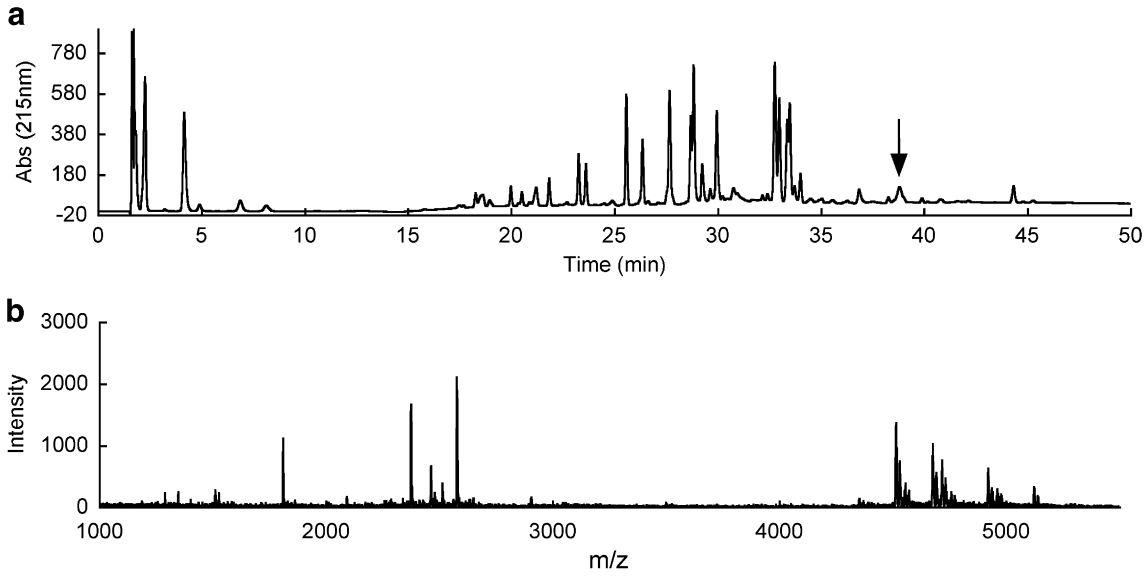

Figure 6. Chromatogram (acidic condition) of tryptic peptides and MS spectrum of a fraction containing glycopeptides from OVA. (a) Reverse-phase chromatogram of tryptic peptides from OVA under acidic conditions; (b) mass spectrum of a marked peak in (a). Glycopeptides signals were observed in the range from *m*/*z* 4500 to 5300.

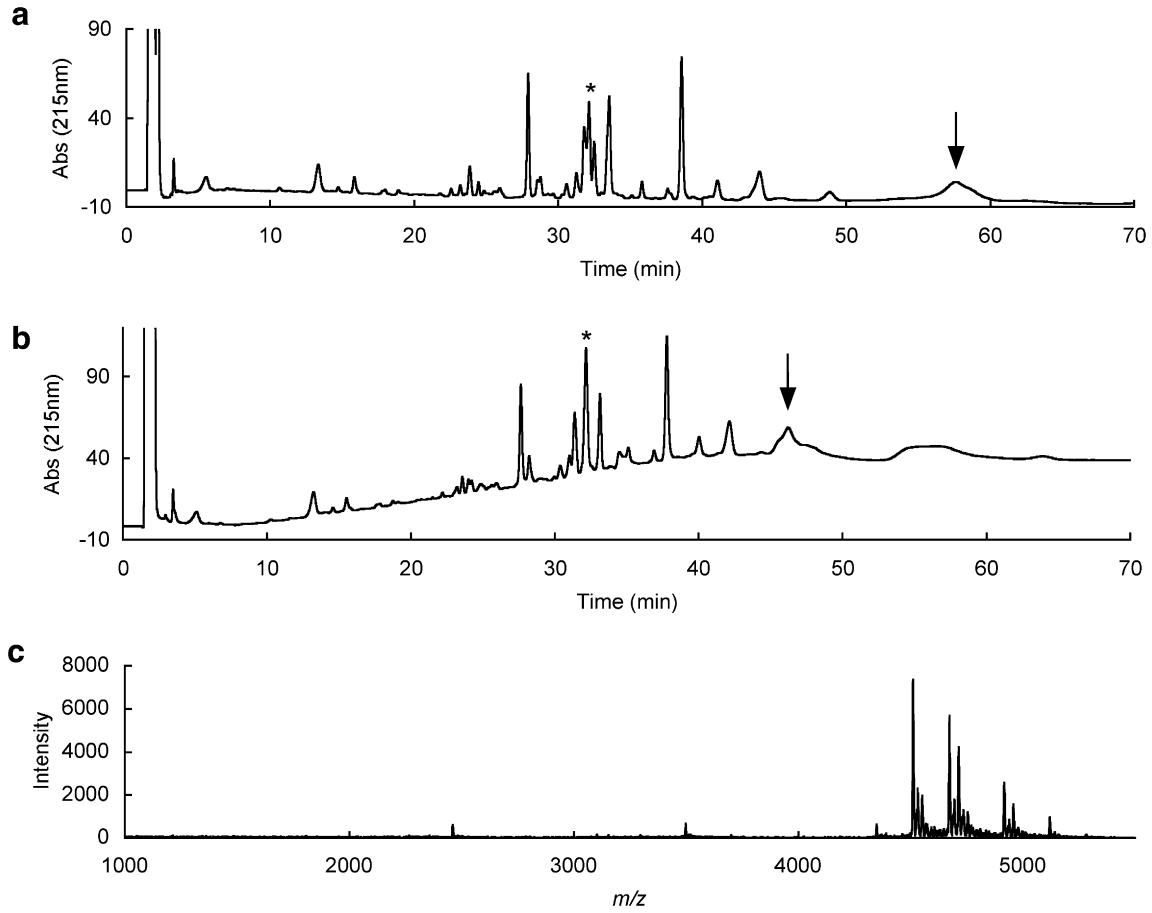

Figure 7. Chromatograms (basic) of tryptic peptides and MS spectrum of a fraction containing glycopeptides from OVA. (a) Chromatogram obtained at pH 9.5 using 40 mM phosphate buffer after separation by Sep-Pak C18; (b) chromatogram obtained at pH 9.5 using 40 mM borate buffer after separation by Sep-Pak C18; (c) mass spectrum of a marked peak in (a) where signals of glycopeptides were observed ranging *m*/*z* 4500–5300. The MS spectrum obtained after HPLC under basic conditions showed more sodiated ions in addition to protonated ones.

glycopeptides were detected by MS at peaks indicated by an arrow. This indicates that the formation of the borate ester changed the retention time of the peaks in RP-HPLC by approximately 11 min compared with the case of phosphate buffer as a mobile phase. To avoid precipitation of salt in 50 mM buffer solutions in 30% acetonitrile, 40 mM buffer solutions were used to elute glycopeptides of OVA, which required an acetonitrile concentration of 30% due to the hydrophobic nature of the analytes. The glycopeptide peaks overlapped with others using borate buffer because of the shorter retention time, the peaks appeared to be well resolved with phosphate buffer. Thus, the fractions of phosphate buffer were analyzed by MS. As a result, it was shown that an extremely clean spectrum showing signals of glycopeptides was obtained (Fig. 7c). The intensities of signals were also improved, which is important if one wants to perform further MS/MS experiments. In addition, the retention time of a peak indicated with star (∗) changed between phosphate and borate. The amino acid sequence of the peak was determined to be VTEQESKPVQMMYQIGLFR (*m*/*z* 2284.3) as a result of MS/MS analysis. It was found that the retention time of certain peptides were somehow affected under these two conditions. According to the sequence of the particular peptide and considering that only two amino acids capable of forming borate esters were present, two methionine residues might have acted as a Lewis base to complex with borate, resulting in the shift. The glycopeptides observed from each protein are listed in Table 3.

The last example is ASF, which is a 37 kDa glycoprotein carrying complex-type oligosaccharides.[27] The glycopeptide fragments obtained from tryptic digestion consist of 15, 28, and 32 amino acids with theoretical average masses of 1742.0, 3018.3, and 3674.0 Da, respectively, all with arginine at the C-terminus. Glycopeptides with *m*/*z* 3364.0 and 3729.1 were obtained as a mixture with other peptides from the 20% acetonitrile fraction by Sep-Pak C18, and glycopeptides with *m*/*z* 4639.1, 5004.0, 5294.2, and 5659.5 were obtained as a mixture with other peptides from the 40% acetonitrile

Table 3. List of identified glycopeptides from model glycoproteins

Observed *m/z*	Theoretical glycan mass	Theoretical peptide mass	Identified glycan structure	Protein
1692.8	1217.0	475.5	$Hex_5HexNAc_2$	RNase B
1855.0	1379.2		$Hex_6HexNAc_2$	
2017.7	1541.3		$Hex_7HexNAc_2$	
2179.3	1703.5		$Hex_8HexNAc_2$	
2341.9	1865.6		$Hex_9HexNAc_2$	
4512.2	1217.0	3295.7	$Hex_5HexNAc_2$	OVA
4550.2	1299.1		$Hex_3HexNAc_4$	
4673.7	1379.2		$Hex_6HexNAc_2$	
4714.0	1420.2		$Hex_5HexNAc_3$	
4755.9	1461.3		$Hex_4HexNAc_4$	
4836.4	1541.3		$Hex_7HexNAc_2$	
4918.0	1623.4		$Hex_5HexNAc_4$	
4958.1	1664.5		$Hex_4HexNAc_5$	
5121.1	1826.6		$Hex_5HexNAc_5$	
5283.8	1988.7		$Hex_6HexNAc_5$	
3364.0	1623.4	1742.0	$Hex_5HexNAc_4$	ASF
3729.1	1988.7		$Hex_6HexNAc_5$	
4639.1	1623.4	3018.3	$Hex_5HexNAc_4$	ASF
5004.0	1988.7		$Hex_6HexNAc_5$	
5294.2	1623.4	3674.0	$Hex_5HexNAc_4$	ASF
5659.5	1988.7		$Hex_6HexNAc_5$	

Protonated signals were shown.

fraction. The former fraction was analyzed in both buffers at 50 mM using the acetonitrile gradient from 10% to 26% and the latter fraction was analyzed under the same conditions used for the analysis of OVA. The peaks of glycopeptides with *m/z* 4639.1 and 5004.0 were overlapped with others using borate buffer as the mobile phase. Thus, the fraction obtained from RP-HPLC with phosphate buffer was used to determine the sequences of the glycopeptides by MS.

As a result of these studies, it was found that identification of all reported glycopeptides based on off-line 'comparative RP-HPLC' MALDI-MS was possible, thus indicating the method has general utility with respect to glycoproteomics.

3.5. MS/MS analysis of glycopeptides

Tandem mass spectrometry (MS/MS) is often necessary to identify a unique fragment that corresponds to the oligosaccharide moiety. It is difficult to analyze molecular species with signals above *m/z* 4000 by MS/MS analysis of glycopeptides with lysine at the C-terminus, as was the case with OVA. However, it was possible to analyze the structure using a purified compound. Figure 8 shows an example of an MS/MS spectrum of glycopeptides containing $Hex_6HexNAc_2$ (*m/z* 4674.9). Y-ion signals that were a result of the cleavage of carbohydrate units from the non-reducing end were observed, and the oligosaccharide sequence was obtained. In addition, we predominantly observed Y_1-ion, which was a product of glycosyl cleavage between two *N*-acetyl-D-glucosamine (GlcNAc) residues. Also, the $^{0,2}X_1$-ion [peptide + (GlcNAc − 120)], which is a cross-ring cleavage product, was found as a descriptive signal. These ion species are known to be very useful for providing the amino acid sequence under MS/MS conditions when the glycosylation sites are not known.[33]

The difficulty in glycoproteomics lies in peak separation, and we have attempted to tackle this issue as described above. As a result of our efforts, MS/MS analysis of larger fragments has become possible. Such examples of MS/MS analysis over *m/z* 5000 are presented in Figure 9. The sequence of glycan part of glycopeptides from ASF containing $Hex_5HexNAc_4$ (*m/z* 5297.4) and $Hex_6HexNAc_5$ (*m/z* 5662.7) could be obtained.

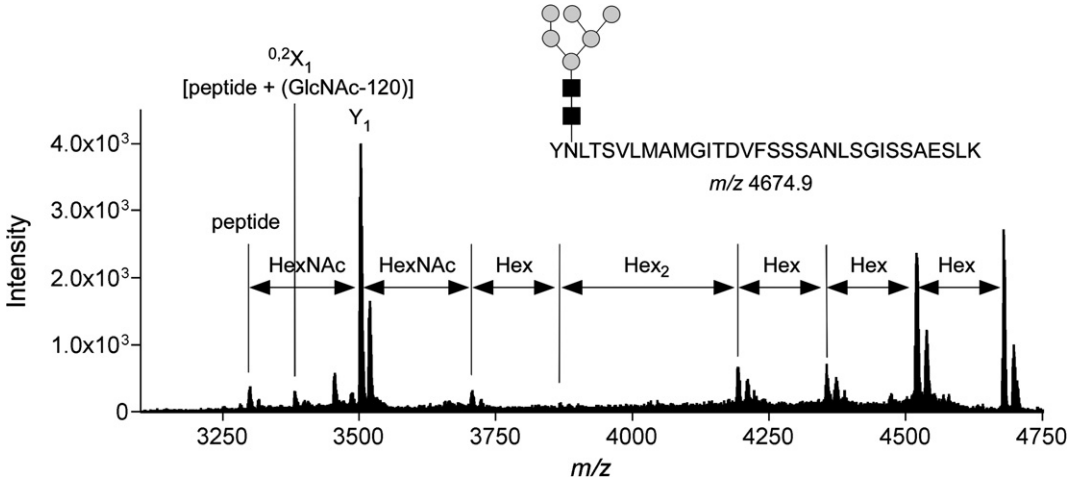

Figure 8. Tandem mass spectrum of glycopeptide with *m/z* 4674.9 $[M+H]^+$ from OVA. Structure: gray circle, Man; black square, GlcNAc.

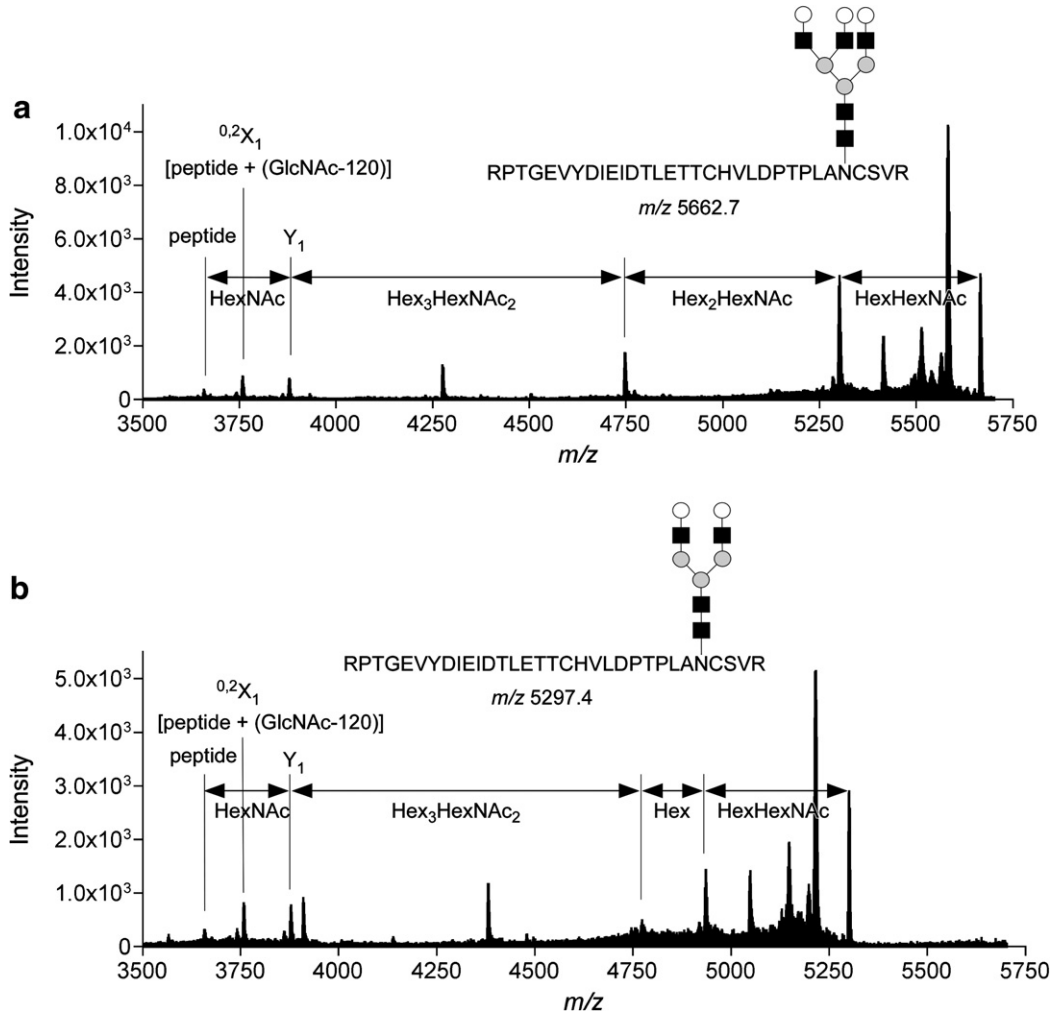

Figure 9. Tandem mass spectra of glycopeptides having same amino acid sequences with (a) *m/z* 5662.7 $[M+H]^+$ and (b) *m/z* 5297.4 $[M+H]^+$ from ASF. Structure: open circle, Gal; gray circle, Man; black square, GlcNAc.

4. Conclusion

Structural analysis of glycoproteins is tedious, time-consuming, and difficult. One of the problems related to the physiochemical properties of glycopeptides and peptides in the ionization process for MS analysis is that the former is difficult to detect in general. Furthermore, ions with higher *m/z* values are more difficult to detect. Thus, it is easy to understand that the detection of glycosylated peptides is problematic. To avoid this problem, one can deglycosylate, then the structures of glycans and protein are separately determined. However, this method does not provide information concerning the sites of glycan attachment when multiple sites exist. Thus, direct analysis is considered to be reasonable. In such cases, a simple solution would be to find a purification method that provides better resolution. We addressed the issue by using a combination of conventional methods, which should be useful for addressing the current problem. Step 1: Rough fractionation using a reverse-phase (RP) cartridge column under acidic conditions. This step decreases the number of peptides in a tryptic digest and contributes to 'better' resolution in the subsequent HPLC separation. Step 2: Comparative RP-HPLC where two chromatograms obtained using phosphate and borate buffers under basic conditions were compared. It has been known that the method based on borate ester formation is useful in carbohydrate analysis; however, the method has never been successfully applied for proteolytic digests. Furthermore, comparison of two different buffer systems provides information regarding the positions of glycopeptides in chromatograms by visual inspection. As a result, glycopeptides thus obtained were used effectively for MS-based structural analysis. To demonstrate the applicability of the method in glycoprotein analysis, we examined various commercially available glycoproteins such as RNase B, IgG1, OVA, and ASF. The results suggest that the method is useful in the analysis of glycoproteins.

References

1. Varki, A. *Glycobiology* **1993**, *3*, 97–130.
2. Helenius, A.; Aebi, M. *Science* **2001**, *291*, 2364–2369.
3. Lowe, J. B. *Cell* **2001**, *104*, 809–812.
4. Dennis, J. W.; Laferté, S.; Waghorne, C.; Breitman, M. L.; Kerbel, R. S. *Science* **1987**, *236*, 582–585.
5. Parekh, R. B.; Dwek, R. A.; Sutton, B. J.; Fernandes, D. L.; Leung, A.; Stanworth, D.; Rademacher, T. W.; Mizuochi, T.; Taniguchi, T.; Matsuta, K.; Takeuchi, F.; Nagano, Y.; Miyamoto, T.; Kobata, A. *Nature* **1985**, *316*, 452–457.
6. Freeze, H. H.; Aebi, M. *Curr. Opin. Struct. Biol.* **2005**, *15*, 490–498.
7. Geyer, H.; Geyer, R. *Biochem. Biophys. Acta* **2006**, *1764*, 1853–1869.
8. Wuhrer, M.; Catalina, M. S.; Deelder, A. M.; Hokke, C. H. *J. Chromatogr., B* **2007**, *849*, 115–128.
9. An, H. J.; Peavy, T. R.; Hedrick, J. L.; Lebrilla, C. B. *Anal. Chem.* **2003**, *75*, 5628–5637.
10. Davies, M.; Smith, K. D.; Harbin, A.-M.; Hounsell, E. F. *J. Chromatogr.* **1992**, *609*, 125–131.
11. Wada, Y.; Tajiri, M.; Yoshida, S. *Anal. Chem.* **2004**, *76*, 6560–6565.
12. Fu, D.; Halbeek, H. V. *Anal. Biochem.* **1992**, *206*, 53–63.
13. García, R.; Rodríguez, R.; Montesino, R.; Besada, V.; González, J.; Cremata, J. A. *Anal. Biochem.* **1995**, *231*, 342–348.
14. Hirabayashi, J.; Arata, Y.; Kasai, K. *Proteomics* **2001**, *1*, 295–303.
15. Geng, M.; Zhang, X.; Bina, M.; Regnier, F. *J. Chromatogr., B* **2001**, *752*, 293–306.
16. Bunkenborg, J.; Pilch, B. J.; Podtelejnikov, A. V.; Wiśniewski, J. R. *Proteomics* **2004**, *4*, 454–465.
17. Qiu, R.; Zhang, X.; Regnier, F. *J. Chromatogr., B* **2007**, *845*, 143–150.
18. Chakraborty, A. B.; Berger, S. J. *J. Biomol. Tech.* **2005**, *16*, 327–335.
19. Toll, H.; Oberacher, H.; Swart, R.; Huber, C. G. *J. Chrmotogr., A* **2005**, *1079*, 274–286.
20. Takahashi, N.; Takahashi, Y.; Ortel, T. L.; Lozier, J. N.; Ishioka, N.; Putnam, F. W. *J. Chromatogr., A* **1984**, *317*, 11–26.
21. Wuhrer, M.; Deelder, A. M.; Hokke, C. H. *J. Chromatogr., B* **2005**, *825*, 124–133.
22. Weitzman, S.; Scott, V.; Keegstra, K. *Anal. Biochem.* **1979**, *97*, 438–449.
23. Honda, S.; Makino, A.; Suzuki, S.; Kakehi, K. *Anal. Biochem.* **1990**, *191*, 228–234.
24. Iwase, H.; Morinaga, T.; Li, Y.-T.; Li, S.-C. *Anal. Biochem.* **1981**, *113*, 93–95.
25. Salt, S. D.; Gander, J. E. *Exp. Mycol.* **1985**, *9*, 9–19.
26. Rothman, R. J.; Warren, L. *Biochim. Biophys. Acta.* **1988**, *955*, 143–153.
27. Rice, K. G.; Rao, N. B. N.; Lee, Y. C. *Anal. Biochem.* **1990**, *184*, 249–258.
28. Fu, D.; Chen, L.; O'Neill, R. A. *Carbohydr. Res.* **1994**, *261*, 173–186.
29. Takahashi, N.; Ishii, I.; Ishihara, H.; Mori, M.; Tejima, S.; Jefferis, R.; Endo, S.; Arata, Y. *Biochemistry* **1987**, *26*, 1137–1144.
30. Takegawa, Y.; Deguchi, K.; Keira, T.; Ito, H.; Nakagawa, H.; Nishimura, S.-I. *J. Chromatogr., A* **2006**, *1113*, 177–181.
31. Lattová, E.; Kapková, P.; Krokhin, O.; Pereault, H. *Anal. Chem.* **2006**, *78*, 2977–2984.
32. Silva, M. L. C. D.; Stubbs, H. J.; Tamura, T.; Rice, K. G. *Arch. Biochem. Biophys.* **1995**, *318*, 465–475.
33. Wuhrer, M.; Hokke, C. H.; Deelder, A. M. *Rapid Commun. Mass Spectrom.* **2004**, *18*, 1741–1748.