ELSEVIER

Structural analysis of a glycoprotein by liquid chromatography–mass spectrometry and liquid chromatography with tandem mass spectrometry
Application to recombinant human thrombomodulin

Satsuki Itoh, Nana Kawasaki*, Miyako Ohta, Takao Hayakawa

Division of Biological Chemistry and Biologicals, National Institute of Health Science, 1-18-1, Kamiyoga, Setagaya-ku, Tokyo 158-8501, Japan

Abstract

Using recombinant human thrombomodulin (rhTM) expressed in Chinese hamster ovary (CHO) cells, we studied the structural analysis of a glycoprotein by liquid chromatography–mass spectrometry (LC–MS) and liquid chromatography with tandem mass spectrometry (LC–MS–MS). First, we analyzed the structure of both the *O*- and *N*-linked glycans in rhTM by oligosaccharide mapping using LC–MS equipped with a graphitized carbon column (GCC–LC–MS). Major *O*- and *N*-linked glycans were determined to be core 1 structure and fucosyl biantennary containing $NeuAc_{0-2}$, respectively. Next, the post-translational modifications and their heterogeneities, including the site-specific glycosylation, were analyzed by mass spectrometric peptide/glycopeptide mapping of trypsin-digested rhTM and precursor-ion scanning. Precursor-ion scanning was successful in the detection of five glycopeptides. Four *N*-glycosylation sites and their site-specific carbohydrate heterogeneity were determined by their mass spectra. *O*-Glycosylation could be estimated on the basis of its mass spectrum. We were able to identify partial β-hydroxylation on Asn324 and Asn439, and *O*-linked glucose on Ser287 from the peptide/glycopeptide map and their mass spectra. We demonstrated that a sequential analysis of LC–MS and LC–MS–MS are very useful for the structural analysis of *O*- and *N*-linked glycans, polypeptides, and post-translational modifications and their heterogeneities, including site-specific glycosylation in a glycoprotein. Our method can be applied to a glycoprotein in biological samples.

1. Introduction

Many methods for performing a structural analysis of a glycoprotein have been reported. LC–MS is one of the most effective tools for the analysis of both protein and carbohydrate moieties. Peptide/glyco-peptide mapping using on-line LC–MS is used to determine the protein sequence and post-translational modifications, including glycosylation [1–5]. Precursor-ion scanning is reported to be useful for the selective detection of glycopeptides in the peptide/

*Corresponding author. Fax: +81-3-3707-6950.
E-mail address: nana@nihs.go.jp (N. Kawasaki).

glycopeptide map [6–8]. This method can trace only glycopeptides in the peptide/glycopeptide map by monitoring the marker ions such as m/z 204 (HexNAc$^+$) and 366 (Hex–HexNAc$^+$) produced from glycopeptides by collision-induced dissociation (CID). We previously demonstrated that LC–MS could be employed for the structural analysis of *N*-linked glycans with the use of a graphitized carbon column (GCC) [3,9–11]. Many *N*-linked glycans could be analyzed by oligosaccharide mapping using GCC-LC–MS. Such analytical approach with GCC-LC–MS can also facilitate the assignment of glycopeptides in a peptide/glycopeptide map and consequently the determination of the glycosylation site and site-specific carbohydrate heterogeneity. Recently, we demonstrated that the use of microbore GCC-LC–MS enabled us to perform the simultaneous microanalysis of various *N*-linked glycans in a glycoprotein [12]. A combination of the mass spectrometric oligosaccharide mapping and peptide/glycopeptide mapping, including the precursor-ion scan, is expected to be a powerful tool for the structural analysis of a small amount of glycoprotein possessing multiple glycosylation sites.

Thrombomodulin (TM) is a glycoprotein that is a receptor for thrombin on the vascular endothelium and plays an essential role as a protein cofactor in thrombin-catalyzed activation of protein C [13,14]. This glycoprotein is present not only on the endothelial cell surface but also in soluble form in plasma and urine [15]. TM on the vascular endothelium consists of a *N*-terminal domain, six epidermal-growth-factor-like (EGF-like) domains, an *O*-glycosylation domain, a transmembrane segment, and a carboxy-terminal cytoplasmic tail. The DNA-predicted sequence suggests the presence of several *O*-glycosylation sites, five *N*-glycosylation sites (Asn29, Asn97, Asn98, Asn364 and Asn391) and two β-hydroxylation sites (Asn324 and Asn439) [16,17]. The *O*-glycosylation sites in a secretable recombinant human TM expressed in 293 cells, which contains all of the extracellular domains, were identified as Ser498, Thr500, Thr504, Thr506 and Ser492 where chondroitin sulfate attached [18]. Linkages of Galβ1–3GalNAc (Core 1 structure), glucose, and sulfated tetrasaccharides to the *O*-glycosylation sites are reported in human urinary TM

[19,20]. Four *N*-glycosylation sites (Asn29, Asn97/98, Asn364 and Asn391) were identified by amino acid sequence analysis [18,20,21]. Structure of *N*-linked oligosaccharides in recombinant TM expressed by C127 cells and urinary TM were studied by using two-dimensional mapping of pyridylaminated oligosaccharides, and are characterized to be mainly fucosyl biantennary and fucosyl triantennary complex-type oligosaccharides [20,21]. β-Hydroxylation at two Asn residues in conserved amino acid sequences (–Cys–X–Asp/Asn–X–X–X–X–Tyr/Phe–X–Cys–X–Cys–) are confirmed in two of six EGF-like domains in bovine TM [16,22,23]. Although various TMs from different sources have been partially characterized by different methods, structural details such as sialylation and site-specific glycosylation of one TM are still unclear.

In this paper, we describe how mass spectrometric oligosaccharide mapping and peptide/glycopeptide

```
        10        a         20              30            40                50
APAEPQPGGS  QCVEHDCFAL  YPGPATFLNA  SQICDGLR/GH  LMTVR/SSVAA
                                    K1             K2

        60        70              80            90               100
DVISLLLNGD  GGVGR/R/R/LWI  GLQLPPGCGD  PK/R/LGPLR/GF  QWVTGDNNTS
            K3                        K6     K8

       110       120             130           140              150
YSR/WAR/LDLN  GAPLCGPLCV  AVSAAEATVP  SEPIWEEQQC  EVK/ADGFLCE
K9   K10                                         K11

       160       170             180           190              200
FHFPATCRPL  AVEPGAAAAA  VSITYGTPFA  AR/GADFQALP  VGSSAAVAPL
                                    K12

       210       220             230           240              250
GLOLMCTAPP  GAVOGHWAR/E  APGAWDCSVE  NGGCEHACNA  IPGAPR/CQCP
                        K13                                   K14

       260       270             280           290              300
AGAALQADGR/  SCTASATQSC  NDLCEHF¦CVP  NPDQPGSY¦SC  MCETGYR/LAA
      K15     K16-1          K16-2        K16-3         K16

       310       320             330           340              350
DQHR/CEDVDD  CILEPSPCPQ  R/CVNTQGGFE  CHCYPNYDLV  DGECVEPVDP
K17                     K18        *

       360       370             380           390              400
CFR/ANCEYOC  OPLNQTSYLC  VCAEGFAPIP  HEPHR/CQMFC  NQTACPADCD
K19                                  K20

       410       420             430           440              450
PNTQASCECP  EGYILDDGFI  CTDIDECENG  GFCSGVCHNL  PGTEECICGP
                                            *

       460       470             480           490              498
DSALVR/HIGT  DCDSGK/VDGG  DSGSGEPPPS  PTPGSTLTPP  AVGLVHSG
K21          K22                                  K23
```
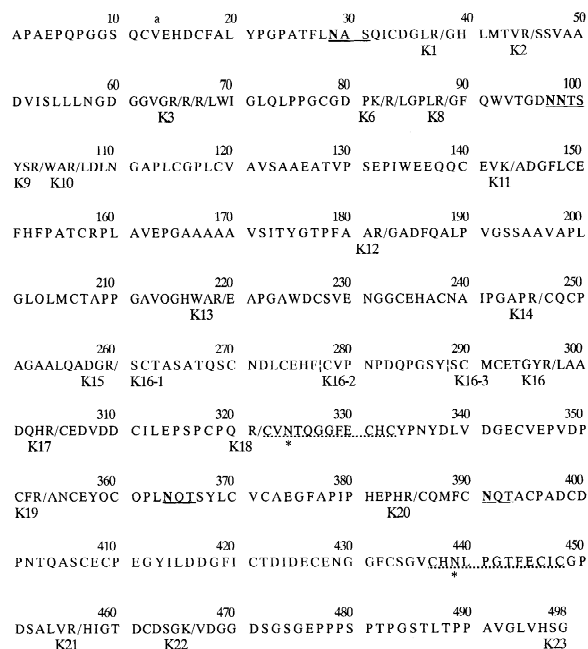
Fig. 1. Amino acid sequence of rhTM. Tryptic cleavage sites are indicated by /, and tryptic peptides are named K1–23. Chymotrypsin cleavage sites in peptide K16 are indicated by ¦ and numbered from K16-1 to K16-3. Underlines are potential *N*-glycosylation sequences. Dashed lines show the conserved amino acid sequence for β-hydroxylation, and asterisks indicate potential β-hydroxylation sites.

mapping with microbore LC–MS are useful for a rapid and convenient analysis of a glycoprotein, using recombinant human TM (rhTM) produced in Chinese hamster ovary (CHO) cells. rhTM used in this study is composed of 498 amino acids, lacking chondroitin sulfate (Fig. 1). We perform oligosaccharide mapping by microbore GCC-LC–MS for the analysis of both *O*- and *N*-linked glycans. Subsequently, peptide/glycopeptide mapping and precursor-ion scanning were carried out to determine the post-translational modifications.

2. Experimental

2.1. Materials

The rhTM expressed by CHO cells was a kind gift from Asahi Chemical Industry (Tokyo, Japan). *N*-Glycosidase F (PNGase F) was obtained from Roche Diagnostics (Mannheim, Germany). Glycopeptidase A was purchased from Seikagaku Kogyo (Tokyo, Japan). Trypsin and chymotrypsin were obtained from Sigma (St. Louis, MO, USA) and Wako Pure Chemical Industries (Osaka, Japan), respectively. All other chemicals used were of the highest purity available.

2.2. Monosaccharide analysis

Monosaccharide analysis was performed by high-pH anion-exchange chromatography with pulsed amperometric detection (HPAEC–PAD) following hydrolysis of the rhTM (50 μg) with 400 μl of 2 *M* trifluoroacetic acid (TFA) for 3 h at 100 °C. The acid-treated sample was lyophilized and then dissolved in water prior to use.

2.3. HPAEC–PAD

HPAEC–PAD was performed with a DX-300 system (Dionex, Sunnyvale, CA, USA) equipped with a CarboPac PA-1 anion-exchange column (250×4 mm, Dionex) and a pulsed electrochemical detector. Analysis of the monosaccharides was carried at an isocratic elution of 19 m*M* NaOH.

2.4. Reduction and S-carboxymethylation of rhTM

Desalted and lyophilized rhTM (360 μg) was dissolved in 360 μl of 0.5 *M* Tris–HCl buffer (pH 8.6) that contained 8 *M* guanidine hydrochloride and 5 m*M* ethylenediaminetetraacetic acid (EDTA). After an addition of 2.6 μl of 2-mercaptethanol, the mixture was allowed to stand at room temperature for 2 h. To this solution, 7.56 mg of monoiodoacetic acid was added, and the resulting mixture was incubated at room temperature for 2 h in the dark. The reaction mixture was applied to PD-10 column (Amersham Pharmacia Biotech, Uppsala, Sweden) to remove the reagents, and the eluate was lyophilized.

2.5. Preparation of O-linked oligosaccharide alditols

Reduced and carboxymethylated rhTM (RCM-rhTM, equivalent to 120 μg of rhTM) was dissolved in 400 μl of 0.1 *M* NaOH containing 0.4 *M* NaBH$_4$ and incubated at room temperature for 24 h [24]. The reaction mixture was neutralized to pH 7.0 by adding diluted acetic acid, and applied to an AG 50W-X8 (Bio-Rad, Hercules, CA, USA) column. The fraction eluted with water was lyophilized, and boric acid was removed by evaporation with methanol.

2.6. Preparation of N-linked oligosaccharides alditols

rhTM (100 μg) was dissolved in 100 μl of 0.1 *M* sodium phosphate buffer, pH 7.2, containing 0.1% Triton X-100 and incubated with 5 U of PNGase F at 37 °C for 2 days. Protein was precipitated with 340 μl of cold ethanol. The supernatant was dried, and the oligosaccharides were dissolved in 100 μl of water. To the oligosaccharide solution, 100 μl of 0.5 *M* NaBH$_4$ was added, and the mixture was left at room temperature for 2 h. Diluted acetic acid was added to the mixture to decompose excess NaBH$_4$ and to adjust it to pH 7.0. The reaction mixture was applied to Supelclean ENVI-Carb (Supelco, Bellefonte, PA, USA) to be desalted. Oligosaccharide alditols were eluted with 30% acetonitrile containing 5 m*M* ammonium acetate and lyophilized. Samples

were dissolved in water and injected to a GCC-LC–MS.

2.7. Trypsin digestion of RCM-rhTM

RCM-rhTM (equivalent to 120 μg of rhTM) was digested in 0.1 *M* Tris–HCl buffer (pH 8.0) containing 2.4 μg of trypsin at 37 °C for 1 h.

2.8. Chymotrypsin digestion of glycopeptide

Glycosylated peptide, peak 10 in Fig. 4, was fractionated and digested in 0.1 *M* Tris–HCl buffer (pH 8.0) containing 1 μg of chymotrypsin at 37 °C for 40 h.

2.9. Glycopeptidase A digestion of glycopeptide

Glycopeptide, peak 16 in Fig. 4, was fractionated and treated in 0.1 *M* sodium acetate (pH 5.0) containing 10 mU of glycopeptidase A at 37 °C for 24 h.

2.10. High-performance liquid chromatography (HPLC) for oligosaccharide alditols prepared from rhTM

HPLC was carried out using a MAGIC 2002 system (Michrom BioResources, Auburn, CA, USA) with a Hypercarb column (5 μm, 150×1.0 mm, Michrom BioResources). The eluents were 5 m*M* ammonium acetate, pH 9.6, containing 2% acetonitrile (pump A); and 5 m*M* ammonium acetate, pH 9.6, containing 80% acetonitrile (pump B). The oligosaccharide alditols were eluted at a flow-rate of 50 μl/min with a gradient of 5–20% of pump B in 80 min for *O*-linked oligosaccharide alditols, and 5–30% of pump B in 80 min for *N*-linked oligosaccharide alditols. The eluate was monitored at 206 nm.

2.11. HPLC of trypsin digested RCM-rhTM

HPLC was carried out using a MAGIC 2002 system with a MAGIC C18 column (150×1.0 mm, Michrom BioResources). The eluents consisted of water containing 2% acetonitrile and 0.05% TFA (pump A); and 80% acetonitrile and 0.05% TFA

Monosaccharide composition analysis of rhTM oligosaccharides[a]
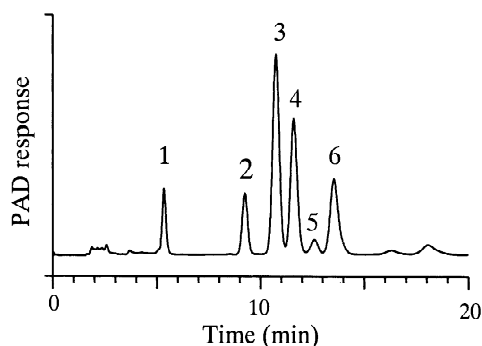
Peak No.	Monosugar	Relative molar proportions	
1	Fuc	1.1[b]	(4.4)[c]
2	GalNH$_2$	0.8[b]	(3.4)[c]
3	GlcNH$_2$	3.9[b]	(15.6)[c]
4	Gal	4.0[b]	(16.2)[c]
5	Glc	0.4[b]	(1.5)[c]
6	Man	3.0[b]	(12.0)[c]

[a] Values represent means of duplicate mesurments.
[b] Data are normalized to 3 mannose residues.
[c] Data are normalized to 12 mannose residues.

Fig. 2. Monosaccharide composition analysis of rhTM. Acid hydrolysate of rhTM was analyzed by HPAEC–PAD.
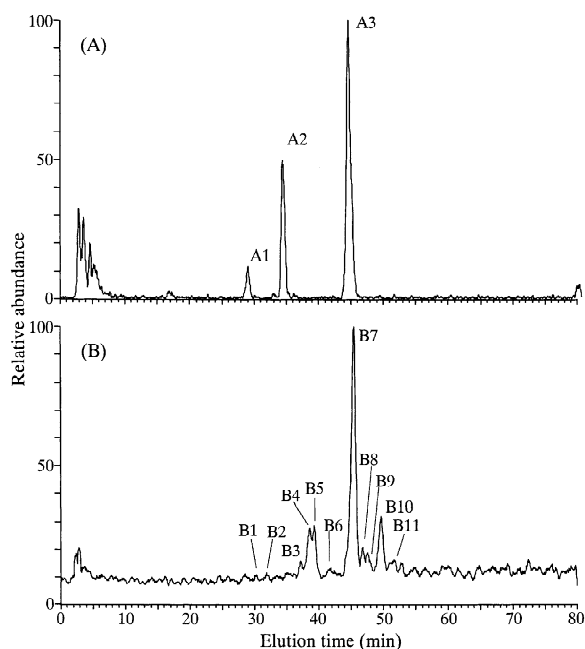

Fig. 3. Extracted ion chromatogram of *O*-linked oligosaccharide alditols (A) and TIC chromatogram of *N*-linked oligosaccharide alditols (B) from rhTM in the positive-ion mode.

(pump B). Trypsin-digested samples of RCM-rhTM were eluted with a liner gradient from 5 to 45% of pump B in 40 min at a flow-rate 50 μl/min. The effluent was monitored at 206 nm.

2.12. Electrospray ionisation (ESI)-MS and Collision-induced dissociation (CID)-ESI-MS–MS

Mass spectra were recorded on a Finnigan TSQ 7000 triple-stage quadrupole mass spectrometer equipped with an electrospray ion (ESI) source (Thermo Finnigan, San Jose, CA, USA). The mass spectrometer was operated in the positive-ion mode. The mass range, m/z 300–1300, was acquired with the scan duration of 2 s for O-linked oligosaccharide alditols, and m/z 700–2400 with the scan duration of 4 s for N-linked oligosaccharide alditols. For trypsin digest of RCM-rhTM, the mass ranges, m/z 400–2400 and m/z 1000–4000, were acquired with the scan duration of 4 and 6 s, respectively. The ESI

voltage was set at 4500 V, and the capillary temperature was 225 °C. The electron multiplier was set at 1200 V. The pressure of the sheath gas was 70 p.s.i., and that of the auxiliary gas was 10 units.

CID-MS–MS precursor-ion scanning, was carried out with argon gas as the collision gas at 2.0 mTorr in the positive-ion mode. A collision energy of −20 eV was used. The electron multiplier was set at 1500 V.

3. Results

3.1. Monosaccharide composition analysis

The monosaccharide composition was determined by the method of Hardy et al. using HPAEC–PAD following acid hydrolysis of rhTM [25]. The result is summarized in Fig. 2. The relative molar ratios of Fuc: GalNH$_2$:GlcNH$_2$:Gal:Glc:Man was 1.1:0.8:3.9:

Table 1
Carbohydrate structure A1–B11 in Fig. 3 and their theoretical masses and observed m/z values

Peak no.	Carbohydrate composition[a]	Sugar type[b]	Theoretical mass[c]	Observed m/z M$^+$	M^{2+}
A1	[Hex][NA]		473.4	474.3	
A2	[Hex][HexNAc][NA]	Core 1	676.6	677.4	
A3	[Hex][HexNAc][NA]$_2$	Core 1	967.9	968.5	
B1	[Fuc][Hex]$_5$[HexNAc]$_4$	FucBi	1789.7	1791.1	895.6
B2	[Fuc][Hex]$_6$[HexNAc]$_5$	FucBiLac, FucTri	2155.0	2156.6	1078.9
B3	[Fuc][Hex]$_6$[HexNAc]$_5$[NA]	FucBiLacNA, FucTriNA	2446.3		1223.8
B4	[Fuc][Hex]$_4$[HexNAc]$_3$	FucBi(1)	1424.3	1424.3	
	[Fuc][Hex]$_5$[HexNAc]$_4$[NA]	FucBiNA	2080.9	2081.7	1041.2
	[Fuc][Hex]$_6$[HexNAc]$_5$[NA]	FucBiLacNA, FucTriNA	2446.3		1223.9
B5	[Fuc][Hex]$_4$[HexNAc]$_3$	FucBi(1)	1424.3	1424.7	
	[Fuc][Hex]$_5$[HexNAc]$_4$[NA]	FucBiNA	2080.9	2081.1	1040.9
	[Fuc][Hex]$_6$[HexNAc]$_5$[NA]	FucBiLacNA, FucTriNA	2446.3		1223.7
B6	[Hex]$_5$[HexNAc]$_4$[NA]$_2$	BiNA$_2$	2226.0	2227.0	1113.5
B7	[Fuc][Hex]$_4$[HexNAc]$_3$[NA]	FucBi(1)NA	1715.6	1715.8	
	[Fuc][Hex]$_5$[HexNAc]$_4$[NA]$_2$	FucBiNA$_2$	2372.2	2372.4	1186.7
	[Fuc][Hex]$_6$[HexNAc]$_5$[NA]$_2$	FucBiLacNA$_2$, FucTriNA$_2$	2737.5		1369.2
B8	[Fuc][Hex]$_6$[HexNAc]$_5$[NA]$_2$	FucBiLacNA$_2$, FucTriNA$_2$	2737.5		1369.3
B9	[Fuc][Hex]$_6$[HexNAc]$_5$[NA]$_2$	FucBiLacNA$_2$, FucTriNA$_2$	2737.5		1369.8
	[Fuc][Hex]$_6$[HexNAc]$_5$[NA]$_3$	FucTriNA$_3$	3028.8		1515.1
B10	[Fuc][Hex]$_6$[HexNAc]$_5$[NA]$_3$	FucTriNA$_3$	3028.8		1515.6
B11	[Fuc][Hex]$_6$[HexNAc]$_5$[NA]$_2$	FucBiLacNA$_2$, FucTriNA$_2$	2737.5		1368.6

[a] Fuc, fucose; Hex, hexose; HexNAc, N-acetylhexosamine; NA, N-acetylneuramic acid.

[b] Bi, biantennary; Tri, triantennary; Tetra, tetraantennary; Lac, N-acetyllactosamine; FucBi(1)

[c] Average mass value.
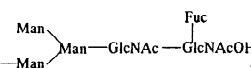
4.0:0.4:3.0. This result suggests that *N*-linked glycans linked to rhTM are mainly complex-type oligosaccharides, that is, fucosylated biantennary oligosaccharides. The existence of *O*-linked glycans was also suggested by the detection of $GalNH_2$ and Glc.

3.2. O-Linked glycans

O-Linked glycans were released by β-elimination with alkaline borohydride and subjected to microbore GCC-LC–MS. Fig. 3A shows the chromatograms of microbore GCC-LC–MS of *O*-linked oligosaccharide alditols prepared from rhTM. *O*-Linked oligosaccharide alditols were successfully separated by GCC-LC–MS with a gradient of 5–20% of pump B in 80 min. Peaks in Fig. 3A are assigned on the basis of their elution time and m/z values in their mass spectra (Table 1). The m/z values of peaks A1, A2, and A3 are consistent with theoretical m/z values of NeuAc–Hex, NeuAc–Hex–HexNAc, and $NeuAc_2$–Hex–HexNAc, respectively. As estimated from monosaccharide analysis, they can be identified as NeuAc–Gal, NeuAc–Gal–GalNAc, and $NeuAc_2$–Gal–GalNAc, respectively. These results suggest that oligosaccharide mapping by GCC-LC–MS is applicable for the analysis of *O*-linked glycans, and major *O*-linked glycans in rhTM produced in CHO cells are a core 1 structure with sialylation.

3.3. N-Linked glycans

N-Linked glycans were released from rhTM with PNGase F and reduced with $NaBH_4$ [3,9]. Enzyme-treated deglycosylation of rhTM was confirmed by SDS–PAGE (data not shown). Fig. 3B shows the total ion current (TIC) chromatogram of *N*-linked oligosaccharide alditols, which were well separated by GCC-LC–MS with a gradient of 5–30% of pump B in 80 min. One of the m/z values of the most abundant peak, peak B7, is consistent with the theoretical m/z value of disialylated fucosyl biantennary. Fucosyl triantennary and fucosyl biantennay lacking *N*-acetyllactosamine (FucBi(1), Table 1) were also detected in rhTM. Most *N*-linked glycans are sialylated, and the heterogeneity in sialylation is observed much as it is in *O*-linked glycans.

3.4. N- and O-Linked glycans at each glycosylation site

To determine the glycosylation sites and site-specific glycosylation, we performed peptide/glyco-peptide mapping of trypsin-digested RCM-rhTM using LC–MS. Peptides expected to be obtained by tryptic digestion of rhTM are shown in Fig. 1 (K1–23). Fig. 4A and B shows a UV chromatogram at 206 nm and a TIC chromatogram of LC–MS of RCM-rhTM trypsin digest in the positive-ion mode, respectively. The peaks of glycopeptides in the TIC chromatogram of the peptide/glycopeptide map were specified by precursor-ion scanning with LC–MS–MS. In this study, we employed m/z 204 ion (HexNAc$^+$) for the marker ion in order to detect both *N*- and *O*-glycosylated peptides. Fig. 4C shows a TIC chromatogram of precursor-ion scanning by LC–MS–MS. By comparing the elution time of peaks in the TIC chromatogram of LC–MS with those of five peaks in precursor-ion scanning, peaks 5, 7, 12, 14 and 16 in Fig. 4B were identified as being those of glycopeptides. Fig. 5 shows mass spectra of peaks 5, 7, 12, 14 and 16 in Fig. 4B. There
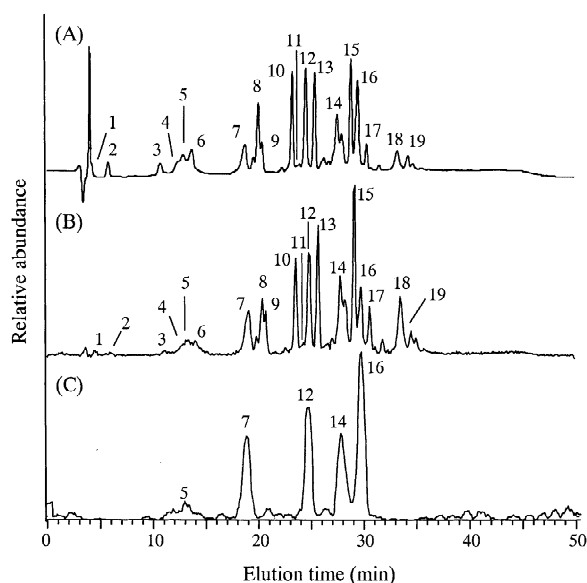

Fig. 4. Peptide/glycopeptide map of rhTM tryptic digest. UV chromatogram at 206 nm (A), TIC chromatogram of LC–MS in the positive-ion mode (B), and TIC chromatogram of LC–MS–MS, precursor-ion scan of m/z 204 (C).
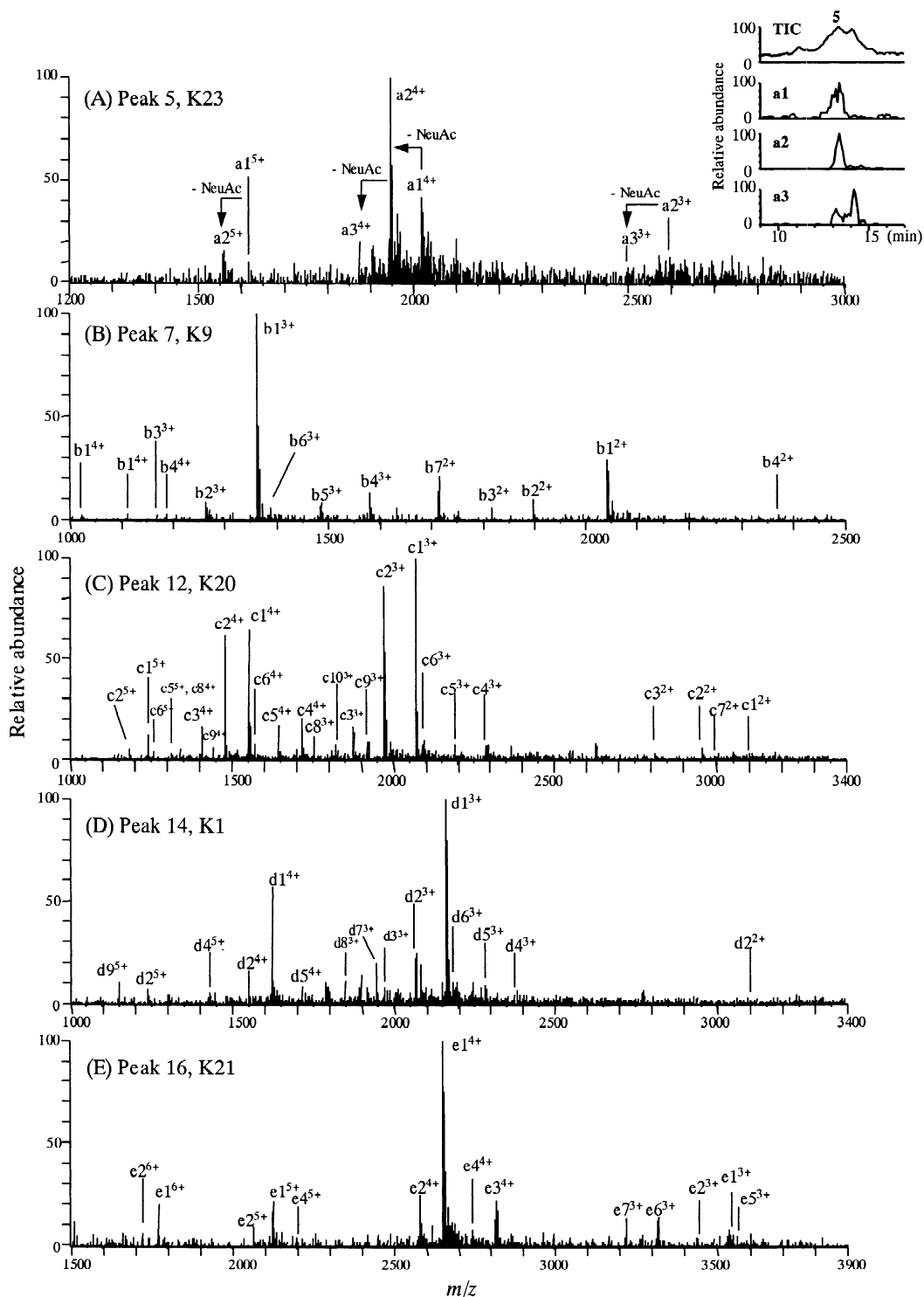
Fig. 5. Mass spectra of glycopeptides in Fig. 4; peak 5 (A), peak 7 (B), peak 12 (C), peak 14 (D), and peak 16 (E). The observed m/z value of each ion is described in Table 3. The inset shows TIC and RIC of glycoforms, a1–3.

are many ions due to the heterogeneity of carbohydrates in these mass spectra. On the basis of the theoretical masses of glycans obtained by oligosaccharide mapping (Table 1) and theoretical masses of peptides, peaks 7, 12, and 16 were assigned to glycopeptide K9, K20 and K21 containing potential *N*-glycosylation sites Asn97 or 98, Asn364 and Asn391, respectively. Peak 14 appeared to consist of peptide K6 and glycopeptide K1 containing Asn29 (Table 2). The carbohydrate structures of individual ions at each *N*-glycosylation site were determined as shown in Table 3. The major *N*-linked glycans, fucosyl biantennary, fucosyl triantennary, and fucosyl biantennay lacking *N*-acetyllactosamine (FucBi(1)), are linked to all glycosylation sites. Biantennary without fucose attaches to Asn29 and Asn364.

Some series of ions are observed in the mass spectra of peak 5 (Fig. 5A). The *m/z* increments among theses ions are 97 (+3), 72 (+4) and 58 (+5), which indicate the existence of glycoforms

with difference in the number of NeuAc. Reconstructed ion chromatograms (RIC) suggest that peak 5 is consisting of glycoforms, a1–3 (Fig. 5A, inset). On the basis of their molecular masses, a1–3 can be estimated to consist of peptide K23, 7Hex, 4HexNAc, and 9–11NeuAc. Together with the results of monosaccharide analysis and oligosaccharide mapping, it is suggested that peptide K23 possesses 7Gal, 4GalNAc, and 9–11NeuAc.

3.5. Glucosylation

Peaks of nonglycosylated peptides, except a hydrophilic small peptide, K22, which is considered to be merely passing through the column, were assigned by comparing the experimental masses with the predicted masses from the amino acid sequence (Fig. 1 and Table 2). Consequently, three posttranslational modified peptides were characterized: glucosylation of peptide K16 and β-hydroxylation of peptide K19 and K21.

Table 2
Tryptic peptide numbers and amino acid residues of peaks in Fig. 4B and their theoretical masses and observed *m/z* values

Peak no.	Peptide no.	Amino acid residues	Theoretical mass[a]	Observed *m/z*				
				M[+]	M[2+]	M[3+]	M[4+]	M[5+]
1	K17	298–304	809.9	810.5	405.7			
2	K10	104–106	431.5	432.3				
3	K2	39–45	813.0	813.6	407.2			
4	K8	84–88	554.7	555.4				
5	K23[b]	467–498	2927.1					
6	K15	247–260	1476.5	1477.3	739.1	493.2		
7	K9[b]	89–103	1731.8					
8	K14	220–246	2886.0		1443.7	963.0	722.6	
9	K18	305–321	2093.1	2093.7	1047.2	698.4		
10	K16[c]	261–297	4516.5		2258.0	1506.2	1130.1	
11	K16	261–297	4354.4		2178.3	1452.8	1089.7	
12	K20[b]	354–385	3852.1					
13	K19[d]	322–353	3916.0		1958.8	1306.2		
	K19	322–353	3900.0		1950.6	1300.6		
14	K6	68–82	1651.9	1652.0	826.7			
	K1[b]	1–38	4135.4					
15	K13	183–219	3704.2		1852.1	1236.2	927.1	741.8
16	K21[b]	386–456	8245.5					
17	K12	144–182	4113.6		2057.1	1371.9	1029.4	823.8
18	K11	107–143	4042.4		2022.3	1348.5	1011.9	809.6
19	K3	46–65	1900.1	1900.3	951.0	634.3		

[a] Average mass value.
[b] Glycopeptide.
[c] Peptide with hexose.
[d] β-Hydroxylated peptide.

Table 3
Structural assignments of glycoforms in Fig. 5, and their theoretical masses and observed m/z values

Glycoform in Fig. 5	Peptide no.	Amino acid residues	Sugar type[a]	Theoretical mass[b]	Observed m/z				
					M^{2+}	M^{3+}	M^{4+}	M^{5+}	M^{6+}
a1	K23	467–498	7Gal, 4GalNAc, 11NA	8078.7			2019.2	1616.8	
a2			7Gal, 4GalNAc, 10NA	7787.5		2596.4	1947.2	1558.4	
a3			7Gal, 4GalNAc, 9NA	7496.2		2499.9	1874.4		
b1	K9	89–103	FucBiNA$_2$	4084.0	2043.0	1362.4	1022.3		
b2			FucBiNA	3792.7	1897.6	1264.9			
b3			FucBi	3501.4	1752.0	1167.9			
b4			FucTriNA$_3$	4740.5	2371.3	1581.7	1186.2		
b5			FucBiLacNA$_2$, FucTriNA$_2$	4449.3		1484.7	1113.2		
b6			FucBiLacNA, FucTriNA	4158.0		1387.1			
b7			FucBi(1)NA	3427.4	1715.3				
c1	K20	354–385	FucBiNA$_2$	6204.3	3104.0	2069.3	1552.7	1242.2	
c2			FucBiNA	5913.0	2957.3	1972.5	1479.3	1183.9	
c3			FucBi	5621.7	2812.1	1874.9	1407.1		
c4			FucTriNA$_3$	6860.8		2287.4	1716.5		
c5			FucBiLacNA$_2$, FucTriNA$_2$	6569.6		2191.3	1643.4	1315.4	
c6			FucBiLacNA, FucTriNA	6278.3		2094.5	1570.3	1257.2	
c7			FucBiLac, FucTri	5987.1	2994.6				
c8			FucBi(1)	5256.4		1753.8	1315.4		
c9			BiNA	5766.9		1923.7	1443.2		
c10			Bi	5475.6		1826.1			
d1	K1	1–38	FucBiNA$_2$	6487.6		2163.9	1623.2		
d2			FucBiNA	6196.3	3100.2	2066.6	1550.2	1240.9	
d3			FucBi	5905.0		1970.2			
d4			FucTriNA$_3$	7144.1		2382.6		1430.3	
d5			FucBiLacNA$_2$, FucTriNA$_2$	6852.9		2285.0	1715.0		
d6			FucBiLacNA, FucTriNA	6561.6		2188.2			
d7			FucBi(1)NA	5831.0		1944.8			
d8			FucBi(1)	5539.7		1848.4			
d9			Bi	5758.9				1152.0	
e1	K21	386–456	FucBiNA$_2$	10 597.7		3534.1	2649.2	2120.1	1767.6
e2			FucBiNA	10 303.7		3436.4	2577.5	2061.4	1718.0
e3			FucTriNA$_3$	11 254.2			2813.6		
e4			FucBiLacNA$_2$, FucTriNA$_2$	10 963.0			2741.0	2193.5	
e5			FucBiLacNA, FucTriNA	10 671.7		3558.3			
e6			FucBi(1)NA	9941.1		3314.2			
e7			FucBi(1)	9649.8		3216.5			

[a] Fuc, fucose; Bi, biantennary; Tri, triantennary; Tetra, tetraantennary; NA, N-acetylneuramic acid; Lac, N-acetyllactosamine; FucBi(1)

```
                              Fuc
          Man                  |
             \      Man—GlcNAc—GlcNAc
              \    /
Gal—GlcNAc—Man/              .
```
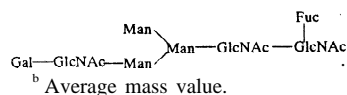
[b] Average mass value.

Fig. 6A shows the mass spectrum of peak 10. The observed mass of peak 10 corresponded to the theoretical mass of peptide K16 plus hexose (Table 2). On the basis of the result of monosaccharide analysis, peptide K16 proved to be linked by glucose (Fig. 2).

Peptide K16, from Ser261 to Arg297, contains eight potential glucosylation sites, five Ser and three Thr residues. To identify the linkage site of glucose, peak 10 was fractionated and treated with chymotrypsin (Fig. 1). Chymotrypsin digest of peak 10 is further analyzed by peptide/glycopeptide mapping.
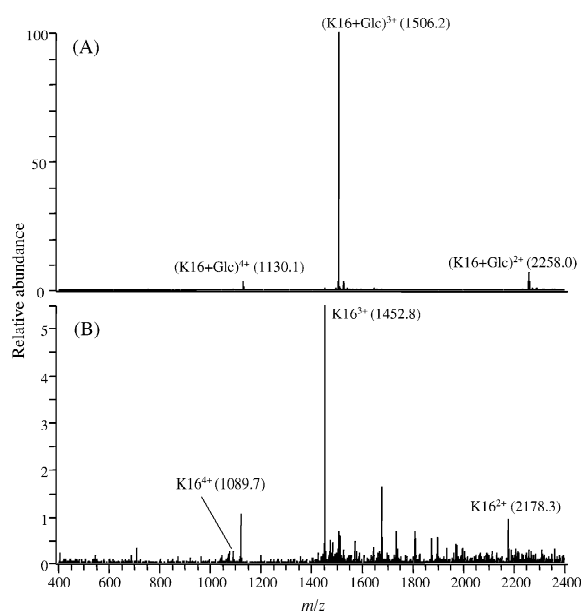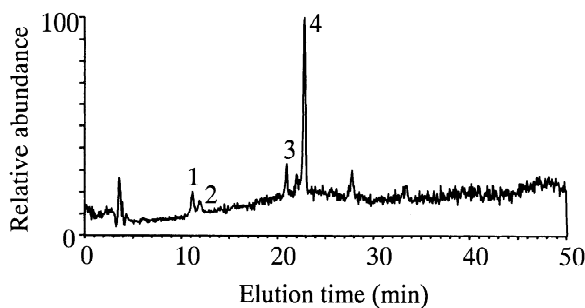
Fig. 6. Mass spectra of peaks 10 and 11 in Fig. 4; glucosylated peptide K16, peak 10 (A), and nonglucosylated peptide K16, peak 11 (B).

Fig. 7 shows a TIC chromatogram of LC–MS of peak 10 chymotrypsin digest in the positive-ion mode. Four peaks were detected, and amino acid residues of each peak were determined by comparing the experimental masses with the predicted mass

Peak No.	Peptide No.	Theoretical mass[a]	Observed mass
1	K16-3	1165.2	1164.7
2	K16-2	1234.3	1396.0
3	K16-1	1991.0	1992.0
4	K16-1 + K16-2	3369.3	3369.3

[a] Average mass value.

Fig. 7. TIC chromatogram of LC–MS of peak 10 chymotrypsin digest in the positive-ion mode.

from the amino acid sequence. These results are summarized in the table in Fig. 7. The observed masses of peaks 1, 3, and 4 in Fig. 7 are consistent with those of peptide K16-3, K16-1, and K16-1 plus K16-2, respectively, which are assumed to be obtained by chymotrypsin digestion of peptide K16. The observed mass of peak 2 in Fig. 7 is not consistent with the mass of peptide K16-2, but with that of peptide K16-2 plus glucose. Therefore, glucose is identified as being linked to Ser287 because the peptide K16-2 contains only one Ser residue, Ser287.

Signals of nonglucosylated peptide K16 ion at m/z 2178.3 (2+), 1452.8 (3+), and 1089.4 (4+) were detected in the mass spectrum of peak 11 (Fig. 6B and Table 2). The signal height ratio of glucosylated peptide K16 ion at m/z 1506.2 (3+) and nonglucosylated peptide K16 ion at 1452.8 (3+) suggested that 95% of Ser287 is glucosylated.

3.6. β-Hydroxylation

Fig. 8 shows the mass spectrum of peak 13. Signals at m/z 1306.2 (3+) and 1958.8 (2+) in Fig. 8 clearly show the β-hydroxylation occurs at peptide K19 that contains potential β-hydroxylation site, Asn324. Ions of nonhydroxylated peptide K19 were also detected at m/z 1300.6 (3+) and 1950.6 (2+) in Fig. 8. The signal height ratio of peptide K19 ion and hydroxylated peptide K19 ion suggests that β-hydroxylation occurred on about 90% of Asn324 (Fig. 8).

Another potential β-hydroxylation site, Asn439, is contained in glycopeptide K21. As shown in Fig. 5E, the mass spectrum of glycopeptide K21 is too

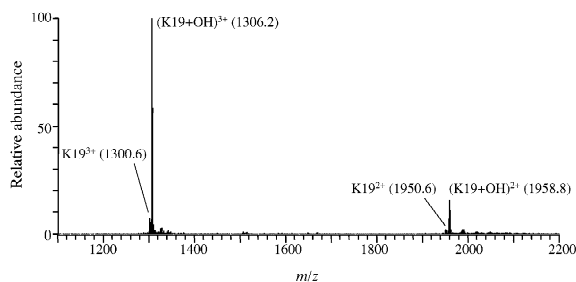
Fig. 8. Mass spectrum of β-hydroxylated peptide K19, peak 13 in Fig. 4.

complicated to clarify the hydroxylation at Asn439. In order to determine whether β-hydroxylation at Asn439 occurs, glycopeptide K21 (peak 16) was fractionated and treated with glycopeptidase A to eliminate *N*-linked glycan. The inset of Fig. 9 shows a TIC chromatogram of LC–MS of deglycosylated peptide K21 in the positive-ion mode. Two close peaks, peak X and Y, are detected. Their mass spectra are shown in Fig. 9A and B, respectively. Observed *m/z* values of peaks X and Y corresponded with theoretical *m/z* values of hydroxylated peptide K21, 2755.2 (3+) and 2066.6 (4+), and peptide K21, 2749.8 (3+), 2062.6 (4+) and 1650.3 (5+), respectively. The ratio of peptide K21 to β-hydroxylated peptide K21 was approximately 75:25. From

these results, Asn439 proved to be partially hydroxylated.

4. Discussion

In this study, we demonstrated the procedure for making a structural analysis of a glycoprotein by mass spectrometric oligosaccharide mapping and peptide/glycopeptide mapping using rhTM as a model glycoprotein. First, we analyzed the structure of both *O*- and *N*-linked glycans released from rhTM by oligosaccharide mapping using GCC-LC–MS. Both *O*- and *N*- linked glycans are well separated and eluted roughly in order of molecular mass. On
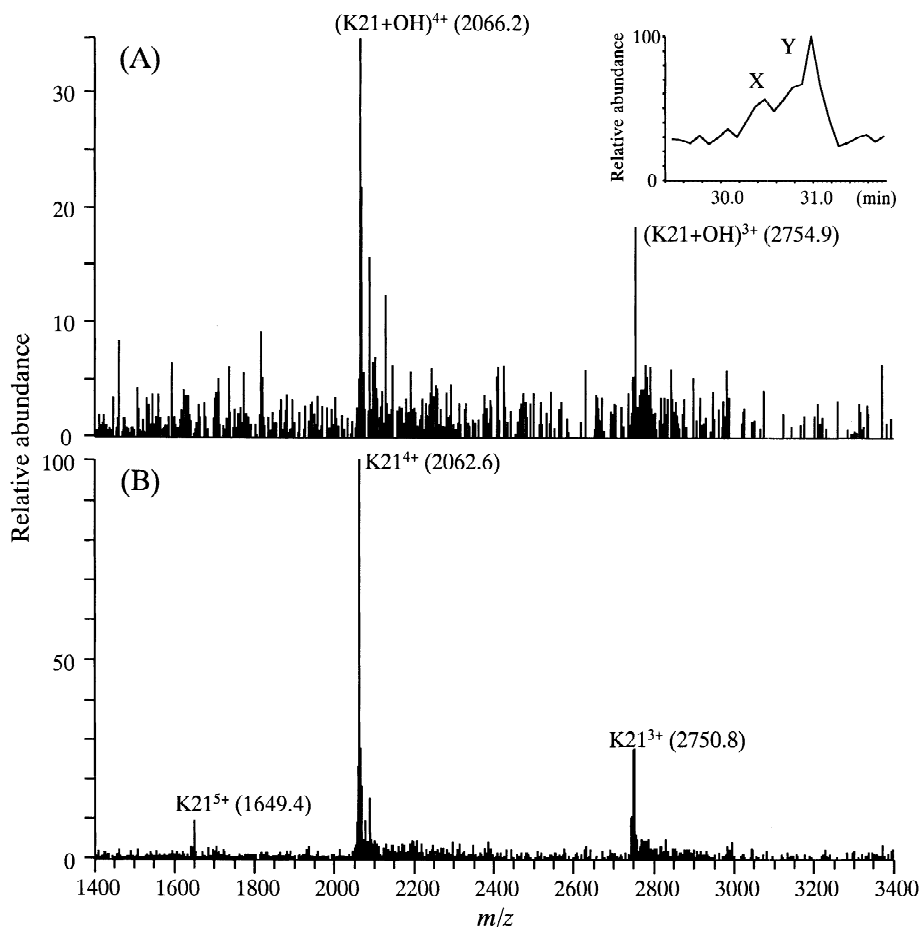

Fig. 9. Mass spectra of glycopeptide K21 after treatment with glycopeptidase A; β-hydroxylated peptide K21, peak X (A), and nonhydroxylated peptide K21, peak Y (B). The inset shows the TIC chromatogram of LC–MS of deglycosylated peptide K21.

the basis of the monosaccharide analysis and molecular mass of oligosaccharides, O-linked glycans were determined to be primarily a core 1 structure, and N-linked glycans were characterized as mainly fucosyl complex-type biantennary. In addition, we can grasp the distribution of sialylated oligosaccharides, which are known to influence the activity and metabolic fate of glycoproteins.

Next, we performed peptide/glycopeptide mapping with LC–MS to identify post-translational modification including glycosylation and hydroxylation. Five glycopeptides could be characterized by precursor-ion scanning. Precursor-ion scanning was also helpful in a case of co-elution of a glycopeptide with some other peptides (peak 14 in Fig. 4; peptide K6 and glycopeptide K1). The N-glycosylation site and site-specific glycosylation could be determined by comparing the experimental masses obtained by oligosaccharide mapping with the theoretical molecular mass of the peptide. Although O-glycosylation sites could not be determined since several O-glycosylation sites exist in the glycopeptide K23, glycans of glycopeptide K23 could be estimated by mass spectrometric analysis together with monosaccharide composition analysis.

In addition to glycosylated peptides, peaks 10, 13, and 16 were characterized to be peptide K16 with glucose and β-hydroxylated peptide K19 and K21, respectively. A glycosylation site was determined by further analysis with LC–MS after digestion of glucosylated peptide K16 with chymotrypsin. Glucosylation of Ser287 was found in rhTM expressed in CHO cells, the same as in native urinary TM [20]. In rhTM expressed in CHO cells, 95% of Ser287 was found to be glucosylated by comparing the signal abundance of peptide K16 ion with glucosylated peptide K16 ion. Two Asn residues, which are a potential β-hydroxylation site in the conserved amino acid sequence, were confirmed to be partially β-hydroxylated on the basis of their mass spectra. Signal abundances of hydroxylated and nonhydroxylated ions suggest that 90% of peptide K19 and 25% of peptide K21 are hydroxylated.

As described above, we can elucidate the carbohydrate structure, post-translational modifications including glycosylation site, and site-specific carbohydrate heterogeneity in a glycoprotein by performing mass spectrometric oligosaccharide mapping, peptide/glycopeptide mapping, and precursor-ion scanning, and comparing the obtained results. Our method can be effective for elucidating the most post-translational modifications, such as phoshorylation and sulfation, and it can be applied for the analysis of a protein in a biological sample.

References

[1] A.W. Guzzetta, L.J. Basa, W.S. Hancock, B.A. Keyt, W.F. Bennett, Anal. Chem. 65 (1993) 2953.
[2] S. Itoh, N. Kawasaki, M. Ohta, M. Hyuga, S. Hyuga, T. Hayakawa, Bull. Natl. Inst. Health Sci. 119 (2001) 65.
[3] N. Kawasaki, M. Ohta, S. Hyuga, M. Hyuga, T. Hayakawa, Anal. Biochem. 285 (2000) 82.
[4] M. Ohta, N. Kawasaki, S. Hyuga, M. Hyuga, T. Hayakawa, J. Chromatogr. A 910 (2001) 1.
[5] V. Ling, A.W. Guzzetta, E. Canova-Davis, J.T. Stults, W.S. Hancock, T.R. Covey, B.I. Shushan, Anal. Chem. 63 (1991) 2909.
[6] R.S. Annan, S.A. Carr, J. Protein Chem. 16 (1997) 391.
[7] S.A. Carr, M.J. Huddleston, M.F. Bean, Protein Sci. 2 (1993) 183.
[8] M.J. Huddleston, M.F. Bean, S.A. Carr, Anal. Chem. 65 (1993) 877.
[9] N. Kawasaki, M. Ohta, S. Hyuga, O. Hashimoto, T. Hayakawa, Anal. Biochem. 269 (1999) 297.
[10] N. Kawasaki, Y. Haishima, M. Ohta, S. Itoh, M. Hyuga, S. Hyuga, T. Hayakawa, Glycobiology 11 (2001) 1043.
[11] N. Kawasaki, M. Ohta, S. Itoh, M. Hyuga, S. Hyuga, T. Hayakawa, Biologicals 30 (2002) 113.
[12] S. Itoh, N. Kawasakai, M. Ohta, M. Hyuga, S. Hyuga, T. Hayakawa, J. Chromatogr. A 968 (2002) 89.
[13] C.T. Esmon, N.L. Esmon, K.W. Harris, J. Biol. Chem. 257 (1982) 7944.
[14] C.T. Esmon, J. Biol. Chem. 264 (1989) 4743.
[15] H. Ishii, P.W. Majerus, J. Clin. Invest. 76 (1985) 2178.
[16] J. Stenflo, A.K. Ohlin, W.G. Owen, W.J. Schneider, J. Biol. Chem. 263 (1988) 21.
[17] K. Suzuki, H. Kusumoto, Y. Deyashiki, J. Nishioka, I. Maruyama, M. Zushi, S. Kawahara, G. Honda, S. Yamamoto, S. Horiguchi, Embo. J. 6 (1987) 1891.
[18] J.F. Parkinson, C.J. Vlahos, S.C. Yan, N.U. Bang, Biochem. J. 283 (1992) 151.
[19] H. Wakabayashi, S. Natsuka, T. Mega, N. Otsuki, M. Isaji, M. Naotsuka, S. Koyama, T. Kanamori, K. Sakai, S. Hase, J. Biol. Chem. 274 (1999) 5436.
[20] H. Wakabayashi, S. Natsuka, M. Honda, M. Naotsuka, Y. Ito, J. Kajihara, S. Hase, J. Biochem. (Tokyo) 130 (2001) 543.
[21] T. Edano, N. Kumai, T. Mizoguchi, M. Ohkuchi, Int. J. Biochem. Cell. Biol. 30 (1998) 77.
[22] R.J. Harris, M.W. Spellman, Glycobiology 3 (1993) 219.
[23] J. Stenflo, A. Lundwall, B. Dahlback, Proc. Natl. Acad. Sci. USA 84 (1987) 368.
[24] D.B. Thomas, R.J. Winzler, J. Biol. Chem. 244 (1969) 5943.
[25] M.R. Hardy, R.R. Townsend, Y.C. Lee, Anal. Biochem. 170 (1988) 54.