Available online at www.sciencedirect.com
SCIENCE DIRECT
ELSEVIER
Analytical Biochemistry 348 (2006) 259–268
ANALYTICAL BIOCHEMISTRY
www.elsevier.com/locate/yabio

Site-specific N-glycosylation analysis of human plasma ceruloplasmin using liquid chromatography with electrospray ionization tandem mass spectrometry

Akira Harazono *, Nana Kawasaki, Satsuki Itoh, Noritaka Hashii, Akiko Ishii-Watabe, Toru Kawanishi, Takao Hayakawa

National Institute of Health Sciences, Division of Biological Chemistry and Biologicals, 1-18-1 Kami-yoga, Setagaya-Ku, Tokyo 158-8501, Japan

Received 8 June 2005
Available online 10 November 2005

Abstract

Ceruloplasmin has ferroxidase activity and plays an essential role in iron metabolism. In this study, a site-specific glycosylation analysis of human ceruloplasmin (CP) was carried out using reversed-phase high-performance liquid chromatography with electrospray ionization tandem mass spectrometry (LC-ESI-MS/MS). A tryptic digest of carboxymethylated CP was subjected to LC-ESI-MS/MS. Product ion spectra acquired data-dependently were used for both distinction of the glycopeptides from the peptides using the carbohydrate B-ions, such as *m/z* 204 (HexNAc) and *m/z* 366 (HexHexNAc), and identification of the peptide moiety of the glycopeptide based on the presence of the b- and y-series ions derived from the peptide. Oligosaccharide composition was deduced from the molecular weight calculated from the observed mass of the glycopeptide and theoretical mass of the peptide. Of the seven potential N-glycosylation sites, four (Asn119, Asn339, Asn378, and Asn743) were occupied by a sialylated biantennary or triantennary oligosaccharide with fucose residues (0, 1, or 2). A small amount of sialylated tetraantennary oligosaccharide was detected. Exoglycosidase digestion suggested that fucose residues were linked to reducing end GlcNAc in biantennary oligosaccharides and to reducing end and/or α1–3 to outer arms GlcNAc in triantennary oligosaccharides and that roughly one of the antennas in triantennary oligosaccharides was α2–3 sialylated and occasionally α1–3 fucosylated at GlcNAc.

Keywords: Ceruloplasmin; Glycopeptide; Liquid chromatography-electrospray tandem mass spectrometry; Product ion spectrum; Exoglycosidase digestion

Ceruloplasmin (CP)[1] is a blue copper serum glycoprotein synthesized in the liver. CP has ferroxidase activity and plays an essential role in iron metabolism [1–4]. The primary structure of human CP has been determined by amino acid sequencing, and it is composed of a single polypeptide chain of 1046 amino acid residues [5]. The amino acid sequence was confirmed from complete cDNA sequence [6]. The major oligosaccharides in human CP were reported to be sialylated bi- and triantennary structures with or without a fucose residue [7,8]. Although four N-glycosylation sites (Asn119, Asn339, Asn378, and Asn743) were identified among seven potential sites [9], the heterogeneity of oligosaccharides was still unknown at each glycosylation site. CP is an acute phase reactant, and the serum concentration increases during inflammation, infection, and trauma [10]. It is known that the patterns of glycosylation are changed by inflammatory cytokines [11]. Several studies have reported that CP is a good diagnostic marker of solid malignant tumors [12,13] and that the CP glycoform might

* Corresponding author. Fax: +81 3 3700 9084.
E-mail address: harazono@nihs.go.jp (A. Harazono).

[1] *Abbreviations used:* CP, ceruloplasmin; LC-ESI-MS, liquid chromatography with electrospray ionization mass spectrometry; Hex, hexose; HexNAc, *N*-acetylhexosamine; LC-ESI-MS/MS, liquid chromatography with electrospray ionization tandem mass spectrometry; EDTA, ethylenediaminetetraacetic acid; TFA, trifluoroacetic acid; Q-TOF, quadrupole time-of-flight; TIC, total ion chromatogram; NeuAc, *N*-acetylneuraminic acid; GlcNAc, *N*-acetylglucosamine; Fuc, fucose.

0003-2697/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.ab.2005.10.036

be a valuable supplement [12]. Thus, it is important to conduct a site-specific glycosylation analysis of normal human CP.

One of the most effective techniques for determining the site-specific carbohydrate heterogeneity of glycoproteins is the mass spectrometric peptide mapping of proteolytic fragments of glycoproteins by liquid chromatography with electrospray ionization mass spectrometry (LC-ESI-MS) [14–19]. The specific detection of glycopeptides in a complex peptide mixture is generally achieved by monitoring specific carbohydrate fragment ions such as *m/z* 204 (HexNAc) and *m/z* 366 (HexHexNAc) produced by cone voltage fragmentation or by precursor ion scanning [15–19]. Because product ion spectra of glycopeptides show high abundant carbohydrate fragment ions and low abundant b- and y-series fragment ions derived from the peptide backbone [20,21], product ion spectra acquired data-dependently in liquid chromatography with electrospray ionization tandem mass spectrometry (LC-ESI-MS/MS) can be used for both the selection from the peptides and the identification of the glycopeptides [22]. MS in combination with specific exoglycosidase digestions allows us to obtain the site-specific information on anomericity and linkage of glycans [23]. In the current study, we conducted a site-specific glycosylation analysis of human CP and successfully determined glycosylation status and glycosylation profile at each N-glycosylation site.

Materials and methods

Materials

Acetonitrile, formic acid, and guanidine hydrochloride were purchased from Wako Pure Chemicals Industries (Osaka, Japan). Purified human CP was purchased from Calbiochem (San Diego, CA, USA). Modified trypsin was purchased from Promega (Madison, WI, USA). α2–3 Neuraminidase (EC 3.2.1.18) of *Macrobdella decora*, a recombinant form, and α1–3,4 fucosidase (EC 3.2.1.51) from *Xanthomonas* sp. were purchased from Calbiochem. α2–3,6,8,9 Neuraminidase (EC 3.2.1.18) of *Arthrobacter ureafaciens*, a recombinant form, and β1–4 galactosidase (EC 3.2.1.23) were purchased from Sigma Chemical (St. Louis, MO, USA). The water used was obtained from a Milli-Q water system (Millipore, Bedford, MA, USA). All other reagents were of the highest quality available.

Reduction and S-carboxymethylation of CP

CP (100 μg) was dissolved in 270 μl of 0.5 M Tris–HCl buffer (pH 8.5) that contained 8 M guanidine hydrochloride and 5 mM ethylenediaminetetraacetic acid (EDTA). After the addition of 2 μl of 2-mercaptoethanol, the mixture was incubated for 2 h at 40 °C. Then 5.67 mg of monoiodoacetic acid was added, and the resulting mixture was incubated for 2 h at 40 °C in the dark. The reaction mixture was applied to a PD-10 column (Amersham Biosciences, Uppsala, Sweden) to remove the reagents, and the eluate was lyophilized.

Trypsin digestion of CP

Reduced and carboxymethylated CP was redissolved in 100 μl of 0.1 M Tris–HCl buffer (pH 8.0). An aliquot of 1 μl of trypsin prepared as 1 μg/μl was added to 50 μl of CP solution (1:50, w/w), and the mixture was incubated for 16 h at 37 °C. The enzyme digestion was stopped by storing at −20 °C before analysis.

HPLC of tryptic digest of CP

Tryptic digests (0.2 and 0.4 μg) of human CP were analyzed by LC-ESI-MS/MS to identify the peptides and glycopeptides, respectively. HPLC was performed on a Paradigm MS 4 (Michrome BioResources, Auburn, CA, USA) equipped with a Magic C18 column (0.2 μ, 50 mm, Michrome BioResources). The eluents consisted of water containing 2% (v/v) acetonitrile and 0.1% (v/v) formic acid (pump A) and 90% acetonitrile and 0.1% formic acid (pump B). Trypsin-digested samples were loaded onto a microtrap (peptide captrap, Michrom BioResources). After a wash with 15 μl H_2O/CH_3CN (98:2) with 0.1% trifluoroacetic acid (TFA), the trapping column was switched into line with the column. Samples were eluted with 5% of B for 10 min, followed by a linear gradient from 5 to 65% of B in 60 min at a flow rate of 2 μl/min.

ESI-Q-TOF-MS/MS

Mass spectrometric analyses were performed using a quadrupole time-of-flight (Q-TOF) mass spectrometer (QSTAR Pulsar, MDS Sciex, Toronto, Canada) equipped with a nano-electrospray ion source. The mass spectrometer was operated in the positive ion mode. The nanospray voltage was set at 2500 V. Mass spectra were acquired at *m/z* 400–2000 or *m/z* 1000–2000 for MS analysis and at *m/z* 100–2000 for MS/MS analysis. After every regular MS acquisition, two MS/MS acquisitions against top two of the multiply charged molecular ions were performed (data-dependent acquisition). The precursor ions with the same *m/z* as acquired previously were excluded for 120 s. The collision energy was varied between 30 and 80 eV depending on the size and charge of the molecular ion. Accumulation times for the spectra were 1.0 and 2.0 s for MS and MS/MS, respectively. All peaks were resolved monoisotopically.

Tandem MS/MS data from LC-ESI-MS/MS runs were submitted to the search engine Mascot to identify the tryptic peptides of CP. One missed cleavage was allowed, and tolerances of 2.0 and 0.8 u mass were used for precursor and product ions, respectively. From the data for LC-ESI-MS/MS at *m/z* 1000–2000, glycopeptide precursor ions were selected manually based on the presence of oligosaccharide oxonium ions such as *m/z* 204 (HexNAc) and *m/z* 366 (HexHexNAc). The glycopeptide ions were assigned based on

the presence of b- and y-series fragment ions of peptides of putative glycopeptides or molecular weight difference of sugar unit. The molecular weight of the carbohydrate in the glycopeptide was calculated from the molecular weights of the glycopeptide and the suggested peptide. The oligosaccharide composition and type were deduced from the molecular weight of the carbohydrate.

Oligosaccharide sequencing by exoglycosidase digestions

Trypsin in the digest of human CP was inactivated by boiling for 5 min at 100 °C. Aliquots of the digest (4 μg) were digested in a volume of 20 μl for 12 h at 37 °C in 50 mM sodium phosphate buffer (pH 5.0) using the following exoglycosidases alone or in combination: α2–3 neuraminidase, 20 mU/ml; α2–3,6,8,9 neuraminidase, 100 mU/ml; α1–3,4 fucosidase, 20 mU/ml; and β1–4 galactosidase, 30 mU/ml. Aliquots (0.08 μg) before and after exoglycosidase digestions were subjected to LC-ESI-MS at *m/z* 700 to 2000 in which MS/MS acquisition was not performed.

Results

Peptide mapping of tryptic digest of human CP (LC-ESI-MS/MS in m/z range of 400–2000)

The amino acid sequence of human CP (National Center for Biotechnology Information protein database: P00450) is shown in Fig. 1. The tryptic peptides, including potential N-glycosylation sites, are shown in bold type. Trypsin can digest human CP into seven glycopeptides containing only one potential N-glycosylation site. To determine the glycosylation state at each glycosylation site, we performed mass spectrometric peptide mapping of the tryptic digest of CP. An aliquot of 0.2 μg of the tryptic digest was analyzed by LC-ESI-MS/MS in the *m/z* range of 400–2000 (data not shown). When molecular ions with more than a single charge were detected, the product ion spectrum was acquired automatically. Peptide identification of each product ion spectrum was done using the Mascot search engine. More than 70% of the amino acid sequence was identified; identified amino acids of CP are underlined in Fig. 1. Three peptides containing the potential N-glycosylation site (Asn208, Asn569, and Asn907 [residues 197–218, 558–579, and 895–917, respectively]) were detected, whereas peptides containing the other N-glycosylation sites were not detected. Thus, Asn119, Asn339, Asn378, and Asn743 might be glycosylated.

KEKHYYIGII ETTWDYASDH GEKKLISVDT EHSNIYLQNG PDRIGRLYKK ALYLQYTDET
FRTTIEKPVW LGFLGPIIKA ETGDKVYVHL KNLASRPYTF HSHGITYYKE HEGAIYPDNT
TDFQRADDKV YPGEQYTYML LATEEQSPGE GDGNCVTRIY HSHIDAPKDI ASGLIGPLII
CKKDSLDKEK EKHIDREFVV MFSVVDENFS WYLEDNIKTY CSEPEKVDKD NEDFQESNRM
YSVNGYTFGS LPGLSMCAED RVKWYLFGMG NEVDVHAAFF HGQALTNKNY RIDTINLFPA
TLFDAYMVAQ NPGEWMLSCQ NLNHLKAGLQ AFFQVQECNK SSSKDNIRGK HVRHYYIAAE
EIIWNYAPSG IDIFTKENLT APGSDSAVFF EQGTTRIGGS YKKLVYREYT DASFTNRKER
GPEEEHLGIL GPVIWAEVGD TIRVTFHNKG AYPLSIEPIG VRFNKNNEGT YYSPNYNPQS
RSVPPSASHV APTETFTYEW TVPKEVGPTN ADPVCLAKMY YSAVDPTKDI FTGLIGPMKI
CKKGSLHANG RQKDVDKEFY LFPTVFDENE SLLLEDNIRM FTTAPDQVDK EDEDFQESNK
MHSMNGFMYG NQPGLTMCKG DSVVWYLFSA GNEADVHGIY FSGNTYLWRG ERRDTANLFP
QTSLTLHMWP DTEGTFNVEC LTTDHYTGGM KQKYTVNQCR RQSEDSTFYL GERTYYIAAV
EVEWDYSPQR EWEKELHHLQ EQNVSNAFLD KGEFYIGSKY KKVVYRQYTD STFRVPVERK
AEEEHLGILG PQLHADVGDK VKIIFKNMAT RPYSIHAHGV QTESSTVTPT LPGETLTYVW
KIPERSGAGT EDSACIPWAY YSTVDQVKDL YSGLIGPLIV CRRPYLKVFN PRRKLEFALL
FLVFDENESW YLDDNIKTYS DHPEKVNKDD EEFIESNKMH AINGRMFGNL QGLTMHVGDE
VNWYLMGMGN EIDLHTVHFH GHSFQYKHRG VYSSDVFDIF PGTYQTLEMF PRTPGIWLLH
CHVTDHIHAG METTYTVLQN EDTKSG

Fig. 1. Primary amino acid sequence of human CP (P00450). The tryptic peptides, including potential N-glycosylation sites, are shown in bold type. Tryptic peptides identified in the LC-ESI-MS/MS analysis are underlined. Cysteine residues are carboxymethylated. Identified N-glycosylation sites are indicated by arrow.

Glycosylation analysis of human CP (LC-ESI-MS/MS in the m/z range of 1000–2000)

N-glycosylated peptides have relatively high molecular weights due to their oligosaccharide moiety. Because ions at lower *m/z* values can be detected in the *m/z* range of 400–2000, glycopeptide ions with higher *m/z* values might be missed to obtain product ion spectra. To detect glycopeptide ions preferentially, another LC-ESI-MS/MS analysis was carried out in the *m/z* range of 1000–2000 using an aliquot of 0.4 μg of the tryptic digest. Fig. 2A shows a total ion chromatogram (TIC) of a TOF-MS scan for the full scan *m/z* 1000–2000. Fig. 2B shows a TIC of the product ion scan. Because product ion spectra of glycopeptide precursor ions have abundant carbohydrate B-ions, *m/z* 204 (HexNAc), *m/z* 186 (HexNAc-H_2O), *m/z* 366 (HexHexNAc), and *m/z* 292 (NeuAc), the extracted ion chromatogram at *m/z* 204.05–204.15 (HexNAc, 204.08) of the product ion scan is illustrated in Fig. 2C. The extracted ion chromatogram at *m/z* 204.05–204.15 of product ion spectra provides a useful indication of the selection of glycopeptide precursor ions. The glycopeptide ions were assigned based on an examination of product ion spectra using the information on amino acid sequences of the peptides containing a putative N-glycosylation site.

Identification of Asn119 glycopeptide

The product ion spectrum of 1366.6 (+3) at 26 min, labeled by A in Fig. 2C, is shown in Fig. 3A. There were abundant oligosaccharide oxonium ions such as *m/z* 204 (HexNAc), *m/z* 366 (HexHexNAc), *m/z* 186 (HexNAc-H_2O), *m/z* 168 (HexNAc-$2H_2O$), *m/z* 274 (NeuAc-H_2O), and *m/z* 292 (NeuAc). Thus, this precursor ion was assigned as a glycopeptide. Several fragment ions consistent with b- and y-series fragment ions [24] derived from the peptide EHEGAIYPDN^{119}TTDFQR (residues 110–125) were detected together with several deamidated (−17) or dehydrogenated (−18) b- and y-series ions and y-series ions with the GlcNAc residue. Thus, the peptide moiety EHEGA IYPDNTTDFQR was suggested. The carbohydrate's molecular weight, 2223.0, was calculated by subtracting the theoretical molecular weight of the peptide (1891.8) from

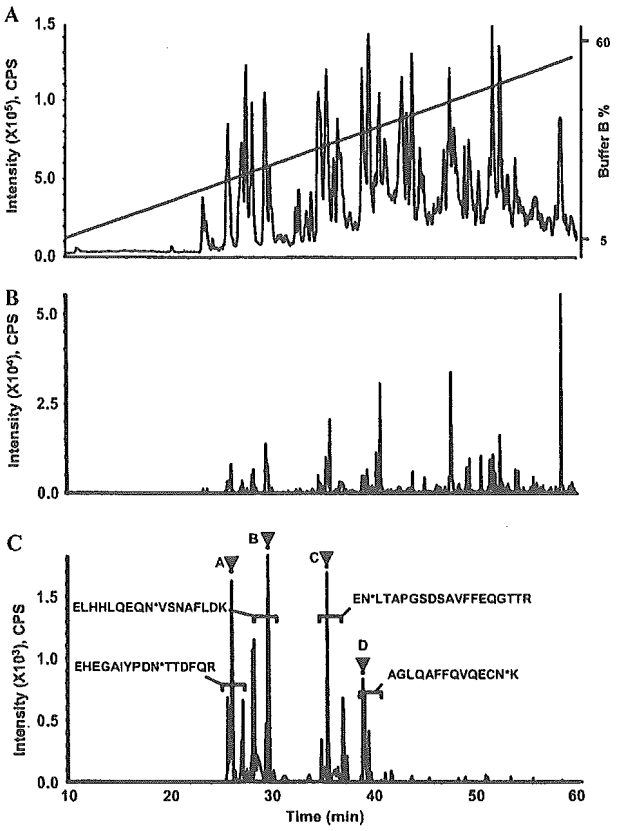

Fig. 2. LC-ESI-MS/MS in the *m/z* range of 1000–2000 of the tryptic digest of human CP. (A) TIC of the TOF-MS scan for the full-scan *m/z* 1000–2000 and the HPLC gradient. (B) TIC of the product ion scan acquired data-dependently. (C) Extracted ion chromatograph at *m/z* 204.05–204.15 of the product ion spectra. Brackets denote glycopeptide fraction and peptide sequences of the glycopeptides. Product ion spectra indicated by A–D are shown in Fig. 3.

the calculated molecular weight of the glycopeptide (4096.7) and adding the molecular weight of H_2O (18.0). The presence of product ions at *m/z* 274 (NeuAc-H_2O) and *m/z* 292 (NeuAc) suggested sialylation of the oligosaccharide. Thus, the carbohydrate's composition, $[HexNAc]_4[Hex]_5[NeuAc]_2$, was deduced.

Identification of Asn743 glycopeptide

The product ion spectrum of 1628.4 (+3) at 29 min, labeled by B in Fig. 2C, is shown in Fig. 3B. This precursor ion was assigned as a glycopeptide due to the presence of abundant oligosaccharide oxonium ions such as *m/z* 204 (HexNAc), *m/z* 366 (HexHexNAc), and *m/z* 292 (NeuAc) in the product ion spectrum. Several fragment ions were consistent with theoretical b- and y-series fragment ions derived from the peptide ELHHLQEQN^{743}VSNAFLDK (residues 735–751). Doubly charged ions of peptide (*m/z* 1011.7), peptide + HexNAc (*m/z* 1113.1), peptide + 2HexNAc (*m/z* 1214.6), peptide + 2HexNAc + Hex (*m/z* 1295.5), peptide + 2HexNAc + 2Hex (*m/z* 1376.7), and peptide + 2HexNAc + 3Hex (*m/z* 1457.5) showed the sequential fragmentation of the pentasaccharide carbohydrate core. The carbohydrate's molecular weight, 2879.1, was calculated from the theoretical molecular weight of the peptide (2021.0) and the calculated molecular weight of the glycopeptide (4882.1). The carbohydrate's composition, $[HexNAc]_5[Hex]_6[NeuAc]_3$, was deduced from the molecular weight.

Identification of Asn378 glycopeptide

The product ion spectrum of 1444.6 (+3) at 35 min, labeled by C in Fig. 2C, is shown in Fig. 3C. Abundant oligosaccharide oxonium ions were detected, as were several fragment ions consistent with b- and y-series fragment ions derived from the peptide EN^{378}LTAPGSDSAVFFEQGT TR (residues 377–391). The carbohydrate's molecular weight, 2222.9, was calculated from the theoretical molecular weight of the peptide (2126.0) and the calculated molecular weight of the glycopeptide (4330.9). Thus, the peptide moiety ENLTAPGSDSAVFFEQGTTR and the carbohydrate's composition, $[HexNAc]_4[Hex]_5[NeuAc]_2$, were suggested.

Identification of Asn339 glycopeptide

The product ion spectrum of 1282.6 (+3) at 39 min, labeled by D in Fig. 2C, is shown in Fig. 3C. The spectrum contains abundant oligosaccharide oxonium ions, and several fragment ions consistent with b- and y-series fragment ions derived from the peptide AGLQAFFQVQECN^{339}K (residues 327–340) were detected. The product ion spectrum contains the ions of the peptide (*m/z* 1640.8) and peptide + HexNAc (*m/z* 1843.9) and several y-series fragment ions of the peptide with a GlcNAc residue. The carbohydrate's molecular weight, 2223.0, was calculated from the theoretical molecular weight of the peptide (1639.7) and the calculated molecular weight of the glycopeptide (3844.7). Thus, the peptide moiety AGLQAFFQVQECNK and the carbohydrate's composition, $[HexNAc]_4[Hex]_5[NeuAc]_2$, were suggested.

Heterogeneity of oligosaccharides at each glycosylation site

Glycopeptides with the potential N-glycosylation sites Asn119, Asn339, Asn378, and Asn743 were detected, whereas no glycopeptides containing the other sites (Asn208, Asn569, and Asn907) could be detected in this LC-ESI-MS/MS analysis. These findings suggest that Asn119, Asn339, Asn378, and Asn743 of human CP are glycosylated and that Asn208, Asn569, and Asn907 are not. Once a glycopeptide was identified, the other glycopeptides with the same peptide could be easily assigned because they were eluted at a similar retention time in the order of the number of NeuAc and had similar product ion spectra and molecular weight difference of sugar units. The oligosaccharide heterogeneity at each four N-glycosylation sites was determined by mass spectrum. For a representative example, the mass spectrum of the glycopeptides containing

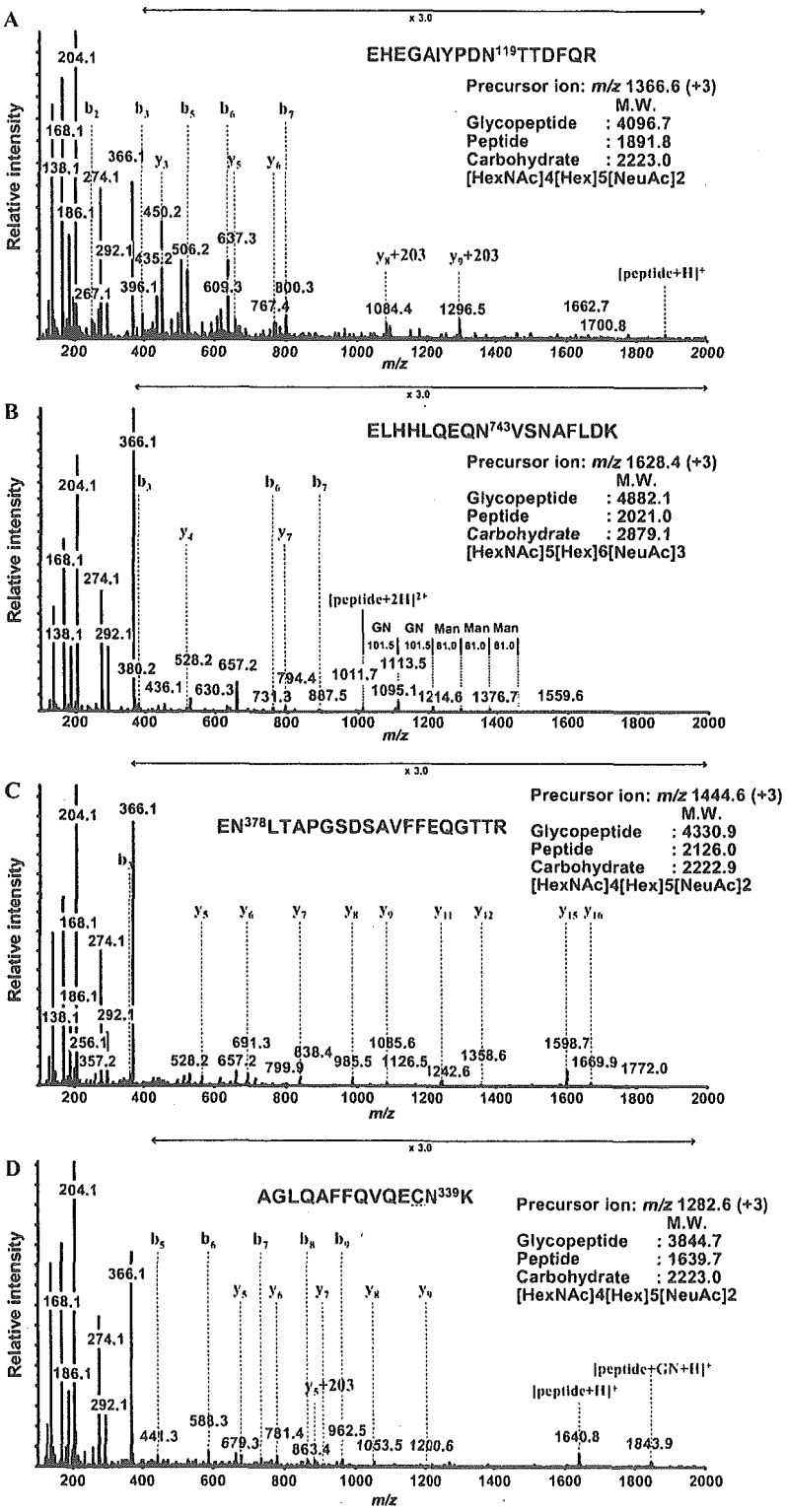

Fig. 3. Product ion spectra of *m/z* 1366.6 (+3) at 26 min (A), *m/z* 1628.4 (+3) at 29 min (B), *m/z* 1444.6 (+3) at 35 min (C), and *m/z* 1282.6 (+3) at 39 min (D) labeled by A, B, C, and D, respectively, in Fig. 2C. These spectra show abundant carbohydrate-derived ions at *m/z* 168 (HexNAc-2H_2O), *m/z* 186 (HexNAc-H_2O), *m/z* 204 (HexNAc), *m/z* 366 (HexHexNAc), *m/z* 274 (NeuAc-H_2O), and *m/z* 292 (NeuAc). The b- and y-series fragment ions [24] derived from the peptide moiety were observed. The molecular weights of the oligosaccharide were calculated from the molecular weights of the glycopeptide and peptide, and the deduced oligosaccharide composition is presented. Cystein residue was carboxymethylated.

Asn743 at 27.5 to 31.5 min is shown Fig. 4. The results of glycosylation analysis are summarized in Table 1. Deduced compositions of the oligosaccharides are estimated based on the calculated molecular weights of the oligosaccharides. Relative peak intensity was calculated by comparing triply charged glycopeptide ions. All glycosylation sites were occupied by at least three kinds of oligosaccharides, namely disialobiantennary structures ($[HexNAc]_4[Hex]_5[NeuAc]_2$), disialobiantennary structures with fucose ($[HexNAc]_4[Hex]_5[NeuAc]_2[Fuc]_1$), and trisialotriantennary structures ($[HexNAc]_5[Hex]_6[NeuAc]_3$). Trisialotriantennary structures with one fucose or two fucoses ($[HexNAc]_5[Hex]_6[NeuAc]_3[Fuc]_{1-2}$) were also detected at Asn119 and Asn743; furthermore, tetrasialotetraantennary structures with no fucose or one fucose ($[HexNAc]_6[Hex]_7[NeuAc]_4[Fuc]_{0-1}$) were detected at Asn743.

Linkage analysis of oligosaccharides by exoglycosydase digestion

To elucidate the oligosaccharide structure in terms of sequence and linkage, aliquots of the tryptic digest were further digested with exoglycosidases. As a representative example, Fig. 5 shows integrated mass spectra during the periods at which Asn119 glycopeptides were eluted in LC-ESI-MS analyses before and after digestion with exoglycosidase arrays. Treatment with α2–3 neuraminidase removed one NeuAc residue from most of the triantennary structures ($[HexNAc]_5[Hex]_6[NeuAc]_3[Fuc]_{0-2}$) and a small amount of biantennary structures ($[HexNAc]_4[Hex]_5[NeuAc]_2[Fuc]_{0-1}$) (Fig. 5B). A minor amount of triantennary structures removed two NeuAc residues. Thus, it appears that most triantennary structures contain one α2–3-linked NeuAc. Treatment with α2–3 neuraminidase + β1–4 galactosidase removed all terminal galactose residues from the desialylated glycans without fucose residues but only partially digested terminal galactoses from the desialylated glycans with fucoses (Fig. 5C). The addition of α1–3,4 fucosidase to α2–3 neuraminidase + β1–4 galactosidase treatment completely digested the remaining terminal galactose by releasing one fucose and one galactose (Fig. 5D). Thus, galactose residues are linked β1–4 to GlcNAc, and undigestion of terminal galactose by β1–4 galactosidase is due to attachment of fucose [25,26]. Because galactose was linked to GlcNAc in the β1–4 position, the fucose removed with α1–3,4 fucosidase may be linked α1–3 to GlcNAc but not α1–4 to GlcNAc. These data strongly suggested that sialyl Lewis X structure was present in human CP. Sialyl Lewis X structure was present predominantly in triantennary oligosaccharides, but a small amount seemed to be present in biantennary oligosaccharides as well. The remaining fucose residue may be linked α1–6 to reducing end GlcNAc (core fucose).

Fig. 6 shows integrated mass spectra of Asn119, Asn743, Asn378, and Asn339 glycopeptides in LC-ESI-MS analysis following digestion with α2–3,6,8,9 neuraminidase + β1–4 galactosidase. Treatment with α2–3,6,8,9 neuraminidase + β1–4 galactosidase removed all NeuAc and then removed terminal galactoses in the outer arms without fucose. Thus, this treatment could differentiate glycoforms based on the location of fucose residues. Fucosylation occurred predominantly at reducing end GlcNAc in biantennary oligosaccharides and occurred at reducing end GlcNAc and/or outer arm GlcNAc in triantennary oligosaccharides. Mass spectra of Asn119 and Asn743 glycopeptides showed higher oligosaccharide heterogeneity, and a minor amount of tetraantennary glycans could be detected. The glycosylation profile

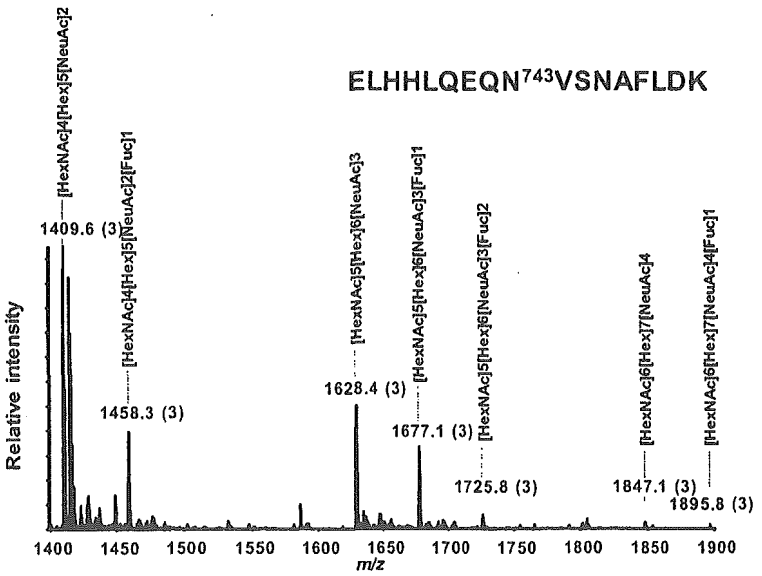

Fig. 4. Mass spectrum of the glycopeptides containing Asn743 eluting at 27.5–31.5 min from Fig. 2A. Deduced composition of the oligosaccharides is indicated based on the molecular weights of the oligosaccharides.

Table 1
Results of site-specific glycosylation analysis of human CP

Retention time (min)	Glycopeptides		Calculated MW	Relative peak intensity[a] (%)	Peptide		Oligosaccharide		
	m/z	Charge			Sequence	Theoretical MW	Calculated MW	Theoretical MW	Composition[b,c]
26	1415.3	+3	4242.8	52	EHEGAIYPDN^{119}TTDFQR	1891.8	2369.0	2368.8	[HexNAc]4[Hex]5[NeuAc]2[Fuc]1
26	1366.6	+3	4096.7	100	EHEGAIYPDN^{119}TTDFQR	1891.8	2223.0	2222.8	[HexNAc]4[Hex]5[NeuAc]2
27	1682.7	+3	5045.1	6	EHEGAIYPDN^{119}TTDFQR	1891.8	3171.3	3171.1	[HexNAc]5[Hex]6[NeuAc]3[Fuc]2
27	1634.0	+3	4899.0	21	EHEGAIYPDN^{119}TTDFQR	1891.8	3025.2	3025.1	[HexNAc]5[Hex]6[NeuAc]3[Fuc]1
27	1225.8	+4	4899.0						
27	1585.3	+3	4753.0	24	EHEGAIYPDN^{119}TTDFQR	1891.8	2879.2	2879.0	[HexNAc]5[Hex]6[NeuAc]3
27	1189.3	+4	4753.0						
28	1458.3	+3	4372.0	35	ELHHLQEQN^{743}VSNAFLDK	2021.0	2369.0	2368.8	[HexNAc]4[Hex]5[NeuAc]2[Fuc]1
28	1409.6	+3	4225.9	100	ELHHLQEQN^{743}VSNAFLDK	2021.0	2222.9	2222.8	[HexNAc]4[Hex]5[NeuAc]2
28	1057.5	+4	4225.9						
29	1725.8	+3	5174.5	5	ELHHLQEQN^{743}VSNAFLDK	2021.0	3171.5	3171.1	[HexNAc]5[Hex]6[NeuAc]3[Fuc]2
29	1294.6	+4	5174.2						
29	1677.1	+3	5028.2	29	ELHHLQEQN^{743}VSNAFLDK	2021.0	3025.3	3025.1	[HexNAc]5[Hex]6[NeuAc]3[Fuc]1
29	1258.1	+4	5028.2						
29	1628.4	+3	4882.1	43	ELHHLQEQN^{743}VSNAFLDK	2021.0	2879.1	2879.0	[HexNAc]5[Hex]6[NeuAc]3
29	1221.5	+4	4882.1						
31[d]	1895.8	+3	5684.4	2	ELHHLQEQN^{743}VSNAFLDK	2021.0	3681.4	3681.3	[HexNAc]6[Hex]7[NeuAc]4[Fuc]1
31[d]	1422.1	+4	5684.3						
31[d]	1847.1	+3	5538.4	3	ELHHLQEQN^{743}VSNAFLDK	2021.0	3535.4	3535.2	[HexNAc]6[Hex]7[NeuAc]4
31	1385.6	+4	5538.3						
35[d]	1493.3	+3	4477.0	6	EN^{378}LTAPGSDSAVFFEQGTTR	2126.0	2369.0	2368.8	[HexNAc]4[Hex]5[NeuAc]2[Fuc]1
35	1444.6	+3	4330.9	100	EN^{378}LTAPGSDSAVFFEQGTTR	2126.0	2222.9	2222.8	[HexNAc]4[Hex]5[NeuAc]2
37	1712.1	+3	5133.2	8	EN^{378}LTAPGSDSAVFFEQGTTR	2126.0	3025.2	3025.1	[HexNAc]5[Hex]6[NeuAc]3[Fuc]1
37	1284.3	+4	5133.2						
37	1663.4	+3	4987.1	23	EN^{378}LTAPGSDSAVFFEQGTTR	2126.0	2879.2	2879.0	[HexNAc]5[Hex]6[NeuAc]3
37	1247.8	+4	4987.2						
39	1331.3	+3	3990.8	14	AGLQAFFQVQECN^{339}K	1639.7	2369.1	2368.8	[HexNAc]4[Hex]5[NeuAc]2[Fuc]1
39	1923.4	+2	3844.7		AGLQAFFQVQECN^{339}K	1639.7	2223.0	2222.8	[HexNAc]4[Hex]5[NeuAc]2
39	1282.6	+3	3844.7	100					
41	1501.3	+3	4500.8	6	AGLQAFFQVQECN^{339}K	1639.7	2879.2	2879.0	[HexNAc]5[Hex]6[NeuAc]3

Note. All masses are monoisotopic. Cysteine residue was carboxymethylated.

[a] Relative peak intensity was calculated by comparing same charge state glycopeptide ions. The intensity of the glycoform with maximum at each glycosylation site was taken as 100%.

[b] The oligosaccharide composition was deduced from the molecular weight of the oligosaccharide.

[c] The glycopeptide ions adducted by NH_4^+ or Na^+ were excluded.

[d] Product ion spectra of these molecular ions were not acquired. However, these were considered glycopeptides because of a molecular weight difference of 146 (Fuc) and the same retention time as other glycopeptides.

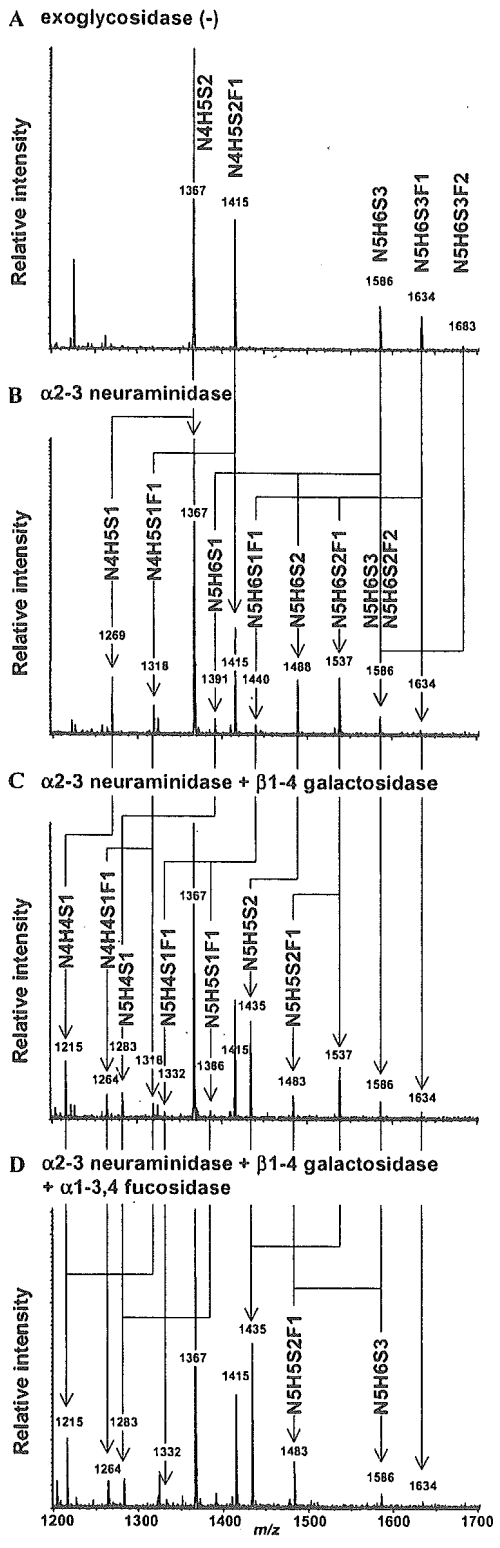

Fig. 5. LC-ESI mass spectra of the glycopeptides containing Asn119 digested with the following exoglycosidases: (A) exoglycosidase (–); (B) α2–3 neuraminidase; (C) α2–3 neuraminidase + β1–4 galactosidase; (D) α2–3 neuraminidase + β1–4 galactosidase + α1–3,4 fucosidase. Arrows between panels A and B, panels B and C, and panels C and D correspond to the digestion of NeuAc, Gal, and Gal+Fuc, respectively. H, hexose; N, *N*-acetylhexosamine; F, fucose; S, *N*-acetylneuraminic acid.

of Asn378 glycopeptides showed lower core fucosylation, and that of Asn339 glycopeptides showed lower branching. These glycosylation profiles provided the heterogeneity of fucose linkage and the number of arms at each glycosylationsite in human CP.

Discussion

A site-specific glycosylation analysis of human CP was conducted using LC-ESI-MS/MS, where product ion spectra were acquired in a data-dependent manner. The collision energy for the product ion scan was adjusted from 30 to 80 eV depending on the size and charge of the precursor ion. Under these conditions, peptide precursor ions were degraded and produced b- and y-series fragment ions derived from the amino acid sequence. Glycopeptide precursor ions produced abundant carbohydrate ions (*m/z* 204, 186, 168, and 366) together with several low intensity b- and y-series fragment ions derived from the amino acid sequence [20,21]. Thus, product ion spectra of glycopeptides are readily distinguishable from those of peptides by such carbohydrate marker ions, and the peptide moiety in the glycopeptide could be deduced from the product ions that were consistent with the expected fragment ions derived from the peptide containing the N-glycosylation site. It is known that the glycopeptide ions are more labile than peptide ions and produce consecutive monosaccharide/polysaccharide losses at much lower collision energy, and this would provide information about branching and fucose location [18]. However, we used relatively high collision energy in this site-specific glycosylation analysis to identify the peptide ions in parallel with the detection and identification of the glycopeptide ions.

Protein coverage of more than 70% in human CP was obtained in the LC-ESI-MS/MS analysis with the *m/z* range of 400–2000 (for peptide mapping). The heterogeneity at four potential N-glycosylation sites was determined in the *m/z* range of 1000–2000 (glycosylation analysis). We could detect all of the potential glycosylation sites as either glycopeptides or nonglycosylated peptides. Peptides containing the potential N-glycosylation site Asn208, Asn569, or Asn907 were detected in nonglycosylated but not glycosylated forms. Peptides with the potential N-glycosylation site Asn119, Asn339, Asn378 or Asn743 were detected in glycosylated but not nonglycosylated forms. These findings indicate that Asn119, Asn339, Asn378, and Asn743 of human CP are glycosylated and that Asn208, Asn569, and Asn907 are not. Human CP was reported to have no O-linked glycosylation [8]. No information on O-glycosylation was obtained from this analysis. These results are consistent with a previous study determining the glycosylation sites of human CP [9].

Heterogeneity of oligosaccharides was determined at each of four glycosylation sites. Disialobiantennary structures with no fucose or one fucose ($[HexNAc]_4$ $[Hex]_5$ $[NeuAc]_2[Fuc]_{0-1}$) and trisialotriantennary structures

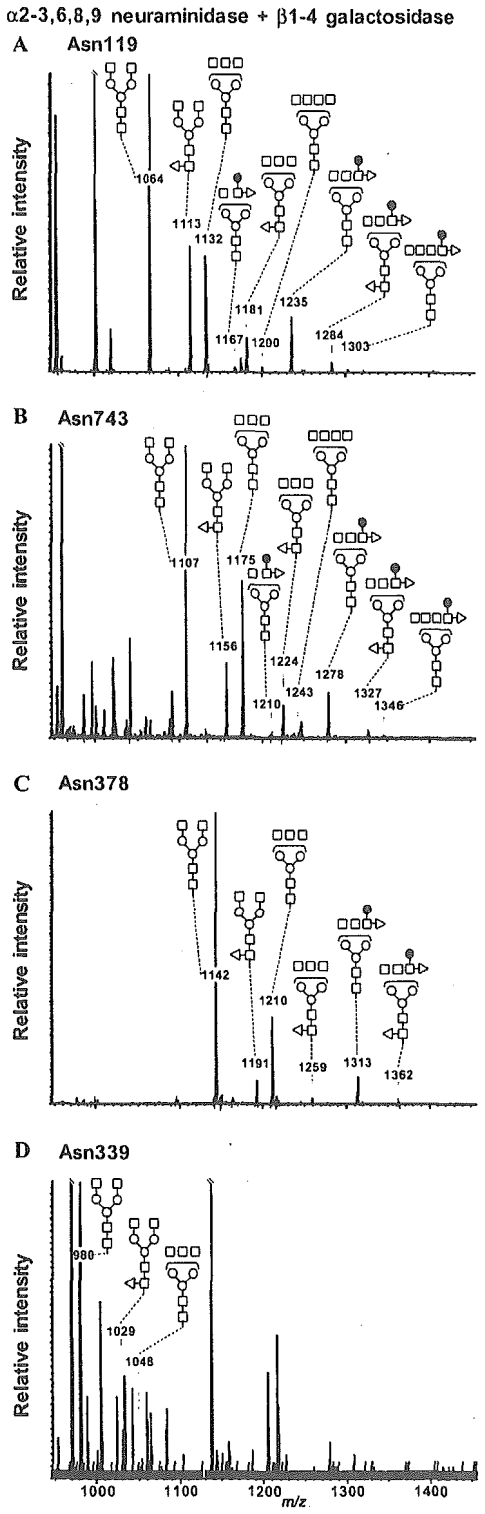

Fig. 6. LC-ESI mass spectra of the glycopeptides containing Asn119 (A), Asn743 (B), Asn378 (C), and Asn339 (D) after digestion with α2–3,6,8,9 neuraminidase + β1–4 galactosidase. Glycosylation profiles showed different degrees of branching and fucosylation at core GlcNAc and outer arm GlcNAc between glycosylation sites. Open circles, mannose; closed circles, galactose; open squares, *N*-acetyl glucosamine; open triangles, fucose.

($[HexNAc]_5[Hex]_6[NeuAc]_3$) were observed at all sites. These dominant oligosaccharides were consistent with structures published previously [7,8]. Furthermore, we detected trisialotriantennary structures with one fucose ($[HexNAc]_5[Hex]_6[NeuAc]_3[Fuc]_1$) at Asn119, Asn378, and Asn743, trisialotriantennary structures with two fucoses ($[HexNAc]_5[Hex]_6[NeuAc]_3[Fuc]_2$) at Asn119 and Asn743, and tetrasialotetraantennary structures with no fucose or one fucose ($[HexNAc]_6[Hex]_7[NeuAc]_4[Fuc]_{0-1}$) at Asn743.

To determine the linkage of fucose and NeuAc, exoglycosidase digestions were performed. Treatment with α2–3 neuraminidase suggested that roughly one antenna of triantennary glycans was linked by NeuAc in the α2–3 position. This is consistent with the previous findings that NeuAc is linked α2–3 to the Galβ1–4GlcNAcβ1–4Manα1–3Manβ1–4GlcNAcβ1–4GlcNAc group in the triantennary glycan in human CP [7,8]. Results from α2–3 neuraminidase + β1–4 galactosidase treatments with or without α1–3,4 fucosidase suggested that fucose residues were linked to reducing end GlcNAc and/or outer arm GlcNAc in the α1–3 position in the antenna where NeuAc is linked to galactose in the α2–3 position. These findings indicated that human CP contains a certain amount of sialyl Lewis X structure in triantennary glycans. Treatment with α2–3,6,8,9 neuraminidase + β1–4 galactosidase reveals the heterogeneity of the location of fucosylation as well as the number of arms. Although relative peak intensity does not express the relative amount of each glycan due to the different ionization efficiencies, the mass spectra showed the difference in fucosylation pattern and number of arms among sites.

No asialo oligosaccharides were detected in this analysis. It is known that desialylated CP is rapidly cleared from the circulation by the asialoglycoprotein receptor within the parenchymal cells of liver [27,28]. It is possible that desialylated CP might be cleared immediately by the liver.

Although the N-linked carbohydrate structures linked to human CP have been studied, only a few carbohydrate structures have been reported and site-specific characterization of these oligosaccharides has not been described. To determine the glycosylation state at each glycosylation site, the tryptic digest was examined by LC-ESI-MS/MS, where product ion spectra were acquired data-dependently. Glycopeptide ions were assigned based on the product ion spectra. Fucose and NeuAc linkages were determined by exoglycosidase digestions. Our data successfully provided comprehensive information on the site-specific N-linked oligosaccharides in human CP. This method is a powerful technique for elucidating the glycosylation of a biological sample.

Acknowledgments

This study was supported by a Grant-in-Aid for Research on Health Sciences focusing on Drug Innovation from the Japan Health Sciences Foundation.

References

[1] S. Osaki, D.A. Johnson, E. Frieden, The possible significance of the ferrous oxidase activity of ceruloplasmin in normal human serum, J. Biol. Chem. 241 (1966) 2746–2751.

[2] K. Yoshida, K. Furihata, S. Takeda, A. Nakamura, K. Yamamoto, H. Morita, S. Hiyamuta, S. Ikeda, N. Shimizu, N. Yanagisawa, A mutation in the ceruloplasmin gene is associated with systemic hemosiderosis in humans, Nat. Genet. 9 (1995) 267–272.

[3] Z.L. Harris, Y. Takahashi, H. Miyajima, M. Serizawa, R.T. MacGillivray, J.D. Gitlin, Aceruloplasminemia: molecular characterization of this disorder of iron metabolism, Proc. Natl. Acad. Sci. USA 92 (1995) 2539–2543.

[4] Z.L. Harris, A.P. Durley, T.K. Man, J.D. Gitlin, Targeted gene disruption reveals an essential role for ceruloplasmin in cellular iron efflux, Proc. Natl. Acad. Sci. USA 96 (1996) 10812–10817.

[5] N. Takahashi, T.L. Ortel, F.W. Putnam, Single-chain structure of human ceruloplasmin: the complete amino acid sequence of the whole molecule, Proc. Natl. Acad. Sci. USA 81 (1984) 390–394.

[6] M.L. Koschinsky, W.D. Funk, B.A. van Oost, R.T. MacGillivray, Complete cDNA sequence of human preceruloplasmin, Proc. Natl. Acad. Sci. USA 83 (1986) 5086–5090.

[7] K. Yamashita, C.J. Liang, S. Funakoshi, A. Kobata, Structural studies of asparagine-linked sugar chains of human ceruloplasmin. Structural characteristics of the triantennary complex type sugar chains of human plasma glycoproteins, J. Biol. Chem. 256 (1981) 1283–1289.

[8] M. Endo, K. Suzuki, K. Schmid, B. Fournet, Y. Karamanos, J. Montreuil, L. Dorland, H. van Halbeek, J.F. Vliegenthart, The structures and microheterogeneity of the carbohydrate chains of human plasma ceruloplasmin: a study employing 500-MHz ^{1}H-NMR spectroscopy, J. Biol. Chem. 257 (1982) 8755–8760.

[9] N. Takahashi, Y. Takahashi, T.L. Ortel, J.N. Lozier, N. Ishioka, F.W. Putnam, Purification of glycopeptides of human plasma proteins by high-performance liquid chromatography, J. Chromatogr. 317 (1984) 11–26.

[10] R.J. Cousins, Absorption, transport, and hepatic metabolism of copper and zinc: special reference to metallothionein and ceruloplasmin, Physiol. Rev. 65 (1985) 238–309.

[11] A. Mackiewicz, M.K. Ganapathi, D. Schultz, I. Kushner, Monokines regulate glycosylation of acute-phase proteins, J. Exp. Med. 166 (1987) 253–258.

[12] J.E. Hansen, J. Iversen, A. Lihme, T.C. Bog-Hansen, Acute phase reaction, heterogeneity, and microheterogeneity of serum proteins as nonspecific tumor markers in lung cancer, Cancer 60 (1987) 1630–1635.

[13] A. Senra Varela, J.J. Lopez Saez, D. Quintela Senra, Serum ceruloplasmin as a diagnostic marker of cancer, Cancer Lett. 121 (1997) 139–145.

[14] V. Ling, A.W. Guzzetta, E. Canova-Davis, J.T. Stults, W.S. Hancock, T.R. Covey, B.I. Shushan, Characterization of the tryptic map of recombinant DNA derived tissue plasminogen activator by high-performance liquid chromatography-electrospray ionization mass spectrometry, Anal. Chem. 63 (1991) 2909–2915.

[15] S.A. Carr, M.J. Huddleston, M.F. Bean, Selective identification and differentiation of N- and O-linked oligosaccharides in glycoproteins by liquid chromatography–mass spectrometry, Protein Sci. 2 (1993) 183–196.

[16] M.J. Huddleston, M.F. Bean, S.A. Carr, Collisional fragmentation of glycopeptides by electrospray ionization LC/MS and LC/MS/MS: methods for selective detection of glycopeptides in protein digests, Anal. Chem. 65 (1993) 877–884.

[17] P.A. Schindler, C.A. Settineri, X. Collet, C.J. Fielding, A.L. Burlingame, Site-specific detection and structural characterization of the glycosylation of human plasma proteins lecithin:cholesterol acyltransferase and apolipoprotein D using HPLC/electrospray mass spectrometry and sequential glycosidase digestion, Protein Sci. 4 (1995) 791–803.

[18] M.A. Ritchie, A.C. Gill, M.J. Deery, K. Lilley, Precursor ion scanning for detection and structural characterization of heterogeneous glycopeptide mixtures, J. Am. Soc. Mass Spectrom. 13 (2002) 1065–1077.

[19] F. Wang, A. Nakouzi, R.H. Angeletti, A. Casadevall, Site-specific characterization of the N-linked oligosaccharides of a murine immunoglobulin M by high-performance liquid chromatography/electrospray mass spectrometry, Anal. Biochem. 314 (2003) 266–280.

[20] J.F. Nemeth, G.P. Hochgesang Jr., L.J. Marnett, R.M. Caprioli, Characterization of the glycosylation sites in cyclooxygenase-2 using mass spectrometry, Biochemistry 40 (2001) 3109–3116.

[21] O. Krokhin, W. Ens, K.G. Standing, J. Wilkins, H. Perreault, Site-specific N-glycosylation analysis: matrix-assisted laser desorption/ionization quadrupole–quadrupole time-of-flight tandem mass spectral signatures for recognition and identification of glycopeptides, Rapid Commun. Mass Spectrom. 18 (2004) 2020–2030.

[22] A. Harazono, N. Kawasaki, T. Kawanishi, T. Hayakawa, Site-specific glycosylation analysis of human apolipoprotein B100 using LC/ESI MS/MS, Glycobiology 15 (2005) 447–462.

[23] C.W. Sutton, J.A. O'Neill, J.S. Cottrell, Site-specific characterization of glycoprotein carbohydrates by exoglycosidase digestion and laser desorption mass spectrometry, Anal. Biochem. 218 (1994) 34–46.

[24] P. Roepstorff, J. Fohlman, Proposal for a common nomenclature for sequence ions in mass spectra of peptides, Biomed. Mass Spectrom. 11 (1984) 601.

[25] K. Maemura, M. Fukuda, Poly-*N*-acetyllactosaminyl *O*-glycans attached to leukosialin: the presence of sialyl Le(x) structures in *O*-glycans, J. Biol. Chem. 267 (1992) 24379–24386.

[26] S. Hemmerich, S.D. Rosen, 6'-Sulfated sialyl Lewis X is a major capping group of GlyCAM-1, Biochemistry 33 (1994) 4830–4835.

[27] C.J. Van Den Hamer, A.G. Morell, I.H. Scheinberg, J. Hickman, G. Ashwell, Physical and chemical studies on ceruloplasmin: IX. The role of galactosyl residues in the clearance of ceruloplasmin from the circulation, J. Biol. Chem. 245 (1970) 4397–4402.

[28] A.G. Morell, G. Gregoriadis, I.H. Scheinberg, J. Hickman, G. Ashwell, The role of sialic acid in determining the survival of glycoproteins in the circulation, J. Biol. Chem. 246 (1971) 1461–1467.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Chromatography A, 1094 (2005) 105–117
JOURNAL OF CHROMATOGRAPHY A
www.elsevier.com/locate/chroma

Characterization of a gel-separated unknown glycoprotein by liquid chromatography/multistage tandem mass spectrometry Analysis of rat brain Thy-1 separated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis

Satsuki Itoh[a], Nana Kawasaki[a,b,*], Akira Harazono[a], Noritaka Hashii[a], Yukari Matsuishi[b], Toru Kawanishi[a], Takao Hayakawa[a]

[a] *Division of Biological Chemistry and Biologicals, National Institute of Health Science, 1-18-1, Kamiyoga, Setagaya-ku, Tokyo 158-8501, Japan*
[b] *CREST, Japan Science and Technology Agency (JST), Japan*

Received 17 May 2005; received in revised form 17 July 2005; accepted 25 July 2005
Available online 5 October 2005

Abstract

We developed an efficient and convenient strategy for protein identification and glycosylation analysis of a small amount of unknown glycoprotein in a biological sample. The procedure involves isolation of proteins by electrophoresis and mass spectrometric peptide/glycopeptide mapping by LC/ion trap mass spectrometer. For the complete glycosylation analysis, proteins were extracted in intact form from the gel, and proteinase-digested glycoproteins were then subjected to LC/multistage tandem MS (MS^n) incorporating a full mass scan, in-source collision-induced dissociation (CID), and data-dependent MS^n. The glycopeptides were localized in the peptide/glycopeptide map by using oxonium ions such as $HexNAc^+$ and $NeuAc^+$, generated by in-source CID, and neutral loss by CID-MS/MS. We conducted the search analysis for the glycopeptide identification using search parameters containing a possible glycosylation at the Asn residue with *N*-acetylglucosamine (203 Da). We were able to identify the glycopeptides resulting from predictable digestion with proteinase. The glycopeptides caused by irregular cleavages were not identified by the database search analysis, but their elution positions were localized using oxonium ions produced by in-source CID, and neutral loss by the data-dependent MS^n. Then, all glycopeptides could be identified based on the product ion spectra which were sorted from data-dependent CID-MS^n spectra acquired around localized positions. Using this strategy, we successfully elucidated site-specific glycosylation of Thy-1, glycosylphosphatidylinositol (GPI)-anchored proteins glycosylated at Asn23, 74, and 98, and at Cys111. High-mannose-type, complex-type, and hybrid-type oligosaccharides were all found to be attached to Asn23, 74 and 98, and four GPI structures could be characterized. Our method is simple, rapid and useful for the characterization of unknown glycoproteins in a complex mixture of proteins.

Keywords: Glycoprotein; LC/MS; Ion trap mass spectrometer; In-source CID; Thy-1

1. Introduction

Glycosylation is one of the most abundant post-translational modifications of proteins [1]. Most glycoproteins exist in heterogeneous forms due to their carbohydrate heterogeneity at multiple glycosylation sites. Because heterogeneity at each glycosylation site can be associated with many biological functions [2,3], it is necessary to analyze the oligosaccharide structures at each glycosylation site.

Mass spectrometric peptide/glycopeptide mapping by liquid chromatography coupled with electrospray ionaization tandem mass spectrometry (LC/ESI-MS/MS) is now used for characterization of glycoproteins [4,5]. Site-specific glycosylation of some gel-separated glycoproteins can be analyzed by in-gel proteinase digestion followed by MS; this method, however, gives unsatisfactory results due to a lower recovery of some glycopeptides from the gel [6–8]. For

* Corresponding author. Tel.: +81 3 3700 1141; fax: +81 3 3707 6950.
E-mail address: nana@nihs.go.jp (N. Kawasaki).

0021-9673/$ – see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.chroma.2005.07.100

complete site-specific glycosylation analysis, all glycopeptide fragments should be recovered from the gel. Hence, the extraction of a whole glycoprotein from the gel before proteinase digestion would be more reasonable than in-gel digestion. Additionally, the poor ionization efficiency of glycopeptides makes it difficult to analyze the glycosylation of glycopeptides in a complex mixture of peptides [6,9]. The glycopeptide-specific method is required for mass spectrometric peptide/glycopeptide mapping.

A precursor ion scan using triple quadrupole-type mass spectrometer is favorably used for analysis of glycopeptides [10–13]. However, this method requires repetitive analysis for the protein identification and glycosylation analysis, as it monitors carbohydrate marker ions such as $HexNAc^+$ and $Hex\text{-}HexNAc^+$ fragmented from glycopeptides by collision-induced dissociation (CID)-MS/MS, and does not provide product ion spectra of non-glycosylated peptides. As such, additional analysis would not be possible for small quantities of proteins, including gel-separated glycoproteins. As an alternative method, we have previously demonstrated peptide/glycopeptide mapping using quadrupole time-of-flight mass spectrometer, by which product ions arise from both peptides and carbohydrates [14]. Using oxonium ions as marker ions, we can sort out product ion spectra of glycopeptides from a number of product ion spectra of peptides, and can determine the amino acid sequences of glycopeptides, glycosylation sites, and monosaccharide composition in a single analysis. Recently, ion trap mass spectrometry (ITMS), which is capable of data-dependent multistage tandem MS (MS^n), has been found to be preferable for use in glycosylation analysis of glycopeptides [15,16]. Glycopeptide-specific detection by precursor ion scan and data-dependent scan cannot be used for glycosylation analysis by ITMS due to the low mass cut-off system. Instead, oxonium ions fragmented by in-source CID are used for the localization of glycopeptides in the peptide/glycopeptide map [3,17]. It has recently been reported that peptide + GlcNAc ions originating from *N*-glycosylated peptides by MS^2 yield peptide b and y ions by further MS^n, and that the peptide sequence and *N*-glycosylation sites can be identified based on the peptide fragment ions [15,16,18]. In addition, another group has reported that glycopeptides can be identified in peptide/glycopeptide map by a search analysis using a database to which all possible cleavage products of the glycopeptides have been added in advance [19]. A combination of peptide/glycopeptide mapping with in-source CID, data-dependent $CID\text{-}MS^n$, and the database search analysis would enable protein identification, glycopeptide selection, and glycosylation analysis of a small amount of glycoprotein.

In the present study, we developed a strategy for the characterization of a small amount of unknown glycoprotein in a biological sample. An unknown glycoprotein was isolated by electrophoresis and extracted from the gel in an intact form. We used sodium dodecyl sulfate (SDS), which is effective for extracting proteins from the gel, and could be easily removed by adding cold acetone. The proteinase-digested glycoprotein was subjected to peptide/glycopeptide mapping, with the sequential scan consisting of a full mass scan, in-source CID, and data-dependent $CID\text{-}MS^n$. Using this method, we carried out site-specific glycosylation analysis of glycosylphosphatidylinositol (GPI)-anchored proteins in rat brain. A computer database search was used for the identification of a GPI-anchored protein and its *N*-glycosylation sites. In-source CID and data-dependent CID-MS/MS were also used for localization of peptides with *N*-glycan and GPI in the peptide/glycopeptide map. On the basis of their product ion spectra, we elucidated *N*-glycosylation at each glycosylation site and the structure of GPIs.

2. Experimental

2.1. Materials

Rat brains were purchased from Nippon SLC (Hamamatsu, Japan). Trypsin-Gold and endoproteinase Asp-N were purchased from Promega (Madison, WI, USA) and Wako Pure Chemical (Osaka, Japan), respectively. Phosphatidylinositol-specific phospholipase C (PIPLC) from *Bacillus cereus* were purchased from Molecular Probes (Eugene, OR, USA). All other chemicals used were of the highest purity available.

2.2. Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) of PIPLC-treated GPI-anchored proteins

PIPLC-treated GPI-anchored proteins were prepared from rat brain utilizing Triton X-114 phase partition and PIPLC digestion [20,21]. Two whole rat brains (2.8 g, Wistar, male, 3 weeks) were homogenized in cold acetone and centrifuged for 10 min at 4 °C. The precipitate was then homogenized in $CHCl_3$: methanol (2:1, v/v) and centrifuged for 10 min at 4 °C. After being washed with methanol, the pellet was homogenized in 50 mM Tris–HCl (pH 7.4) containing 150 mM NaCl, 1 mM ethylenediaminetetraacetic acid (EDTA), and 1 mM phenylmethylsulfonyl fluoride (PMSF), and centrifuged at 10,000 × *g* at 4 °C for 20 min. The pellet was resuspended in the same buffer with an additional 2% Triton X-114 (v/v), and stirred at 4 °C for 16 h. After centrifugation at 10,000 × *g* at 4 °C for 20 min, the supernatant was subjected to Triton X-114 phase-partitioning at 37 °C for 10 min. The detergent phase was resuspended with an equal volume of 50 mM Tris–HCl (pH 7.4) containing 150 mM NaCl. Solubilized membrane proteins in the detergent phase were precipitated with cold acetone and were resuspended in 400 μl of 50 mM Tris–HCl (pH 7.4). Following the addition of PIPLC (1 U), the suspension was incubated at 37 °C for 18 h. The suspension was resubjected to Triton X-114 phase-partitioning, and PIPLC-treated GPI-anchored proteins were precipitated with cold acetone from the aqueous phase. PIPLC-treated GPI-anchored proteins obtained from

50 mg of rat brain were separated by SDS-PAGE (12.5%) after carboxyamidomethylation [22].

2.3. Extraction and digestion of gel-separated proteins

The protein in gel band was extracted with 20 mM Tris–HCl containing 1% SDS by shaking vigorously overnight after breaking down the gel into small bits. The extract was filtered with Ultrafree-MC (0.22 μm, Millipore, Bedford, USA), and the protein was precipitated by adding cold acetone. The precipitate was digested with trypsin (1 μg) in 20 μl of 0.1 M Tris–HCl (pH8.0) at 37 °C for 16 h, or with Asp-N (0.4 μg) in 20 μl of 5 mM Tris–HCl (pH 7.5) at 37 °C overnight.

2.4. LC/MS^n

Proteolytic peptides were separated by a Magic C18 column (50 mm × 0.2 mm, 3 μm, Michrom BioResources, Auburn, CA, USA) with a Paradaim MS4 HPLC system (Michrom BioResources Inc., Auburn, CA, USA) consisting of pump A: 0.1% formic acid and 2% acetonitrile, and pump B: 0.1% formic acid and 90% acetonitrile. Separation was performed with a linear gradient of 5–65% of pump B in 40 min after 5% in 10 min of pump B at a flow rate 3 μl/min. Mass spectra were recorded by Finnigan LTQ (Thermo Electron, San Jose, CA, USA) with the sequential scan: a full mass scan (*m/z* 300–2000), a full mass scan with in-source CID (*m/z* 80–500, collision energy: 50 V), and data-dependent CID-MS^n for most intense ions at each scan with dynamic exclusion for 30 s. Scan time (*m/z* 300–2000) is approximately 0.1 s. The operating condition used for LC/ITMS was as follows: tube lens offset of 130 V, capillary voltage of 2.0 kV, capillary temperature of 200 °C.

2.5. Computer database search analysis

All product ions obtained by LC/ITMS were subjected to the computer database search analysis with the TurboSEQUEST search engine (Thermo Electron, San Jose, CA, USA). We used the NCBI database (rat, updated at February 2003) and following search parameters: a static modification of carboxyamidomethylation (57 Da) at Cys, a possible modification of GlcNAc (203 Da) at Asn, and trypsin used for digestion.

3. Results

3.1. Extraction of whole proteins from the gel

Rat brain PIPLC-treated GPI-anchored proteins were separated by SDS-PAGE (Fig. 1), and the most noticeable band at 20–25 kDa was cut off from the gel and crushed. The gel pieces were shaken vigorously in 1% SDS, and the extracted protein was precipitated with cold acetone to remove SDS.

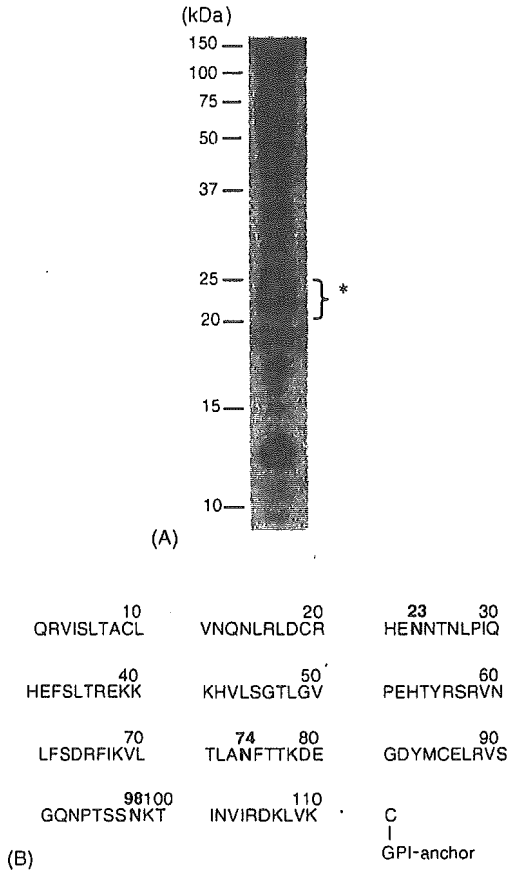

Fig. 1. (A) SDS-PAGE of PIPLC-treated GPI-anchored proteins from rat brain. (B) Amino acid sequence of rat Thy-1. *N*-Glycosylation sites are indicated by bold face. The protein at 20–25 kDa indicated by asterisk was subjected to the glycosylation analysis in this study.

We checked the recovery of the protein at 20–25 kDa by comparing the fluorescence intensity (Ex 633 nm/Em 670 nm) of the proteins at 20–25 kDa visualized by Coomassie staining before and after extraction. Approximately 55% of the protein at 20–25 kDa could be recovered from the gel (data not shown). The protein was digested with trypsin and subjected to the sequential scan consisting of full mass scans with and without in-source CID and data-dependent MS^n by LC/ITMS for protein identification and glycosylation analysis.

3.2. Database search analysis

Fig. 2(A) shows the peptide/glycopeptide map of the trypsin-digested protein at 20–25 kDa. First, all product ions generated by data-dependent MS^n were used for the database search analysis. Using search parameters described in Section 2.5, the protein was identified as Thy-1, a glycoprotein containing three *N*-glycosylation sites at Asn23, 74, and 98, and a GPI attachment site at Cys111. The search analysis also suggested the glycosylation at Asn74 and 98, with elution positions of 34 min (peak T6, Val69-Lys78) and 3.5 min (peak T1, Val89-Lys99), respectively (Fig. 2(A)). Although

Table 1
Glycosylation analysis of rat brain Thy-1

Glycosylation site	Carbohydrate composition[a]				Theoretical carbohydrate mass[b]	Trypsin					Asp-N				
	dHex	Hex	HexNAc	NeuAc		Observed *m/z*	Charge state	Peak number	Amino acid residue[c]	Theoretical *m/z*[b]	Observed *m/z*	Charge state	Peak number	Amino acid residue[c]	Theoretical *m/z*[b]
Asn23	0	5	2	0	1234.43	845.33	3	T3	H21-H31 (1315.63)	845.03	799.47	3	A3	E22-H31 (1178.57)	799.34
						937.27	3	T4	H21-F33 (1591.74)	937.06	1198.69	2	A3	E22-H31 (1178.57)	1198.50
						1405.81	2	T4	H21-F33 (1591.74)	1405.09	949.60	3	A3	N23-T36 (1626.80)	948.75
											1423.40	2	A3	N23-T36 (1626.80)	1422.62
	0	6	2	0	1396.49	899.49	3	T3	H21-H31 (1315.63)	899.04	853.54	3	A3	E22-H31 (1178.57)	853.36
						1348.35	2	T3	H21-H31 (1315.63)	1348.06	1279.66	2	A3	E22-H31 (1178.57)	1279.53
						942.16	3	T3	H21-E32 (1444.67)	942.06	1211.52	2	A5	E22-Q30 (1041.51)	1211.00
						991.37	3	T5	H21-F33 (1591.74)	991.08					
						1486.40	2	T5	H21-F33 (1591.74)	1486.12					
	0	3	5	0	1519.57						1044.43	3	A3	N23-T36 (1626.80)	1043.79
	0	7	2	0	1558.54	996.41	3	T3	H21-E32 (1444.67)	996.07	907.59	3	A3	E22-H31 (1626.80)	907.37
											1361.03	2	A3	E22-H31 (1178.57)	1360.56
	1	3	5	0	1665.62						1093.32	3	A3	N23-T36 (1626.80)	1092.48
	1	5	4	0	1786.65						1133.80	3	A3	N23-T36 (1626.80)	1132.82
	0	6	3	1	1890.66						1168.51	3	A4	N23-T36 (1626.80)	1167.49
	1	6	4	0	1948.70						1187.21	3	A3	N23-T36 (1626.80)	1186.84
Asn74	1	2	3	0	1405.52	1026.26	3	T6	V59-F75 (1996.12)	1026.18					
	0	5	2	0	1234.43	949.52	2	T2	A73-K78 (680.35)	949.39	995.71	3	A7	D64-K78 (1766.01)	995.15
						1162.72	2	T6	V69-K78 (1106.62)	1162.53					
	1	4	3	0	1421.52	1043.45	2	T2	A73-K78 (680.35)	1042.94					
	1	3	4	0	1462.54	1063.68	2	T2	A73-K78 (680.35)	1063.45					
						1276.61	2	T6	V69-K78 (1106.62)	1276.58					
	1	5	3	0	1583.57	1124.18	2	T2	A73-K78 (680.35)	1123.96					
						1337.64	2	T6	V69-K78 (1106.62)	1337.10					
	1	3	5	0	1665.62	1165.13	2	T2	A73-K78 (680.35)	1164.99	1272.15	2	A4	T71-K78 (894.48)	1272.05
						919.71	3	T6	V69-K78 (1106.62)	919.09	1139.79	3	A7	D64-K78 (1766.01)	1138.88
						1378.19	2	T6	V69-K78 (1106.62)	1378.12					
	2	5	3	0	1729.63	1197.53	2	T2	A73-K78 (680.35)	1196.99					
	2	4	4	0	1770.66	1217.67	2	T2	A73-K78 (680.35)	1217.51					
						1430.71	2	T6	V69-K78 (1106.62)	1430.64					
	1	5	4	0	1786.65	1226.38	2	T2	A73-K78 (680.35)	1225.50					
						960.01	3	T6	V69-K78 (1106.62)	959.43					
						1439.05	2	T6	V69-K78 (1106.62)	1438.64					
	1	4	5	0	1827.68	973.79	3	T6	V69-K78 (1106.62)	973.10					
	2	5	4	0	1932.71	1298.75	2	T2	A73-K78 (680.35)	1298.53	1228.18	3	A7	D64-K78 (1766.01)	1227.91
						1008.71	3	T6	V69-K78 (1106.62)	1008.11					
						1512.21	2	T6	V69-K78 (1106.62)	1511.67					
	1	6	4	0	1948.70	1306.81	2	T2	A73-K78 (680.35)	1306.53					
						1013.98	3	T6	V69-K78 (1106.62)	1013.45					
	0	4	5	1	1972.71						1241.91	3	A7	D64-K78 (1766.01)	1241.25
	2	4	5	0	1973.73	880.04	3	T2	A73-K78 (680.35)	879.70	1241.84	3	A7	D64-K78 (1766.01)	1241.59
						1319.24	2	T2	A73-K78 (680.35)	1319.04					
						1022.43	3	T6	V69-K78 (1106.62)	1021.79					
						1532.49	2	T6	V69-K78 (1106.62)	1532.49					

	1	6	3	1	2036.72	1043.24	3	T7	V69-K78 (1106.62)	1042.78					
	1	5	4	1	2077.75	1057.08	3	T7	V69-K78 (1106.62)	1056.46					
	2	6	4	0	2094.76	1062.65	3	T6	V69-K78 (1106.62)	1062.13	1282.43	3	A7	D64-K78 (1766.01)	1281.93
	1	4	5	1	2118.77	1070.38	3	T7	V69-K78 (1106.62)	1070.14					
	2	5	5	0	2135.79	1076.32	3	T6	V69-K78 (1106.62)	1075.81					
	1	6	4	1	2239.80	1110.97	3	T7	V69-K78 (1106.62)	1110.48					
	3	5	5	0	2281.85	1124.78	3	T6	V69-K78 (1106.62)	1124.49					
	1	7	5	0	2313.84	851.76	4	T6	V69-K78 (1106.62)	851.62					
	1	4	6	1	2321.85	1138.16	3	T7	V69-K78 (1106.62)	1137.83					
	2	5	6	1	2629.96	1241.13	3	T7	V69-K78 (1106.62)	1240.53					
Asn98	0	5	2	0	1234.43	1168.34	2	T1	V89-K99 (1117.54)	1167.99	912.24	3	A2	E86-K99 (1515.76)	911.74
											1367.31	2	A2	E86-K99 (1515.76)	1367.10
											930.26	3	A3	Q92-R105 (1570.84)	930.10
											1021.28	3	A3	E86-N102 (1843.94)	1021.13
											1011.53	3	A4	T95-K110 (1815.06)	1010.81
											1516.43	2	A4	T95-K110 (1815.06)	1516.75
											1144.21	3	A5	E86-R105 (2212.18)	1143.88
	0	3	4	0	1316.49	1209.79	2	T1	V89-K99 (1117.54)	1209.02					
	0	3	5	0	1519.57	1310.68	2	T1	V89-K99 (1117.54)	1310.56	1007.10	3	A2	E86-K99 (1515.76)	1006.78
											1106.39	3	A4	T95-K110 (1815.06)	1106.55
											1239.21	3	A5	E86-R105 (2212.18)	1238.92
	2	3	4	0	1608.60						951.28	3	A4	G91-N102 (1259.61)	951.08
											1426.22	2	A4	G91-N102 (1259.61)	1426.11
	1	4	4	0	1624.60	951.37	3	T3	S96-D106 (1245.67)	951.76					
	0	5	4	0	1640.59	1371.84	2	T1	V89-K99 (1117.54)	1371.07	960.04	4	A5	E86-R105 (2212.18)	959.70
											1279.69	3	A5	E86-R105 (2212.18)	1279.26
	0	5	3	1	1728.61						1176.27	3	A4	T95-K110 (1815.06)	1176.23
	0	5	4	0	1640.59						1146.70	3	A4	T95-K110 (1815.06)	1146.89
	1	6	3	0	1745.62						1082.52	3	A2	E86-K99 (1515.76)	1082.13
											1181.80	3	A4	T95-K110 (1815.06)	1181.90
											986.38	4	A5	E86-R105 (2212.18)	985.96
											1314.72	3	A5	E86-R105 (2212.18)	1314.27
	1	5	4	0	1786.65						1195.47	3	A4	T95-K110 (1815.06)	1195.57
											996.54	4	A5	E86-R105 (2212.18)	996.21
											1328.49	3	A5	E86-R105 (2212.18)	1327.95
	1	5	3	1	1874.67						1018.89	4	A6	E86-R105 (2212.18)	1018.22
	0	6	3	1	1890.66	1496.18	2	T1	V89-K99 (1117.54)	1496.10	1230.24	3	A4	T95-K110 (1815.06)	1230.24
											1022.70	4	A6	E86-R105 (2212.18)	1022.22
	2	5	4	0	1932.71	1517.35	2	T1	V89-K99 (1117.54)	1517.13					
	1	6	4	0	1948.70	1525.78	2	T1	V89-K99 (1117.54)	1525.12	1150.13	3	A2	E86-K99 (1515.76)	1149.83
											1249.75	3	A4	T95-K110 (1815.06)	1249.59
											1259.45	3	A3	E86-N102 (1843.94)	1259.22
											1037.27	4	A5	E86-R105 (2212.18)	1036.73
											1382.34	3	A5	E86-R105 (2212.18)	1381.97
	1	6	3	1	2036.72						1059.07	4	A6	E86-R105 (2212.18)	1058.73
	0	6	4	1	2093.74						1298.06	3	A4	T95-K110 (1815.06)	1297.94
											1073.37	4	A6	E86-R105 (2212.18)	1072.99

[a] dHex, deoxyhexose; Hex, hexose; HexNAc, *N*-acetylhexosamine; NeuAc, *N*-acetylneuramic acid.
[b] Monoisotopic value.
[c] Theoretical peptide mass is indicated in parenthesis.

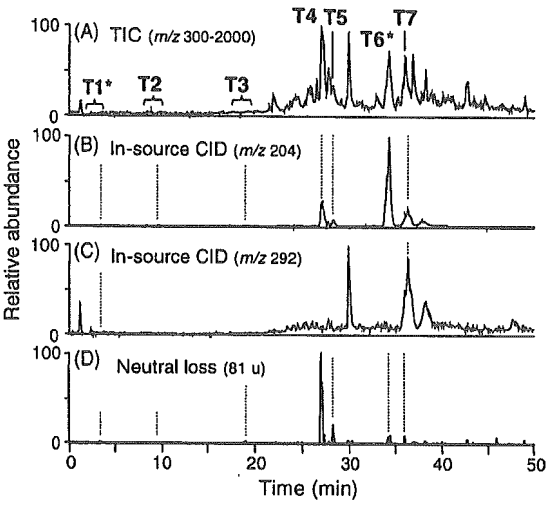

Fig. 2. Total ion chromatogram (TIC) of trypsin-digested protein at 20–25 kDa (*m/z* 300–2000) (A), mass chromatograms from TIC with in-source CID of *m/z* 204 (B) and 292 (C), and neutral loss chromatogram of 81 u by data-dependent CID-MS/MS (D). Asterisks mean the peak of glycopeptides identified by the database search analysis.

the glycopeptide Val89-Lys99 contains two Asn residues, Asn93 and 98, only Asn98 was identified as a glycosylation site because of detection of b and y ions modified with GlcNAc at Asm98.

Next, to study the site-specific glycosylation at Asn74 and 98, product ion spectra of glycopeptides were sorted from the numbers of product ion spectra acquired around peak T6 and T1. We sorted out product ion spectra of glycopeptides using B series ions, such as $Hex_1HexNAc_1^+$ and $Hex_2HexNAc_1^+$ (*m/z* 366 and 528) originated from glycopeptides by CID-MS/MS, as marker ions [23]. We could sort out 14 product ion spectra originated from glycopeptide Val69-Lys78 around peak T6. The monosaccharide compositions of *N*-glycans at Val69-Lys78 were calculated as $dHex_{0-3}Hex_{2-7}HexNAc_{2-5}$ on the basis of the *m/z* values of their molecular ions and the theoretical mass of the peptide. Likewise, seven product ion spectra originated from glycopeptide Val89-Lys99 were sorted from those around peak T1, and their monosaccharide compositions were estimated as $dHex_{0-2}Hex_{3,5,6}HexNAc_{2-5}NeuAc_{0,1}$ (Table 1). Glycosylation at Asn74 and 98 were elucidated by a detailed examination of these product ion spectra as follows.

3.2.1. Analysis of the glycosylation at Asn74 of peptide Val69-Lys78

Fig. 3(A) shows a product ion spectrum of the glycopeptide Val69-Lys78 at 34.52 min. Its precursor ion is the doubly charged ion at *m/z* 1512.2. Many product ions generated by cleavages of glycosidic linkages can be observed in this product ion spectrum. The most intense ion at *m/z* 1311 is assigned to a peptide bearing the reducing end of GlcNAc, which was caused by glycosidic linkage cleavage of *N*-linked oligosaccharide. Fig. 3(B) is the product ion spectrum of the peptide + GlcNAc ion at *m/z* 1311. The b and y ions generated by cleavages of the peptide backbone prove that this glycopeptide is the peptide Val69-Lys78 glycosylated at Asn74.

The molecular weight of the carbohydrate moiety can be calculated as 1933.8 Da by subtracting the theoretical mass of the peptide (1106.6 Da) from the calculated glycopeptide mass (3022.4 Da). Consequently, the monosaccharide composition can be estimated as $dHex_2Hex_5HexNAc_4$. In the product ion spectrum (Fig. 3(A)), B ions corresponding to $dHex_1Hex_1HexNAc_1$ ($B_{2\alpha}$) and $dHex_1Hex_2HexNAc_1$ ($B_{3\alpha}$) were detected at *m/z* 512 and 674, respectively. These results indicate that one of two dHex, which are likely to be Fuc, attaches to Gal-GlcNAc at the non-reducing end in a similar manner as the Lewis a/x antigen (Gal-(Fuc-)GlcNAc-), or the blood group H-determinant (Fuc-Gal-GlcNAc-). The product ion at *m/z* 350 produced from the triply charged precursor ion at *m/z* 1008.7 corresponded to $dHex_1HexNAc_1$ (data not shown), suggesting that Fuc attaches to GlcNAc like the Lewis a/x antigen (Gal-(Fuc-)GlcNAc-). The attachment site of the other Fuc can be deduced at inner trimannosyl core GlcNAc from the observation of Y ions at *m/z* 1457, 1660, and 1822, which correspond to Val69-Lys78 plus $dHex_1HexNAc_1$ ($Y_{1\alpha}$), $dHex_1HexNAc_2$ ($Y_{2\alpha}$), and $dHex_1Hex_1HexNAc_2$ ($Y_{3\alpha/3\beta/3\gamma}$), respectively. In addition, the product ion at *m/z* 1411 resulting from the precursor ion at *m/z* 1512.2 by loss of 101.6 u (HexNAc), suggests a linkage of non-substituted HexNAc at the non-reducing terminal end. Together with detection of the product ion at *m/z* 940 ($Y_{3\alpha/1\beta/3\beta}^{+}$, [GlcNAc-Man-GlcNAc-GlcNAc-peptide+H]$^{2+}$), it can be deduced that this HexNAc is a bisecting GlcNAc attached to the core mannose residue via a β1–4 linkage. From these product ions, we could deduce two oligosaccharide structures. One is the structure indicated in Fig. 3(A), inset, and the other is one containing a Gal-Gal-(Fuc-)GlcNAc-Man-branch instead of a Gal-(Fuc-)GlcNAc-Man-branch. Detection of Gal-(Fuc-)GlcNAc-Man$^+$ at *m/z* 674 but not Gal-Gal-(Fuc-)GlcNAc-Man$^+$ at *m/z* 836 suggests that this oligosaccharide structure can be assigned to the structure indicated in Fig. 3(A), inset.

The carbohydrate structures of the other glycopeptide Val69-Lys78 detected around peak T6 can be characterized as the high-mannose-type oligosaccharide (M5), complex-type oligosaccharides containing some partial structures such as inner core Fuc, bisecting GlcNAc, the Lewis a/x antigen, and blood group H-determinant, and hybrid-type oligosaccharides (Table 1).

3.2.2. Analysis of the glycosylation at Asn98 of peptide Val89-Lys99

Fig. 4 shows one of the product ion spectra of the glycopeptide Val89-Lys99 at 3.47 min. Its precursor ion is the doubly charged ion at *m/z* 1525.8. The monosaccharide composition, $dHex_1Hex_6HexNAc_4$, can be estimated based on the calculated mass of the carbohydrate moiety (1950.0 Da) obtained by subtracting the mass of the theo-

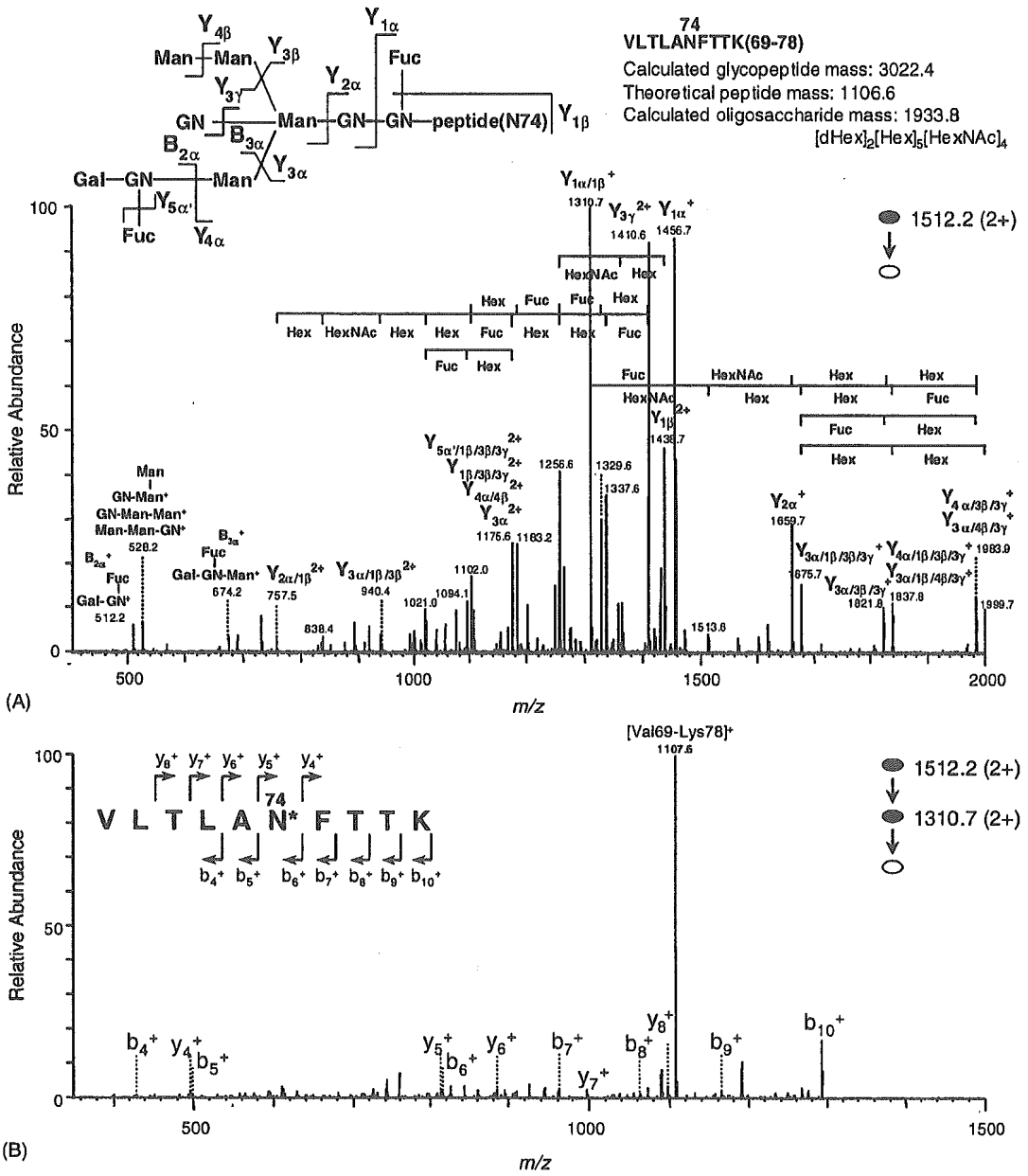

Fig. 3. (A) Product ion spectrum (MS^2) of the doubly charged glycopeptide precursor ion at *m/z* 1512.2 in peak T6. The glycopeptide Val69-Lys78 is glycosylated with oligosaccharide, $dHex_2Hex_5HexNAc_4$ at Asn74, and the inset is the deduced oligosaccharide structure. (B) MS^3 product ion spectrum derived from the doubly charged glycopeptide precursor ion at *m/z* 1512.2, followed by further fragmentation of the product ion at *m/z* 1310.7.

retical typtic peptide mass (1117.5 Da) from the calculated glycopeptide mass (3049.5 Da). Y ions corresponding to Val89-Lys99 plus $dHex_1HexNAc_1$ ($Y_{1\alpha}$), $dHex_1HexNAc_2$ ($Y_{2\alpha}$), and $dHex_1Hex_1HexNAc_2$ ($Y_{3\alpha/3\beta/3\gamma}$) detected at *m/z* 1468, 1671, and 1833, respectively, reveals that one Fuc residue is linked to the inner trimannosyl core GlcNAc. Additionally, the product ion at *m/z* 1424 suggests a linkage of non-substituted HexNAc at the non-reducing terminal end. Together with the product ions at *m/z* 945 and 1890, it can be deduced that this HexNAc is a bisecting GlcNAc that attaches to a core mannose residue via a β1–4 linkage. On the basis of the product ions at *m/z* 487, 528 and 1380, corresponding to Hex_3 ($B_{2\beta}$), $Hex_2HexNAc_1$ ($B_{3\alpha}$), and $Hex_6HexNAc_2$ ($B_{4\alpha}$), the oligosaccharide structure was characterized as a hybrid-type oligosaccharide (Fig. 4, inset).

The carbohydrate structures of the other glycopeptide Val89-Lys98 detected around peak T1 are characterized as high-mannose-type oligosaccharide (M5), complex-type, and hybrid-type oligosaccharides, which include bisecting GlcNAc and Lewis a/x structures (Table 1).

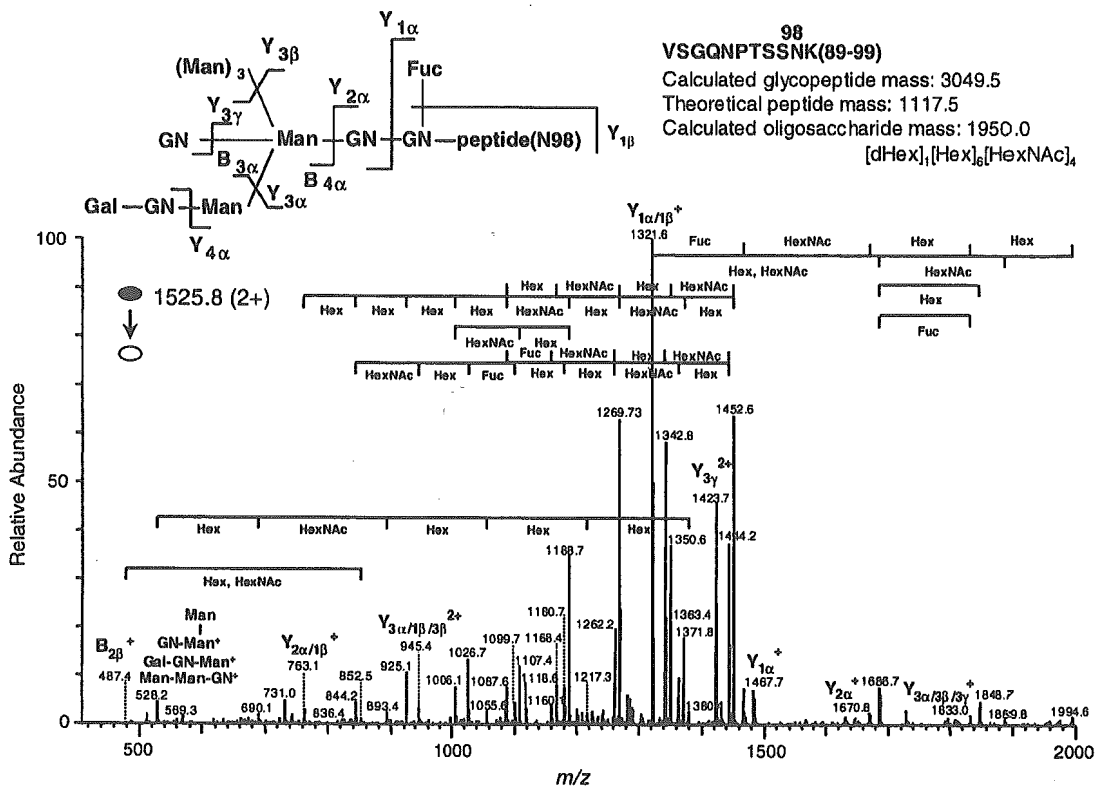

Fig. 4. Product ion spectrum of the doubly charged glycopeptide precursor ion at *m/z* 1525.8 in peak T1. The glycopeptide Val89-Lys99 is glycosylated with the oligosaccharide, $dHex_1Hex_6HexNAc_4$ at Asn98, and the inset is the deduced oligosaccharide structure.

3.3. Detection of glycopeptides by in-source CID and CID-MS/MS

Glycopeptides containing Asn23 could not be identified by the database search analysis. Therefore, we first localized all glycopeptides in the peptide/glycopeptide map using oxonium marker ions generated by in-source CID. Fig. 2(B and C) shows mass chromatograms of oxonium marker ions, $HexNAc^+$ (*m/z* 204) and $NeuAc^+$ (*m/z* 292), respectively. The mass chromatogram of *m/z* 204 indicates that the glycopeptides were localized around 3.7, 9.7, 19.1, 27.2, 28.4 34.3, 36.3, and 37.8 min. The mass chromatogram of *m/z* 292 suggests that the glycopeptides bearing NeuAc were localized around 3.7, 30.0, 36.4, and 38.2 min. In addition to the localization of glycopeptides by in-source CID, we monitored neutral loss caused by data-dependent CID-MS/MS. The neutral loss chromatogram of 81 u indicates the localization of doubly charged glycopeptides ions with Hex at the non-reducing ends. The elution positions of the localized glycopeptides by neutral loss are almost identical to those by in-source CID. Second, for confirmation of the elution position of glycopeptides and characterization of the carbohydrate moiety, we sorted the product ion spectra of glycopeptides from enormous numbers of data-dependently acquired product ion spectra around localized glycopeptides by using oligosaccharide oxonium ions as marker ions. Consequently, the locations of glycopeptides were confirmed in peak T1-6 (Fig. 2(A)). The peaks T1 and 6 correspond to the location of glycopeptides identified by the database search as Val89-Lys99 and Val69-Lys78, respectively. Four glycopeptide peaks were newly sorted by in-source CID and data-dependent CID-MS/MS. Structural assignment of the glycopeptides in these peaks was carried out using their MS^n spectra as follows.

3.3.1. Analysis of the glycosylation at Asn23 of peptide His21-Phe33

Fig. 5(A) shows one of the product ion spectra of the glycopeptide His21-Phe33 in peak T4. Its precursor ion is the triply charged ion at *m/z* 937.3. The intense product ion at *m/z* 899 is assigned to a doubly charged ion of peptide plus GlcNAc on the basis of Y series ions. The region of His21-Phe33 containing Asn23 in Thy-1 was suggested as the peptide moiety of this glycopeptide, 1593.3 Da, by the FindPept tool available on the internet (ExPASY Proteomics tools, Swiss Institute of Bioinformatics, http://us.expasy.org/tools/findpept.html). We examined the data-dependently acquired product ion spectrum of the precursor ion at *m/z* 899 and found that the *m/z* values of b and y ions in the product ion spectrum were identical to those of predictable product ions originating from the peptide His21-Phe33 modified with HexNAc at Asn23 (Fig. 5(B)). From the calculated oligosaccharide mass (1235.1 Da) obtained by subtracting the theoretical typtic

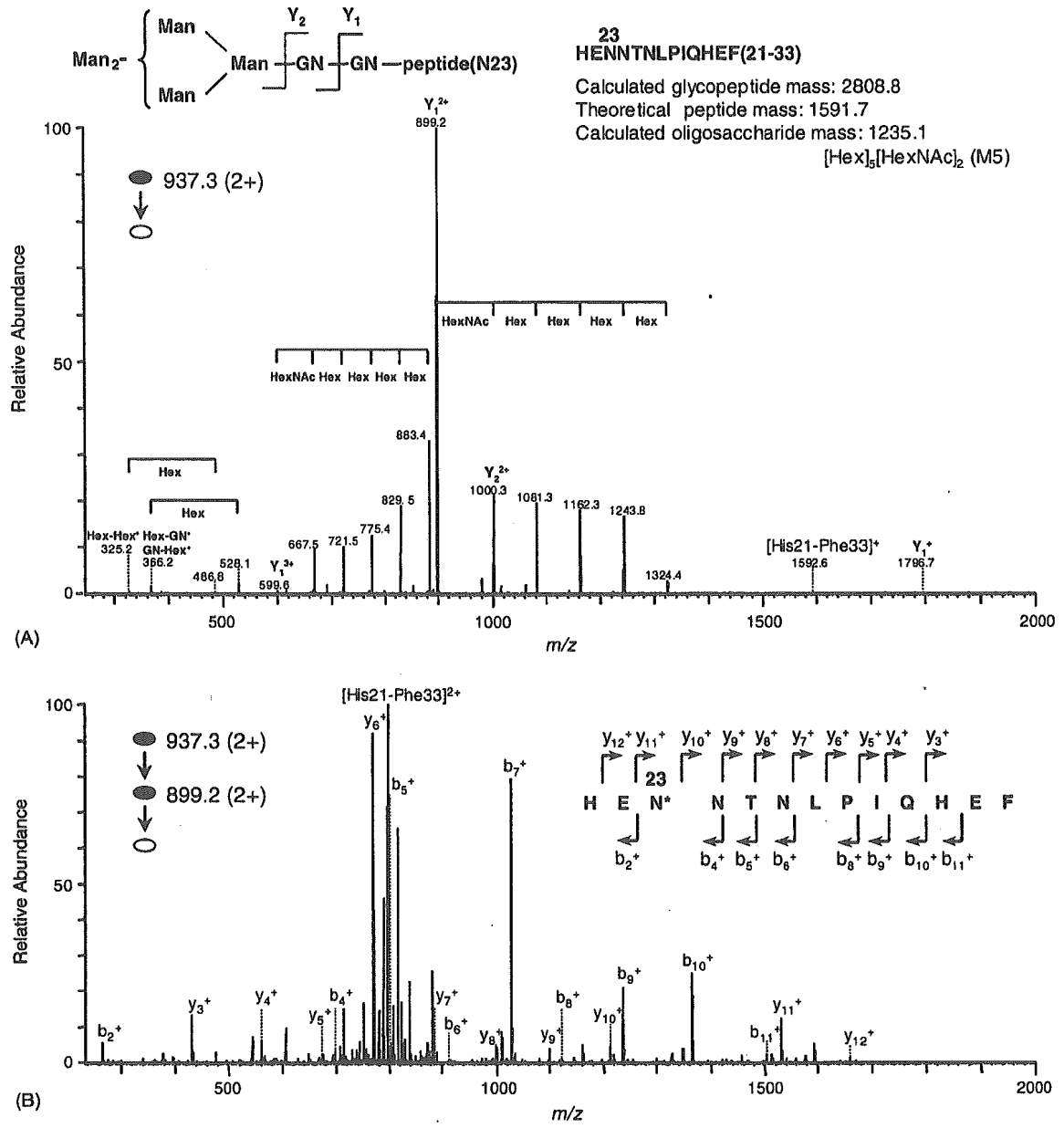

Fig. 5. (A) Product ion spectrum (MS^2) of the doubly charged glycopeptide precursor ion at *m/z* 937.3 in peak T4. The glycopeptide His21-Phe33 is glycosylated with oligosaccharide, $Hex_5HexNAc_2$ at Asn23, and the inset is the deduced oligosaccharide structure. (B) MS^3 product ion spectrum derived from a doubly charged glycopeptide precursor ion at *m/z* 937.3, followed by further fragmentation of the product ion at *m/z* 899.2.

peptide mass (1591.7 Da) from the calculated glycopeptide mass (2808.8 Da) together with product ions at *m/z* 366 and 528, it is indicated that this peptide carries $Hex_5HexNAc_2$, i.e. high-mannose-type oligosaccharide, M5. All product ion spectra in peak T4 revealed that peptides His21-Phe33 contain only high-mannose-type oligosaccharide (M5).

3.3.2. *Analysis of glycopeptides in peaks T2, 3, 5, and 7*

Similarly, product ion spectra of glycopeptides around peaks T2, 3, 5, and 7 were sorted by using oligosaccharide oxonium marker ions generated by MS/MS. In product ion spectra sorted out from around peak T2, the intense ion at *m/z* 884 was detected and assigned to a singly charged ion of a peptide plus GlcNAc. The peptide was suggested to be Ala73-Lys78 containing Asn74 by the FindPept tool. The monosaccharide composition can be estimated from the calculated mass of oligosaccharide moiety obtained by subtracting the theoretical mass of Ala73-Lys78 (680.35 Da) from calculated glycopeptide mass. Oligosaccharide structure of the glycopeptides is characterized based on their product

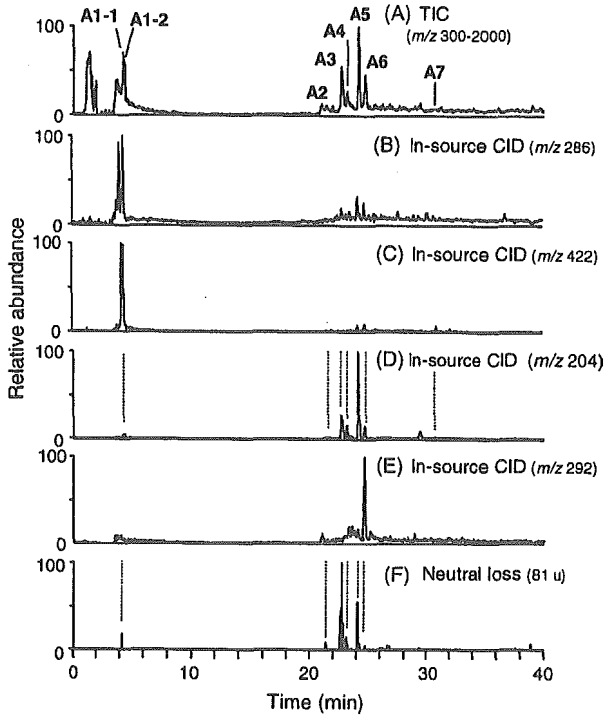

Fig. 6. Total ion chromatogram (TIC) of Asp-N digested protein at 20–25 kDa (*m/z* 300–2000) (A), mass chromatograms from TIC with in-source CID of *m/z* 286 (B), 422 (C), 204 (D), and 292 (E), and neutral loss chromatogram of 81 u by data-dependent CID-MS/MS (F).

ion spectra. Glycopeptides in peak T2 were characterized as Ala73-Lys78 glycosylated at Asn74 with *N*-glycans consisting of $dHex_{0-2}Hex_{3-6}HexNAc_{2-5}$. These *N*-glycans can be identified as high-mannose-type oligosaccharide (M5), and complex-type and hybrid-type oligosaccharides containing Fuc attached to inner trimannosyl core GlcNAc. Their structural assignments are summarized in Table 1. Glycopeptides in peak T3 can be identified as a mixture of peptide His21-His31 and His21-Glu32 glycosylated at Asn23, and Ser96-Asp106 glycosylated at Asn98. Asn23 was attached by high-mannose-type oligosaccharides, M5, 6, and 7, and Asn98 was occupied by *N*-glycan consisting of $dHex_1Hex_4HexNAc_4$ with a Lewis a/x structure as a partial structure. Glycopeptides in peak T5 were characterized as peptide His21-Phe33 glycosylated at Asn23 with high-mannose-type oligosaccharide, M6. Glycopeptides in peak T7 were assigned to be Val69-Lys78 glycosylated at Asn74 with *N*-glycans composed of $dHex_{1-2}Hex_{4-6}HexNAc_{3-6}NeuAc$.

3.4. Analysis of the GPI moiety of rat Thy-1

Since trypsin digestion provided Cys-GPI, which could not be retained on the C_{18} column, Asp-N digestion was also performed to obtain more hydrophobic peptides attached by GPI (GPI-peptides). Fig. 6(A) shows the peptide/glycopeptide map obtained by LC/ITMS of Asp-N digested Thy-1. We localize the GPI-peptides using marker ions, EtN-PO_4-Man$^+$ at *m/z* 286 and GlcN-inositol-PO_4^+ at *m/z* 422, originating from the core structure of the GPI moiety by in-source CID (EtN, ethanolamine; GlcN, glucosamine). Mass chromatograms of *m/z* 286 and 422 suggest the locations of the GPI-peptides to be around 4.2 (peak A1-1) and 4.4 min (peak A1-2) (Fig. 6(B and C)). Using product ions originated from GPI moiety, such as GlcN-inositol-PO_4^+ and PO_4-Man-GlcN$^+$ (*m/z* 422 and 404), as marker ions, four product ion spectra of GPI-peptides were sorted out from all product ion spectra around peaks A1-1 and 1-2. Their precursor ions were doubly charged ions at *m/z* 1132 and 1213 (peak A1-1), 1051 and 1151 (peak A1-2). Based on these product ion spectra, we characterized GPI-peptides as the peptide Asp106-Cys111 with a GPI core structure plus Hex_{0-2}, $HexNAc_{1-2}$ and PO_4-EtN.

Fig. 7(A) shows the product ion spectrum of the doubly charged GPI-peptide ion at *m/z* 1051 in peak A1-2. In addition to product ions at *m/z* 422, those originating from the GPI moiety were detected at *m/z* 404 (PO_4-Man-GlcN$^+$), 447 (EtN-PO_4-Man-GlcN$^+$), 650 (EtN-PO_4-(HexNAc-)Man-GlcN$^+$), 787 (peptide-EtN$^+$), 868 (peptide-EtN-PO_4^+), 1191 (peptide-EtN-PO_4-Man-Man$^+$), 1477 (peptide-EtN-PO_4-Man-Man-(EtN-PO_4-)Man$^+$), 1638 (peptide-EtN-PO_4-Man-Man-(EtN-PO_4-)Man-GlcN$^+$), and 1898 (peptide-EtN-PO_4-Man-Man-(EtN-PO_4-)Man-GlcN-inositol-PO_4^+). From these fragments, it can be deduced that this peptide is Asp106-Cys111 carrying the GPI, as indicated in the inset in Fig. 7(A).

The other GPI-peptide in peak A1-1 was characterized as having side chains; -Hex attached to M1, -PO_4-EtN and -HexNAc attached to M3, based on the product ion spectrum of the doubly charged precursor ion at *m/z* 1132 (data not shown). These two GPI structures are identical to those that have been previously reported [24].

Product ion spectra of doubly charged ion at *m/z* 1151 and 1213 suggested that they contained GPI which bear one HexNAc or two Hex in addition to GPI in Fig. 7(A) respectively. Fig. 7(B) shows the product ion spectra of the doubly charged precursor ions at *m/z* 1151 in peak A1-2. In addition to *m/z* 422, we detected product ions at *m/z* 366 (HexNAc-Man$^+$), 447 (EtN-PO_4-Man-GlcN$^+$), 650 (EtN-PO_4-(HexNAc-)Man-GlcN$^+$), 1229 (peptide-EtN-PO_4-(HexNAc-)Man$^+$), 1391 (peptide-EtN-PO_4-(HexNAc-)Man-Man$^+$), 1676 (peptide-EtN-PO_4-(HexNAc-)Man-Man-(EtN-PO_4-)Man$^+$), 1838 (peptide-EtN-PO_4-(HexNAc-)Man-Man-(EtN-PO_4-)Man-GlcN$^+$), and 1880 (peptide-EtN-PO_4-(HexNAc-)Man-Man-(EtN-PO_4-)(HexNAc-)Man$^+$). These fragment ions suggest the attachment of -HexNAc to Man1, and -PO_4-EtN and -HexNAc to Man3 as indicated in the inset of Fig. 7(B). Similarly, product ion spectra of the doubly charged precursor ion at *m/z* 1213 indicate the attachment of 2Hex and HexNAc to Man1 and Man3-PO_4-EtN (data not shown). To our knowledge, this is the first report of these two GPI structures in Thy-1.