Trends in Glycoscience and Glycotechnology
Vol. 20 No. 112 (May 2008) pp. 97-116

Profile of the Authors

Nana Kawasaki graduated from the department of Pharmaceutical Sciences, Hokkaido University in 1984. She is a member of the division of Biological Chemistry and Biologicals at National Institute of Health Sciences. Her current research focuses on a development of evaluation methods for glycoprotein products by mass spectrometry.

Satsuki Itoh graduated from the Department of Chemistry, Faculty of Science, Osaka University in 1995 (MS). She has worked at the Division of Biological Chemistry & Biologicals, the National Institute of the Health Sciences since 2000. She received her Ph.D. degree from Osaka University in 2007. She investigates the glycosylation of glycoproteins by mass spectrometry.

Noritaka Hashii is a research scientist of National Institute of Health Sciences (NIHS). He received his Ph.D. from Josai University under the supervision of Prof. Seiichi Kondo in 2003. His current research interests are structural analyses of carbohydrates such as N-, O-linked oligosaccharides and glycosaminoglycans by mass spectrometry.

Akira Harazono graduated from the department of Physiological Chemistry, Kyoto University in 1995 and received a Ph.D. at Kyoto University in 2002. His recent research focuses on study of quality test for peptide/protein products.

Daisuke Takakura graduated in the Bioscience Course in 2002 at the Azabu University and received his PhD in 2007 from the same university. From 2007, he held a post-doctoral position at the Division of Biological Chemistry Biologicals in National Institute of Health Sciences.

Trends in Glycoscience and Glycotechnology
Vol. 20 No. 112 (May 2008) pp. 97-116

Profile of the Authors

Dr. Teruhide Yamaguchi graduated from Biology at Kobe University in 1976. He is the director of the division of Biological Chemistry and Biologicals at National Institute of Health Sciences, and is engaged in investigating the safety, quality and efficacy of biologics. He is member of the Pharmaceutical and Food Affairs Council in Japan and is contributing to review quality aspects of registration applications of new drugs. He is also contributing to ICH activity as the member of Gene Therapy EWG.

Journal of Chromatography B, 869 (2008) 20–30

Contents lists available at ScienceDirect

Journal of Chromatography B

journal homepage: www.elsevier.com/locate/chromb

Simultaneous glycosylation analysis of human serum glycoproteins by high-performance liquid chromatography/tandem mass spectrometry

Akira Harazono[a,*], Nana Kawasaki[a,b], Satsuki Itoh[a], Noritaka Hashii[a], Yukari Matsuishi-Nakajima[a,b], Toru Kawanishi[a], Teruhide Yamaguchi[a]

[a] Division of Biological Chemistry and Biologicals, National Institute of Health Sciences, 1-18-1 Kami-yoga, Setagaya-Ku, Tokyo 158-8501, Japan
[b] Core Research for Evolutional Science and Technology (CREST) of Japan Science and Technology Agency (JST), Kawaguchi Center Building, 4-1-8 Hon-cho, Kawaguchi, Saitama 332-0012, Japan

ARTICLE INFO

Article history:
Received 19 December 2007
Accepted 5 May 2008
Available online 10 May 2008

Keywords:
Glycosylation analysis
Human serum
Glycopeptide
LC/MS
LC/MS/MS
Immunoglobulin G
Haptoglobin
Transferrin
Ceruloplasmin

ABSTRACT

Changes in the glycosylation of some serum proteins are associated with certain diseases. In this study, we performed simultaneous site-specific glycosylation analysis of abundant serum glycoproteins by LC/Qq-TOF MS of human serum tryptic digest, the albumin of which was depleted. The glycopeptide peaks on the chromatogram were basically assigned by database searching with modified peak-list text files of MS/MS spectra and then based on mass differences of glycan units from characterized glycopeptides. Glycopeptide of IgG, haptoglobin and ceruloplasmin were confirmed by means of a comparison of their retention times and *m/z* values with those obtained by LC/MS of commercially available glycoproteins. Mass spectrometric carbohydrate heterogeneity in the assigned glycopeptides was analyzed by an additional LC/MS. We successfully demonstrated site-specific glycosylation of 23 sites in abundant serum glycoproteins.

1. Introduction

Glycosylation of proteins is a common post-translational modification of proteins [1], and most proteins in serum are glycosylated [2]. Changes in the oligosaccharide moieties of certain serum glycoproteins are associated with human diseases. Oligosaccharides lacking galactose residues in immunoglobulin G (IgG) are increased in rheumatoid arthritis [3,4] and Crohn's syndrome [5]. Congenital disorders of glycosylation (CDG) are genetic disorders in the N-linked glycosylation processing pathway [6], and can be diagnosed by glycosylation analysis of serum glycoproteins [7], such as transferrin and haptoglobin. Significant increases in fucose levels and oligosaccharide branches in haptoglobin have been found to be associated with ovarian cancer [8,9], lung cancer [10–12], pancreatic cancer [13] and hepatocellular carcinoma [14]. Changes in glycosylation are also found in acute-phase proteins, such as alpha-1-acid glycoprotein and ceruloplasmin, in lung cancer [15]. These findings suggest the potential of the glycosylation analysis of serum glycoproteins in diagnosis of some diseases and an investigation of new biomarkers. At present the glycosylation of each protein is examined individually, therefore simultaneous analysis of serum glycoproteins has been required for rapid diagnosis with a limited amount of sample.

Mass spectrometry (MS) is known as a powerful tool for the glycosylation analysis of serum proteins. For the glycosylation analysis of serum glycoproteins, the enrichment of glycopeptides by lectin-affinity or hydrophilic chromatography is useful due to their low ionization efficiency, ionization suppression effects, and microheterogeneity [16–19]. There are still concerns about the loss of some glycopeptides during the preparation procedure, biased recoveries toward certain glycan structures, and low reproducibility of recovery. Liquid chromatography/mass spectrometry (LC/MS) is effective for the separation of glycopeptides and for the simultaneous glycosylation analysis.

Abbreviations: ESI, electrospray ionization; Fuc, fucose; GlcNAc, N-acetylglucosamine; Hex, hexose; HexNAc, N-acetylhexosamine; HPLC, high-performance liquid chromatography; IgG, immunoglobulin G; MS, mass spectrometry; MS/MS, tandem mass spectrometry; NeuAc, N-acetylneuraminic acid; Qq-TOF, quadrupole–quadrupole time-of-flight mass spectrometry; TIC, total ion chromatogram; EIC, extracted ion chromatogram.

* Corresponding author. Tel.: +81 3 3700 9074; fax: +81 3 3700 9084.
E-mail addresses: harazono@nihs.go.jp (A. Harazono), nana@nihs.go.jp (N. Kawasaki).

1570-0232/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.jchromb.2008.05.006

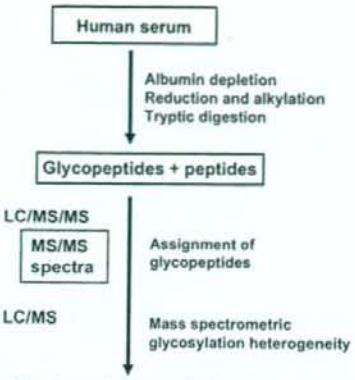

Fig. 1. Strategy for glycosylation analysis of abundant glycoproteins in serum. Human serum in which albumin was roughly removed was reduced and alkylated at cysteine residues. A mixture of peptides resulting from trypsin digestion was subjected to LC/MS/MS and LC/MS. Glycopeptides were assigned by elucidating MS/MS spectra (database searching). Mass spectrometric heterogeneity at each glycosylation site was analyzed by an additional LC/MS.

Recent progress in MS/MS and multiple-stage MS (MS^n) of glycopeptides allows for the characterization of both peptide and oligosaccharide moieties based on fragment ions [17,20–27]. Previously it was demonstrated that the Qq-TOF type mass spectrometer provides highly abundant carbohydrate-related ions at lower *m/z* values such as *m/z* 204 $[HexNAc+H]^+$ and 366 $[HexHexNAc+H]^+$, glycopeptide-related ions with sequentially lost saccharide units, including $[peptide+H]^+$ and $[peptide+GlcNAc+H]^+$ at higher *m/z* values, and *b*- and *y*-ions derived from peptide backbone [20,23,26,28]. These fragment ions could be used in database search to deduce peptide of glycopeptide.

In this study we demonstrated LC/MS(/MS) of human serum digest for the simultaneous glycosylation analysis of abundant serum proteins. Fig. 1 shows the strategy for the glycosylation analysis. Human serum, the albumin of which was depleted, was carboxymethylated and digested with trypsin. LC/MS/MS of the digest was performed by using the LC/Qq-TOF MS instrument. Glycopeptide ions were basically assigned by database searching with modified peak-list text files. Mass spectrometric heterogeneity at each glycosylation site was analyzed by an additional LC/MS, in which the acquisition of MS/MS was not allowed. By LC/MS of albumin-depleted human serum digest, we were successful in the site-specific glycosylation analysis of abundant serum glycoproteins.

2. Experimental

2.1. Materials

Pooled normal human serum was purchased from Sigma (St. Louis, MO, USA). Human haptoglobin and polyclonal immunoglobulin G, which were purified from normal human serum, were purchased from Calbiochem (San Diego, CA, USA). Modified trypsin was purchased from Promega (Madison, WI, USA). The water used was obtained from a Milli-Q water system (Millipore, Bedford, MA). All other reagents were of the highest quality available.

2.2. Sample preparation

Human serum (5 μl) was depleted of albumin using the Montage Albumin Depletion Kit (Millipore, Bedford, MA, USA) according to the manufacturer's protocol. Lyophilized albumin-depleted sample and each of the glycoproteins (100 μg) were dissolved in 50 μl of 0.5 M Tris–HCl buffer (pH 8.5) that contained 7 M guanidine hydrochloride and 5 mM EDTA. After the addition of 2 μl of 1 M dithiothreitol, the mixture was incubated for 30 min at 56 °C. Then, 4.7 μl of 1 M sodium monoiodoacetate was added, and the resulting mixture was incubated for 30 min at room temperature in the dark. The reaction mixture was applied to a PD-10 column (GE Healthcare, Little Chalfont, UK) to remove the reagents, and a fraction of the carboxymethylated proteins was dried. The sample was redissolved in 50 μl of 50 mM Tris–HCl buffer (pH 8.0). An aliquot of 1 μl of modified trypsin prepared as 1 μg/μl was added, and then the mixtures were incubated for 12 h at 37 °C. The enzyme digestions were stopped by boiling for 10 min and stored at −20 °C before analysis.

2.3. LC/MS and LC/MS/MS

The tryptic digests corresponding to 0.01–0.3 μl of human serum or a tryptic digest of purified commercially available glycoprotein (0.1–1.0 μg) was loaded onto a nanotrap (AMR Inc., Tokyo, Japan). After a wash with 10 μl of 2% (v/v) acetonitrile containing 0.1% (v/v) TFA, the trapping column was switched into line with the column. HPLC was performed on a Paradigm MS 4 (Michrome BioResources, Auburn, CA, USA) equipped with a MonoCap High Resolution 750 column (0.2 mm × 750 mm, GL Sciences Inc., Tokyo, Japan) at a flow rate of about 2 μl/min. 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). Samples were eluted with 5% of B for 2.5 or 5.0 min followed by a linear gradient from 5 to 90% of pump B in 85 min or by linear gradients from 5 to 25% for 80 min, 25–45% for the next 60 min, 45–65% for the next 40 min and 60–90% for the next 20 min (total 205 min).

Mass spectrometric analyses were performed using a QSTAR Pulsar i Qq-TOF mass spectrometer (AB/MDS Sciex, Toronto, Canada) equipped with a nano-electrospray ion source. The mass spectrometer was operated in positive ion mode. The nano-spray voltage was set at 1700 V. Mass spectra were acquired over *m/z* 1000–2000 for MS, and *m/z* 100–2000 for MS/MS. After every regular MS acquisition, MS/MS acquisitions were performed against the top two multiply charged ions by a data-dependent acquiring method. The precursor ions with the same *m/z* as previously acquired were excluded for 60 or 90 s. The collision energy was varied between 30 and 70 eV depending on the size and charge of the molecular ion. The accumulation time of the spectra was 1.0 s for MS, and 2.0 or 5.0 s for MS/MS. All signals were monoisotopically resolved.

2.4. Assignment of glycopeptide peaks by database search

Detection and assignment of glycopeptide ions from LC/ESI MS/MS data were performed by elucidating MS/MS spectra or database search. Briefly, glycopeptide ions were selected manually based on presence of oligosaccharide oxonium ions such as *m/z* 204 and 366 in their MS/MS spectra. The information of *m/z* values and charge states of peptides in the glycopeptides was deduced by sequential degradation pattern at N-glycan core structure in their MS/MS spectra. The MS/MS spectra of glycopeptides were converted to peak-list text files, and then oligosaccharide-related ions (*m/z* 168, 186, 204, 274, 292 and 366 or ions under *m/z* 370)

and the ions larger than peptide ion were deleted. Modified peak-list text files were submitted to against the nonredundant human Swiss–Prot protein database (version 48.2) using Mascot search engine with following parameters: a specified trypsin enzymatic cleavage with two possible missed cleavage, peptide tolerance of 1.2 Da, fragment ion tolerance of 0.8 Da, and variable modifications of cysteine (carboxymethylation) or cystein (carboxymethylation) and methionine (oxidation). Suggested peptides were validated by manual inspection of the spectra, and the presence of more than four consecutive fragments of amino acid sequence was used as criteria for peptide identification.

3. Results

3.1. Locating glycopeptides in the chromatogram

Mass spectrometric glycosylation analysis of human serum was performed by LC/Qq-TOF MS of tryptic digest using in a positive ion mode. In this method, all serum glycoproteins should be completely digested by trypsin. When the tryptic digest was subject to LC/MS/MS with the MS range m/z 400–2000, results of Mascot database search using 3 missed cleavage sites suggested that most peptides were completely digested (missed cleavage <1) and few incompletely digested peptides (missed cleavage 3) were present. Many missed cleavage sites were present at N- or C-terminal, or adjacent to two or more acidic amino acid residues (D, E or carboxymethylated C) (data not shown). Fig. 2A shows the total ion chromatogram (TIC) obtained by LC/MS/MS with MS range m/z 1000–2000 of tryptic digest (corresponding to approximately 0.3 μl of serum) using a reversed phase MonoCap High Resolution 750 column (0.2 mm × 750 mm) with a gradient of 5–90% of B in 205 min. In order to locate the glycopeptide peaks and determine m/z and charge state, the intensity of the oxonium HexNAc$^+$ (m/z 204.05–204.15) that arose by data-dependent MS/MS was depicted as the extracted ion chromatogram (Fig. 2B). We confirmed that most of these MS/MS spectra were of glycopeptides by the presence of abundant carbohydrate-derived ions, such as m/z 204 ([HexNAc+H]$^+$), 186 ([HexNAc+H−H_2O]$^+$), 292 ([NeuAc+H]$^+$), 274 ([NeuAc+H−H_2O]$^+$) and 366 ([HexHexNAc+H]$^+$).

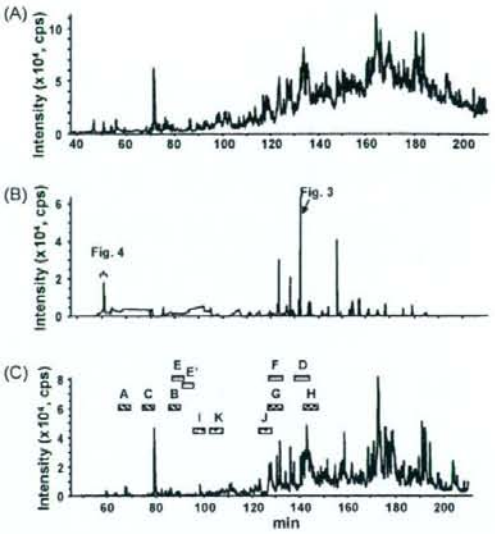

Fig. 2. LC/MS/MS and LC/MS of tryptic digest of human serum. (A) TIC (m/z 1000–2000) obtained by the LC/MS/MS. (B) EIC (m/z 204.05–204.15) obtained by the data-dependent MS/MS. (C) TIC obtained by the additional LC/MS in which data-dependent MS/MS was not allowed. Peak assignment: A, IgG1; B, IgG2; C, IgG3/IgG4; D–F, haptoglobin; G and H, transferrin; I–K, ceruloplasmin. Mass spectra of fractions A–K are shown in Fig. 7.

3.2. Assignment of glycopeptide peaks by a database search

Glycopeptides were assigned by manual database searching. As a representative example, the MS/MS spectrum acquired from [M+4H]$^{4+}$ (m/z 1221.8 (4+)) at 133 min is shown in Fig. 3. There are some abundant ions derived from carbohydrates, such as m/z 204, 186, 292, 274 and 366 in the lower m/z region. Degradation pattern and mass difference of 203 u between the fragment ions at m/z 1340.2 (2+) and those at 1441.7 (2+) in the higher m/z region suggests that the ions are [peptide+2H]$^{2+}$ and [peptide+HexNAc+2H]$^{2+}$, respectively. Based on these m/z values the molecular mass of

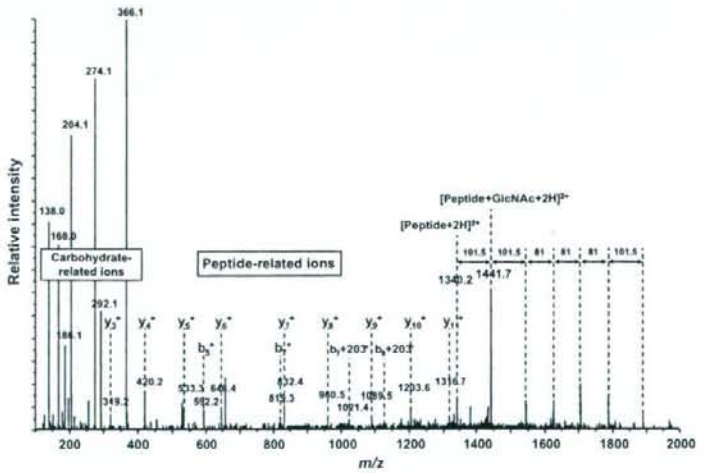

Fig. 3. MS/MS spectrum acquired from m/z 1221.8 (4+) by data-dependent LC/MS/MS of trypsin-digested human serum. Mascot database search using m/z 1340.2 (2+) of peptide and fragment ions (m/z 370–1300) suggested peptide sequence MVSHHN^{184}LTTGATLINEQWLLTTAK in haptoglobin (P00738).

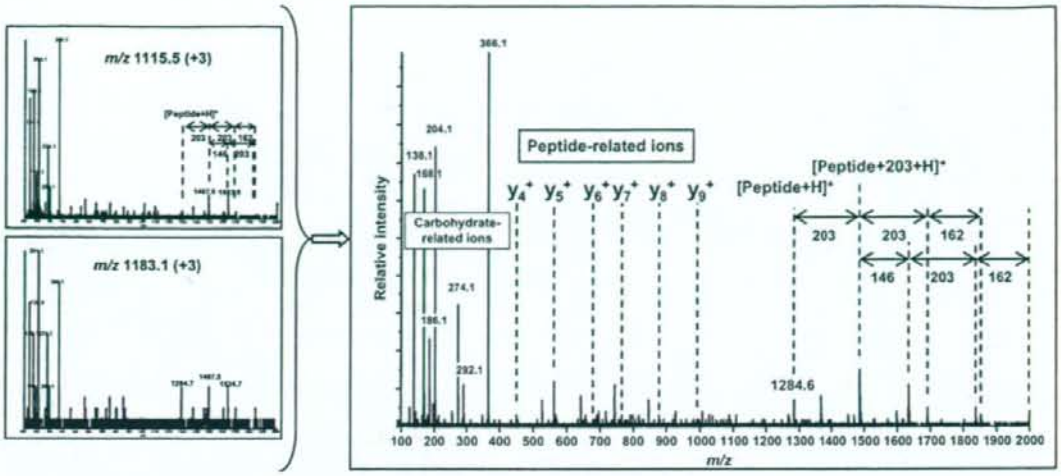

Fig. 4. Integrated MS/MS spectrum of *m/z* 1115.5 (3+) and 1183.1 (3+) at 52–53 min that show similar fragment patterns. Mascot database search using *m/z* 1284.5 (1+) of peptide and fragment ions (*m/z* 370–1280) suggested YKN[46]NSDISSTR in Ig mu chain C region (P01871).

the peptide was calculated to be 2678.4. For the peptide identification, a database search requires the *m/z* values and charge state of the peptide precursor ions and fragment ions but not of the carbohydrate- and glycopeptide-related ions. We deleted the carbohydrate-related ions in the lower *m/z* region (under *m/z* 370) and the glycopeptide-related ions in the higher *m/z* region (over *m/z* 1340) from the peak-list text files, and then submitted the modified peak-list text files for a Mascot database search of the human Swiss–Prot database with 1 missed cleavage, peptide tolerance of 1.2 Da, fragment ion tolerance of 0.8 Da and variable modifications of cysteine (carboxymethylation). The peptide suggested was MVSHHN[184]LTTGATLINEQWLLTTAK in human haptoglobin (P00738). As shown in Fig. 3, many ions were consistent with *b*- and *y*-series peptide fragment ions derived form MVSHHNLTTGATLINEQWLLTTAK. The molecular mass of the carbohydrate moiety was calculated to be 2204.7, which suggests the carbohydrate composition of $HexNAc_4Hex_5NeuAc_2$.

3.3. Assignment of glycopeptide peaks by a database search with integrated spectra

Glycopeptides that have the same peptide backbone show quite similar fragment patterns in the case of Qq-TOF MS. When glycopeptides showed insufficient peptide fragment ions in the CID-MS/MS spectra due to low peak intensity, we integrated the similar MS/MS spectra into one spectrum, and the integrated spectrum was submitted for a database search. As a representative example, the spectrum obtained by integrating two spectra of *m/z* 1115.5 (3+) and 1183.1 (3+) acquired around 60 min is shown in Fig. 4. Mascot database search using the information of *m/z* 1284.5 (1+) of peptide which was deduced by sequential degradation pattern at N-glycan core structure, and modified peak-list text files between *m/z* 370 and 1250 suggests that the peptide moiety is YKN[46]NSDISSTR in Ig mu chain C region (P01871).

By elucidating MS/MS spectra, 19 tryptic glycopeptides (20 N-glycosylation sites) in 14 glycoproteins were determined (Table 1). The ions, which were confirmed as glycopeptides by data-dependent MS/MS, were underlined. Other glycoforms, whose MS/MS spectra were not acquired, were assigned based on their mass difference of saccharide units from characterized glycopeptides. Since high intensity glycopeptide ions showed high quality of MS/MS spectra and were subjected to data-dependent MS/MS several times, many of them could be assigned. Low intensity glycopeptide ions showed poor MS/MS spectra for detection of peptide fragment ions, about 20% of MS/MS spectra of glycopeptides could not be assigned (data not shown).

3.4. Confirmation of glycopeptide peaks using commercially available glycoproteins

We conducted peptide mapping of commercially available polyclonal IgG and haptoglobin, and then *m/z* values and charge states of the glycopeptides were used for confirmation of assignment of glycopeptides and assignment of undetected glycopeptides. Glycosylation data of ceruloplasmin in previous report [28] was also utilized.

Tryptic digest (0.2 μg and 0.4 μg) of commercially available human polyclonal IgG was analyzed by LC/ESI MS/MS at *m/z* 400–2000 and 1000–2000 with a gradient of 5–90% of B in 85 min. The MS data were submitted for database searching against the human Swiss–Prot database using the computer program Mascot. Polypeptides of IgG heavy chain C region of IgG1 (P01857), IgG2 (P01859), IgG3 (P01860) and IgG4 (P01861) and light chain C region of Kappa (P01834) and Lambda (P01842) chain and other proteins were identified (data not shown). Fig. 5A and A′ show TIC of LC/MS/MS at *m/z* 1000–2000 of polyclonal IgG and EIC of data-dependent MS/MS at *m/z* 204.05–205.15, respectively. It was found that glycopeptide ions were eluted at 7–12 min (fraction A), 15–17 min (fraction C) and 18–21 min (fraction B) based on the presence of the oligosaccharide-related ions in their MS/MS spectra and mass differences of saccharide units. MS/MS spectra after 25 min were not of glycopeptides. The glycopeptide peaks from fraction A and fraction B were assigned as the glycopeptides containing Fc-glycosylation site in IgG1 (EEQYNSTYR) and IgG2 (EEQFNSTFR) based on data-dependent MS/MS spectra, respectively (data not shown). Data-dependent MS/MS spectra from fraction C suggested molecular mass of 1171.5 Da for the peptide, but could not suggest amino acid sequence due to low abundance of peptide fragment ions (data not shown). Based on the molecular mass of the peptide, the glycopeptide peaks from fraction C would be EEQYNSTFR from IgG3 (CAA67886) and/or EEQFNSTYR from IgG4 (P01861), which are attached to core-fucosylated agalacto-

Table 1
Summary of analysis of serum glycoproteome with higher ion intensities

Glycopeptide					Oligosaccharide	Protein (Protein ID)	Theoretical MW
Retention time (min)	m/z[a]	Charge	Observed MW	Relative peak intensity[b]	Observed MW	Glycopeptide Peptide sequence Deduced oligosaccharide composition[c]	Peptide Oligosaccharide
						Ig gamma-1 chain C region (P01857)	1[d]
						EEQYNSTYR[A]	1188.50
67.3	1479.6	2+	2957.1	13.1	1768.6	[HexNAc]4[Hex]5[Fuc]1	1768.64
67.3	1297.0	2+	2592.0	3.0	1403.5	[HexNAc]3[Hex]4[Fuc]1	1403.51
67.4, 67.6	1398.5	2+	2795.1	33.1	1606.5	[HexNAc]4[Hex]4[Fuc]1	1606.59
67.4, 67.7	1216.0	2+	2429.9	4.3	1241.4	[HexNAc]3[Hex]3[Fuc]1	1241.45
67.7	1317.5	2+	2633.0	27.8	1444.5	[HexNAc]4[Hex]3[Fuc]1	1444.53
67.9	1500.1	2+	2998.1	2.3	1809.6	[HexNAc]5[Hex]4[Fuc]1	1809.67
	1000.4	3+	2998.0				
67.9	1581.1	2+	3160.2	0.2	1971.7	[HexNAc]5[Hex]5[Fuc]1	1971.72
68.0	1406.6	2+	2811.1	0.9	1622.6	[HexNAc]4[Hex]5	1622.58
68.2, 68.4	1325.5	2+	2649.0	2.8	1460.5	[HexNAc]4[Hex]4	1460.53
68.2	1419.1	2+	2836.1	2.9	1647.6	[HexNAc]5[Hex]3[Fuc]1	1647.61
68.5	1244.5	2+	2486.9	1.4	1298.4	[HexNAc]4[Hex]3	1298.48
69.1	1625.1	2+	3248.1	2.2	2059.6	[HexNAc]4[Hex]5[NeuAc]1[Fuc]1	2059.73
	1083.7	3+	3248.1				
69.6	1544.1	2+	3086.2	0.5	1897.6	[HexNAc]4[Hex]4[NeuAc]1[Fuc]1	1897.68
						EEQYNSTYRVVSVLTVLHQDWLNGK[E]	2977.49
161.9, 162.2	1147.0	4+	4584.1		1606.6	[HexNAc]4[Hex]4[Fuc]1	1606.59
162.4	1106.5	4+	4422.1		1444.6	[HexNAc]4[Hex]3[Fuc]1	1444.53
						EEQYNSTYRVVSVLTVLHQDWLNGKEYK[E]	3397.69
156.4	1034.3	5+	5166.4		1768.7	[HexNAc]4[Hex]5[Fuc]1	1768.64
156.6, 157.1	1001.8	5+	5004.0		1606.3	[HexNAc]4[Hex]4[Fuc]1	1606.59
						Ig gamma-2 chain C region (P01859)	1[d]
						EEQFNSTFR[B]	1156.51
85.4	1463.6	2+	2925.1	6.3	1768.6	[HexNAc]4[Hex]5[Fuc]1	1768.64
85.5	1281.0	2+	2560.0	1.4	1403.5	[HexNAc]3[Hex]4[Fuc]1	1403.51
85.7, 86.3	1382.5	2+	2763.1	19.4	1606.5	[HexNAc]4[Hex]4[Fuc]1	1606.59
85.7, 86.4	1200.0	2+	2397.9	2.8	1241.4	[HexNAc]3[Hex]3[Fuc]1	1241.45
85.7	1565.1	2+	3128.2	0.0	1971.7	[HexNAc]5[Hex]5[Fuc]1	1971.72
86.0	1484.1	2+	2966.2	1.0	1809.6	[HexNAc]5[Hex]4[Fuc]1	1809.67
86.5	1301.5	2+	2601.0	21.8	1444.5	[HexNAc]4[Hex]3[Fuc]1	1444.53
86.5	1390.5	2+	2779.0	0.0	1622.5	[HexNAc]4[Hex]5	1622.58
86.9	1403.0	2+	2804.1	1.8	1647.5	[HexNAc]5[Hex]3[Fuc]1	1647.61
87.0, 87.5	1309.5	2+	2617.0	0.1	1460.5	[HexNAc]4[Hex]4	1460.53
87.6	1228.5	2+	2454.9	0.2	1298.4	[HexNAc]4[Hex]3	1298.48
89.4	1609.1	2+	3216.2	1.6	2059.7	[HexNAc]4[Hex]5[NeuAc]1[Fuc]1	2059.73
	1073.1	3+	3216.2				
90.0	1528.1	2+	3054.2	1.5	1897.6	[HexNAc]4[Hex]4[NeuAc]1[Fuc]1	1897.68
	1019.1	3+	3054.1				
						Gamma 3 immunoglobulin constant heavy chain (CAA67886)	2[d]
						EEQYNSTFR[C]	1172.51
						Ig gamma-4 chain C region (P01861)	1[d]
						EEQFNSTYR[C]	1172.51
76.4	1471.6	2+	2941.1	1.0	1768.6	[HexNAc]4[Hex]5[Fuc]1	1768.64
76.5	1289.0	2+	2576.0	0.2	1403.5	[HexNAc]3[Hex]4[Fuc]1	1403.51
76.6, 76.8	1390.6	2+	2779.1	3.4	1606.6	[HexNAc]4[Hex]4[Fuc]1	1606.59
76.5, 76.8	1208.0	2+	2413.9	0.4	1241.4	[HexNAc]3[Hex]3[Fuc]1	1241.45
76.7	1492.1	2+	2982.1	0.2	1809.6	[HexNAc]5[Hex]4[Fuc]1	1809.67
76.9	1309.5	2+	2617.0	3.6	1444.5	[HexNAc]4[Hex]3[Fuc]1	1444.53
77.0	1398.6	2+	2795.1	0.1	1622.6	[HexNAc]4[Hex]5	1622.58
76.9	1317.5	2+	2633.0	0.1	1460.5	[HexNAc]4[Hex]4	1460.53
77.0	1411.1	2+	2820.1	0.3	1647.6	[HexNAc]5[Hex]3[Fuc]1	1647.61
77.4	1236.4	2+	2470.8	0.1	1298.3	[HexNAc]4[Hex]3	1298.48
78.5	1617.1	2+	3232.1	0.3	2059.6	[HexNAc]4[Hex]5[NeuAc]1[Fuc]1	2059.73
79.1	1536.1	2+	3070.1	0.1	1897.6	[HexNAc]4[Hex]4[NeuAc]1[Fuc]1	1897.68
						Haptoglobin (P00738)	4[d]
						MVSHHNLTTGATLINEQWLLTTAK[D]	2678.39
137.9	1531.7	3+	4592.1		1913.7	[HexNAc]4[Hex]5[NeuAc]1	1913.68
	1149.0	4+	4592.0	30.1	1913.6		

Table 1 (*Continued*)

Glycopeptide					Oligosaccharide	Protein (Protein ID)	Theoretical MW
Retention time (min)	m/z[a]	Charge	Observed MW	Relative peak intensity[b]	Observed MW	Glycopeptide Peptide sequence Deduced oligosaccharide composition[c]	Peptide Oligosaccharide
141.4	1221.8	4+	4883.1	88.7	2204.7	[HexNAc]4[Hex]5[NeuAc]2	2204.77
						M(O)VSHHNLTTGATLINEQWLLTTAK	2694.38
137.3	1153.0	4+	4608.1	28.9	1913.7	[HexNAc]4[Hex]5[NeuAc]1	1913.68
140.5, 141.1	1225.8	4+	4899.1	64.3	2204.7	[HexNAc]4[Hex]5[NeuAc]2	2204.77
	1634.1	3+	4899.1		2204.8		
						NLFLNHSENATAK[E]	1457.73
86.3	1395.0	4+	5576.0	0.1	4118.3	[HexNAc]8[Hex]10[NeuAc]3	4118.45
87.0	1650.3	4+	6597.3	0.3	5139.6	[HexNAc]10[Hex]12[NeuAc]4	5139.81
87.6	1595.6	4+	6378.3	0.9	4920.6	[HexNAc]9[Hex]11[NeuAc]4[Fuc]1	4920.73
87.9	1559.1	4+	6232.3	2.7	4774.5	[HexNAc]9[Hex]11[NeuAc]4	4774.68
88.6	1504.3	4+	6013.0	0.1	4555.3	[HexNAc]8[Hex]10[NeuAc]4[Fuc]1	4555.60
88.9	1467.8	4+	5867.1	4.1	4409.4	[HexNAc]8[Hex]10[NeuAc]4	4409.54
90.4	1759.6	4+	7034.5	0.2	5576.8	[HexNAc]10[Hex]12[NeuAc]5[Fuc]1	5576.96
90.7	1723.1	4+	6888.5	0.5	5430.8	[HexNAc]10[Hex]12[NeuAc]5	5430.90
91.5	1668.3	4+	6669.4	0.3	5211.6	[HexNAc]9[Hex]11[NeuAc]5[Fuc]1	5211.83
91.7	1631.8	4+	6523.3	0.4	5065.6	[HexNAc]9[Hex]11[NeuAc]5	5065.77
87.8	1124.7	3+	3371.2	-	1913.5	[HexNAc]4[Hex]5[NeuAc]1	1913.68
91.6	1221.7	3+	3662.2	-	2204.5	[HexNAc]4[Hex]5[NeuAc]2	2204.77
						VVLHPNYSQVDIGLIK[F]	1794.00
127	1358.6	3+	4072.8	2.2	2278.8	[HexNAc]5[Hex]6[NeuAc]1	2278.81
128.2	1236.9	3+	3707.7	3.1	1913.7	[HexNAc]4[Hex]5[NeuAc]1	1913.68
131.4	1455.6	3+	4363.8	4.8	2569.8	[HexNAc]5[Hex]6[NeuAc]2	2569.90
	1092.0	4+	4363.8				
131.8	1333.9	3+	3998.7	89.2	2204.7	[HexNAc]4[Hex]5[NeuAc]2	2204.77
	1000.7	4+	3998.7		2204.7		
134.1	1552.7	3+	4655.0	7.5	2860.9	[HexNAc]5[Hex]6[NeuAc]3	2861.00
134.1	1164.7	4+	4654.9		2860.9		
						Transferrin (P02787)	2[d]
						CGLVPVLAENYNK[G]	1476.73
126.1	1252.8	3+	3755.5	0.9	2278.8	[HexNAc]5[Hex]6[NeuAc]1	2278.81
127.0	1131.1	3+	3390.4	1.7	1913.7	[HexNAc]4[Hex]5[NeuAc]1	1913.68
129.8	1349.9	3+	4046.7	1.6	2569.9	[HexNAc]5[Hex]6[NeuAc]2	2569.90
130.6	1228.2	3+	3681.5	46.8	2204.7	[HexNAc]4[Hex]5[NeuAc]2	2204.77
133.1	1446.9	3+	4337.8	0.8	2861.0	[HexNAc]5[Hex]6[NeuAc]3	2861.00
						QQQHLFGSNVTDCSGNFCLFR[H]	2516.08
143.8	1623.3	3+	4866.9	4.1	2350.8	[HexNAc]4[Hex]5[NeuAc]2[Fuc]1	2350.83
143.9	1181.2	4+	4720.9		2204.8	[HexNAc]4[Hex]5[NeuAc]2	2204.77
	1574.6	3+	4720.9	50.5	2204.8		
146.0	1842.0	3+	5523.0	0.9	3007.0	[HexNAc]5[Hex]6[NeuAc]3[Fuc]1	3007.06
146.2	1793.4	3+	5377.3	1.6	2861.2	[HexNAc]5[Hex]6[NeuAc]3	2861.00
						Ceruloplasmin (P00450)	4[d]
						EHEGAIYPDNTTDFQR[I]	1891.83
95.6	1415.2	3+	4242.7	0.1	2350.8	[HexNAc]4[Hex]5[NeuAc]2[Fuc]1	2350.83
	1061.7	4+	4242.6		2350.8		
96.3	1366.5	3+	4096.6	6.0	2204.8	[HexNAc]4[Hex]5[NeuAc]2	2204.77
	1025.2	4+	4096.7		2204.9		
98.1, 98.5	1633.9	3+	4898.7	0.4	3006.9	[HexNAc]5[Hex]6[NeuAc]3[Fuc]1	3007.06
	1226.0	4+	4899.9		3008.0		
98.8	1585.2	3+	4752.7	0.4	2860.9	[HexNAc]5[Hex]6[NeuAc]3	2861.00
	1189.2	4+	4752.6		2860.8		
						ENLTAPGSDSAVFFEQGTTR[J]	2125.99
127.0	1493.2	3+	4476.6	0.0	2350.6	[HexNAc]4[Hex]5[NeuAc]2[Fuc]1	2350.83
127.4	1444.5	3+	4330.6	2.8	2204.6	[HexNAc]4[Hex]5[NeuAc]2	2204.77
129.3	1663.3	3+	4987.0	0.2	2861.0	[HexNAc]5[Hex]6[NeuAc]3	2861.00
						ELHHLQEQNVSNAFLDK[K]	2021.00
104.3	1093.9	4+	4371.7	1.4	2350.7	[HexNAc]4[Hex]5[NeuAc]2[Fuc]1	2350.83
	1458.3	3+	4371.8	0.2	2350.8		
105.2	1057.4	4+	4225.7	4.5	2204.7	[HexNAc]4[Hex]5[NeuAc]2	2204.77
	1409.6	3+	4225.6	1.2	2204.6		

Table 1 (*Continued*)

Glycopeptide					Oligosaccharide	Protein (Protein ID)	Theoretical MW
Retention time (min)	m/z[a]	Charge	Observed MW	Relative peak intensity[b]	Observed MW	Glycopeptide Peptide sequence Deduced oligosaccharide composition[c]	Peptide Oligosaccharide
106.6	1294.6	4+	5174.2	1.2	3153.2	[HexNAc]5[Hex]6[NeuAc]3[Fuc]2	3153.12
106.8, 107.4	1258.0	4+	5027.9	2.0	3006.9	[HexNAc]5[Hex]6[NeuAc]3[Fuc]1	3007.06
107.7	1221.5	4+	4881.8	1.9	2860.8	[HexNAc]5[Hex]6[NeuAc]3	2861.00
						Alpha-1-antitrypsin (P01009)	3[d]
						YLGNATAIFFLPDEGK	1754.89
154.6	1369.6	3+	4105.7	2.2	2350.8	[HexNAc]4[Hex]5[NeuAc]2[Fuc]1	2350.83
154.8	1320.9	3+	3959.7	140.6	2204.8	[HexNAc]4[Hex]5[NeuAc]2	2204.77
						Alpha-2-HS-glycoprotein (P02765)	2[d]
						VCQDCPLLAPLNDTR	1772.81
136.9	1326.9	3+	3977.7		2204.8	[HexNAc]4[Hex]5[NeuAc]2	2204.77
						Alpha-2-macroglobulin (P01023)	8[d]
						VSNQTLSLFFTVLQDVPVR	2162.17
187.9, 188.8	1505.3	3+	4512.7	5.1	2350.6	[HexNAc]4[Hex]5[NeuAc]2[Fuc]1	2350.83
188.3	1456.7	3+	4367.0	22.5	2204.8	[HexNAc]4[Hex]5[NeuAc]2	2204.77
						Beta-2-glycoprotein 1 (P02749)	4[d]
						VYKPSAGNNSLYR	1467.75
83.5	1273.8	3+	3818.5	1.5	2350.8	[HexNAc]4[Hex]5[NeuAc]2[Fuc]1	2350.83
83.6	1225.2	3+	3672.5	6.9	2204.8	[HexNAc]4[Hex]5[NeuAc]2	2204.77
85.2	1492.6	3+	4474.6	0.3	3006.9	[HexNAc]5[Hex]6[NeuAc]3[Fuc]1	3007.06
85.4	1443.9	3+	4328.6	0.5	2860.8	[HexNAc]5[Hex]6[NeuAc]3	2861.00
						LGNWSAMPSCK	1250.54
109.7	1152.8	3+	3455.3		2204.7	[HexNAc]4[Hex]5[NeuAc]2	2204.77
						Complement C3 (P01024)	3[d]
						TVLTPATNHMGNVTFTIPANR	2254.15
121.0	1265.9	3+	3794.6	5.4	1540.4	[HexNAc]2[Hex]7	1540.53
121.2	1211.8	3+	3632.5	47.8	1378.4	[HexNAc]2[Hex]6	1378.48
121.6	1157.8	3+	3470.4	10.2	1216.3	[HexNAc]2[Hex]5	1216.42
						Hemopexin (P02790)	5[d]
						SWPAVGNCSSALR	1404.65
115.3	1252.8	3+	3755.5	0.7	2350.8	[HexNAc]4[Hex]5[NeuAc]2[Fuc]1	2350.83
115.8	1204.1	3+	3609.4	10.3	2204.7	[HexNAc]4[Hex]5[NeuAc]2	2204.77
						ALPQPQNVTSLLGCTH	1735.86
115.8	1314.5	3+	3940.4	10.3	2204.5	[HexNAc]4[Hex]5[NeuAc]2	2204.77
						Ig alpha-1 chain C region (P01876)	2[d]
						Ig alpha-2 chain C region (P01877)	4[d]
						LSLHRPALEDLLLGSEANLTCTLTGLR	2963.58
165.2, 165.7	1157.8	4+	4627.2	8.0	1663.6	[HexNAc]5[Hex]4	1663.61
165.8	1117.3	4+	4465.0	15.4	1501.4	[HexNAc]5[Hex]3	1501.56
165.9	1046.0	4+	4180.0	5.6	1216.4	[HexNAc]2[Hex]5	1216.42
169.2	1220.3	4+	4877.2	48.8	1913.6	[HexNAc]4[Hex]5[NeuAc]1	1913.68
168.8, 169.4	1256.9	4+	5023.5	1.3	2059.9	[HexNAc]4[Hex]5[NeuAc]1[Fuc]1	2059.73
169.9	1179.8	4+	4715.3	4.0	1751.7	[HexNAc]4[Hex]4[NeuAc]1	1751.62
170.0	1169.6	4+	4674.2	5.1[f]	1710.6	[HexNAc]3[Hex]5[NeuAc]1	1710.60
169.0	1271.2	4+	5080.8	21.0[f]	2117.2	[HexNAc]5[Hex]5[NeuAc]1	2116.76
	1017.2	5+	5081.0		2117.4		
169.9	1230.4	4+	4917.6	11.8	1954.0	[HexNAc]5[Hex]4[NeuAc]1	1954.70
173.1	1293.2	4+	5168.8	9.0	2205.2	[HexNAc]4[Hex]5[NeuAc]2	2204.77
						PALEDLLLGSEANLTCTLTGLR[e]	2357.21
174.4	1287.2	3+	3858.6	14.3	1501.4	[HexNAc]5[Hex]3	1501.56
176.7	1424.6	3+	4270.9	65.8	1913.7	[HexNAc]4[Hex]5[NeuAc]1	1913.68
176.4	1492.5	3+	4474.5	20.5	2117.3	[HexNAc]5[Hex]5[NeuAc]1	2116.76
						Ig alpha-2 chain C region (P01877)	
						TPLTANITK	957.55
84.1	1006.8	3+	3017.2	3.7	2059.7	[HexNAc]4[Hex]5[NeuAc]1[Fuc]1	2059.73
	1509.6	2+	3017.2		2059.6		
84.1	1074.4	3+	3220.3	4.8	2262.8	[HexNAc]5[Hex]5[NeuAc]1[Fuc]1	2262.81
	1611.2	2+	3220.3		2262.8		
87.2	1103.8	3+	3308.3	1.3	2350.8	[HexNAc]4[Hex]5[NeuAc]2[Fuc]1	2350.83
						Ig mu chain C region (P01871)	5[d]
						YKNNSDISSTR[e]	1283.61

Table 1 (*Continued*)

Glycopeptide					Oligosaccharide	Protein (Protein ID)	Theoretical MW
Retention time (min)	m/z[a]	Charge	Observed MW	Relative peak intensity[b]	Observed MW	Glycopeptide Peptide sequence Deduced oligosaccharide composition[c]	Peptide Oligosaccharide
59.7	1115.5	3+	3343.4	40.1	2059.7	[HexNAc]4[Hex]5[NeuAc]1[Fuc]1	2059.73
60.3	1183.1	3+	3546.4	16.4	2262.8	[HexNAc]5[Hex]5[NeuAc]1[Fuc]1	2262.81

[a] Underlines indicated that these ions were assigned by elucidating data-dependent MS/MS of LC/ESI MS/MS of human serum digest.
[b] Centroid peak intensity (count per sec) in integrated MS spectra during glycopeptide eluting period.
[c] Oligosaccharide compositions were deduced from molecular weights.
[d] Number of potential N-glycosylation sites.
[e] Missed cleavage or unexpected digestion.
[f] Other ions with same m/z overlapped.
[A–K] Mass spectra were shown in Fig. 7A–K.
All masses are monoisotopic. Cysteine residue was carboxymethylated. Potential N-glycosylation sites were underlined. M(O), oxidized methionine; Fuc, fucose; Hex, hexose; HexNAc, N-acetylhexosamine; NeuAc, N-acetylneuraminic acid

or mono/diagalacto-biantennary complex-type glycan. Integrated mass spectra of fraction A, B and C were shown in Fig. 5B–D.

Haptoglobin has four potential N-glycosylation sites. We performed peptide mapping using a tryptic digest of haptoglobin under a chromatographic condition similar to that of human serum. Fig. 6A and A′ show TIC obtained by LC/MS/MS with mass range m/z 1000–2000 and EIC of data-dependent MS/MS at m/z 204.05–205.15, respectively. Glycopeptides for four potential glycosylation sites were assigned by elucidating MS/MS spectra (spectra were not shown). Glycopeptides of NLFLN^{207}HSEN^{211}ATAK containing two N-glycosylation sites were eluted in fraction E as two glycosylated forms (Fig. 6B) and fraction E′ as one glycosylated forms (Fig. 6C). The former glycosylated form was more abundant than the later form. These glycosylation sites could not be characterized separately by trypsin digestion. Glycopeptides of VVLHPN^{241}YSQVDIGLIK and MVSHHN^{184}LTTGATLINEQWLLTTAK were eluted in fractions F and D, respectively (Fig. 6D and E). From the molecular masses of oligosaccharides we inferred that a majority of oligosaccharides in haptoglobin are di-, tri-, and tetra-antennary forms and that some oligosaccharides were not fully saturated with NeuAc, and few glycans were fucosylated.

Using the data of relative retention times, accurate m/z values and charge states obtained by peptide mapping of commercially available glycoproteins, we confirmed already assigned glycopeptides and further assigned undetected glycopeptides (IgG3/IgG4 and two sites of ceruloplasmin), with the exceptions of one of the glycopeptides from ceruloplasmin, intensity of which was only noise levels.

3.5. Site-specific glycosylation analysis

To analyze the heterogeneity of glycosylation at each site, we performed an additional LC/MS in which switching to MS/MS was not allowed (Fig. 2C). Utilizing the information of retention time, accurate m/z and charge state of assigned glycopeptides by LC/MS/MS, corresponding glycopeptides were assigned in LC/MS data by mass chromatogram. When two or more glycoforms were detected, mass spectrometric heterogeneity was calculated using

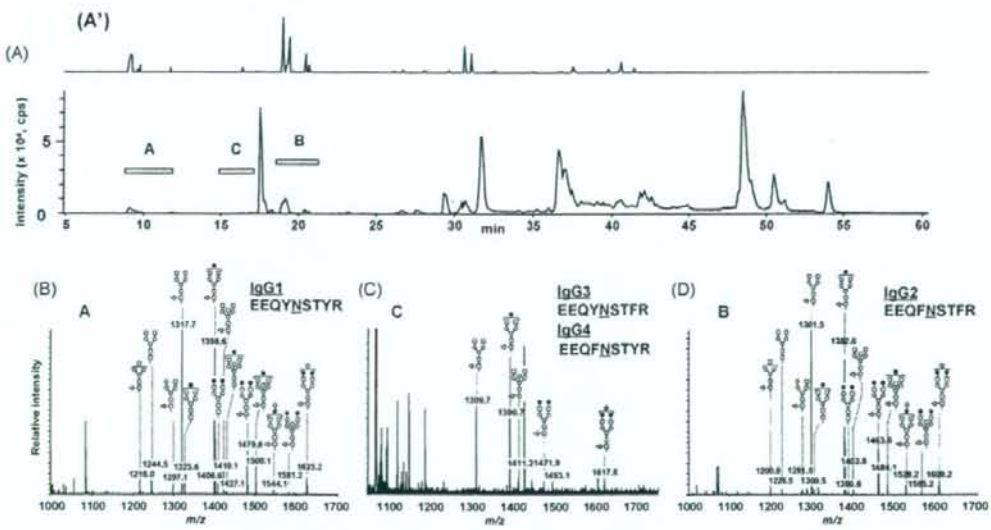

Fig. 5. Peptide map of commercially available human polyclonal IgG. (A) TIC (m/z 1000–2000) obtained by LC/MS/MS of trypsin-digested IgG. (A′) EIC (m/z 204.05–204.15) obtained by data-dependent MS/MS. (B) Mass spectrum of peak A, which was assigned as glycopeptides of EEQYNSTYR of IgG1 (P01857). (C) Mass spectrum of peak C, which would be glycopeptides of EEQYNSTFR of IgG3 (CAA67886) and/or EEQFNSTYR of IgG4 (P01861). (D) Mass spectrum of peak B, which was assigned as glycopeptides of EEQFNSTFR of IgG2 (P01859).

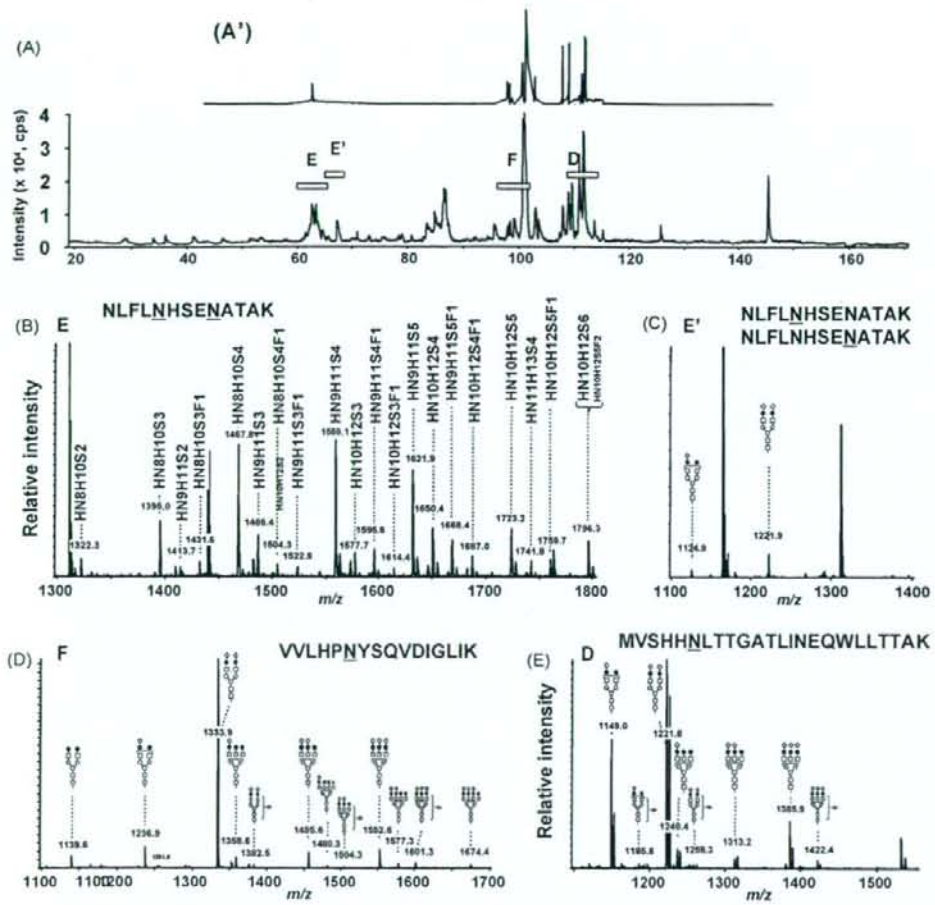

Fig. 6. Peptide map of commercially available human haptoglobin. (A) TIC (*m/z* 1000–2000) obtained by LC/MS/MS of trypsin-digested haptoglobin. (A') EIC (*m/z* 204.05–204.15) obtained by data-dependent MS/MS. (B) Mass spectrum of peak E, which was assigned as glycopeptides of NLFLN^{207}HSEN^{211}ATAK attached to two N-glycan. (C) Mass spectrum of peak E', which was assigned as glycopeptides of NLFLN^{207}HSEN^{211}ATAK attached to one N-glycan. (D) Mass spectrum of peak F, which was identified as VVLHPN^{241}YSQVDIGLIK. (E) Mass spectrum of peak D, which was identified as MVSHHN^{184}LTTGATLINEQWLLTTAK. H, hexose; HN, *N*-acetylhexosamine; S, *N*-acetylneuraminic acid; F, fucose.

integrated mass spectra during the periods eluting the glycopeptides with same peptide. In Fig. 7, we show integrated mass spectra of fraction A–K (Fig. 2C) as the mass spectrometric heterogeneity of glycosylation in IgG1 (Fig. 7A), IgG2 (Fig. 7B), IgG3/IgG4 (Fig. 7C), haptoglobin (Fig. 7D–F), transferrin (Fig. 7G and H) and ceruloplasmin (Fig. 7I–K). Centroid ion intensity (count/sec) of each glycopeptide at the most intense isotope distribution was used as relative peak intensity. The mass spectrometric heterogeneity of the Fc-glycosylation sites of IgG1 (Fig. 7A) and IgG2 (Fig. 7B) was consistent with those of the commercially available polyclonal IgG (Fig. 5B and D) and previous reports [29]. The glycosylation pattern of haptoglobin at each site was similar to that of the commercially available haptoglobin except that peak intensities of minor glycoforms were noise level (Figs. 6B–E and 7D–F). The glycosylation of transferrin (Fig. 7G and H) at each site was consistent with previous reports [29]. Three glycopeptides of the four expected ones derived from ceruloplasmin could be assigned on the chromatogram of the serum sample (Fig. 7I–K), and their glycosylation patterns were in agreement with those in our previous reports [28]. Table 1 summarized LC retention time, *m/z* and charge, relative peak intensities of assigned glycopeptides in LC/MS. No O-glycosylated peptides were detected in this study. It would be due to low amount of O-glycosylation in serum and huge sample complexity.

4. Discussion

Alteration of glycans in several serum glycoproteins is a potential marker for several diseases. Several glycomic approaches to the diagnosis using mass spectrometric techniques have been proposed. The most common procedure involves analyzing the liberated glycans by MALDI-TOF MS or LC/ESI-MS, but this method provides no information on the glycosylation sites or protein sources. Another approach involves mass spectrometric analysis of glycopeptides resulting from proteolytic digestion. The enrichment of glycopeptides is useful due to their low ionization efficiency, but loss of glycopeptides cannot be avoidable. In the present study, we performed LC/MS(/MS) with high resolution separation to obtain

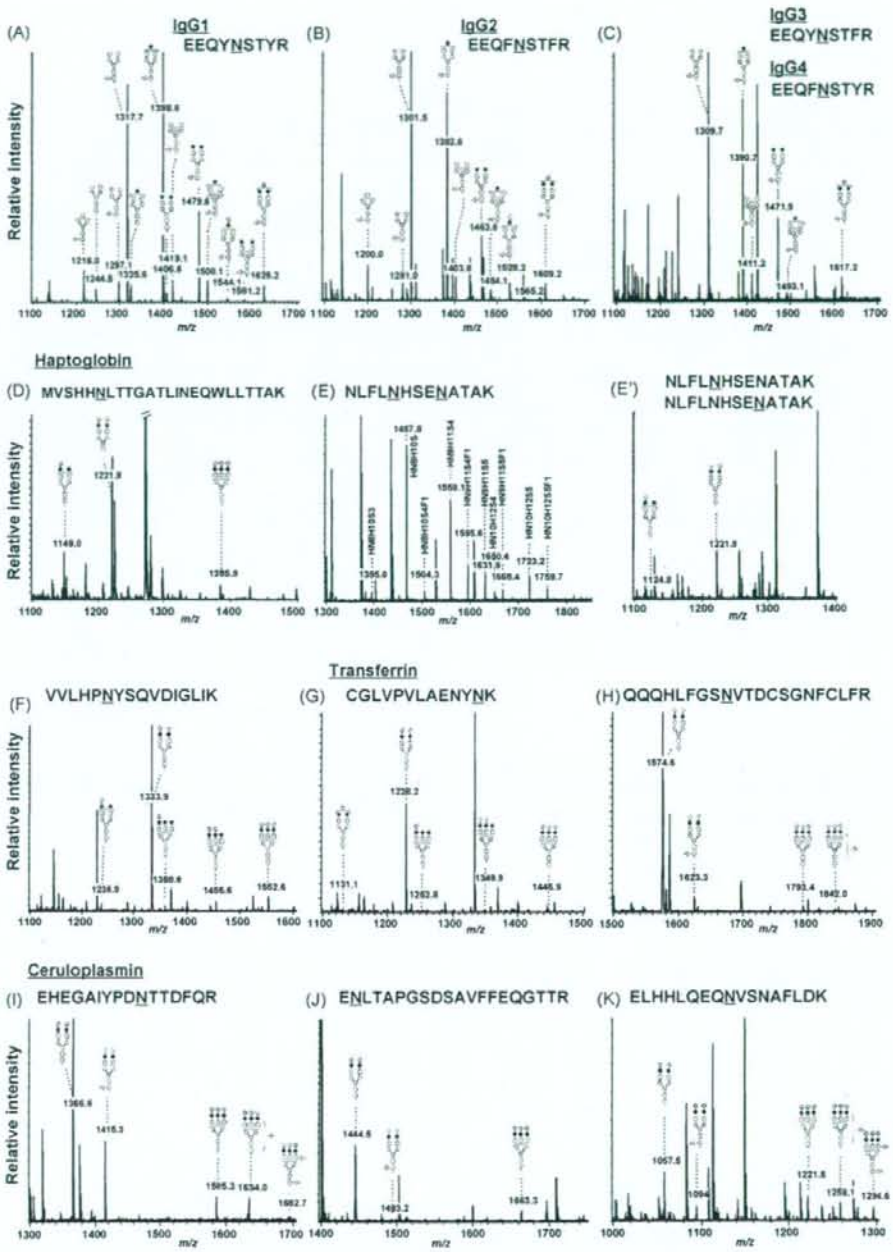

Fig. 7. Mass spectrometric heterogeneity of glycosylation in abundant serum glycoproteins. Integrated mass spectra obtained by the additional LC/MS of human serum digest. (A) IgG1; (B) IgG2; (C) IgG3/IgG4; (D–F) haptoglobin ((E) diglycosylated; (E'), monoglycosylated); (G and H) transferrin; (I–K) ceruloplasmin.

mass spectrometric glycosylation profiles at each glycosylation site of abundant glycoproteins in human serum.

MS/MS spectra are useful for detection and assignment of glycopeptide ions. Because MS/MS spectra of glycopeptide precursor ions have abundant carbohydrate B-ions, such as m/z 204 ([HexNAc+H]$^+$), and 366 ([HexHexNAc+H]$^+$), presence of these ions is a useful indication of the selection of glycopeptide precursor ions. MS/MS spectra of glycopeptide also contain ions of peptide and peptide plus glycans and several *b*- and *y*-series fragment ions of peptide backbone when using Qq-TOF mass spectrometer. These allow us to differentiate the glycopeptide ions with different peptide backbone and further to deduce peptide containing

N-linked glycosylation sites by database search. When MS/MS spectra of the glycopeptides obtained in a data-dependent manner were poor for detection of peptide fragment ions, improvement of MS/MS spectra quality by integrating several similar MS/MS spectra into one spectrum was effective. Composition of attached glycan can be deduced from molecular weight of glycan. Utilizing the data of site-specific glycosylation analysis of commercial glycoproteins (IgG, haptoglobin and ceruloplasmin) allowed us assign the corresponding glycopeptides in complex LC/MS(/MS) chromatogram.

We preliminary performed LC/MS/MS of serum tryptic digest to locate glycopeptides and assign by data-dependent MS/MS. Using LC retention time, accurate m/z and charge state of assigned glycopeptides, we successfully determined mass spectrometric heterogeneity of 23 glycosylation sites in 15 glycoproteins by LC/MS analyze using digest corresponding 0.3 μl of serum. Although there have been many reports on the analysis of human serum digest to show the glycosylation sites of abundant serum glycoproteins [30–33], less has been reported on their glycosylation. Glycopeptides detected in this study were those derived from glycoproteins which are present at about 0.2–5 mg/ml in human serum, and only glycopeptides with higher ionization efficiency were detected. Thus, it was suggested that detection limit of our method without sample enrichment procedure would be >0.2 mg/ml. It was thought that sample comlexicity, ionization suppression of low abundant glycopeptides and necessity of high quality of MS/MS spectrum for database searching reduced sensitivity. In order to characterize more glycosylation sites, combination of glycopeptide enrichment and depletion of abundant serum proteins is needed.

Acknowledgements

This study was supported in part by a Grant-in-Aid from the Ministry of Health, Labor, and Welfare, a Grant-in-Aid from the Ministry of Education, Culture, Sports and Technology, and Core Research for the Evolutional Science and Technology Program (CREST), Japan Science and Technology Corp (JST).

References

[1] A. Varki, Glycobiology 3 (1993) 97.
[2] Z. Yang, W.S. Hancock, J. Chromatogr. A 1053 (2004) 79.
[3] R.B. Parekh, R.A. Dwek, B.J. Sutton, D.L. Fernandes, A. Leung, D. Stanworth, T.W. Rademacher, T. Mizuochi, T. Taniguchi, K. Matsuta, et al., Nature 316 (1985) 452.
[4] L.A. Omtvedt, L. Royle, G. Husby, K. Sletten, C.M. Radcliffe, D.J. Harvey, R.A. Dwek, P.M. Rudd, Arthr. Rheum 54 (2006) 3433.
[5] R. Dube, G.A. Rook, J. Steele, R. Brealey, R. Dwek, T. Rademacher, J. Lennard-Jones, Gut 31 (1990) 431.
[6] H.H. Freeze, Glycobiology 11 (2001) 129R.
[7] M. Ferens-Sieczkowska, K. Zwierz, A. Midro, I. Katnik-Prastowska, Arch. Immunol. Ther. Exp. (Warsz) 50 (2002) 67.
[8] S. Thompson, E. Dargan, G.A. Turner, Cancer Lett. 66 (1992) 43.
[9] R. Saldova, L. Royle, C.M. Radcliffe, U.M. Abd Hamid, R. Evans, J.N. Arnold, R.E. Banks, R. Hutson, D.J. Harvey, R. Antrobus, S.M. Petrescu, R.A. Dwek, P.M. Rudd, Glycobiology 17 (2007) 1344.
[10] Y. Otake, I. Fujimoto, F. Tanaka, T. Nakagawa, T. Ikeda, K.K. Menon, S. Hase, H. Wada, K. Ikenaka, J. Biochem. (Tokyo) 129 (2001) 537.
[11] B. Kossowska, M. Ferens-Sieczkowska, R. Gancarz, E. Passowicz-Muszynska, R. Jankowska, Clin. Chem. Lab. Med. 43 (2005) 361.
[12] G.A. Turner, Adv. Exp. Med. Biol. 376 (1995) 231.
[13] N. Okuyama, Y. Ide, M. Nakano, T. Nakagawa, K. Yamanaka, K. Moriwaki, K. Murata, H. Ohigashi, S. Yokoyama, H. Eguchi, O. Ishikawa, T. Ito, M. Kato, A. Kasahara, S. Kawano, J. Gu, N. Taniguchi, E. Miyoshi, Int. J. Cancer 118 (2006) 2803.
[14] M.A. Comunale, M. Lowman, R.E. Long, J. Krakover, R. Philip, S. Seeholzer, A.A. Evans, H.W. Hann, T.M. Block, A.S. Mehta, J. Proteome Res. 5 (2006) 308.
[15] J.E. Hansen, J. Iversen, A. Lihme, T.C. Bog-Hansen, Cancer 60 (1987) 1630.
[16] R.R. Drake, E.E. Schwegler, G. Malik, J. Diaz, T. Block, A. Mehta, O.J. Semmes, Mol. Cell Proteomics 5 (2006) 1957.
[17] Y. Wada, M. Tajiri, S. Yoshida, Anal. Chem. 76 (2004) 6560.
[18] M. Tajiri, S. Yoshida, Y. Wada, Glycobiology 15 (2005) 1332.
[19] P. Hagglund, J. Bunkenborg, F. Elortza, O.N. Jensen, P. Roepstorff, J. Proteome Res. 3 (2004) 556.
[20] J.F. Nemeth, G.P. Hochgesang Jr., L.J. Marnett, R.M. Caprioli, Biochemistry 40 (2001) 3109.
[21] J.P. Hui, T.C. White, P. Thibault, Glycobiology 12 (2002) 837.
[22] K. Hakansson, H.J. Cooper, M.R. Emmett, C.E. Costello, A.G. Marshall, C.L. Nilsson, Anal. Chem. 73 (2001) 4530.
[23] O. Krokhin, W. Ens, K.G. Standing, J. Wilkins, H. Perreault, Rapid Commun. Mass Spectrom. 18 (2004) 2020.
[24] U.M. Demelbauer, M. Zehl, A. Plematl, G. Allmaier, A. Rizzi, Rapid Commun. Mass Spectrom. 18 (2004) 1575.
[25] M. Wuhrer, C.H. Hokke, A.M. Deelder, Rapid Commun. Mass Spectrom. 18 (2004) 1741.
[26] A. Harazono, N. Kawasaki, T. Kawanishi, T. Hayakawa, Glycobiology 15 (2005) 447.
[27] S. Itoh, N. Kawasaki, A. Harazono, N. Hashii, Y. Matsuishi, T. Kawanishi, T. Hayakawa, J. Chromatogr. A 1094 (2005) 105.
[28] A. Harazono, N. Kawasaki, S. Itoh, N. Hashii, A. Ishii-Watabe, T. Kawanishi, T. Hayakawa, Anal. Biochem. 348 (2006) 259.
[29] Y. Wada, P. Azadi, C.E. Costello, A. Dell, R.A. Dwek, H. Geyer, R. Geyer, K. Kakehi, N.G. Karlsson, K. Kato, N. Kawasaki, K.H. Khoo, S. Kim, A. Kondo, E. Lattova, Y. Mechref, E. Miyoshi, K. Nakamura, H. Narimatsu, M.V. Novotny, N.H. Packer, H. Perreault, J. Peter-Katalinic, G. Pohlentz, V.N. Reinhold, P.M. Rudd, A. Suzuki, N. Taniguchi, Glycobiology 17 (2007) 411.
[30] J. Bunkenborg, B.J. Pilch, A.V. Podtelejnikov, J.R. Wisniewski, Proteomics 4 (2004) 454.
[31] T. Liu, W.J. Qian, M.A. Gritsenko, D.G. Camp 2nd, M.E. Monroe, R.J. Moore, R.D. Smith, J. Proteome Res. 4 (2005) 2070.
[32] Y. Wang, S.L. Wu, W.S. Hancock, Glycobiology 16 (2006) 514.
[33] P. Hagglund, R. Matthiesen, F. Elortza, P. Hojrup, P. Roepstorff, O.N. Jensen, J. Bunkenborg, J. Proteome Res. 6 (2007) 3021.

Glycosylation Analysis of IgLON Family Proteins in Rat Brain by Liquid Chromatography and Multiple-Stage Mass Spectrometry

Satsuki Itoh, Akiko Hachisuka, Nana Kawasaki, Noritaka Hashii, Reiko Teshima, Takao Hayakawa, Toru Kawanishi, and Teruhide Yamaguchi

Division of Biological Chemistry and Biologicals, National Institute of Health Sciences, 1-18-1, Kamiyoga, Setagaya-ku, Tokyo 158-8501, Japan, Core Research for Evolutional Science and Technology of Japan Science and Technology Agency, Kawaguchi Center Building, 4-1-8 Hon-cho, Kawaguchi, Saitama 332-0012, Japan, and Pharmaceutical Research and Technology Institute, Kinki University, 3-4-1 Kowakae, Higashi-Osaka 577-8502, Japan

Biochemistry®

Reprinted from
Volume 47, Number 38, Pages 10132–10154

Glycosylation Analysis of IgLON Family Proteins in Rat Brain by Liquid Chromatography and Multiple-Stage Mass Spectrometry[†]

Satsuki Itoh,[‡] Akiko Hachisuka,[‡] Nana Kawasaki,[*,‡,§] Noritaka Hashii,[‡] Reiko Teshima,[‡] Takao Hayakawa,[∥] Toru Kawanishi,[‡] and Teruhide Yamaguchi[‡]

Division of Biological Chemistry and Biologicals, National Institute of Health Sciences, 1-18-1, Kamiyoga, Setagaya-ku, Tokyo 158-8501, Japan, Core Research for Evolutional Science and Technology of Japan Science and Technology Agency, Kawaguchi Center Building, 4-1-8 Hon-cho, Kawaguchi, Saitama 332-0012, Japan, and Pharmaceutical Research and Technology Institute, Kinki University, 3-4-1 Kowakae, Higashi-Osaka 577-8502, Japan

Received May 23, 2008; Revised Manuscript Received July 17, 2008

ABSTRACT: IgLON family proteins, including limbic-associated membrane protein (LAMP), opioid-binding cell adhesion molecule (OBCAM), neurotrimin, and Kilon, are immunoglobulin (Ig) superfamily cell adhesion molecules. These molecules are composed of three Ig domains and a glycosylphosphatidylinositol (GPI) anchor and contain six or seven potential N-glycosylation sites. Although their glycosylations are supposed to be associated with the development of the central nervous system like other Ig superfamily proteins, they are still unknown because of difficulty in isolating individual proteins with a high degree of homology in performing carbohydrate analysis. In this study, we conducted simultaneous site-specific glycosylation analysis of rat brain IgLON proteins by liquid chromatography and multiple-stage mass spectrometry ($LC-MS^n$). The rat brain GPI-linked proteins were enriched and separated by sodium dodecyl sulfate−polyacrylamide gel electrophoresis. The four proteins were extracted from the gel, and subjected to $LC-MS^n$ after proteinase digestions. A set of glycopeptide MS data, including the mass spectrum, the mass spectrum in the selected ion monitoring mode, and the product ion spectra, was selected from all data based on carbohydrate-related ions in the MS/MS spectrum. The peptide portion and the carbohydrate structure were identified on the basis of peptide-related ion and carbohydrate-related ions, and the accurate mass. The site-specific glycosylations of four proteins were elucidated as follows. *N*-Glycans near the N-terminal were disialic acid-conjugated complex- and hybrid-type oligosaccharides. The first Ig domains were occupied by Man-5-9. Diverse oligosaccharides, including Lewis a/x-modified glycans, a brain-specific glycan known as BA-2, and Man-5, were found to be attached to the third Ig domain. Three common structures of glycans were found in the GPI moiety of LAMP, OBCAM, and neurotrimin.

Cell adhesion molecules on cell surfaces are involved in several biological events, such as cell−cell interaction, signaling, and cellular traffic. In the central nervous system, cell adhesion molecules are associated with the differentiation and migration of neurons, and neurite outgrowth. The immunoglobulin (Ig) superfamily, which contains one or more Ig-like domains, is known as one of the cell adhesion molecule families in the central nervous system (*1*). The Ig superfamily includes various proteins, such as P0, Thy-1, myelin-associated glycoprotein (MAG), neural cell adhesion molecule (NCAM), L1, contactin, and IgLON family proteins. Glycosylation of the Ig superfamily proteins is known to be involved in cell−cell interactions (*2*−*4*). Polysialylated glycans in the fifth domain of NCAM are thought to inhibit the interaction of NCAM with other molecules and to promote neural plasticity through a repulsive interaction (*5*, *6*). The HNK-1 epitope in the third and fifth domains of NCAM is known to mediate molecular recognition in the nervous system (*7*).

The IgLON superfamily includes the limbic-associated membrane protein (LAMP),[1] the opioid-binding cell adhesion molecule (OBCAM), neurotrimin, and Kilon (*8*−*14*), and

[†] This work was supported in part by a Grant-in-Aid from the Ministry of Health and Labor and Welfare, and Core Research for Evolutional Science and Technology Program (CREST) of the Japan Science and Technology Agency (JST).

* To whom correspondence should be addressed: Division of Biological Chemistry and Biologicals, National Institute of Health Sciences, 1-18-1, Kamiyoga, Setagaya-ku, Tokyo 158-8501, Japan. Telephone: +81-3-3700-9074. Fax: +81-3-3707-6950. E-mail: nana@nihs.go.jp.

[‡] National Institute of Health Sciences.

[§] Core Research for Evolutional Science and Technology of Japan Science and Technology Agency.

[∥] Kinki University.

[1] Abbreviations: LC, liquid chromatography; MS, mass spectrometry; MS^n, multiple-stage mass spectrometry; LAMP, limbic-associated membrane protein; OBCAM, opioid-binding cell adhesion molecule; GlcNAc, *N*-acetylglucosamine; GPI, glycosylphosphatidylinositol; PI-PLC, phosphatidylinositol-specific phospholipase C; PNGase F, peptide *N*-glycosidase F; IT-MS, ion trap mass spectrometer; FT ICR-MS, Fourier transform ion cyclotron resonance mass spectrometer; GCC, graphitized carbon column; TIC, total ion chromatogram; CID, collision-induced dissociation; SIM, selected ion monitoring; dHex, deoxyhexose; Hex, hexose; HexNAc, *N*-acetylhexosamine; Fuc, fucose; Man, mannose; Gal, galactose; GlcNAc, *N*-acetylglucosamine; GlcN, glucosamine; NeuAc, *N*-acetylneuraminic acid; $EtNH_2$, ethanolamine; Ino, inositol; BA-2, brain-specific sugar chain, GlcNAcβ1−2Manα1−6(GlcNAcβ1−4)(GlcNAcβ1−2Manα1−3)Manβ1−4GlcNAcβ1−4(Fucα1−6)GlcNAc; SDS−PAGE, sodium dodecyl sulfate−polyacrylamide gel electrophoresis.

10.1021/bi8009778 CCC: $40.75
Published on Web 08/26/2008

FIGURE 1: Amino acid sequence and potential N-glycosylation sites (in bold) of IgLON family proteins. Their accession numbers in Swiss-prot database are shown in parentheses after their names. The C-terminal amino acids in the proteins are predicted GPI attachment sites.

these proteins are distributed differently in the central nervous system during the development of neurons in a brain (*11, 13*−*18*). The IgLON family proteins consist of three Ig domains, the third of which is attached to a glycosylphosphatidylinositol (GPI) anchor. Each of the IgLON family proteins includes six or seven consensus N-glycosylation sites (Figure 1), and the glycosylation is presumed to play essential roles in the neural circuit formation like other Ig superfamily proteins (*2*−*4*). However, since the high degree of homology of their amino acid sequences makes it difficult to isolate the individual proteins of this family to perform carbohydrate analysis, their glycosylation features are still unknown with the exception of a linkage of *N*-glycans in OBCAM and Kilon and of high mannose-type and hybrid-type oligosaccharides in LAMP (*9, 18, 19*).

Recently, liquid chromatography and mass spectrometry (LC−MS) and liquid chromatography and multiple-stage mass spectrometry (LC−MS^n) have been widely applied to the site-specific glycosylation analysis of a glycoprotein (*20*−*24*). Generally, a tryptic digest of an isolated glycoprotein is separated with a reversed-phase or normal-phase column, and the separated glycopeptides are directly subjected to MS and MS^n (*25*−*27*). The site-specific glycosylation is deduced from the mass spectra of the glycopeptides, and the sequences of both the peptide and carbohydrate portions are deduced from the fragment ions in the MS^n spectra. Using this technique, we previously performed a site-specific glycosylation analysis of rat brain Thy-1, which contains three N-glycosylation sites and a GPI anchor (*28*). GPI-anchored proteins enriched via phase partitioning with Triton X-114 and PIPLC digestion were separated by SDS−PAGE, and the Thy-1 protein extracted from the gel was digested with trypsin or endoproteinase Asp-N. The Thy-1 glycopeptides were separated and analyzed by using a liquid chromatography and ion trap mass spectrometer (IT-MS) equipped with a C18 column. The peptide portions of glycopeptides were identified on the basis of the *m/z* values of the peptide-related ions and the b- and y-ions that arose from the peptide backbone. The carbohydrate structures at each glycosylation site and in the GPI moiety were successfully determined from fragment ions in the MS/MS spectra. This result suggests that LC−MS^n can be effectively utilized for site-specific glycosylation analysis of each glycoprotein in the mixture of several glycoproteins simultaneously. .

In this study, we conducted site-specific glycosylation analyses of rat LAMP, OBCAM, neurotrimin, and Kilon using LC−MS^n. The GPI-linked proteins in the rat brains were separated by SDS−PAGE, and the IgLON family proteins were extracted from a gel band (45−70 kDa). The mixture of proteins was digested with proteinases, and the site-specific glycosylation analysis of the four proteins was performed by using an ion trap-Fourier transform ion cyclotron resonance mass spectrometer (IT-MS-FT ICR-MS), which is capable of acquiring the accurate mass as well as the MS^n spectra. We successfully elucidated the site-specific glycosylation and the structure of the GPI moieties of LAMP, OBCAM, neurotrimin, and Kilon. This is the first report of the simultaneous site-specific glycosylation analysis of four similar glycoproteins.

EXPERIMENTAL PROCEDURES

Materials. The rat brains (Wister, male, 3 weeks old) were purchased from Nippon SLC (Hamamatsu, Japan). Phosphatidylinositol-specific phospholipase C (PIPLC) from *Bacillus cereus* was obtained from Molecular Probes (Eugene, OR). Trypsin-Gold was purchased from Promega (Madison, WI). PNGase F and endoproteinase Glu-C were purchased from Roche Diagnostics (Mannheim, Germany). SimplyBlue SafeStain was obtained from Invitrogen (Carlsbad, CA). All other chemicals were of the highest available purity.

SDS−PAGE of Enriched Lipid-Free GPI-Linked Proteins. Lipid-free GPI-linked proteins were enriched from rat brain as reported previously (*28, 29*). Briefly, the homogenate of two rat brains (total wet weight of 1.4 g) was defatted and solubilized with 2% Triton X-114 at 4 °C overnight (*29, 30*). After centrifugation, the supernatant was subjected to Triton X-114 phase partitioning at 37 °C. Cold acetone was added to the detergent phase containing solubilized membrane proteins, and the resulting precipitate was digested with PIPLC. After the PIPLC digest mixture had been subjected to Triton X-114 phase partitioning, lipid-free GPI-linked proteins in the aqueous phase were precipitated via addition of cold acetone. These proteins were separated by SDS−PAGE (12.5%) (brain wet weight of 50 mg/lane) after carboxyamidomethylation (*31*) and detected after being stained with Coomassie Brilliant Blue G-250 using SimplyBlue SafeStain.

Protein Identification. Gel-separated proteins were extracted after in-gel trypsin digestion as previously reported (*32*) and subjected to LC−MS/MS with a Paradigm MS4 HPLC system (Michrom BioResources, Inc., Auburn, CA) consisting of pump A with 0.1% formic acid and 2% acetonitrile and pump B with 0.1% formic acid and 90% acetonitrile. Peptides were separated with a Magic C18 column (50 mm × 0.2 mm, 3 μm; Michrom BioResources Inc.) with a linear gradient from 5 to 65% of pump B over

20 min at a flow rate of 3 μL/min. Mass spectra were recorded with a Finnigan LTQ system (Thermo Fisher Scientific, Waltham, MA) using sequential scan events: MS (*m/z* 450−2000) followed by data-dependent MS/MS on the IT-MS for the most intense ions in positive ion mode. For protein identification, all obtained product ions were subjected to a computer database search analysis with the TurboSEQUEST search engine (Thermo Fisher Scientific) using the Swiss-Prot database and search parameters: a static modification of carboxyamidomethylation (57 Da) at Cys and trypsin for digestion.

Extraction and Proteinase Digestion of the 45−70 kDa Proteins Separated by SDS−PAGE. The gel-separated proteins were extracted as previously reported (*28*). The proteins were extracted with 20 mM Tris-HCl containing 1% SDS by being shaken vigorously overnight after the gel had been broken down into small bits. The extract was filtered with Ultrafree-MC (0.22 μm; Millipore, Bedford, MA), and the proteins were precipitated via addition of cold acetone. The resulting precipitate was digested with endoproteinase Glu-C (3.75 μg) in 30 μL of 0.1 M ammonium acetate (pH 8.0) at 37 °C for 4 days, followed by incubation with additional trypsin (1 μg) at 37 °C overnight.

LC−MS^n. Proteolytic peptides were separated by reversed-phase columns, Magic C30 and C18 (50 mm × 0.1 mm, 3 μm; Michrom BioResources), and a graphitized carbon column (GCC), Hypercarb 5 μ (150 mm × 0.2 mm; Thermo Fisher Scientific), with a Paradigm MS4 HPLC system consisting of pump A with 0.1% formic acid and 2% acetonitrile and pump B with 0.1% formic acid and 90% acetonitrile. For analysis of glycopeptides, separation was performed with a linear gradient from 5 to 50% pump B over 100 min followed by a 50 to 95% B gradient over 10 min and 95% B over 10 min at a flow rate of 0.5 μL/min, and mass spectra were recorded with a Finnigan LTQ-FT system (Thermo Fisher Scientific) using sequential scan events: MS (*m/z* 1000−2000 or 700−2000) with the IT-MS followed by MS with the IT-MS-FT ICR-MS in selected ion monitoring (SIM) mode and data-dependent MS^n with the IT-MS for the most intense ions. The LC−MS^n runs were performed with a C30 column and scan range of *m/z* 1000−2000 (condition A), twice, with a C30 column and scan range of *m/z* 700−2000 (condition B), once, and with a C18 column and scan range of *m/z* 1000−2000 (condition C), once. For analysis of GPI-linked peptides, separation was performed with a linear gradient from 5 to 60% pump B over 100 min at a flow rate of 2 μL/min for a GCC, and mass spectra were recorded with a Finnigan LTQ system using sequential scans: a single mass scan (*m/z* 700−2000) with the IT-MS followed by data-dependent MS^n scans with the IT-MS for the most intense ions, twice. LC−MS^n was performed using a capillary voltage of 1.8 kV and a capillary temperature of 200 °C.

RESULTS

Preparation of Lipid-Free IgLON Glycopeptides. Figure 2 illustrates the experimental procedure for the glycosylation analysis of IgLON family proteins. Lipid-free GPI-linked proteins in a rat brain tissue sample were enriched via phase partitioning with Triton X-114 and PIPLC digestion. The enriched proteins were separated by SDS−PAGE and stained

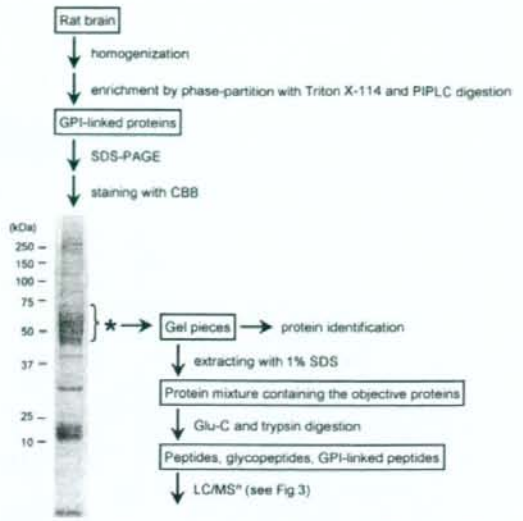

FIGURE 2: Experimental procedure for site-specific glycosylation analysis of IgLON family proteins and SDS−PAGE (12.5%) of lipid-free GPI-linked proteins which were enriched from rat brain. The asterisk indicates the gel band containing IgLON family proteins.

with Coomassie Brilliant Blue. The presence of LAMP, OBCAM, neurotrimin, and Kilon in the gel band at 45−70 kDa was confirmed by in-gel trypsin digestion followed by LC−MS/MS. The IgLON proteins were extracted with other comigrated proteins from 45−70 kDa bands in other lanes by being shaken in 1% SDS. After SDS had been removed, the mixture of proteins was digested with endoproteinase Glu-C and trypsin. Most of the resulting glycopeptides contained only a single N-glycosylation site, with the exception of LGTTN^{270}ASLPLNPPSTAQYGITG^{287} in Kilon, which included a predicted GPI attachment site at Gly287 in addition to a potential N-glycosylation site at Asn270 (Figure 1). The glycopeptides from IgLON family proteins was separated by using three different columns: a reversed-phase column, a C30 and a C18 column for hydrophobic glycopeptides, and a GCC for hydrophilic glycopeptides, including GPI-linked peptides.

Glycosylation Analysis of LAMP. LC−MS analysis was performed via MS on the IT-MS and data-dependent MS in SIM mode on the FT ICR-MS, and data-dependent MS/MS and MS/MS/MS were performed on the IT-MS in the positive ion mode (Figure 3). After MS data acquisition, the MS/MS spectrum (scan *n*) of a glycopeptide was selected manually from all MS data on the basis of the existence of carbohydrate distinctive fragments, such as $Hex_1HexNAc_1^+$ (*m/z* 366) and $Hex_1HexNAc_1NeuAc^+$ (*m/z* 657). Then a set of the glycopeptide's MS data consisting of the mass spectrum (scan $n-2$), the mass spectrum in SIM on the FT ICR-MS (scan $n-1$), the MS/MS spectrum (scan *n*), and the MS/MS/MS spectrum (scan $n+1$) was selected from all the MS data (step 1). The carbohydrate structure was deduced from the fragment ions appearing in the MS/MS spectrum (scan *n*), and the peptide portion was estimated from the peptide-related ions (step 2). The sequences of some peptides were confirmed by the b- and y-ions that arose from Y_1 ([peptide + HexNAc + H]$^+$) in MS/MS/MS (scan $n+$

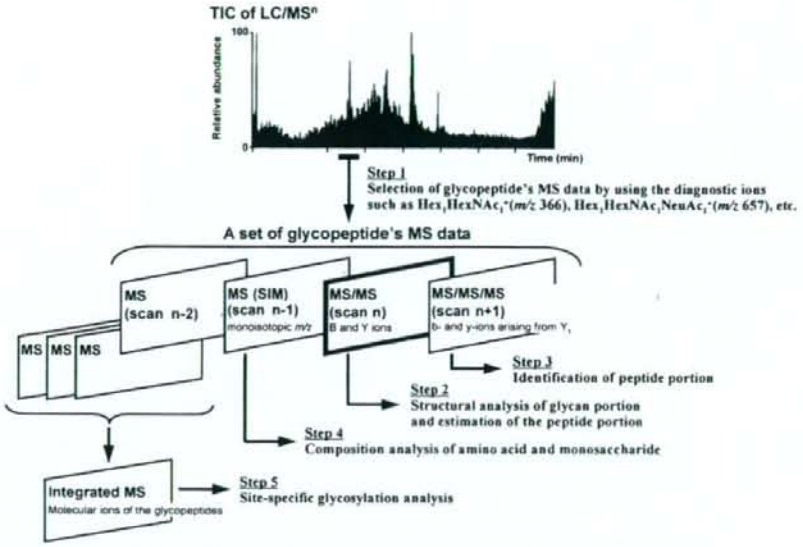

FIGURE 3: Methods used for LC−MS^n and data analysis.

1) (step 3). The accurate molecular mass that was calculated from the monoisotopic *m/z* value and the charge state acquired by FT ICR-MS in SIM mode (scan $n - 1$) was used to corroborate the assignment of the peptide and glycan moieties (step 4). The mass spectra acquired at the elution position, where the glycopeptides that yielded identical Y_1 ions in the MS/MS and/or MS/MS/MS spectra, were integrated, and the site-specific glycosylation was elucidated on the basis of the distribution of molecular ions in the integrated mass spectra (step 5). As a representative separation pattern, a total ion chromatogram (TIC) obtained by LC−MS^n with a C30 column (scan range of *m/z* 1000−2000) is shown in Figure 4A. The MS/MS spectra containing the diagnostic ions at *m/z* 366 and 657 were picked out from all the MS data, and the peptides eluted at positions 1−25 were determined to be the glycopeptides on the basis of the carbohydrate-related ions. The 19% of spectra acquired at elution time, including positions 1−25, could be traced back to the glycopeptides of IgLON family proteins.

As for LAMP, it has seven potential N-glycosylation sites at Asn12, -38, -108, -120, -251, -259, and -272, and Asn287 is the predicted site of GPI linkage. On the basis of the presence of the peptide-related ions ([peptide + HexNAc + H]$^+$, Y_1 or $Y_{1\alpha/1\beta}$; or [peptide + dHex-HexNAc + H]$^+$, $Y_{1\alpha}$), glycopeptides that were eluted at the positions 1, 11, 14, 12, 4, and 24 were estimated to be the glycopeptides containing Asn12, -38, -108, -251, -259, and -272, respectively. The MS/MS spectra of the glycopeptide containing Asn120 (GS**N**^{120}VTLVCMANGRPE) were not acquired in any of the runs. However, glycosylation at Asn120 was confirmed by the detection of the peptide substituted with Asp (GS**D**^{120}VTLVCMANGRPEPVITWR) after PNGase F digestion (data not shown). Panels A1−F1 of Figure 5 show the representative MS/MS and MS/MS/MS spectra acquired at positions 11, 1, 14, 12, 4, and 24, respectively. The integrated mass spectra of the glycopeptides containing Asn38, -12, -108, -251, -259, and -272 are shown in panels A2−F2 of Figure 5, respectively. The feature of the glycosylation at each glycosylation site was elucidated on the basis of these MS spectra.

(i) Asn38 (Asn43 in OBCAM and Asn38 in neurotrimin). Panel A1 of Figure 5 shows one of the MS/MS spectra acquired at position 11. The peptide portion, VAWL(GlcNAc-)N^{38}R, was confirmed on the basis of the b- and y-ions that arose from Y_1 (*m/z* 961.5) in the MS/MS/MS spectrum (panel A1″ of Figure 5). A series of doubly charged Y ions with an *m/z* spacing pattern, 81 *m/z* units (Hex), suggests the linkage of Man-7 to this peptide. The attachment of Man-7 to VAWLN^{38}R, whose theoretical monoisotopic *m/z* value ([M + 2H]$^{2+}$) is 1149.983, was ascertained by the observed monoisotopic *m/z* value (1149.986) acquired in SIM mode on the FT ICR-MS (panel A1′ of Figure 5). Panel A2 of Figure 5 shows the integrated mass spectrum which was obtained from the mass spectra of glycopeptides that yielded Y_1 (*m/z* 961.5) via MS/MS. Four noticeable ion peaks (peaks a-1−a-4) appearing with the differences of 81 *m/z* units are assigned to VAWLN^{38}R glycosylated with Man-6-9 (Table 1A). The MS/MS spectra of DKNSKVAWLN^{38}R and CVVEDKNSKVAWLN^{38}R, which were picked out from positions 9 and 15, also revealed that Man-5, -7, and -8 were attached to Asn38.

(ii) Asn12. Panel B1 of Figure 5 shows the representative MS/MS spectrum of glycopeptide, GTDN^{12}ITVR, which was selected from position 1. From the $Y_{1\alpha}$ ion (*m/z* 1224.5) together with monoisotopic *m/z* value of the molecular ion (*m/z* 1173.132) and a series of doubly charged Y ions with an *m/z* spacing pattern, 146 (NeuAc), 101 (HexNAc), and 81 *m/z* units (Hex), the carbohydrate portion was estimated to be dHex$_1$Hex$_5$HexNAc$_4$NeuAc$_4$. Furthermore, a complex-type oligosaccharide, to which one branch of disialic acid was attached, was deduced from the presence of $B_{4\alpha}/Y_{5\alpha'}$ (*m/z* 495.3), $B_{2\alpha}$ (*m/z* 582.7), $B_{3\alpha}$ (*m/z* 744.9), $B_{4\alpha}/Y_{5\alpha''}$ and $B_{4\alpha}/Y_{7\alpha'}$ (*m/z* 948.2), and $B_{4\alpha}$ (*m/z* 1239.5) (inset of panel B1 of Figure 5). The integrated mass spectrum at position 1 suggests that the majority of the glycans at Asn12 are hybrid- and complex-type oligosaccharides containing disialic acids

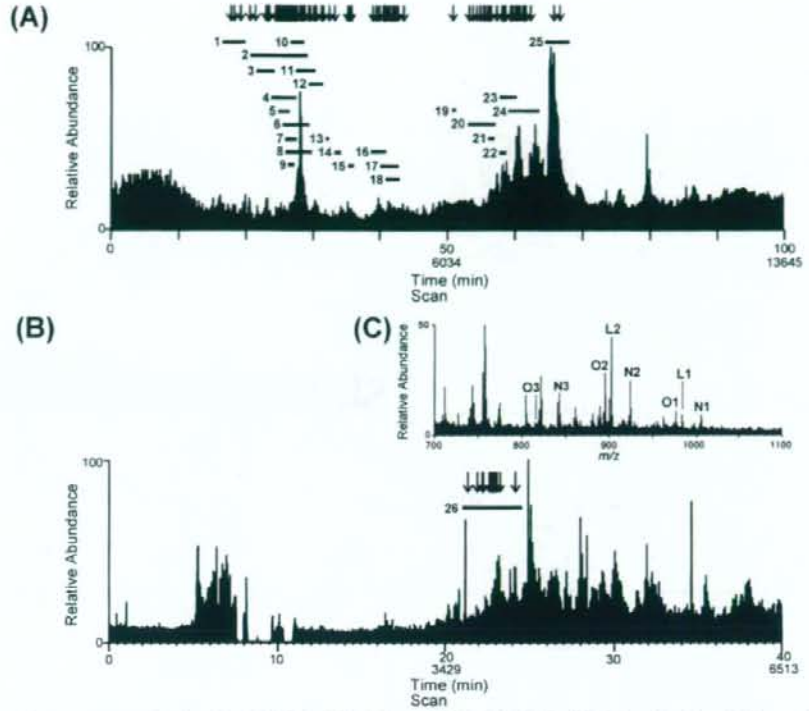

FIGURE 4: Total ion chromatograms obtained by C30-LC−MS^n (A) and GCC-LC−MS^n (B). Lines 1−25 and 26 are the elution positions of glycopeptides and GPI-linked peptides, respectively. The down arrow denotes the extracted position of the MS/MS spectra. (C) Integrated mass spectrum obtained from elution position 26. L1 and L2 are molecular ions of GPI-linked peptides from LAMP, N1−N3 those from neurotrimin, and O1−O3 those from OBCAM.

(panel B2 of Figure 5 and Table 1B). In addition, the partial glycosylation at Asn12 was indicated by the detection of nonglycosylated GTDN^{12}ITVR.

(iii) Asn108. The MS/MS spectrum of glycosylated ISN^{108}ISSDVTVNE ($Y_{1\alpha/1\beta}$, *m/z* 1480.6) acquired at position 14 is shown in panel C1 of Figure 5. The attachment of a Lewis a/x [Le$^{a/x}$, Gal-(Fuc-)GlcNAc-] or H antigen (Fuc-Gal-GlcNAc-) motif to the bisected complex-type oligosaccharide was deduced from the monosaccharide composition ($dHex_2Hex_4HexNAc_5$) and the Le$^{a/x}$ and H antigen-related ion (*m/z* 512.1) and $Y_{1\beta/3\alpha/3\beta}{}^{2+}$ (*m/z* 1024.3) (panel C1 of Figure 5, peak c-1 in panel C2 of Figure 5). The alternative LC−MS^n run with the C30 column (scan range of *m/z* 1000−2000) suggested that ISN^{108}ISSD is also occupied by sialyl Le$^{a/x}$ (sLe$^{a/x}$)-modified or core-fucosylated hybrid-type oligosaccharides based on the presence of NeuAc-Hex-(dHex-)HexNAc$^+$ (*m/z* 803.1), Hex-(dHex-)HexNAc$^+$ (*m/z* 512.3), NeuAc-Hex$^+$ (*m/z* 454.2), and [peptide + dHex + HexNAc + H]$^+$ (*m/z* 1084.3) (data not shown, Table 1C).

(iv) Asn251. The representative MS/MS spectrum of the glycopeptide containing GQSSLTVTN^{251}VTE ($Y_{1\alpha/1\beta}$, *m/z* 1438.6; elution position 12) is shown in panel D1 of Figure 5. From the monoisotopic mass and the Le$^{a/x}$-related ions (*m/z* 350.3 and 512.2), the carbohydrate structure was estimated to be a complex-type oligosaccharide to which the Le$^{a/x}$ motif was attached ($dHex_2Hex_4HexNAc_5$; inset of panel D1 of Figure 5). Other glycans at Asn251 were characterized as complex-type oligosaccharides containing sLe$^{a/x}$ or Lewis b/y [Le$^{b/y}$, Fuc-Gal-(Fuc-)GlcNAc-] based on the molecular ions in the integrated mass spectrum (peaks d-1−6 in panel D2 of Figure 5), the sLe$^{a/x}$-related ions (*m/z* 803, 657, and 512), and the Le$^{b/y}$-related ions (*m/z* 658.2, 512.1, and 350.2) acquired by the alternative run with the C30 column (scan range of *m/z* 700−2000) (Table 1D).

(v) Asn259. Panel E1 of Figure 5 shows the product ion spectra of HYGN^{259}YTCVAANK linked by $dHex_1Hex_3HexNAc_5$, which was deduced from the $Y_{1\alpha/1\beta}$ ion (*m/z* 1600.6) and the monoisotopic mass acquired at position 4. The BA-2, which is a core-fucosylated and agalacto-biantennary oligosaccharide with bisecting GlcNAc, and known as a brain-specific carbohydrate, was suggested by the product ions at *m/z* 1085.3 (bisecting GlcNAc) and 1746.6 (core-fucosylation) (inset of panel E1 of Figure 5). The majority of other glycans at Asn259 were characterized as Le$^{a/x}$-modified complex and hybrid types. Man-5 was suggested to be a minor glycan (panel E2 of Figure 5 and Table 1E).

(vi) Asn272. Panel F1 of Figure 5 shows the MS/MS and MS/MS/MS spectra of glycopeptide LGVTN^{272}ASLVLFR ($Y_{1\alpha/1\beta}$, *m/z* 1492.8), which were acquired at position 24. The monosaccharide composition ($dHex_2Hex_4HexNAc_5$) and the presence of $Y_{3\alpha/3\beta}{}^{2+}$ (*m/z* 1103.8) and Le$^{a/x}$-related ion suggested the attachment of a Le$^{a/x}$ or H antigen motif to the bisected and core-fucosylated complex-type oligosaccharide (inset of panel F1 of Figure 5). The MS/MS spectra of the LGVTN^{272}ASLVLFRPGSVR glycopeptides ($Y_{1\alpha/1\beta}{}^{2+}$, *m/z* 1069) were also picked out at position 24 (data not shown). The *m/z* values of molecular ions appearing in the

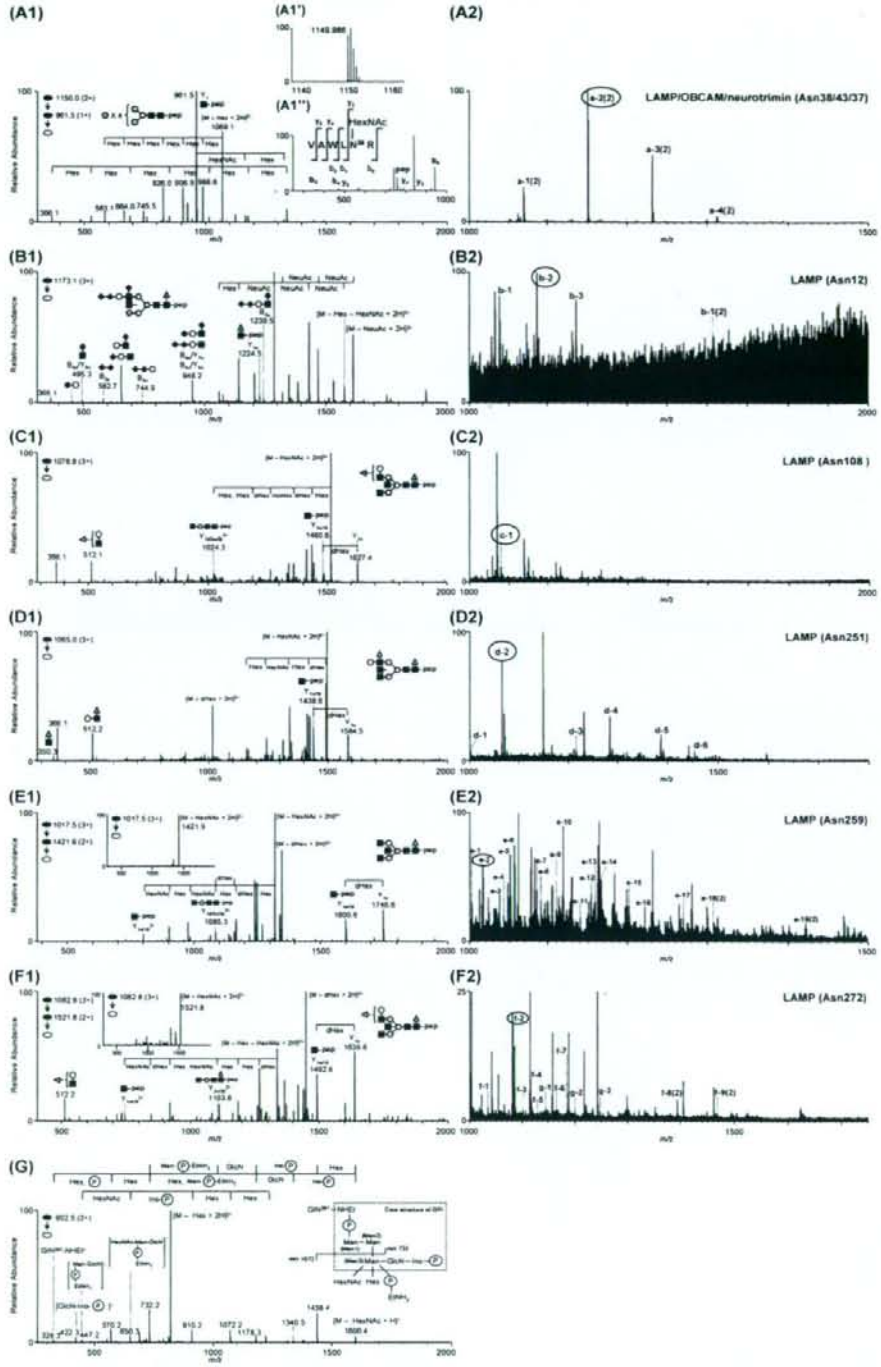

FIGURE 5: MS spectra of LAMP glycopeptides. (A1) MS/MS spectrum of glycopeptide VAWLN^{38}R; elution position, 11; precursor ion, $[M + 2H]^{2+}$ (m/z 1150.0). (A1′) Mass spectrum on the FT ICR-MS in SIM mode. (A1″) MS/MS/MS spectrum acquired from Y_1 (m/z 961.5). (A2) Integrated mass spectrum obtained from position 11. (B1) MS/MS spectrum of glycopeptide GTDN^{12}ITVR; elution position, 1; precursor ion, $[M + 3H]^{3+}$ (m/z 1173.1). (B2) Integrated mass spectrum at position 1. (C1) MS/MS spectrum of glycopeptide ISN^{108}ISSDVTVNE; elution position, 14; precursor ion, $[M + 3H]^{3+}$ (m/z 1078.8). (C2) Integrated mass spectrum at position 14. (D1) MS/MS spectrum of glycopeptide GQSSLTVTN^{251}VTE; elution position, 12; precursor ion, $[M + 3H]^{3+}$ (m/z 1065.0). (D2) Integrated mass spectrum at position 12. (E1) MS/MS and MS/MS/MS spectra of glycopeptide HYGN^{259}YTCVAANK; elution position, 4; precursor ion, $[M + 3H]^{3+}$ (m/z 1017.5). (E2) Integrated mass spectrum at position 4. (F1) MS/MS and MS/MS/MS spectra of glycopeptide LGVTN^{272}ASLVLFR; elution position, 24; precursor ion, $[M + 3H]^{3+}$ (m/z 1082.8). (F2) Integrated mass spectrum at position 24. (G) MS/MS spectrum of GPI-linked GIN^{287}; elution position, 26; precursor ion, $[M + 2H]^{2+}$ (m/z 902.5). Symbols are as in Figure 9.