Key words: apolipoprotein B100/*N*-linked oligosaccharide/glycopeptide/liquid chromatography electrospray mass spectrometry/product ion spectrum

Introduction

Low-density lipoprotein (LDL) is the main cholesterol carrier in human plasma, and a high serum level of LDL is involved in the development of atherosclerosis. LDL is originally secreted as very low-density lipoprotein (VLDL). VLDL is converted to LDL, then, removed from the circulation. Apolipoprotein B100 (apoB100) is the only protein component of LDL and is the ligand recognized by the LDL receptor. The amino acid sequence of human apoB100 has been deduced by analysis of the apoB100 cDNA sequence (Chen *et al.*, 1986; Knott *et al.*, 1986; Law *et al.*, 1986; Yang *et al.*, 1986). Mature apoB100 consists of 4536 amino acids and its molecular weight has been calculated to be 513 kd. ApoB100 has 19 potential *N*-glycosylation sites (Asn-X-Ser/Thr), of which 16 asparagine residues are found to be glycosylated (Yang *et al.*, 1989). The carbohydrate moieties were linked to asparagine residues at the following 16 positions: 158, 956, 1341, 1350, 1496, 2752, 2955, 3074, 3197, 3309, 3331, 3384, 3438, 3868, 4210, and 4404. The carbohydrate structures of the *N*-linked sugar chains of human apoB100 were reported to be high mannose, hybrid, and mono- and disialylated complex type oligosaccharides (Taniguchi *et al.*, 1979; Garner *et al.*, 2001). The role of carbohydrate moieties of apoB100 has been investigated by several laboratories. The *N*-linked oligosaccharides at the amino terminus of human apoB100 are important for the assembly and secretion of VLDL (Vukmirica *et al.*, 2002). Seven of the *N*-glycans are predicted to occur close to the LDL-receptor binding region of apoB100 and seem to have an important role (Yang *et al.*, 1986, 1989). The carbohydrate composition of apoB100, particularly sialylation, has been considered to contribute to the atherogenic properties of LDL. However, Shireman and Fisher (1979) reported that they do not appear to play a significant role in the binding of apoB100 to the LDL receptor. Furthermore, the distribution and diversity of human apoB100 oligosaccharides isolated from normolipidemic, hypercholesterolemic, and hypertriglyceridemic diabetic subjects were highly

conserved even though characterized in LDL subfractions (Garner *et al.*, 2001). The potential function of apoB100 carbohydrates post hepatic secretion is not well understood. Glycoproteins have a variety of sugar chains at each glycosylation site. Because of the individual functions at each glycosylation site, a comparison of glycosylation among various sites is important. Therefore, in order to investigate the role of carbohydrate moieties of apoB100, we attempted to determine the carbohydrate heterogeneity site-specifically.

In order to determine the site-specific carbohydrate heterogeneity of glycoproteins, the glycoprotein must be digested into peptides and glycopeptides, and then both the peptide and sugar chain of each glycopeptide analyzed. One of the most effective techniques for mapping proteolytic fragments of glycoproteins is liquid chromatography coupled with electrospray ionization mass spectrometry (Ling *et al.*, 1991; Duffin *et al.*, 1992; Carr *et al.*, 1993; Kawasaki *et al.*, 2004). The specific detection of glycopeptides can be achieved by monitoring specific diagnostic sugar oxonium ions such as *m/z* 204 (HexNAc) and 366 (HexHexNAc) produced by cone voltage fragmentation, or by precursor ion scanning (Duffin *et al.*, 1992; Carr *et al.*, 1993). However, when many *N*-glycosylation sites are present within a glycoprotein, the chromatogram becomes extremely complex and assignment of the glycopeptide ions is very difficult. We presented here an alternative strategy for the site-specific glycosylation analysis of a peptide and glycopeptide mixture using LC/ESI/MS/MS, where we acquired the product ion spectrum for all significant molecular ions in a data-dependent manner. Product ion spectra of molecular ions allow the specific detection of glycopeptides from a complex mixture of peptides based on the presence of diagnostic sugar oxonium ions of oligosaccharides. Furthermore, this method allows confirmation of the amino acid sequence of a glycopeptide by the presence of b- and y-series fragment ions of the peptide. Using this method, we identified one previously unidentified *N*-glycosylated site of ApoB100 and determined the oligosaccharide heterogeneity of each of 17 *N*-glycosylation sites. Our findings provide

information on the structure of apoB100 that will be useful to future studies on the structure, function and metabolism of plasma LDL.

Results

Enzyme digestion

To determine the oligosaccharide heterogeneity at each glycosylation site, reduced and carboxymethylated apoB100 was digested into peptides and glycopeptides. Table I shows the amino acid sequences of the tryptic or chymotryptic peptides including the putative *N*-glycosylation sites. The putative glycosylation sites were numbered (G1-G19). Bold type indicates the previously reported *N*-glycosylation sites (G2-6 and G9-19). When apoB100 is digested by trypsin, potential *N*-glycosylation sites, Asn1341 (G4) and Asn1350 (G5), belong to the same peptide. Because chymotrypsin can cleave apoB100 into glycopeptides containing one glycosylation site, we attempted to analyze both proteolytic fragments from trypsin digestion and chymotrypsin digestion to identify the site-specific glycosylation.

LC/ESI/MS/MS analysis of tryptic digest of apoB100

The schema of a site-specific glycosylation analysis of apoB100 is shown in Figure 1. A mixture of tryptic peptides and glycopeptides was subjected to LC/ESI/MS/MS with a reversed-phase column. Figure 2A shows a total ion chromatogram (TIC) of a TOF MS scan for the full scan *m/z* 1000-2000. When double or higher charged molecular ions were detected, the product ion spectrum was automatically acquired. Figure 2B shows a TIC of the product ion scan. The collision energy at the second quadrupole for the product ion

scan was adjusted from 50 to 80 eV depending on the size and charge of the precursor ion. Under these conditions, peptide precursor ions produced b- and y-series fragment ions derived from its amino acid sequence (data not shown), and glycopeptide precursor ions produced abundant carbohydrate-specific ions, *m/z* 204, 186, 168, and 366 (described later). The intensity of ions at *m/z* 204.05-204.15 (HexNAc, 204.08) in each product ion scan are illustrated in Figure 2C. The extracted ion chromatogram at *m/z* 204 (Figure 2C) and 366 (data not shown) provides useful information on the selection of glycopeptide precursor ions. The product ion spectra of glycopeptides show a very characteristic pattern (Figure 3, 4 and 6). There were intense oligosaccharide-derived peaks of *m/z* 204 (HexNAc), 366 (HexHexNAc), 186 (HexNAc - H_2O), and 168 (HexNAc - $2H_2O$), and if present, 163 (Hex), 292 (Neu5Ac) and 274 (Neu5Ac - H_2O). Therefore, we can very easily distinguish the glycopeptide precursor ions from peptide ions. As expected, many parent ions having 204 and 366 fragment ions in the product ion spectrum were detected and most of these precursor ions were found as glycopeptides.

The glycopeptides were assigned based on an examination of product ion spectra using the information on the peptides containing a putative *N*-glycosylation site. Figures 3A, 3B and 3C show the product ion spectra of 1412.1 (+2) at 18 min, 1160.4 (+3) at 20 min and 1271.1 (+3) at 22 min for the glycopeptides. There were intense carbohydrate B^+ ions such as *m/z* 204 (HexNAc), 366 (HexHexNAc) and 186 (HexNAc - H_2O) and other weak peaks in the product ion spectra. These product ion spectra were very similar to each other (Figure 3A, 3B and 3C). Careful examination of these product ion spectra for the glycopeptides revealed that several fragment ions were consistent with b- and y-series fragment ions derived from the peptide FVEGSHNSTVSLTTK (residue 3378-3392). The deduced b- and y-series fragment ions of the peptide FVEGSHNSTVSLTTK were listed, and the fragment ions detected in the product ion spectrum of 1160.4 (+3) are underlined in the table (Figure 3D). The molecular ions of the peptide (*m/z* 1606) and peptide + GlcNAc (*m/z* 1809) were also detected in the product ion spectra (Figure

3A, 3B and 3C). These results suggest that these glycopeptides have the same peptide, FVEGSHNSTVSLTTK, including the *N*-glycosylation site Asn3384 (G15). Carbohydrate molecular weight was calculated by subtracting the theoretical molecular weight of the peptide (1605.8) from the calculated molecular weight of the glycopeptide and adding the molecular weight of H_2O (18.0). The oligosaccharide structure was deduced based on the molecular weight and previously reported oligosaccharides of apoB100. The presence of product ions at *m/z* 274 (Neu5Ac-H_2O) and 292 (Neu5Ac) suggested that those at *m/z* 1160.4 (+3) and 1271.1 (+3) were glycopeptide ions having sialylated oligosaccharides. Thus, the carbohydrate compositions, $[HexNAc]_2[Hex]_5$, $[HexNAc]_3[Hex]_6[Neu5Ac]_1$ and $[HexNAc]_4[Hex]_5[Neu5Ac]_2$, were deduced from the carbohydrate molecular weights, 1234.4, 1890.4 and 2222.5, respectively.

Figures 4 shows the product ion spectra of 1294.8 (+3) at 55 min and 1152.7 (+3) at 35 min for other glycopeptides. There are intense carbohydrate B^+ ions in the product ion spectra. Several ions consisting of b- and y-series fragment ions from the peptide TIHDLHLFIENIDHNK (residue 2198-2213) were found in the product ion spectrum of 1294.8 (+3) (Figure 4A), and detected ions are underlined in the table. The molecular ions of the peptide (*m/z* 1968.9) were also detected in the product ion spectra. The carbohydrate molecular weight was calculated from the molecular weight of the peptide, 1968.0, and the molecular weight of the glycopeptide, 3881.4. Carbohydrate composition was deduced from the carbohydrate molecular weight (1931.4) and presence of Neu5Ac. Thus, the peptide moiety TIHDLHLFIENIDHNK and carbohydrate composition $[HexNAc]_4[Hex]_5[Neu5Ac]_1$ were suggested.

Many ions in the product ion spectrum of 1152.7 (+3) were consistent with the b- and y-series fragment ions derived from the peptide YSFNSSMLYSTAK (Figure 4B). However, the deduced peptide ion *m/z* at 1526.7 and peptide + GlcNAc at 1729.8 were not detected. The difference of 203 between the product ions at *m/z* 1542.5 and 1745.7 suggests that the molecular weight of the peptide moiety may be 1541.5, and an

increase in mass of 16 Da suggests that the methionine residue of YSFNSSMLYSTAK (residue 3435-3447, molecular weight 1525.7) was oxidized. The deduced b- and y-series fragment ions of the peptide, YSFNSSMLYSTAK, with the oxidized methionine were listed and detected peptide fragment ions were underlined. Thus, the product ions at *m/z* 1542.5 and 1745.7 were considered the peptide and peptide + GlcNAc ions, respectively. Our method identified unexpected oxidization of methionine residue (Fig 4B). The carbohydrate molecular weight was calculated, and the carbohydrate composition, $[HexNAc]_4[Hex]_5[Neu5Ac]_1$, was deduced from the carbohydrate molecular weight, 1931.4, and presence of Neu5Ac.

Results of site-specific glycosylation analysis from tryptic digest are summarized in Table II. In order to avoid miss-assignment, only ions that were confirmed as glycopeptides by the product ion spectra or co-eluting ions with glycopeptides were listed. We determined 13 of the 19 potential *N*-glycosylation sites and the oligosaccharide heterogeneity at each site in a site-specific glycosylation analysis of the tryptic digest of apoB100. The type of oligosaccharide was deduced based on the oligosaccharide composition. Glycopeptides containing *N*-glycosylation sites Asn7, 956, 1341, 1350, 2533 and 3074 (G1, 3, 4, 5, 8 and 11) could not be detected. The relative peak intensity does not accurately express the relative amount of glycoforms, because of the different ionization efficiency of each glycoform, different detection sensitivity at *m/z*, and simultaneous acquision of MS and MS/MS spectra. However, the relative peak intensity of each glycopeptide would provide an indication of the distribution in glycoforms.

When each product ion spectrum of the peptide ions in this LC/ESI/MS/MS analysis was identified by the computer program Mascot, the sequence coverage of apoB100 was 38 %. The ions, *m/z* 1177.9 (+3) at 64 min, 1289.0 (+3) at 91 min, 1053.2 (+3) at 84 min were identified as TIHDLHLFIENIDFN^{2212}KSGSSTASWIQNVDTK containing the potential *N*-glycosylation site Asn2212 (G7), SSVITLNTNAELFN^{3331}QSDIVAHLLSSSSSVIDALQYK, containing Asn3331 (G14),

and DFHSEYIVSASN^{4404}FTSQLSSQVEQFLHR containing Asn4404 (G19), respectively (data not shown). These results indicate that some parts of these glycosylation sites were not glycosylated. There were many unexplained peptides and glycopeptides in the digest (data not shown). This may be due to the unexpected digestion or nonspecific cleavage of apolipoprotein B100 as well as the multiple isoforms of the proteins.

LC/ESI/MS/MS analysis of chymotryptic digest of apoB100

In order to determine the carbohydrate at undetected glycosylation sites in the tryptic digest including Asn1341 and 1350 (G4 and G5), which belong to the same tryptic peptide, the chymotrypsin digest was analyzed by LC/ESI/MS/MS using the same methodology. Figure 5A shows TIC of the TOF MS scan for the full scan *m/z* 700-2000. The collision energy was adjusted at 40 to 80 eV depending on the precursor ions. TIC of the product ion scan and extracted ion chromatogram at *m/z* 204.05-204.15 (HexNAc) are presented in Figure 5B and 5C, respectively.

Figure 6A shows the product ion spectrum of 768.4 (+2) at 14 min for the chymotryptic glycopeptide. The carbohydrate B^+ ions, y1 and b2 ions of peptide NW (residue 1341-1342), and peptide + GlcNAc ion were found in the product ion spectrum. The carbohydrate composition, $[HexNAc]_2[Hex]_5$, was deduced from the calculated carbohydrate molecular ion, 1234.6. Figure 6B shows the product ion spectrum of 1444.1 (+2) at 9 min for the glycopeptide. The carbohydrate B^+ ions, peptide and peptide + GlcNAc ions, and peptide fragment ions from the peptide SGGNTSTDHF (residue 1347-1356) were detected in the product ion spectrum. Carbohydrate molecular weight, 1882.8, was calculated and the oligosaccharide composition, $[HexNAc]_2[Hex]_6$, was deduced from the molecular weight. The peptide fragment ions were also detected in the product ion spectrum for the chymotryptic glycopeptides as tryptic glycopeptides. The

peptide and peptide + GlcNAc ions were detected in product ion spectra. These ions helped us to determine the peptide moiety of the glycopeptide ion.

Results of the site-specific analysis of glycosylation of the chymotryptic digest are summarized in Table III. The oligosaccharide heterogeneity at each of 13 *N*-glycosylation sites was determined by LC/MS/MS from the chymotryptic digest of apoB100 (Table III).

Carbohydrate diversity of each site

Results for the tryptic digest and chymotrypsin digest of apoB100 are listed in Table IV. The oligosaccharide composition and type were deduced based on the molecular weight and previously reported oligosaccharide structures of apolipoprotein B100. No information on glycosylation at Asn7 and 2533 (G1 and 8) was obtained from the analysis of the tryptic or chymotryptic digest. When the tryptic digest of apoB100 was analyzed by LC/ESI/MS/MS with the MS range *m/z* 400-2000, the sequence coverage of apoB100 was 41 % and tryptic peptides containing Asn7, 2212, 2533 or 2955 (G1, 7, 8 or 10) were detected (data not shown). Together with the result of LC/ESI/MS/MS with the MS range *m/z* 1000-2000, Asn7 and 2533 (G1 and 8) were not glycosylated or glycosylated only under detection sensitivity, and Asn2212, 2955, 3331 and 4404 (G7, 10, 14 and 19) was partially glycosylated. These findings indicated that 17 of 19 potential *N*-glycosylation sites in apoB100 were glycosylated.

The most heterogeneous oligosaccharides were found at Asn3384 (G15). Asn3384 possessed neutral or monosialylated hybrid and monoantennary complex type, and mono- or disialylated biantennary complex type oligosaccharides as well as one high-mannose type oligosaccharide. Asn158, 1341, 1350, 3309 and 3331 (G2, 4, 5, 13 and 14) were occupied by high-mannose type oligosaccharides, while Asn956, 1496, 2212, 2752, 2955, 3074, 3197, 3438, 3868, 4210 and 4404 (G3, 6, 7, 9-12, and16-19) were predominantly

occupied by monosialylated or disialylated biantennary complex type oligosaccharides (Figure 7).

Discussion

Although the role of the carbohydrate structures in LDL and/or apoB100 has been examined in several studies (Attie *et al.*, 1979; Filipovic *et al.*, 1979; Shireman and Fisher, 1979; Orekhov *et al.*, 1989; Fujioka *et al.*, 2000), it is still unknown. It is necessary to elucidate the diversity of the oligosaccharides at each *N*-glycosylation site. This is the first report on the characterization of *N*-linked oligosaccharides in apoB100 at each glycosylation site. The protein was initially carboxymethylated and digested with an enzyme (trypsin or chymotrypsin), and then the complex mixtures of peptides and glycopeptides were subjected to LC/ESI/MS/MS analysis. Product ion scan of each precursor ion was carried out in a data-dependent manner. The glycopeptide molecular ions were easily distinguished from peptide ions by the presence of carbohydrate-related oxonium ions such as *m/z* 204 (HexNAc), 186 (HexNAc - H_2O), 168 (HexNAc - H_2O), 366 (HexHexNAc), and others in product ion spectra. Furthermore, product ion spectra provided information for the elucidation of the amino acid sequence of the glycopeptides. The oligosaccharide structure could be deduced based on the calculated molecular weight of the oligosaccharide moiety. The glycopeptide precursor ion was assigned using three strageties. 1) By comparing the product ions of the glycopeptide with the expected fragment ions derived from the peptide containing the *N*-linked glycosylation site, we could directly deduce the peptide moiety. The molecular weight of the oligosaccharide moiety was calculated from the observed molecular weight of the glycopeptide and the theoretical molecular weight of the identified peptide. The carbohydrate composition and structure were deduced from the calculated molecular weight of the oligosaccharide. 2) There were relatively intense peaks of the peptide and peptide + GlcNAc ions in the glycopeptide product

ion spectrum. Thus, the *m/z* difference of 203 between fragment ions in the product ion spectrum could suggest the molecular weight of the peptide moiety. The peptide was determined from this suggested molecular weight and the molecular weight of the peptides containing the putative *N*-glycosylation site. The molecular weight of the carbohydrate was calculated, and the carbohydrate composition and structure were deduced. 3) Possible glycopeptide masses were calculated from the peptide masses containing the *N*-linked glycosylation site and possible *N*-linked oligosaccharide masses. The possible glycopeptide mass with the measured mass of the glycopeptide was identified. Assignment of peptide moiety was confirmed by the presence of the fragment ions derived from the peptide in the product ion spectrum.

The elution time as well as mass of a glycopeptide is also helpful to elucidate the oligosaccharide structure. The glycopeptides were eluted following reversed-phase HPLC based on the peptide and further separated based on the structure of the attached oligosaccharide (Kawasaki *et al.*, 2004). The glycopeptides having the same amino acid sequence were eluted in order of the number of Neu5Ac. Our results show that LC/ESI/MS/MS with high sensitivity and high detection resolution is a powerful technique for the site-specific glycosylation analysis of glycoprotein.

Our study revealed that seventeen of the 19 potential *N*-glycosylation sites in apoB100 were glycosylated, and the diversity of oligosaccharides at each of these *N*-glycosylation sites was determined. The deduced oligosaccharide structures in the present study were consistent with the structures previously identified in apoB100 (Garner *et al.*, 2001). Asn2212, which was reported to be unglycosylated (Yang *et al.*, 1989), could be glycosylated. The *N*-glycan structures and patterns are very different at each site. Asn158, 1341, 1350, 3309 and 3331 were occupied by high-mannose type oligosaccharides. The other sites except Asn1496 and 3384 (G6 and G15) were predominantly occupied by monosialylated or disialylated biantennary complex type oligosaccharides and no neutral oligosaccharides were detected. These sialylated glycans may play an important biological role. Asn1496 and 3384 were occupied by

high-mannose, hybrid and complex type *N*-linked oligosaccharides. Hybrid type oligosaccharides were found only at these two sites. The oligosaccharides at Asn 3384 are most heterogeneous and at least 12 different oligosaccharide structures were present. Neutral complex type and neutral hybrid type oligosaccharides were detected only at this site. It is unlikely that this oligosaccharide heterogeneity is due to the fact that the apoB100 used in this study was extracted from the pooled serum of normolipidemic subjects, because no hybrid type oligosaccharides were detected except at Asn1496 and 3384 in this study and it was reported that the diversity of the oligosaccharides of apoB100 was highly conserved among subjects (Taniguchi *et al.*, 1989; Garner *et al.*, 2001). It may be suggested that the diversity of the oligosaccharides at each glycosylation site was also conserved among subjects.

The relationship between sialylation and LDL-receptor binding has been examined. Desialylation of LDL increased the internalization of LDL by aortic smooth muscle cells (Filipovic *et al.*, 1979) and macrophage (Fujioka *et al.*, 2000) aortic intimal cells (Orekhov *et al.*, 1989), but had no effect on degradation in hepatocytes (Attie *et al.*, 1979). These findings appear controversial. Asn3309, 3331 and 3384 are located near the LDL receptor-binding site in apoB100 (residues 3359-3369). Our data showed that these glycosylation sites were populated by high-mannose type (at Asn3309 and 3331) or a variety of oligosaccharides including neutral or sialylated oligosaccharides (at Asn3384). These findings may indicate that sialic acid residues of apoB100 did not play a significant role in LDL-receptor binding and that desialylated LDL might be internalized by another mechanism. Shireman and Fisher (1979) reported that the removal of carbohydrate from LDL did not alter its binding to fibroblasts. Thus, the carbohydrate moieties of LDL might not have a significant role in LDL-receptor binding.

The most interesting observation was that the most heterogeneous oligosaccharides were found at the *N*-glycosylation site (Asn3384) nearest to the LDL receptor-binding site. ApoB100 enwraps the very low-density lipoprotein (VLDL) and LDL particle. The C-terminal crosses over near the LDL receptor

binding site and inhibits binding of VLDL to the LDL receptor (Boren *et al.*, 1998). Conversion of VLDL to smaller LDL allows interaction with the LDL receptor. It is likely that the size of the VLDL/LDL particle could affect the 3D conformation around here. Thus, the variety of oligosaccharide at Asn3384 may reflect the local 3D conformation of the VLDL particle and accessibility of trimming and glycosyl transferase enzymes.

The procedure described in this report provides an easy and efficient method for the identification of glycosylation sites and oligosaccharide heterogeneity of glycoproteins. Site-specific glycosylation analysis of apoB100 revealed that the diversity of oligosaccharide was distinct at each site. These data provide information to understand the role of oligosaccharides of apoB100 in LDL particles

Materials and methods

Materials

Acetonitrile, formic acid, chymotrypsin and guanidine hydrochloride were from Wako Pure Chemicals Industries (Osaka, Japan). Tosylphenylalanine chloromethane-treated trypsin (TPCK-treated trypsin) was from Sigma (St. Louis, MO). Human apolipoprotein B100 (ApoB100) was purchased from MP Biomedicals Inc. (Irvine, CA). This product is derived from pooled human plasma, which is not particularly high fat plasma. The water used was obtained from a Milli-Q water system (Millipore, Bedford, MA). All other reagents were of the highest quality available.

Reduction and S-carboxymethylation of apoB100

Apolipoprotein B100 (500 μg) was dissolved in 810 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 6 μl of 2-mercaptoethanol, the mixture was incubated for 2 h at 40 °C. Then, 17 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 Pharmacia Biotech, Uppsala, Sweden) to remove the reagents, and the eluate was lyophilized.

Enzyme digestion of apoB100

Reduced and carboxymethylated apoB100 was redissolved in 500 μl of 0.1 M Tris-HCl buffer (pH 8.0). Half of the reduced and carboxymethylated apoB100 was incubated with 0.02 μg/μl of TPCK-treated trypsin (1/50 w/w) for 2 h at 37 °C and the rest was incubated with 0.04 μg/μl of chymotrypsin (1/25 w/w)

for 72 h at 37 °C. The enzyme digestions were stopped by storing at -20 °C before analysis.

HPLC of trypsin or chymotrypsin-digested apoB100

Tryptic digest (4 μg, about 8 pmol) and chymotryptic digest (2 μg, about 4 pmol) were analyzed by LC/ESI/MS/MS. HPLC was performed on a Paradigm MS 4 equipped with a Magic C18 column (0.2 × 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 or chymotrypsin-digested samples of apoB100 were eluted with 5 % of B for 10 min followed by a liner gradient from 5 to 70 % of pump B in 130 min at a flow-rate of 2 μl/min.

ESI Q-TOF MS/MS

Mass spectrometric analyses were performed using a QSTAR Pulsar i quadrupole time of flight (Q-TOF) mass spectrometer (AB/MDS Sciex, Toronto, Canada) equipped with a nano-electrospray ion source. The mass spectrometer was operated in the positive ion mode. The nano spray voltage was set at 2500 V. Mass spectra for MS analysis were acquired over *m/z* 1000-2000 and 700-2000 for tryptic and chymotryptic digests, respectively, and for MS/MS analysis, over *m/z* 100-2000. After every regular MS acquision, MS/MS acquision was performed against multiple charged ions. The molecular ions were selected by data-dependent acquiring in the quadrupole analyzer and fragmented in the hexapole collision cell. The collision energy was varied between 40 and 80 eV depending on the size and charge of the molecular ion. All signals were monoisotopically resolved. Accumulation time of spectra is 1.0 and 2.0 sec for MS and MS/MS, respectively.

Acknowledgments

We thank Dr. Nishimaki-Mogami for helpful suggestions. This study was supported by a grant-in-aid for Research on Health Sciences focusing on Drug Innovation from the Japan Health Sciences Foundation.

Abbreviations

apoB100, apolipoprotein B100; ESI, electrospray ionization; LC, liquid chromatography; LDL, low density lipoprotein; MS, mass spectrometry; VLDL, very low density lipoprotein; TIC, total ion chromatogram; TOF, time of flight

References

Attie, A.D., Weinstein, D.B., Freeze, H.H., Pittman, R.C., and Steinberg, D. (1979) Unaltered catabolism of desialylated low-density lipoprotein in the pig and in cultured rat hepatocytes. *Biochem. J.*, **180**, 647-654.

Boren, J., Lee, I., Zhu, W., Arnold, K., Taylor, S., and Innerarity, T.L. (1998) Identification of the low density lipoprotein receptor-binding site in apolipoprotein B100 and the modulation of its binding activity by the carboxyl terminus in familial defective apo-B100. *J. Clin. Invest.*, **101**, 1084-1093.

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

Chen, S.H., Yang, C.Y., Chen, P.F., Setzer, D., Tanimura, M., Li, W.H., Gotto, A.M.Jr, and Chan, L. (1986) The complete cDNA and amino acid sequence of human apolipoprotein B-100. *J. Biol. Chem.*, **261**, 12918-12921.

Duffin, K.L., Welply, J.K., Huang, E., and Henion, J.D. (1992) Characterization of *N*-linked oligosaccharides by electrospray and tandem mass spectrometry. *Anal. Chem.* **64**, 1440-1448.

Filipovic, I., Schwarzmann, G., Mraz, W., Wiegandt, H., and Buddecke, E. (1979) Sialic-acid content of low-density lipoproteins controls their binding and uptake by cultured cells. *Eur. J. Biochem.*, **93**, 51-55.

Fujioka, Y., Taniguchi, T., Ishikawa, Y., Yokoyama, M. (2000) Significance of acidic sugar chains of apolipoprotein B-100 in cellular metabolism of low-density lipoproteins. *J. Lab. Clin. Med.*, **136**, 355-362.

Garner, B., Harvey, D.J., Royle, L., Frischmann, M., Nigon, F., Chapman, M.J., and Rudd, P.M. (2001) Characterization of human apolipoprotein B100 oligosaccharides in LDL subfractions derived from normal and hyperlipidemic plasma: deficiency of alpha-*N*-acetylneuraminyllactosyl-ceramide in light and small dense LDL particles. *Glycobiology*, ***11***, 791-802.

Kawasaki, N., Ohta, M., Itoh, S., and Hayakawa, T. (2004) Analyses of glycopeptides and glycoproteins by liquid chromatography-mass spectrometry and liquid chromatography-tandem mass spectrometry. *Methods Mol. Biol.*, **251**, 263-274.

Knott, T.J., Pease, R.J., Powell, L.M., Wallis, S.C., Rall, S.C.Jr, Innerarity, T.L., Blackhart, B., Taylor, W.H., Marcel, Y., Milne, R., Johnson, D., Fuller, M., Lusis, J., McCarthy, B.J., Mahley, R.W., Levy-Wilson, B., and Scott, J. (1986) Complete protein sequence and identification of structural domains of human apolipoprotein B. *Nature*, **323**, 734-738.

Law, S.W., Grant, S.M., Higuchi, K., Hospattankar, A., Lackner, K., Lee, N., and Brewer, H.B.Jr. (1986) Human liver apolipoprotein B-100 cDNA: complete nucleic acid and derived amino acid sequence. *Proc. Natl. Acad. Sci.*, **83**, 8142-8146.

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

Orekhov, A.N., Tertov, V.V., Mukhin, D.N., and Mikhailenko, I.A. (1989) Modification of low density lipoprotein by desialylation causes lipid accumulation in cultured cells: discovery of desialylated lipoprotein with altered cellular metabolism in the blood of atherosclerotic patients. *Biochem. Biophys. Res. Commun.*, **162**, 206-211.

Segrest, J.P., Jones, M.K., Mishra, V.K., Anantharamaiah, G.M., and Garber, D.W. (1994) ApoB-100 has a pentapartite structure composed of three amphipathic alpha-helical domains alternating with two amphipathic beta-strand domains. Detection by the computer program LOCATE. *Arterioscler Thromb.*, **14**, 1674-1685.

Shireman, R.B., and Fisher, W.R. (1979) The absence of a role for the carbohydrate moiety in the binding of apolipoprotein B to the low density lipoprotein receptor. *Biochim. Biophys. Acta.*, **572**, 537-540.

Taniguchi, T., Ishikawa, Y., Tsunemitsu, M., and Fukuzaki, H. (1989) The structures of the asparagine-linked sugar chains of human apolipoprotein B-100. *Arch. Biochem. Biophys.*, **273**, 197-205.

Vukmirica, J., Nishimaki-Mogami, T., Tran, K., Shan, J., McLeod, R.S., Yuan, J., and Yao, Z. (2002) The N-linked oligosaccharides at the amino terminus of human apoB are important for the assembly and secretion of VLDL. *J. Lipid. Res.*, **43**, 1496-1507.

Yang, C.Y., Chen, S.H., Gianturco, S.H., Bradley, W.A., Sparrow, J.T., Tanimura, M., Li, W.H., Sparrow, D.A., DeLoof, H., Rosseneu, M., Lee, F.S., Gu, Z.W., Gotto, A.M.Jr., and Chan, L. (1986) Sequence, structure, receptor-binding domains and internal repeats of human apolipoprotein B-100. *Nature*, **323**, 738-742.

Yang, C.Y., Gu, Z.W., Weng, S.A., Kim, T.W., Chen, S.H., Pownall, H.J., Sharp, P.M., Liu, S.W., Li, H.W., Gotto, A.M.Jr., and Chan, L. (1989) Structure of apolipoprotein B-100 of human low density lipoproteins. *Arteriosclerosis*, **9**, 96-108.

Table I. The amino acid sequences of the tryptic or chymotryptic peptides including the putative *N*-glycosylation sites in apolipoproteir

N-Glycosylation site		Tryptic digests	Theoretical	Chymotryptic digests
residue	ID	Sequence	Mass[b]	Sequence
Asn[7]	G1	EEBMLEN[7]VSLVCPK	1677.8	EN[7]VSL
Asn[158]	**G2**	QVLFLDTVYGN[158]CSTHFTVK	2229.1	GN[158]CSTHF
Asn[956]	**G3**	QVTPGLNYCTSGAYSN[956]ASSTDSASYYPLTGDTR	3550.5	SN[956]ASSTDSASY
Asn[1341]	**G4**	LYQLQVPLLGVLDLSTNVYSNLYN[1341]WSASYSGGN[1350]TS	4692.3	N[1341]W
Asn[1350]	**G5**			SGGN[1350]TSTDHF
Asn[1496]	**G6**	FN[1496]SSYLQGTNQITGR	1684.8	N[1496]SSY
Asn[2212]	G7	TIHDLHLFIENIDFN[2212]K	1968.0	N[2212]KSGSSTASW
Asn[2533]	G8	N[2533]LTDFAEQYSIQDWAK	1928.9	AAKN[2533]L
Asn[2752]	**G9**	IQSPLFTLDANADIGN[2752]GTTSANEAGIAASITAK	3231.6	DANADIGN[2752]GTTSANEAGIAASITAI
Asn[2955]	**G10**	VNQNLVYESGSLN[2955]FSK	1797.9	N[2955]F
Asn[3074]	**G11**	YNQN[3074]FSAGNNENIMEAHVGINGEANLDFLNIPLTIPEMF	4359.1	NQN[3074]F
Asn[3197]	**G12**	SYN[3197]ETK	740.3	N[3197]ETKIKF
Asn[3309]	**G13**	ELCTISHIFIPAMGN[3309]ITYDFSFK	2704.3	IPAMGN[3309]ITY
Asn[3331]	**G14**	SSVITLNTNAELFN[3331]QSDIVAHLLSSSSSVIDALQYK	3864.0	N[3331]QSDIVAHL
Asn[3384]	**G15**	FVEGSHN[3384]STVSLTTK	1605.8	VEGSHN[3384]STVSL
Asn[3438]	**G16**	YDFN[3438]SSMLYSTAK	1525.7	N[3438]SSML
Asn[3868]	**G17**	FEVDSPVYN[3868]ATWSASLK	1912.9	N[3868]ATW
Asn[4210]	**G18**	VHN[4210]GSEILFSYFQDLVITLPFELR	2836.5	SKVHN[4210]GSEIL
Asn[4404]	**G19**	DFHSEYIVSASN[4404]FTSQLSSQVEQFLHR	3155.5	IVSASN[4404]F

Human apolipoprotein B100 amino acid sequence (NP_000375. apolipoprotein B [gi:4502153]) was obtained fr

[a] Potential *N*-glycosylation sites were identified with the consensus sequence NXS/T, where X is any amino aci

[b] Monoisotopic mass value.

Bold type indicates previously reported *N*-glycosylation sites.

Cystein residue was carboxymethylated and carboxymethylated cystein was underlined.