Glycobiology vol. 15 no. 12 pp. 1332–1340, 2005
doi:10.1093/glycob/cwj019
Advance Access publication on July 21, 2005

Differential analysis of site-specific glycans on plasma and cellular fibronectins: application of a hydrophilic affinity method for glycopeptide enrichment

Michiko Tajiri[2], Shumi Yoshida[3] and Yoshinao Wada[1,2,3,4]

[2]CREST, Japan Science and Technology Agency, 4-1-8 Honcho Kawaguchi, Saitama 332-0012, Japan; [3]Graduate School of Medicine, Osaka University, 2-2 Yamadaoka, Suita 565-0871, Japan; and [4]Department of Molecular Medicine, Osaka Medical Center and Research Institute for Maternal and Child Health, 840 Murodo-cho Izumi, Osaka 594-1101, Japan

Received on May 17, 2005; revised on July 16, 2005; accepted on July 18, 2005

Isolation of glycopeptides utilizing hydrogen bonding between glycopeptide glycans and a carbohydrate-gel matrix in the organic phase is useful for site-specific characterization of oligosaccharides of glycoproteins, when combined with mass spectrometry. In this study, recovery of glycopeptides was improved by including divalent cations or increasing the organic solvent in the binding solution, without losing specificity, whereas it was still less effective for those with a long peptide backbone exceeding 50 amino acid residues. The method was then applied to the analysis of glycan heterogeneities at seven N-glycosylation sites in each of the plasma and cellular fibronectins (FNs). There was a remarkable site-specific difference in fucosylation between these isoforms; Asn1244 selectively escaped the global fucosylation of cellular FN, whereas only Asn1007 and Asn2108 of the plasma isoform underwent modification. In addition, a new O-glycosylation site was identified at Thr279 in the connecting segment between the fibrin- and heparin-binding domain and the collagen-binding domain, and the glycopeptide was reactive to a peanut agglutinin lectin. Considering that another mucin-type O-glycosylation site lies within a different connecting segment, the O-glycosylation of FN was suggested to play a significant role in segregating the neighboring domains and thus maintaining the topology of FN and the domain functions. In addition, the method was applied to apolipoprotein B-100 (apoB100) whose *N*-glycan structures at 17 of 19 potential sites have been reported, and characterized the remaining sites. The results also demonstrated that the enriched glycopeptide provides resources for site-specific analysis of oligosaccharides in glycoproteomics.

Key words: apolipoprotein B-100/fibronectin/matrix-assisted laser desorption ionization mass spectrometry/N-linked oligosaccharide/O-linked oligosaccharide

Introduction

Methodological development is emerging as an essential issue in glycoproteomics, a field combining proteomics and glycomics (Varki, 1993; Banks *et al*., 2000; Harvey, 2001; Spiro, 2002). Recent demands for glycoprotein analysis have focused on site-specific features, namely protein sequence, glycosylation sites and glycan structures, because these are required to understand carbohydrate function in the local protein folding relevant to the overall glycoprotein functions (Rudd and Dwek, 1997; Kuster *et al*., 2001). To this end, digestion with specific endoproteinases is employed to separate the glycosylation sites on individual peptides, and reversed-phase high-performance liquid chromatography (HPLC) coupled with electrospray ionization (ESI) mass spectrometry (MS) is then used for characterization of both glycan and peptide structures of these glycopeptides (Carr *et al*., 1993; Huddleston *et al*., 1993). A striking feature of this technique is that carbohydrate-specific ions are generated by collision-induced dissociation (CID) in the electrospray source region or in the tandem MS setup and can be utilized to detect glycopeptides with high sensitivity during the chromatographic separation of complex digest mixtures, even when unmodified peptides are more abundant than glycosylated peptides. This technique was first reported by Carr and co-workers and soon came to be widely used (Burlingame, 1996). However, in cases in which many peptides with the same elution time are introduced into the mass spectrometer from HPLC, the mass spectrum is too complicated, because of multiple-charged ions, to characterize oligosaccharide structures. Moreover, the ionization of glycopeptides is suppressed in the presence of unmodified peptides, leading to underrepresentation, or complete absence of glycopeptide peaks.

Taking a different approach, a variety of methods has been used to enrich glycopeptides from complex peptide mixtures. Porous graphitized carbon has an affinity for oligosaccharides and is especially effective in separating oligosaccharide isomers (Davies *et al*., 1993; Koizumi, 1996; Larsen *et al*., 2005). This method can be applied to isolation of glycopeptides with smaller peptide backbones but does not produce sufficiently selective enrichment of tryptic glycopeptides. Lectin-affinity chromatography is used to isolate glycopeptides with specific oligosaccharide structures but cannot be used to prepare a glycopeptide pool suitable for characterization of total glycan heterogeneity (Fu and van Halbeek, 1992; Garcia *et al*., 1995; Bunkenborg *et al*., 2004). Recently, we reported a simple method utilizing the hydrogen bonding between carbohydrates of glycopeptides and carbohydrate-based gels such as cellulose or related substrates (Wada *et al*., 2004). Obviously, the method was

[1]To whom correspondence should be addressed; e-mail: waday@mch.pref.osaka.jp

a derivative of paper chromatography (Betts, 1964). The enriched glycopeptide mixture can be directly analyzed by MS for oligosaccharide structures at individual glycosylation sites. Alternatively, the glycopeptides are subjected to HPLC or LC/MS, especially when there are numbers of glycosylation sites in the protein. This strategy reduces the load on the resulting mass spectrum and chromatogram, as compared to the direct infusion of an enzymatic digest as described above. This method is potentially effective for *O*-glycans, but its application has not yet been reported.

Fibronectin (FN) is a large glycoprotein whose primary role is to attach cells to a variety of extracellular matrices (Ruoslahti, 1988). FN is a dimer of two similar polypeptides linked at their COOH-terminals by two disulfide bonds, and each chain is ~2350 amino acids long. Each chain contains six domains harboring specific binding sites for sulfated proteoglycans, fibrin, denatured forms of collagen, DNA, and cell-surface integrins (Hynes, 1985). There are multiple isoforms of FN (Yamada, 1983; Mosher, 1984). One, called plasma FN, is soluble and circulates in the blood and other body fluids, where it is thought to enhance blood clotting, wound healing, and phagocytosis. All other forms assemble on the cell surface and are deposited in the extracellular matrix as highly insoluble FN fibrils, thus regulating the shape of cells and the organization of the cytoskeleton and playing essential roles in migration and cellular differentiation of many cell types. There are seven potential N-glycosylation sites in human FN, and these *N*-glycans are largely responsible for the carbohydrate content of this molecule. The oligosaccharide structures have been reported in some detail; the *N*-glycan of plasma FN is composed of complex-type biantennary oligosaccharides and is largely sialylated, whereas fibroblast-derived cellular FN contains fucose linked to the innermost *N*-acetylglucosamine (GlcNAc) and is less sialylated (Fukuda and Hakomori, 1979; Takasaki *et al.*, 1979; Fukuda *et al.*, 1982). Additionally, a difference in the linkage of sialic acid between these FN isoforms were also recognized (Fukuda and Hakomori, 1979), and oncofetal or tissue-specific alterations of glycan structures have been reported as well (Wagner *et al.*, 1981; Isemura *et al.*, 1984; Zhu *et al.*, 1984; Krusius *et al.*, 1985). However, these data were not site-specific but rather were obtained on a global basis, and the O-glycosylation has not been adequately studied.

Herein, the glycopeptide enrichment method was improved in recovery, and applied to the analysis of site-specific *N*- and *O*-glycans of plasma and cellular FN isoforms. In addition, the *N*-glycan data of apolipoprotein B-100 (apoB100), whose *N*-glycan structures were very recently characterized at 17 of 19 potential sites (Harazono *et al.*, 2005), were completed.

Results and Discussion

Efficiency of hydrophilic affinity isolation of glycopeptides

Isolation of glycopeptides from an enzymatic digest of glycoproteins with cellulose or Sepharose is based on the hydrogen bonding between carbohydrate oxygens of the gel-matrix and glycopeptides (Wada *et al.*, 2004). In our previous study, the binding was carried out in a solution of 1-butanol/ethanol/H_2O (4:1:1, v/v) according to a cellulose

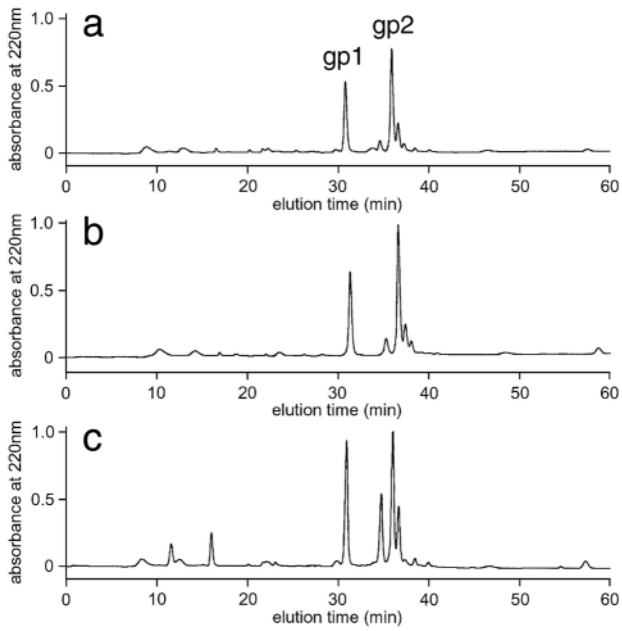

Fig. 1. Reversed-phase chromatograms of a glycopeptide fraction isolated from a tryptic digest of human transferrin by the hydrophilic affinity method with different binding solvents. A tryptic digest of carbamidomethylated human transferrin (100 μg) was mixed with Sepharose CL4B in different binding solutions: (**a**) 1-butanol/ethanol/H_2O (4:1:1, v/v), (**b**) 1-butanol/ethanol/H_2O (5:1:1), (**c**) 1-butanol/ethanol/H_2O (5:1:1, v/v) containing 1 mM $MnCl_2$. The bound components were recovered and analyzed. The major peaks in the chromatogram are two glycopeptides, CGLVPVLAENYNK (gp1) and QQQHLFGSNVTDCSGNFCLFR (gp2). Recovery was evaluated according to the intensity of these signals.

chromatography method for oligosaccharides (Shimizu *et al.*, 2001), and the recovery of glycopeptides utilizing a gel-matrix Sepharose CL4B was 30% for a few hundred microgram tryptic digest of human transferrin. In our attempts to improve the method, a higher concentration of 1-butanol as 1-butanol/ethanol/H_2O (5:1:1, v/v) and addition of divalent cations improved recovery by 20–40% without a substantial increase in the contamination of unglycosylated peptides, as demonstrated by the glycopeptides from human transferrin bearing biantennary oligosaccharides (Figure 1). Alternatively, addition of divalent cations especially of manganese in the binding solution afforded a reproducible improvement, whereas other elements such as Mg, Ca, Ba, Co(II), and Ni(II) were also effective. This was because divalent cations form complexes with saccharides (Das *et al.*, 2001; Davis and Brodbelt, 2005) and thereby enhance the binding of glycopeptides for isolation.

With respect to the limit of peptide backbone length, our method was effective up to ~40–50 amino acids for N-linked glycopeptides as described below. Peptides containing few amino acid residues with a hydroxyl side group, such as serine, threonine, aspartic acid, and glutamic acid, tend to enter the glycopeptide fraction (Wada *et al.*, 2004).

Site-specific N*-glycans of plasma and cellular FNs*

Plasma FN is secreted into the circulation by hepatocytes, and the insoluble cellular form is synthesized by fibroblasts

and other cell types. Although alternative mRNA splicing of a common transcript gives rise to structural polymorphism between various forms of FN as well as between different subunits within each type, seven potential N-glycosylation sites are conserved among all isoforms. The major *N*-glycan of cellular and plasma FNs is a biantennary complex-type structure, although significant variation has been detected between these isoforms (Fukuda and Hakomori, 1979; Takasaki *et al.*, 1979). Specifically, Fucα1→6 linked to the innermost GlcNAc residue of the chitobiosyl core is found only in cellular FN, but is nearly absent from plasma one (Fukuda and Hakomori, 1979; Fukuda *et al.*, 1982). On the other hand, the plasma isoform undergoes extensive α2→6 sialylation at the terminal galactose residue of the biantennary side chain, whereas the cellular isoform is less sialylated (Takasaki *et al.*, 1979; Fukuda *et al.*, 1982). In addition, other structures of multiantennary and polylactosamine types are found in the cellular FN derived from fetal tissues or transformed cells (Wagner *et al.*, 1981; Zhu *et al.*, 1984). These variations have been described on a global basis but not site specifically.

FN contains nearly 200 basic amino acid residues, and complete cleavage with trypsin will yield 64 peptides (more than two amino acids long) from an isoform with a maximum of 2355 residues. First, to unravel the site-specific differences among attached glycans, tryptic glycopeptides were enriched by the method described above. The glycopeptides were separated by reversed-phase HPLC (Figure 2a), because direct analysis of glycopeptide mixtures, which was effective for glycoproteins containing a few glycosylation sites, was too complex to characterize the individual sites. Major peaks in the chromatogram corresponded to glycopeptides containing Asn430, Asn528, Asn542, Asn1007, or Asn1244 as determined by MS. On the other hand, the large glycopeptides containing Asn877 and Asn2108 were barely detected, but were easily identified in the reversed-phase chromatogram of the digest, in which only several peptides of large size and thus with long retention times were present (Figure 2b). These peptides were collected and digested with endoproteinase Asp-N (AspN), and subjected to MS.

The identification and characterization of glycopeptides were carried out using matrix-assisted laser desorption ionization mass spectrometry (MALDI MS), because it allows the generated ions to be scrutinized by repetitive analyzes of the same sample spot. Two different modes of ion separation were employed in this study; linear time-of-flight mass spectrometry (TOF MS) produced an outline of glycan heterogeneity and relative abundances, and quadrupole ion trap (QIT) TOF MS enabled us to perform multiple-stage tandem MS (MS^n) which characterizes individual ions in detail (Koy *et al.*, 2003; Demelbauer *et al.*, 2004; Wada *et al.*, 2004).

The MALDI linear TOF mass spectra of the glycopeptides containing Asn1007 are shown in Figure 3, in which differences in sialylation and fucosylation were evident between plasma and cellular isoforms, and triantennary chains were present at this site. The illustrated structures in the figures were derived from MALDI QIT MS. Taking Asn542 as an example (Figure 4), each of the ions in the high mass region (>*m*/*z* 3000) generated the same series of product ions in the low mass region (*m*/*z* 1800 to *m*/*z* 2400),

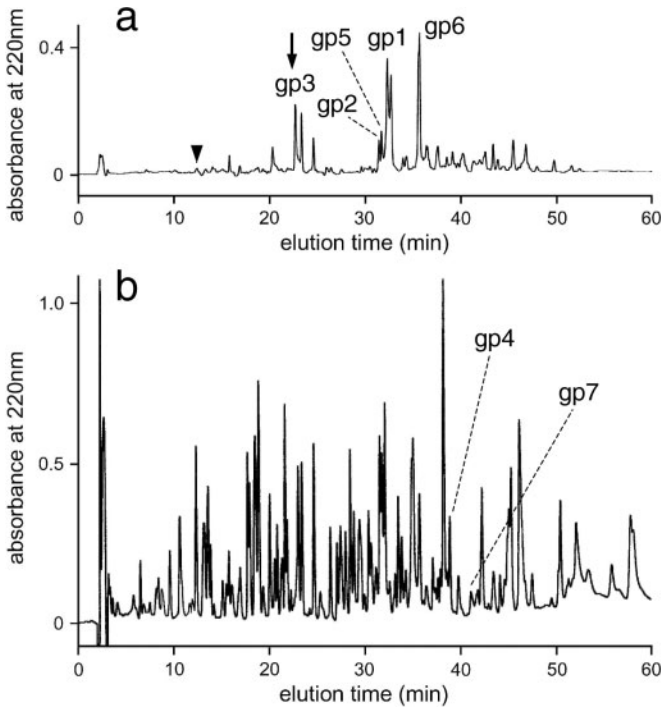

Fig. 2. Reversed-phase chromatograms of tryptic peptides from plasma fibronectin (FN). (**a**) The Sepharose-binding fraction from a tryptic digest of carbamidomethylated human FN (100 μg) was recovered and analyzed. (**b**) Total digest. FN has seven *N*-glycosylation sites, and the derived glycopeptides are numbered herein from the NH_2-terminal. *N*-Glycosylation sites are as follows: gp1, Asn430; gp2, Asn528; gp3, Asn542; gp4, Asn877; gp5, Asn1007; gp6, Asn1244; gp7, Asn2108. Large glycopeptides, gp4 and gp7, were identified in the total digest. Arrow and arrowhead denote the O-glycosylated peptides containing Thr279 and Thr 2064/2065, respectively.

indicating all of these signals to be derived from the same glycopeptide with a 1931.8 Da peptide backbone. In the succeeding CID, the ion at *m*/*z* 1932.8 generated product ions that specified the peptide sequence (data not shown). The fucosylation at the innermost GlcNAc of the chitobiose core was determined by the product ion representing the intact peptide moiety with retaining fucosyl GlcNAc (i.e., the signal at *m*/*z* 2282.1 in Figure 4b), whereas the assumed anomeric α1→6 linkage structure was based on previous studies (Fukuda and Hakomori, 1979; Fukuda *et al.*, 1982).

In the linear TOF mass spectrum, signal intensities of each glycoform on the same peptide backbone approximately represented the relative abundances of the individual glycopeptide molecules (Mortz *et al.*, 1996). This is because the protonation occurs at the peptide moiety and the efficiency of ionization is thus scarcely affected by oligosaccharide structures except for those with charged groups such as sialic acid. However, the outline of sialylation can be estimated by a linear TOF mass spectrum, even though it is likely to be eliminated in-source under strong laser irradiation (Mortz *et al.*, 1996; Papac *et al.*, 1996).

The glycan structures are summarized in Table I. Obviously, Asn1244 was excluded from the global fucosylation

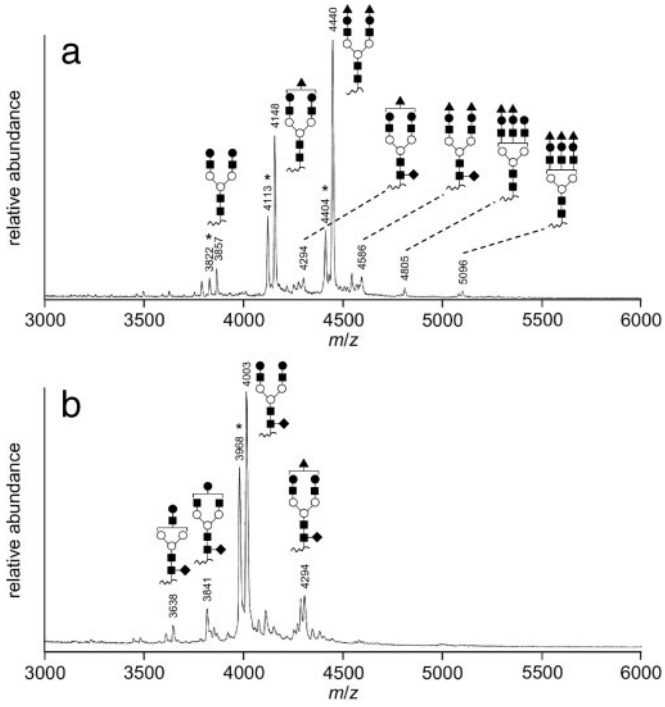

Fig. 3. Matrix-assisted laser desorption ionization (MALDI) linear TOF mass spectra of the glycopeptide containing Asn1007 from human fibronectin (FN). (**a**) Plasma FN. (**b**) Cellular FN. The intensity of the signals roughly represents the relative abundances of the molecules. The signals indicated by asterisks are derived from an Asn528-containing glycopeptide. ■, *N*-acetylglucosamine; ○, mannose; ●, galactose; ◆, fucose; ▲, *N*-acetylneuraminic acid.

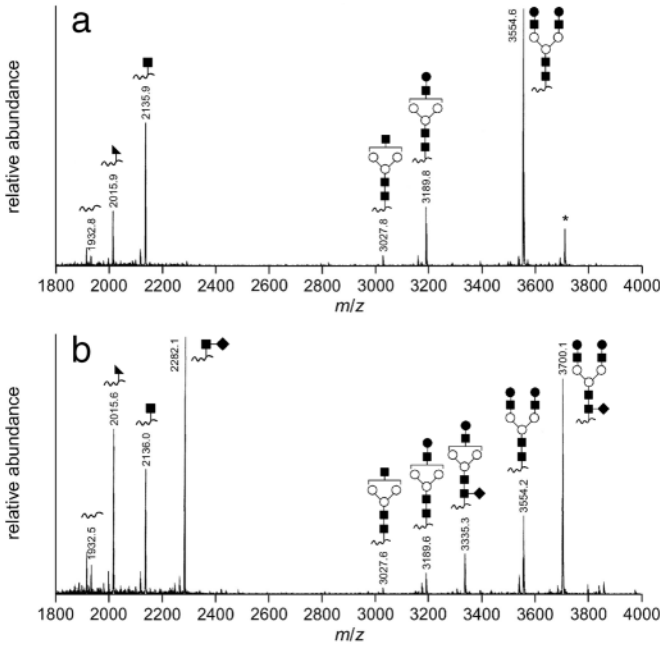

Fig. 4. Matrix-assisted laser desorption ionization quadrupole ion trap time-of-flight (MALDI QIT TOF) mass spectra of the glycopeptide containing Asn542 from human fibronectin (FN). (**a**) Plasma FN. (**b**) Cellular FN. Symbols are as described in the legend to Fig. 3. The fragment ion at *m/z* 2282.1 indicates a fucose linked to the innermost GlcNAc residue of the chitobiosyl core. The ion marked by an asterisk corresponds to the peptide with an additional Arg at the NH_2-terminal of the peptide at *m/z* 3554.6.

of cellular FN, whereas only Asn1007 and Asn2108 of the plasma isoform were fucosylated. Lactosamine elongation, which was reported in the collagen-binding domain (Asn430, Asn528, and Asn542) of cellular FN from placenta or amniotic fluid (Zhu *et al*., 1984; Krusius *et al*., 1985), was quite minor, if any was present, in the preparation studied (Figure 3b). This was consistent with a relatively low carbohydrate content of FNs from adult plasma and from cultured normal skin fibroblasts (Fukuda and Hakomori, 1979; Takasaki *et al*., 1979).

The results also suggest the capacity of the this method for glycopeptide enrichment. As shown in the reversed-phase chromatogram (Figure 2a), glycopeptides 16–44 amino acids in length were efficiently recovered, whereas those containing Asn877 and Asn2108 were not, probably because of their peptide backbone sizes of 61 and 58 residues, respectively.

O-Glycosylation of FN

In contrast to N-glycosylation, O-glycosylation of FN has not been adequately studied. In 1985, Krusius *et al*. (1985) reported the presence of 4.2 and 0.9 mol of core one Galβ1→3GalNAc residue in FNs from amniotic fluid and plasma, respectively, and that these glycans are sialylated. Subsequently, this type of O-glycosylation was detected by a specific antibody in both N- and C-terminal halves of the cellular and plasma FNs (Nichols *et al*., 1986). The latter site of plasma FN was located at either Thr2064 or Thr2065 in connecting strand 3, or IIICS, between heparin- and fibrin-binding domains (Ichihara-Tanaka *et al*., 1990; Tressel *et al*., 1991). In synovial fluid (cellular) FN, another O-glycosylation site was suggested within a region immediately COOH-terminal to the collagen-binding domain, based on lectin blotting (Carsons *et al*., 1987; Carsons, 2002).

O-Glycosylated peptides were undetectable in the first screening of tryptic glycopeptides. However, two peptides probably bearing the O-linked oligosaccharides were identified in the glycopeptide-enriched fraction from a chymotryptic digest. A representative MALDI QIT mass spectrum is shown in Figure 5; the spaces between the peaks suggested O-glycosylation, and tandem MS of the ions at *m/z* 2330.89 and *m/z* 2603.91 yielded the same product ion at *m/z* 1965.82, which represented the protonated molecule of a peptide, KCERHTSVQTTSSGSGPF (positions 269–286). Furthermore, the b-ion series of product ions in the low mass region determined the glycosylation at Thr279 (Figure 5b). The mucin-type structure, Galβ1→3GalNAc, was confirmed by the affinity of the glycopeptide to peanut agglutinin lectin (data not shown) and cellular FN was occupied by the same structure at the site. In a similar way, a single mucin-type *O*-glycan was also identified in a chymotryptic peptide RRTTPPTTATPIRHRPRPYPPNVGEEIQIGHIPREDVDY (positions 2058–2096) of connecting strand 3, or IIICS, in which a single galactosamine-based carbohydrate group had been reported at either position Thr2064 or Thr2065 (Tressel *et al*., 1991). Retrospectively, the corresponding peptides containing Thr279 and Thr2064/

Table I. Site-specific *N*-glycans of plasma and cellular fibronectins (FNs)

	Plasma FN	Cellular FN
Asn430	[HexNAc]4[Hex]5[NeuAc]2	[HexNAc]4[Hex]5[NeuAc]1[Fuc]1
	[HexNAc]4[Hex]5[NeuAc]1[a]	[HexNAc]4[Hex]5[Fuc]1[a]
	[HexNAc]4[Hex]5	[HexNAc]4[Hex]4[Fuc]1
		[HexNAc]3[Hex]4[Fuc]1
Asn528	[HexNAc]4[Hex]5[NeuAc]2	[HexNAc]4[Hex]5[NeuAc]1[Fuc]1
	[HexNAc]4[Hex]5[NeuAc]1[a]	[HexNAc]4[Hex]5[Fuc]1[a]
	[HexNAc]4[Hex]5	[HexNAc]4[Hex]4[Fuc]1
		[HexNAc]3[Hex]4[Fuc]1
Asn542	[HexNAc]4[Hex]5[NeuAc]2	[HexNAc]4[Hex]5[NeuAc]1[Fuc]1
	[HexNAc]4[Hex]5[NeuAc]1[a]	[HexNAc]4[Hex]5[Fuc]1[a]
	[HexNAc]4[Hex]5	[HexNAc]4[Hex]4[Fuc]1
		[HexNAc]3[Hex]4[Fuc]1
		[HexNAc]4[Hex]5[NeuAc]1
		[HexNAc]4[Hex]5
		[HexNAc]4[Hex]4
		[HexNAc]3[Hex]4
Asn877	[HexNAc]4[Hex]5[NeuAc]2[a]	[HexNAc]4[Hex]5[NeuAc]1[Fuc]1
	[HexNAc]4[Hex]5[NeuAc]1	[HexNAc]4[Hex]5[Fuc]1
	[HexNAc]4[Hex]5	[HexNAc]4[Hex]5[NeuAc]1
	[HexNAc]4[Hex]4	[HexNAc]4[Hex]5[a]
Asn1007	[HexNAc]4[Hex]5[NeuAc]2[a]	[HexNAc]4[Hex]5[NeuAc]1[Fuc]1
	[HexNAc]4[Hex]5[NeuAc]1	[HexNAc]4[Hex]5[Fuc]1[a]
	[HexNAc]4[Hex]5	[HexNAc]4[Hex]4[Fuc]1
	[HexNAc]4[Hex]5[NeuAc]2[Fuc]1	[HexNAc]3[Hex]4[Fuc]1
	[HexNAc]4[Hex]5[NeuAc]1[Fuc]1	
	[HexNAc]4[Hex]5[Fuc]1	
	[HexNAc]5[Hex]6[NeuAc]3	
	[HexNAc]5[Hex]6[NeuAc]2	
	[HexNAc]5[Hex]6[NeuAc]1	
	[HexNAc]5[Hex]6[NeuAc]3[Fuc]1	
	[HexNAc]5[Hex]6[NeuAc]2[Fuc]1	
	[HexNAc]5[Hex]6[NeuAc]1[Fuc]1	
Asn1244	[HexNAc]4[Hex]5[NeuAc]2	[HexNAc]4[Hex]5[NeuAc]1
	[HexNAc]4[Hex]5[NeuAc]1[a]	[HexNAc]4[Hex]5[a]
	[HexNAc]4[Hex]5	[HexNAc]4[Hex]4[NeuAc]1
		[HexNAc]4[Hex]4
		[HexNAc]4[Hex]3
		[HexNAc]3[Hex]4
		[HexNAc]3[Hex]3
Asn2108	[HexNAc]4[Hex]5[NeuAc]2	[HexNAc]4[Hex]5[NeuAc]2[Fuc]1
	[HexNAc]4[Hex]5[NeuAc]1[a]	[HexNAc]4[Hex]5[NeuAc]1[Fuc]1
	[HexNAc]4[Hex]5	[HexNAc]4[Hex]5[Fuc]1[a]
	[HexNAc]4[Hex]5[NeuAc]2[Fuc]1	
	[HexNAc]4[Hex]5[NeuAc]1[Fuc]1	
	[HexNAc]4[Hex]5[Fuc]1	

[a]The most abundant species.

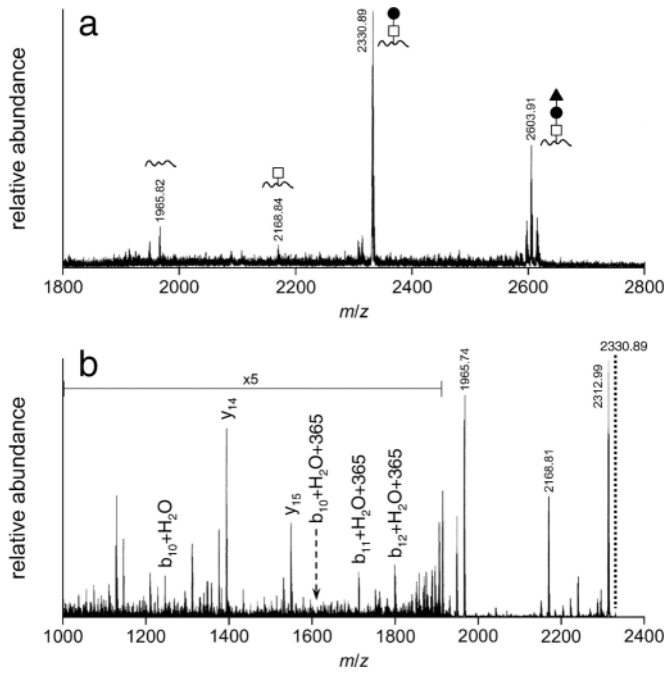

Fig. 5. Matrix-assisted laser desorption ionization quadrupole ion trap time-of-flight (MALDI QIT TOF) mass spectra of the glycopeptide containing a new O-glycosylation site, Thr279, from plasma fibronectin (FN). (**a**) MS^1 spectrum. A sialylated O-glycan is dehydrated in the MALDI QIT TOF mass spectrum, as observed at *m/z* 2603.91. (**b**) MS/MS spectrum of the ion at *m/z* 2330.89. The product ion mass spectrum showed an ion at *m/z* 1965.74 representing the peptide moiety, and the ions bearing a HexNAc-Hex disaccharide derived from dissociation of the peptide backbone. The product ion b_{11} contains the oligosaccharide residue but b_{10} does not, indicating the glycosylation site to be the 11th residue from the NH_2-terminal of this peptide.

2065 could be found in the enriched glycopeptides from a tryptic digest (Figure 2a). An additional O-glycosylation site was reported in the COOH-terminal segment of synovial fluid FN (Carsons *et al*., 1987; Carsons, 2002), but was not found in either plasma or the cellular isoform in this study.

In an earlier study, the reactivity of a monoclonal antibody recognizing the Galβ1→3GalNAc residue was abolished by digestion with thermolysin, suggesting that *O*-glycans of this type are located within the proteolytically sensitive regions between domains (Nichols *et al*., 1986). Indeed, a new O-glycosylation site was located in the connecting segment between the NH_2-terminal (fibrin- and heparin-binding) and collagen-binding domains in this study, and another O-glycosylation site was within IIICS between the heparin-binding and COOH-terminal (fibrin-binding) domains. Light-scattering and nuclear magnetic resonance (NMR) studies have demonstrated that mucin-type glycosylation, even with the attachment of only a single innermost GalNAc residue, causes a global conformational change in the peptide core, significantly extends chain dimensions and increases chain stiffness (Gerken *et al*., 1989; Shogren *et al*., 1989; Coltart *et al*., 2002). Taking these findings into account, O-glycosylation might play a role in segregating the tightly folded domains of FN and thus maintaining domain topology. The amino acid sequences encompassing these O-glycosylation sites are presented in Table II. The region is not quite homologous among species, considering the very high conservation of this molecule. More interestingly, Thr279 is an exception to the common characteristic of O-glycosylation occurring in proximity to proline residues.

New N-glycosylation site on apoB100

ApoB100 is a large glycoprotein comprised of 4536 amino acid residues. In a recent LC/ESI MS/MS study, the *N*-glycan structures at 17 of 19 potential sites were characterized, and the remaining sites, Asn34 (Asn7 for the translated protein) and Asn2560 (Asn2533), were suggested to be glycosylated at undetectable levels if at all (Harazono *et al*., 2005).

In this study, using the method of glycopeptide enrichment followed by MALDI MS, the oligosaccharide structures at 18 of 19 sites including Asn2560 were elucidated. Asn2560 was occupied by a sialylated biantennary type of oligosaccharide as shown in Figure 6. This site might have escaped detection in the previous study, presumably because so many different peptides, most of which were not glycopeptides, having the same retention time prevented the LC/ESI MS technology from yielding significant intensities of glycopeptide-derived ions. The glycan structures at other sites were largely consistent with those of a previous report except for a small difference, for example, hybrid-type oligosaccharides were present at Asn3465, as well as Asn3411, near the low-density lipoprotein (LDL)-receptor binding site.

Asn34, which is the seventh residue from the NH_2-terminus of mature apoB100 protein, was most likely unmodified, because close observation of the tryptic digest identified only an unglycosylated peptide, and the corresponding glycopeptide was not found in the glycopeptide-enriched fraction. Several N-glycosylation sites are partially modified in apoB100 (Harazono *et al*., 2005), and this is the case with Asn2560. The minimal distance between an Asn-X-Ser/Thr acceptor site and the end of a transmembrane segment of an integral membrane protein required for half-maximal glycosylation is ~10 residues (Nilsson and von Heijne, 1993). Therefore,

Table II. Sequences including the *O*-glycan attachment sites

Homo sapiens	(272) RH▼TSVQTTSSGSGPF (286)	(2058) RRTTPPTTATPIRHR (2072)
Bos taurus	(241) RH TSLQTTSAGSGSF (255)	(1937) RRTTPPTTATPVRHR (1951)
Mus musculus	(273) RH ALQSASAGSGSF (286)	(2148) RRTTPPTAATPVRLR (2162)
Rattus norvegicus	(273) RH VLQSASAGSGSF (286)	(2148) RRTTPPTAATPVRLR (2162)

The sequence is derived from the SWISS-PROT database. Arrowhead indicates the endpoint of the fibrin- and heparin-binding domain. O-Glycosylation sites are underlined.

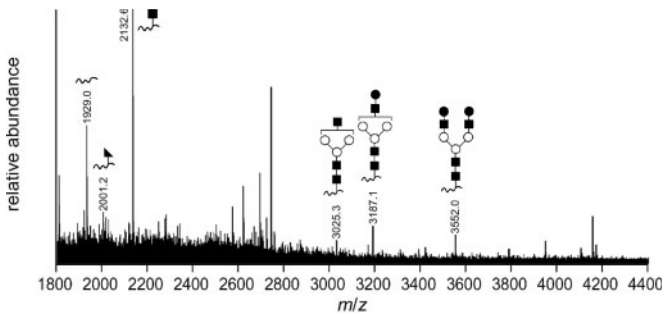

Fig. 6. Matrix-assisted laser desorption ionization quadrupole ion trap time-of-flight (MALDI QIT TOF) mass spectrum of the glycopeptide containing Asn2560 from human apoB100. The ions derived from the glycopeptide containing Asn2560 are indicated. Each of the labeled ions in the high mass region (*m/z* 3025.3 to *m/z* 3552.0) generated the same series of product ions as observed in the low mass region (*m/z* 1929.0 to *m/z* 2132.6), indicating that all of these ions were derived from the Asn2560-containing peptide with 1928.0 Da. ■, *N*-acetylglucosamine; ○, mannose; ●, galactose; □, fucose; ▲, *N*-acetylneuraminic acid.

depressed levels of glycosylation on apoB100 may underlie the null modification of Asn34 near the signal peptide, although a mechanism allowing the addition of core oligosaccharides to the acceptor site after cleavage of a signal peptide was suggested in *Saccharomyces cerevisiae* (Chen *et al.*, 2001).

In conclusion, glycopeptide enrichment was improved by divalent cations and/or an increased butanol concentration in the binding solution, and this method was successfully applied to site-specific characterization of the *N*-glycans of FN and apoB100. The α1→6 fucosylation of the innermost GlcNAc revealed a remarkable site-specific difference in N-glycosylation between plasma and cellular FNs. Glycopeptide enrichment was effective for glycopeptides with mucin-type oligosaccharides as well, and a new O-glycosylation site in the interdomain segment of FN suggested a role of this type of modification in maintaining the conformation of this molecule. Although LC/MS technology is an efficient method of detecting glycopeptides in complex peptide mixtures and of characterizing the site-specific oligosaccharide structures, a strategy utilizing glycopeptide enrichment followed by MALDI MS is also effective for the same purposes and even allows repetitive analyzes. Alternatively, glycopeptide enrichment and LC/ESI-MS/MS may be a powerful combination facilitating efforts to meet the rapidly growing demand for site-specific analysis in glycoproteomics.

Materials and methods

Materials

Sepharose CL4B was purchased from Amersham Bioscience (Piscataway, NJ). Human transferrin, iodoacetamide, lysylendopeptidase (*Achromobacter* protease I), and recrystallized 2,5-dihydroxybenzoic acid (DHB) were purchased from Wako Pure Chemical (Osaka, Japan). AspN, trypsin (Sequence Grade Modified Trypsin, from porcine pancreas), and α-chymotrypsin (TLCK-treated, from bovine pancreas) were from Roche Diagnostics (Mannheim, Germany), Promega (Madison, WI), and Sigma (St. Louis, MO), respectively. FN from human foreskin fibroblasts (cellular FN) and human apoB100 were purchased from Sigma and Calbiochem (La Jolla, CA), respectively. Plasma FN was purified from adult volunteer plasma by gelatin-affinity chromatography (Engvall and Ruoslahti, 1977).

Hydrophilic affinity separation of glycopeptides

Transferrin, FNs, and apoB100 were reduced with dithiothreitol followed by carbamidomethylation, and the alkylated samples were digested with a mixture of trypsin and lysylendopeptidase, or with chymotrypsin. The digestion was carried out in 0.1 M Tris HCl, pH 8.0.

Enrichment of glycopeptides from digests was carried out according to a method described previously (Wada *et al.*, 2004). Briefly, a typical 100 μg digest was mixed with a 15 μL packed volume of Sepharose CL4B in 1 mL of an organic solvent of 1-butanol/ethanol/H_2O (5:1:1, v/v) containing 1 mM each of $MnCl_2$, $CaCl_2$, $CoCl_2$, $NiSO_4$, $CuCl_2$, or $ZnCl_2$. After gentle shaking for 45 min, the gel was washed twice with the same organic solvent. The gel was then incubated with an aqueous solvent of ethanol/H_2O (1:1, v/v) for 30 min, and the solution phase was recovered and dried using a vacuum concentrator. After drying the eluate, the glycopeptides were subjected to MS or separated by HPLC. HPLC was carried out on a C18 column (150 × 1.0 mm) with a linear gradient elution of acetonitrile (1–30%, v/v) in 0.1% (v/v) trifluoroacetic acid (TFA).

The glycopeptides isolated from HPLC were digested with AspN in the buffer described above, and the dried digest was directly analyzed by MS.

MS

Two MALDI TOF mass spectrometers were employed for the structural analysis of glycopeptides. MALDI linear TOF measurements were carried out on a Voyager DE Pro MALDI TOF mass spectrometer with a nitrogen pulsed laser (337 nm) (Applied Biosystems, Foster City, CA). Mass spectrometric peptide fragmentation and sequencing were performed on a MALDI multiple-stage tandem mass spectrometer AXIMA-QIT (Shimadzu, Kyoto, Japan) with a nitrogen pulsed laser (337 nm). Argon was used as the collision gas for CID. Typically in these measurements, 0.1–1 pmol of a glycopeptide mixture was dissolved in 1 μL of 10 mg/mL of DHB dissolved in a 0.1% (v/v) TFA and 30% (v/v) acetonitrile solution on a MALDI sample target and dried. Measurements were carried out in positive ion mode.

Protein database

The positions of amino acid residues and the domain names were according to TRFE_HUMAN (transferrin), FINC_HUMAN (FN), and APOB_HUMAN (apoB100) in the SWISS-PROT database.

Acknowledgments

The authors thank Shimadzu Corporation for instrument support. This work was supported by the 21st Century Center of Excellence (COE) Program, Osaka University.

Abbreviations

apoB100, apolipoprotein B-100; CID, collision-induced dissociation; ESI, electrospray ionization; FN, fibronectin; GlcNAc, *N*-acetylglucosamine; HPLC, high-performance liquid chromatography; MALDI MS, matrix-assisted laser desorption ionization mass spectrometry; QIT, quadrupole ion trap; TOF MS, time-of-flight mass spectrometry.

References

Banks, R.E., Dunn, M.J., Hochstrasser, D.F., Sanchez, J.C., Blackstock, W., Pappin, D.J., and Selby, P.J. (2000) Proteomics: new perspectives, new biomedical opportunities. *Lancet*, **356**, 1749–1756.

Betts, T.J. (1964) One-phase solvent systems for paper chromatography. *J. Pharm. Sci.*, **53**, 794–797.

Bunkenborg, J., Pilch, B.J., Podtelejnikov, A.V., and Wisniewski, J.R. (2004) Screening for N-glycosylated proteins by liquid chromatography mass spectrometry. *Proteomics*, **4**, 454–465.

Burlingame, A. (1996) Characterization of protein glycosylation by mass spectrometry. *Curr. Opin. Biotechnol.*, **7**, 4–10.

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.

Carsons, S. (2002) Enhanced expression of a peanut agglutinin reactive O linked oligosaccharide on fibronectins from the synovial fluid of patients with rheumatic disease: quantitation, domain localization, and functional significance. *J. Rheumatol.*, **29**, 896–902.

Carsons, S., Lavietes, B.B., Slomiany, A., Diamond, H.S., and Berkowitz, E. (1987) Carbohydrate heterogeneity of fibronectins. Synovial fluid fibronectin resembles the form secreted by cultured synoviocytes but differs from the plasma form. *J. Clin. Invest.*, **80**, 1342–1349.

Chen, X., VanValkenburgh, C., Liang, H., Fang, H., and Green, N. (2001) Signal peptidase and oligosaccharyltransferase interact in a sequential and dependent manner within the endoplasmic reticulum. *J. Biol. Chem.*, **276**, 2411–2416.

Coltart, D.M., Royyuru, A.K., Williams, L.J., Glunz, P.W., Sames, D., Kuduk, S.D., Schwarz, J.B., Chen, X.T., Danishefsky, S.J., and Live, D.H. (2002) Principles of mucin architecture: structural studies on synthetic glycopeptides bearing clustered mono-, di-, tri-, and hexasaccharide glycodomains. *J. Am. Chem. Soc.*, **124**, 9833–9844.

Das, T.M., Rao, C.P., and Kolehmainen, E. (2001) Interaction of metal ions with N-glycosylamines: isolation and characterization of the products of 4,6-O-benzylidene-N-(o-carboxyphenyl)-beta-D-glucopyranosylamine with different metal ions. *Carbohydr. Res.*, **335**, 151–158.

Davies, M.J., Smith, K.D., Carruthers, R.A., Chai, W., Lawson, A.M., and Hounsell, E.F. (1993) Use of a porous graphitised carbon column for the high-performance liquid chromatography of oligosaccharides, alditols and glycopeptides with subsequent mass spectrometry analysis. *J. Chromatogr.*, **646**, 317–326.

Davis, B.D. and Brodbelt, J.S. (2005) LC-MSn methods for saccharide characterization of monoglycosyl flavonoids using postcolumn manganese complexation. *Anal. Chem.*, **77**, 1883–1890.

Demelbauer, U.M., Zehl, M., Plematl, A., Allmaier, G., and Rizzi, A. (2004) Determination of glycopeptide structures by multistage mass spectrometry with low-energy collision-induced dissociation: comparison of electrospray ionization quadrupole ion trap and matrix-assisted laser desorption/ionization quadrupole ion trap reflectron time-of-flight approaches. *Rapid Commun. Mass Spectrom.*, **18**, 1575–1582.

Engvall, E. and Ruoslahti, E. (1977) Binding of soluble form of fibroblast surface protein, fibronectin, to collagen. *Int. J. Cancer*, **20**, 1–5.

Fu, D. and van Halbeek, H. (1992) N-glycosylation site mapping of human serotransferrin by serial lectin affinity chromatography, fast atom bombardment-mass spectrometry, and 1H nuclear magnetic resonance spectroscopy. *Anal. Biochem.*, **206**, 53–63.

Fukuda, M. and Hakomori, S. (1979) Carbohydrate structure of galactoprotein a, a major transformation-sensitive glycoprotein released from hamster embryo fibroblasts. *J. Biol. Chem.*, **254**, 5451–5457.

Fukuda, M., Levery, S.B., and Hakomori, S. (1982) Carbohydrate structure of hamster plasma fibronectin. Evidence for chemical diversity between cellular and plasma fibronectins. *J. Biol. Chem.*, **257**, 6856–6860.

Garcia, R., Rodriguez, R., Montesino, R., Besada, V., Gonzalez, J., and Cremata, J.A. (1995) Concanavalin A- and wheat germ agglutinin-conjugated lectins as a tool for the identification of multiple N-glycosylation sites in heterologous protein expressed in yeast. *Anal. Biochem.*, **231**, 342–348.

Gerken, T.A., Butenhof, K.J., and Shogren, R. (1989) Effects of glycosylation on the conformation and dynamics of O-linked glycoproteins: carbon-13 NMR studies of ovine submaxillary mucin. *Biochemistry*, **28**, 5536–5543.

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

Harvey, D.J. (2001) Identification of protein-bound carbohydrates by mass spectrometry. *Proteomics*, **1**, 311–328.

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

Hynes, R. (1985) Molecular biology of fibronectin. *Annu. Rev. Cell Biol.*, **1**, 67–90.

Ichihara-Tanaka, K., Titani, K., and Sekiguchi, K. (1990) Recombinant carboxyl-terminal fibrin-binding domain of human fibronectin expressed in mouse L cells. *J. Biol. Chem.*, **265**, 401–407.

Isemura, M., Yamaguchi, Y., Munakata, H., Aikawa, J., Kan, M., Yamane, I., and Yosizawa, Z. (1984) Isolation and characterization of human placenta fibronectin. *J. Biochem. (Tokyo)*, **96**, 163–169.

Koizumi, K. (1996) High-performance liquid chromatographic separation of carbohydrates on graphitized carbon columns. *J. Chromatogr. A*, **720**, 119–126.

Koy, C., Mikkat, S., Raptakis, E., Sutton, C., Resch, M., Tanaka, K., and Glocker, M.O. (2003) Matrix-assisted laser desorption/ionization- quadrupole ion trap-time of flight mass spectrometry sequencing resolves structures of unidentified peptides obtained by in-gel tryptic digestion of haptoglobin derivatives from human plasma proteomes. *Proteomics*, **3**, 851–858.

Krusius, T., Fukuda, M., Dell, A., and Ruoslahti, E. (1985) Structure of the carbohydrate units of human amniotic fluid fibronectin. *J. Biol. Chem.*, **260**, 4110–4116.

Kuster, B., Krogh, T.N., Mortz, E., and Harvey, D.J. (2001) Glycosylation analysis of gel-separated proteins. *Proteomics*, **1**, 350–361.

Larsen, M.R., Hojrup, P., and Roepstorff, P. (2005) Characterization of gel-separated glycoproteins using two-step proteolytic digestion combined with sequential microcolumns and mass spectrometry. *Mol. Cell Proteomics*, **4**, 107–119.

Mortz, E., Sareneva, T., Julkunen, I., and Roepstorff, P. (1996) Does matrix-assisted laser desorption/ionization mass spectrometry allow analysis of carbohydrate heterogeneity in glycoproteins? A study of natural human interferon-gamma. *J. Mass Spectrom.*, **31**, 1109–1118.

Mosher, D.F. (1984) Physiology of fibronectin. *Annu. Rev. Med.*, **35**, 561–575.

Nichols, E.J., Fenderson, B.A., Carter, W.G., and Hakomori, S. (1986) Domain-specific distribution of carbohydrates in human fibronectins and the transformation-dependent translocation of branched type 2 chain defined by monoclonal antibody C6. *J. Biol. Chem.*, **261**, 11295–11301.

Nilsson, I.M. and von Heijne, G. (1993) Determination of the distance between the oligosaccharyltransferase active site and the endoplasmic reticulum membrane. *J. Biol. Chem.*, **268**, 5798–5801.

Papac, D.I., Wong, A., and Jones, A.J. (1996) Analysis of acidic oligosaccharides and glycopeptides by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *Anal. Chem.*, **68**, 3215–3223.

Rudd, P.M. and Dwek, R.A. (1997) Glycosylation: heterogeneity and the 3D structure of proteins. *Crit. Rev. Biochem. Mol. Biol.*, **32**, 1–100.

Ruoslahti, E. (1988) Fibronectin and its receptors. *Annu. Rev. Biochem.*, **57**, 375–413.

Shimizu, Y., Nakata, M., Kuroda, Y., Tsutsumi, F., Kojima, N., and Mizuochi, T. (2001) Rapid and simple preparation of N-linked

oligosaccharides by cellulose-column chromatography. *Carbohydr. Res.*, **332**, 381–388.

Shogren, R., Gerken, T.A., and Jentoft, N. (1989) Role of glycosylation on the conformation and chain dimensions of O-linked glycoproteins: light-scattering studies of ovine submaxillary mucin. *Biochemistry*, **28**, 5525–5536.

Spiro, R.G. (2002) Protein glycosylation: nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds. *Glycobiology*, **12**, 43R–56R.

Takasaki, S., Yamashita, K., Suzuki, K., Iwanaga, S., and Kobata, A. (1979) The sugar chains of cold-insoluble globulin. A protein related to fibronectin. *J. Biol. Chem.*, **254**, 8548–8553.

Tressel, T., McCarthy, J.B., Calaycay, J., Lee, T.D., Legesse, K., Shively, J.E., and Pande, H. (1991) Human plasma fibronectin. Demonstration of structural differences between the A- and B-chains in the III CS region. *Biochem. J.*, **274**, 731–738.

Varki, A. (1993) Biological roles of oligosaccharides: all of the theories are correct. *Glycobiology*, **3**, 97–130.

Wada, Y., Tajiri, M., and Yoshida, S. (2004) Hydrophilic affinity isolation and MALDI multiple-stage tandem mass spectrometry of glycopeptides for glycoproteomics. *Anal. Chem.*, **76**, 6560–6565.

Wagner, D.D., Ivatt, R., Destree, A.T., and Hynes, R.O. (1981) Similarities and differences between the fibronectins of normal and transformed hamster cells. *J. Biol. Chem.*, **256**, 11708–11715.

Yamada, K.M. (1983) Cell surface interactions with extracellular materials. *Annu. Rev. Biochem.*, **52**, 761–799.

Zhu, B.C., Fisher, S.F., Pande, H., Calaycay, J., Shively, J.E., and Laine, R.A. (1984) Human placental (fetal) fibronectin: increased glycosylation and higher protease resistance than plasma fibronectin. Presence of polylactosamine glycopeptides and properties of a 44-kilodalton chymotryptic collagen-binding domain: difference from human plasma fibronectin. *J. Biol. Chem.*, **259**, 3962–3970.