Original Article

Label-Free Analysis of *O*-Glycosylation Site-Occupancy Based on the Signal Intensity of Glycopeptide/Peptide Ions

Yoshinao Wada

Department of Molecular Medicine, Osaka Medical Center and Research Institute for Maternal and Child Health, Osaka, Japan

Mucin-type *O*-glycosylation is a major posttranslational modification of proteins. The level of *O*-glycosylation at a site could be useful in terms of evaluating various disease conditions. To address the feasibility of measuring *O*-glycosylation levels based on the glycopeptide ion intensity in a mass spectrum, apolipoprotein CIII (apoC3), a protein that contains a single core-1 *O*-glycan Gal-GalNAc disaccharide was analyzed by matrix-assisted laser desorption ionization (MALDI) time-of-flight (TOF) mass spectrometry (MS). The intensity of protonated ions for an equimolar mixture of desialylated and deglycosylated apoC3s were the same in linear TOF measurements. No substantial in-source decay, including the cleavage of the protein-sugar linkage was observed. The glycopeptide derived from apoC3 and the unglycosylated counterpart, when analyzed by MALDI reflectron TOF MS indicated that post-source decay was minimal. These collective findings demonstrate the feasibility of label-free quantitation of *O*-glycan occupancy by MS when the glycans are small and neutral. This method provides a tool for use in glycoproteomics as a complement of our previous report (DOI: 10.1021/pr900913k) for calculating the saccharide composition of *O*-glycans.

Keywords: *O*-glycan, apolipoprotein CIII, matrix-assisted laser desorption/ionization, label-free quantitation, in-source decay

INTRODUCTION

Glycosylation is a common post-translational modification that occurs in more than half of all secretory and cellular proteins. Glycans in glycoproteins refer to any form of mono-, oligo-, or polysaccharide that is covalently linked to a polypeptide backbone, usually *via N-* or *O*-linkages. Glycans attached to a specific site of proteins usually have structural diversity or "microheterogeneity," and the structure of each glycan is called a "glycoform." The attachment of a specific glycoform often changes the physicochemical and biological properties of a protein and frequently results in the transformation of cellular phenotypes as observed in cancer.[1–3] This is the reason why glycosylation has attracted attention as a biomarker. There are two types of altered glycosylations, namely a change in glycoform profile or a change in the level of modification. These alterations are typically found in a class of pediatric diseases, Congenital Disorders of Glycosylation (CDGs), most of which are due to enzymatic defects in the *N*-glycan synthesis pathway. CDGs are classified into CDG-I and CDG-II according to the type of abnormality described above, *i.e.*, decreased glycosylation levels and a change in glycoform profiles, respectively.[4,5]

Glycan profiling involves the elucidation of the relative abundance of each glycoform attached to a protein or a specific site. Liquid chromatography (LC) of glycans, which are released from glycoproteins and then labeled with chromophores at the reducing end to enhance sensitivity, has long been the standard method of profiling. The Human Proteome Organisation Human Disease Glycomics/Proteome Initiative (HGPI) coordinated a multi-institutional studies for evaluating methodologies that are widely used for defining *N-* and *O*-glycan content in glycoproteins.[6,7] The most important message from these studies was that mass spectrometry (MS)-based strategies provide an effective means of both the identification and quantitation of glycans. It should also be noted that the matrix-assisted laser desorption/ionization (MALDI) MS of permethylated glycans was reliable, as evidenced by the fact that only minor variations were found for data from different laboratories. Another approach to glycan profiling is glycopeptide analysis, which allows the identification of site-specific glycan structures to delineate the role of glycans in protein folding and related functions. Site occupancy can also be assessed by the analysis of glycopeptides but not of the released glycans.

Incomplete *N*-glycosylation is found in certain types of disorders such as chronic alcohol abuse or hepatic dysfunction and CDG-I,[8–10] and the (un)occupancy level can be a useful marker for the detection or evaluation of these conditions. Indeed, the identification of unglycosylated proteins from patient's serum by MS is currently employed for the diagnosis of CDG-I in Japan.[11] Regarding the quantitation of occupancy, however, it is generally believed that the relative abundance of the ions derived from an glycosylated species to those of the unglycosylated counterpart poorly correlates with the real ratio of these molecules. This is typically the case that glycans bearing acidic residues such as neuraminic acid and sulfated sugars can affect ionization efficiency. In addition, when the size of *N*-glycans usually exceeds 1,000 Da, this causes glycosylated proteins or peptides to give smaller signals compared with unglycosylated species simply due to the increased molecular mass. This problem

Correspondence to: Yoshinao Wada, *Department of Molecular Medicine, Osaka Medical Center and Research Institute for Maternal and Child Health, 840 Murodo-cho, Izumi, Osaka 594–1101, Japan*, e-mail: waday@mch.pref.osaka.jp

is partly solved by *N*-glycosidase F treatment which transforms the glycan-attached asparagine to aspartic acid and enables the discrimination of the unglycosylated asparagine from the transformed aspartic acid by 1 Da in the mass spectrum. Furthermore, the difference can be increased by introducing ^{18}O during the hydrolysis.[12,13]

We previously reported on the label-free quantitation of saccharide compositions of mucin-type *O*-glycans and its application to IgA nephropathy and rheumatoid arthritis.[14] Mucin-type *O*-glycosylation is the major type of modification to serine (Ser) and threonine (Thr) residues of mammalian glycoproteins. This involves the covalent binding of an *N*-acetylgalactosamine (GalNAc) moiety *via* an *α-O* glycosidic linkage to the –OH of Ser or Thr. The core structure of a typical mucin-type *O*-glycan is comprised of up to three monosaccharides and is smaller than that of *N*-glycans, although GalNAc may be extended with sugars including galactose (Gal), *N*-acetylglucosamine, fucose, or *N*-acetyl neuraminic acid (NeuAc). It can be speculated that such a small glycan allows the determination of occupancy levels based on the signal intensity in the mass spectrum. To address this issue in more detail, the difference in ionization efficiency between glycosylated and unglycosylated peptides/proteins as well as the dissociation of the glycan moiety was investigated, using a small *O*-glycosylated protein, apolipoprotein C-III (apoC3).

MATERIALS AND METHODS

Isolation of the very low density lipoprotein (VLDL) fraction

The VLDL fraction was separated from serum of a healthy volunteer by sucrose density gradient ultracentrifugation, as described previously.[15] Briefly, a serum volume of 200 μL was incubated with 8 μL of Sudan Black B (Sigma-Aldrich, St. Louis, MO) dissolved in dimethyl sulfoxide (1% w/v) at 37°C for 30 min. The solution was then diluted with the same volume of distilled water and layered over 2.0 mL of 20% sucrose in an open-top polycarbonate tube. The tube was spun at 100,000 rpm (356,000 g) at 20°C for 10 h in an Optima TLX ultracentrifuge (Beckman-Coulter, Brea, CA) using a Beckman TLA-100.3 rotor. After separation, the VLDL layer was collected by aspiration.

Enzymatic treatment

Desialylation and deglycosylation was enzymatically performed as follows. The VLDL fraction (200 μL) was diluted with the same volume of 100 mM sodium acetate, pH 5.0, and the solution was incubated with 2 mU of neuraminidase from *Arthrobacter ureafaciens* (Nacalai, Kyoto, Japan) or a mixture of 2 mU neuraminidase and 5 mU *O*-glycosidase from *Diplococcus pneumoniae* (Roche, Indianapolis, IN) at 37°C for 3 h. After confirmation of the removal of NeuAc or *O*-glycan by MS, the reaction was stopped by adding 0.5% (final concentration) trifluoroacetic acid (TFA). The relative concentrations of apoC3 in these VLDL fractions were determined by high performance liquid chromatography (HPLC) using a reversed phase column (Develosil 300C8-HG-5, 1.0×150 mm, Nagoya, Japan) with an acetonitrile gradient in 0.1% TFA and detection by UV absorbance at 220 nm.

Following desialylation or deglycosylation, an equimolar mixture of deglycosylated and desialylated apoC3 was digested with lysylendopeptidase (Wako, Osaka, Japan) in 0.2 M Tris–HCl, pH 8.5 at 37°C for 6 h to yield a peptide (residues 61–79) containing the *O*-glycosylation site. An equal peptide/glycopeptide ratio was verified by HPLC, as described above.

Mass spectrometry

The protein and peptide samples were desalted by solid-phase extraction using ZipTipC18 (Millipore, Bedford, MA) and eluted with 10 μL of 0.1% TFA and 70% acetonitrile. Matrix-assisted laser desorption/ionization (MALDI) MS was carried out on a Voyager DE Pro MALDI-TOF mass spectrometer (AB Sciex, Framingham, MA) fitted with a nitrogen pulsed laser (337 nm). The detectors for ions were a hybrid detector consisting of a single microchannel plate, a fast scintillator and a photomultiplier for linear mode and a multichannel plate for the reflector mode. The desalted sample solution was mixed with the same volume of 2% (w/v) of 2,5-dihydroxybenzoic acid dissolved in 50% (v/v) acetonitrile on a stainless sample target. The measurements were carried out in the positive ion and linear or reflectron time-of-flight (TOF) mode. The mass spectra acquired by at least 200 laser shots were accumulated.

Nano-electrospray ionization MS was carried out using an LTQ XL ion trap mass spectrometer (Thermo Fisher Scientific, Waltham, MA) with a dual conversion dynode detector. Peptide samples were dissolved in a 0.1% formic acid and 20% (v/v) acetonitrile solution and directly infused into the mass spectrometer using a nano-electrospray tip.

RESULTS AND DISCUSSION

The mature form of apoC3 (exact mass=8759.2173; relative molecular mass=8764.6) is comprised of 79 amino acid residues and has a single core 1 *O*-glycan (Gal*β*1-3GalNAc) linked to Thr near the C-terminal end but no *N*-glycans. The *O*-glycan of apoC3 is terminally modified by up to two sialic acids (NeuAc), and significant levels of unsialylated species are not present in serum.[16] In the MALDI linear TOF mass spectrum, the major ions of apoC3 are the mono- and di-sialylated species observed at *m/z* 9422.2 and *m/z* 9713.5, respectively, and small peaks corresponded to the molecule without an attached glycan (*m/z* 8765.6) or with a Gal-GalNAc disaccharide (*m/z* 9130.9), as reported previously[15] (Fig. 1a). The mass spectrum was obtained in the positive ion mode, and the relative abundance of the sialylated apoC3 ions underestimates the actual content of these species, indicating the unglycosylated or the Gal-GalNAc disaccharide-bearing isoform to be the quite minor components in total apoC3 molecules.

Next, to address the feasibility of the label-free quantitation of site occupancy based on a comparison of the signals derived from unglycosylated and asilao apoC3s in the mass spectrum, unglycosylated (deglycosylated) and desialylated apoC3 samples were prepared (Scheme 1 for workflow). After their concentrations were determined by HPLC, these samples were mixed at varying ratios and MS spectra were obtained. The mass spectra of apoC3 after desialylation and after deglycosylation are presented in Figs. 1b and 1c, respectively. Interestingly, no peak for the unglycosylated species (*m/z* 8765.6) was found in Fig. 1b, demonstrating that
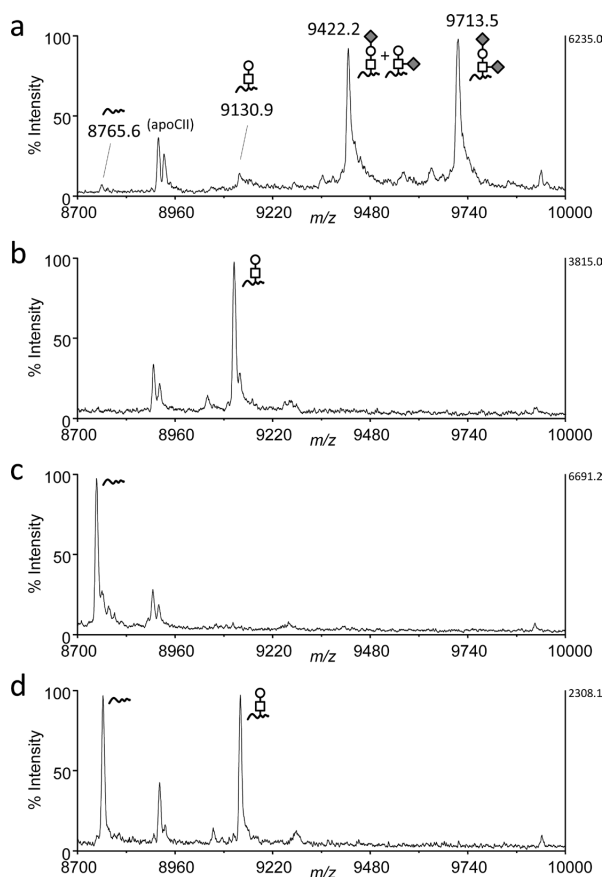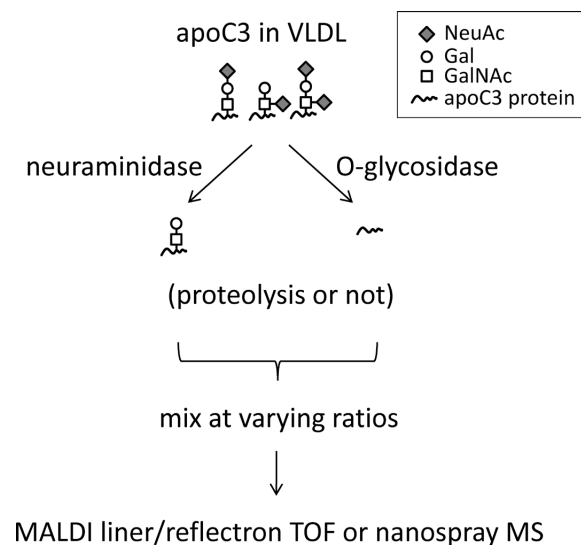
Fig. 1. MALDI linear TOF mass spectra of apoCIII in VLDL.

Proteins in the VLDL fraction from a healthy individual were analyzed by MALDI linear TOF MS. The mass region of *m/z* 8,700–10,000 covering all apoCIII glycoforms is presented. The average mass of the protonated molecular ions of each isoform is indicated in parentheses. The glycoform is given at the corresponding peak. square, GlaNAc; circle, Gal; diamond, NeuAc; short wavy line, unglycosylated apoC3. (a) before enzymatic treatment. (b) after desialylation. (c) after deglycosylation. (d) equimolar mixture of desialylated and deglycosylated samples.

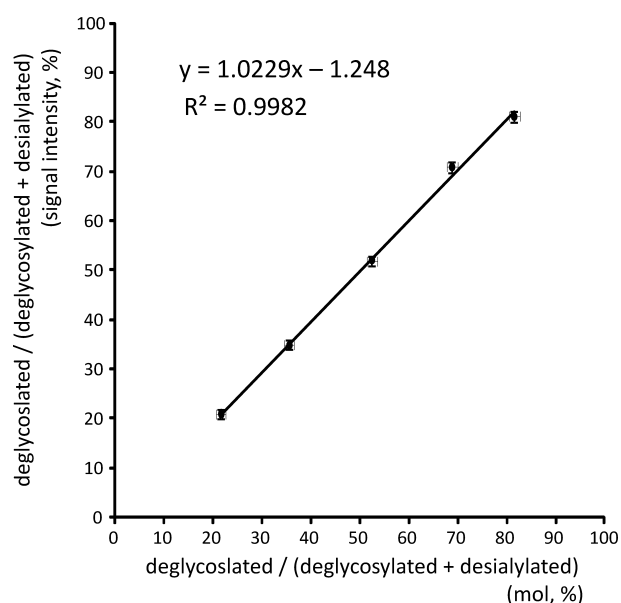Scheme 1. Workflow of experiments.

Fig. 2. Correlation of the relative signal intensity in a MALDI mass spectrum and real molar ratio.

The relative concentrations of apoC3 in deglycosylated and desialylated VLDL samples were measured by HPLC, and these samples were mixed at varying molar ratios and analyzed by MALDI linear TOF MS. The relative signal intensities were plotted against the molar ratios (mean±S.D., *n*=5).

in-source dissociation at the protein-sugar linkage during ionization is negligible and that the samples are sufficiently pure of use in subsequent experiments. The lack of in-source decay was also confirmed by the MALDI linear TOF MS of glycopeptides produced by the lysylendopeptidase digestion of desialylated apoC3 (data not shown).

The deglycosylated and desialylated apoC3s were mixed at the same molar ratio, and subjected to MALDI TOF MS. Gal-GalNAc modification gives only a 4% increase to the unmodified apoC3, and the peak height of the ions for these species was the same in the MALDI linear TOF mass spectrum (Fig. 1d). As shown in Fig. 2, analyses of mixtures of samples at various ratios demonstrated a good correlation between the relative signal intensity and the real ratio. These collective findings suggest that the site occupancy of *O*-glycosylation can be calculated from the mass spectrum when the glycans are neutral.

To test whether this label-free quantitation is valid for glycopeptides, the C-terminal peptide (sequence 61–79 of mature apoC3, exact mass=2136.02) was analyzed by MALDI MS in the linear and reflectron modes. As shown

in Fig. 3, the peak height of the protonated molecular ions of *O*-glycosylated peptide at *m/z* 2502.2 was 35±5% or 35± 1% (mean±S.D., *n*=3) smaller than that of the unglycosylated peptide in the linear or reflectron mode, respectively. This difference in signal intensity was probably due to a 17% increase in the mass corresponding to the glycan moiety of the glycopeptide ions. Considering a general trend for the decreased intensity for other peptide ions according to an increase in their mass, the ionization efficiency of *O*-glycopeptides is presumably comparable to that for unglycosylated species. In addition, no signals suggesting post-source decay were found in the reflectron TOF mass spectrum (Fig.
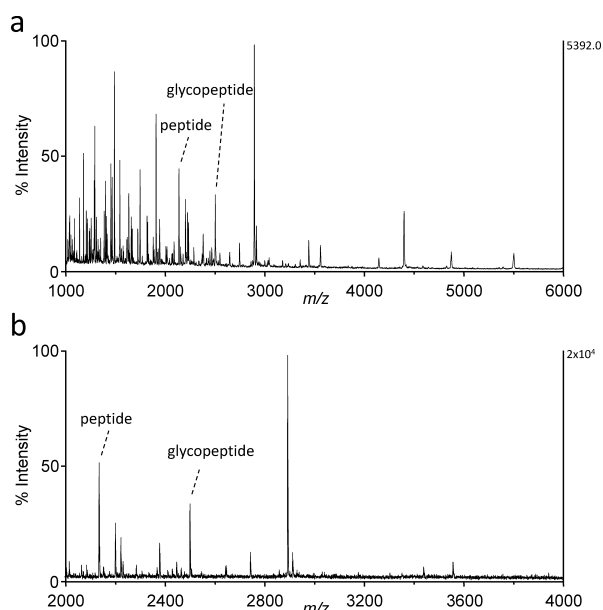
Fig. 3. MALDI TOF mass spectra of a lysylendopeptidase digest of VLDL.

The ions of peptide (sequence 61–79 of apoC3, 2136.0 Da) and the corresponding glycopeptide containing Gal–GalNAc disaccharide (2501.2 Da) are indicated. Other signals are derived from apoC3 or other VLDL proteins. (a) linear TOF mass spectrum, (b) reflectron TOF mass spectrum.
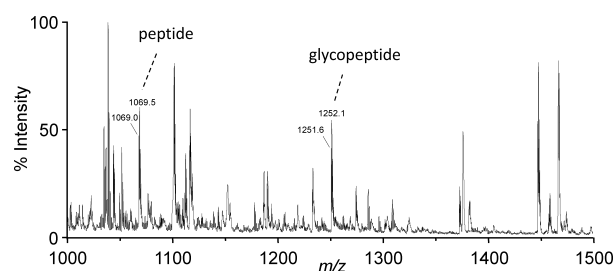

Fig. 4. Nano-electrospray ionization mass spectrum of a lysylendopeptidase digest of VLDL.

The sample shown in Fig. 3 was analyzed by nano-electrospray ionization MS with an ion-trap mass spectrometer. The doubly charged ions for peptide and glycopeptide derived from apoC3 are indicated.

3b). Finally, the same peptide sample was analyzed by nano-electrospray ionization MS. As shown in Fig. 4, the signal intensities for the glycosylated and unglycosylated peptides were comparable with each other.

A decreased *O*-glycan occupancy of apoC3 was reported in a patient with autosomal recessive cutis laxa syndrome type-2 (ARCL2) [MIM 219200].[15] ARCL2 is caused by loss-of-function mutations in the a2 subunit of the vesicular ATPase H^+-pump,[17] which contributes to the retrograde vesicular transport from Golgi to the endoplastic reticulum (ER).[18] It can be speculated that the absence of *O*-glycosylation in ARCL2 is due to a decreased concentration of GalNAc transferase in the ER, because *O*-glycosylation is regulated by the Gogi-to-ER relocation of the initiating GalNAc transferase enzyme.[19] The redistribution of GalNAc transferase *via* the retrograde vesicular transport system is regulated by various growth factors, and thus would be expected to occur in other diseases and various cellular conditions.

In conclusion, the label-free determination of occupancy levels described herein constitutes a complement to our previous report on the MS-based calculation of saccharide composition.[14] Furthermore, previously unaddressed concerns regarding the MALDI MS of *O*-glycans, *i.e.* ionization efficiency or dissociation during ionization, have been appropriately assessed. The set of label-free analyses of saccharide composition and site occupancy would be a useful glycoproteomic tool serving biomarker discovery as well as for the evaluation of various diseases.

Acknowledgements

REFERENCES

1) J. N. Arnold, R. Saldova, U. M. Hamid, P. M. Rudd. Evaluation of the serum *N*-linked glycome for the diagnosis of cancer and chronic inflammation. *Proteomics* 8: 3284–3293, 2008.

2) D. H. Dube, C. R. Bertozzi. Glycans in cancer and inflammation—potential for therapeutics and diagnostics. *Nat. Rev. Drug Discov.* 4: 477–488, 2005.

3) K. Ohtsubo, J. D. Marth. Glycosylation in cellular mechanisms of health and disease. *Cell* 126: 855–867, 2006.

4) J. Jaeken, G. Matthijs. Congenital disorders of glycosylation: a rapidly expanding disease family. *Annu. Rev. Genomics Hum. Genet.* 8: 261–278, 2007.

5) H. H. Freeze. Congenital Disorders of Glycosylation: CDG-I, CDG-II, and beyond. *Curr. Mol. Med.* 7: 389–396, 2007.

6) Y. Wada, P. Azadi, C. E. Costello, A. Dell, R. A. Dwek, H. Geyer, R. Geyer, K. Kakehi, N. G. Karlsson, N. 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. Comparison of the methods for profiling glycoprotein glycans—HUPO Human Disease Glycomics/Proteome Initiative multi-institutional study. *Glycobiology* 17: 411–422, 2007.

7) Y. Wada, A. Dell, S. M. Haslam, B. Tissot, K. Canis, P. Azadi, M. Bäckström, C. E. Costello, G. C. Hansson, Y. Hiki, M. Ishihara, H. Ito, K. Kakehi, N. Karlsson, C. E. Hayes, N. Kato, N. Kawasaki, K. H. Khoo, K. Kobayashi, D. Kolarich, A. Kondo, C. Lebrilla, M. Nakano, H. Narimatsu, J. Novak, M. V. Novotny, E. Ohno, N. H. Packer, E. Palaima, M. B. Renfrow, M. Tajiri, K. A. Thomsson, H. Yagi, S. Y. Yu, N. Taniguchi. Comparison of methods for profiling *O*-glycosylation: Human Proteome Organisation Human Disease Glycomics/Proteome Initiative multi-institutional study of IgA1. *Mol. Cell. Proteomics* 9: 719–727, 2010.

8) P. Gravel, C. Walzer, C. Aubry, L. P. Balant, B. Yersin, D. F. Hochstrasser, J. Guimon. New alterations of serum glycoproteins in alcoholic and cirrhotic patients revealed by high resolution two-dimensional gel electrophoresis. *Biochem. Biophys. Res. Commun.* 220: 78–85, 1996.

9) H. Stibler. Carbohydrate-deficient transferrin in serum: a new marker of potentially harmful alcohol consumption reviewed. *Clin. Chem.* 37: 2029–2037, 1991.

10) Y. Wada, A. Nishikawa, N. Okamoto, K. Inui, H. Tsukamoto, S. Okada, N. Taniguchi. Structure of serum transferrin in carbohydrate-deficient glycoprotein syndrome. *Biochem. Biophys. Res.*

Commun. 189: 832–836, 1992.

11) Y. Wada. Mass spectrometry for congenital disorders of glycosylation, CDG. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* 838: 3–8, 2006.

12) J. Gonzalez, T. Takao, H. Hori, V. Besada, R. Rodriguez, G. Padron, Y. Shimonishi. A method for determination of *N*-glycosylation sites in glycoproteins by collision-induced dissociation analysis in fast atom bombardment mass spectrometry: identification of the positions of carbohydrate-linked asparagine in recombinant alpha-amylase by treatment with peptide-*N*-glycosidase F in ¹⁸O-labeled water. *Anal. Biochem.* 205: 151–158, 1992.

13) Z. Liu, J. Cao, Y. He, L. Qiao, C. Xu, H. Lu, P. Yang. Tandem ¹⁸O stable isotope labeling for quantification of *N*-glycoproteome. *J. Proteome Res.* 9: 227–236, 2010.

14) Y. Wada, M. Tajiri, S. Ohshima. Quantitation of saccharide compositions of *O*-glycans by mass spectrometry of glycopeptides and its application to rheumatoid arthritis. *J. Proteome Res.* 9: 1367–1373, 2010.

15) Y. Wada, M. Kadoya, N. Okamoto. Mass spectrometry of apolipoprotein C-III, a simple analytical method for mucin-type *O*-glycosylation and its application to an autosomal recessive cutis laxa type-2 (ARCL2) patient. *Glycobiology* 22: 1140–1144, 2012.

16) S. Wopereis, S. Grünewald, E. Morava, J. M. Penzien, P. Briones, M. T. García-Silva, P. N. Demacker, K. M. Huijben, R. A. Wevers. Apolipoprotein C-III isofocusing in the diagnosis of genetic defects in *O*-glycan biosynthesis. *Clin. Chem.* 49: 1839–1845, 2003.

17) U. Kornak, E. Reynders, A. Dimopoulou, J. van Reeuwijk, B. Fischer, A. Rajab, B. Budde, P. Nürnberg, F. Foulquier, D. Lefeber, Z. Urban, S. Gruenewald, W. Annaert, H. G. Brunner, H. van Bokhoven, R. Wevers, E. Morava, G. Matthijs, L. Van Maldergem, S. Mundlos; ARCL Debré-type Study Group. Impaired glycosylation and cutis laxa caused by mutations in the vesicular H⁺-ATPase subunit ATP6V0A2. *Nat. Genet.* 40: 32–34, 2008.

18) V. Hucthagowder, E. Morava, U. Kornak, D. J. Lefeber, B. Fischer, A. Dimopoulou, A. Aldinger, J. Choi, E. C. Davis, D. N. Abuelo, M. Adamowicz, J. Al-Aama, L. Basel-Vanagaite, B. Fernandez, M. T. Greally, G. Gillessen-Kaesbach, H. Kayserili, E. Lemyre, M. Tekin, S. Türkmen, B. Tuysuz, B. Yüksel-Konuk, S. Mundlos, L. Van Maldergem, R. A. Wevers, Z. Urban. Loss-of-function mutations in ATP6V0A2 impair vesicular trafficking, tropoelastin secretion and cell survival. *Hum. Mol. Genet.* 18: 2149–2165, 2009.

19) D. J. Gill, J. Chia, J. Senewiratne, F. Bard. Regulation of *O*-glycosylation through Golgi-to-ER relocation of initiation enzymes. *J. Cell Biol.* 189: 843–858, 2010.