Analytical Methods

High-Throughput Protein Glycomics: Combined Use of Chemoselective Glycoblotting and MALDI-TOF/TOF Mass Spectrometry**

Shin-Ichiro Nishimura, Kenichi Niikura, Masaki Kurogochi, Takahiko Matsushita, Masataka Fumoto, Hiroshi Hinou, Ryousuke Kamitani, Hiroaki Nakagawa, Kisaburo Deguchi, Nobuaki Miura, Kenji Monde, and Hirosato Kondo*

Glycosylation is one of the posttranslational modifications of proteins in eukaryotes. This step is thought to modulate a wide range of protein functions both on the cellular surfaces and within the cells.[1] This highly dynamic process in which rapid changes in the carbohydrate structures occur in response to cellular signals or cellular stages result in the key informational markers of some serious human diseases. For example, it is known that carbohydrates of IgG can be drastically altered to form unusual structures in patients with rheumatoid arthritis[2] and specific carbohydrates are used as tumor-associated markers in pancreatic and colon cancers.[3] Recently, it has also been reported that the oligosaccharides of the therapeutic human IgG play a critical role of enhancing antibody-dependent cellular cytotoxicity.[4] Glycosylation of nuclear and cytoplasmic proteins with a GlcNAc residue is ubiquitous in nearly all eukaryotes and is a crucial event in posttranslational modifications for regulating protein functions within the cells.[1b] Dynamic perturbations in *O*-GlcNAc regulation have been implicated in the etiology of type II diabetes, cancer, and neurodegenerative diseases.[5]

Although there have been substantial advances in our understanding of the effects of glycosylation on some biological systems, we still do not fully understand the specific functional roles of carbohydrates and the relationship between their structures and functions. The major difficulty in carbohydrate sequencing is a consequence of the fact that the purification of trace amounts of oligosaccharides requires extremely tedious multistep processes. This is because crude sample mixtures prepared by enzymatic digestion from cells, organs, serum, etc. usually contain large amounts of impurities such as peptides, lipids, and salts. These technical problems in the sequencing of carbohydrates make it impossible to achieve high-throughput protein glycomics. We report herein that the combined use of two novel techniques, chemoselective glycoblotting and MALDI-TOF/TOF mass spectrometry,[6] allows for both facile purification and precise analysis of common oligosaccharides and glycopeptides from native glycoproteins.

The concept of the "sugar family tree," constructed by Fischer,[7] motivated us to use the chemoselective "glycoblotting" technique, which has become a key feature in the efficient isolation of carbohydrates in this study. Fischer found that the reaction of glucose, mannose, and major oligosaccharides with phenylhydrazine proceeds smoothly to give the corresponding stable phenylhydrazone derivatives under mild basic and aqueous conditions. Once oligosaccharides are released from glycoconjugates, they can be regarded as a general class of the compound library which have an aldehyde or ketone group at each reducing terminal. As indicated in Scheme 1a, aldehydes preferentially react with reagents bearing hydrazine-like functional groups. Fischer-type reagents do not need any catalyst or reducing agent to accelerate the reaction with sugars and the reactions usually proceed under mild conditions. However, the reactions of aldehydes with primary amino groups require some activating reagent for the formation of stable products. As a result, carbohydrates preferentially react with Fischer-type reagents even in the presence of large amounts of peptides or amino acids with primary amino groups.

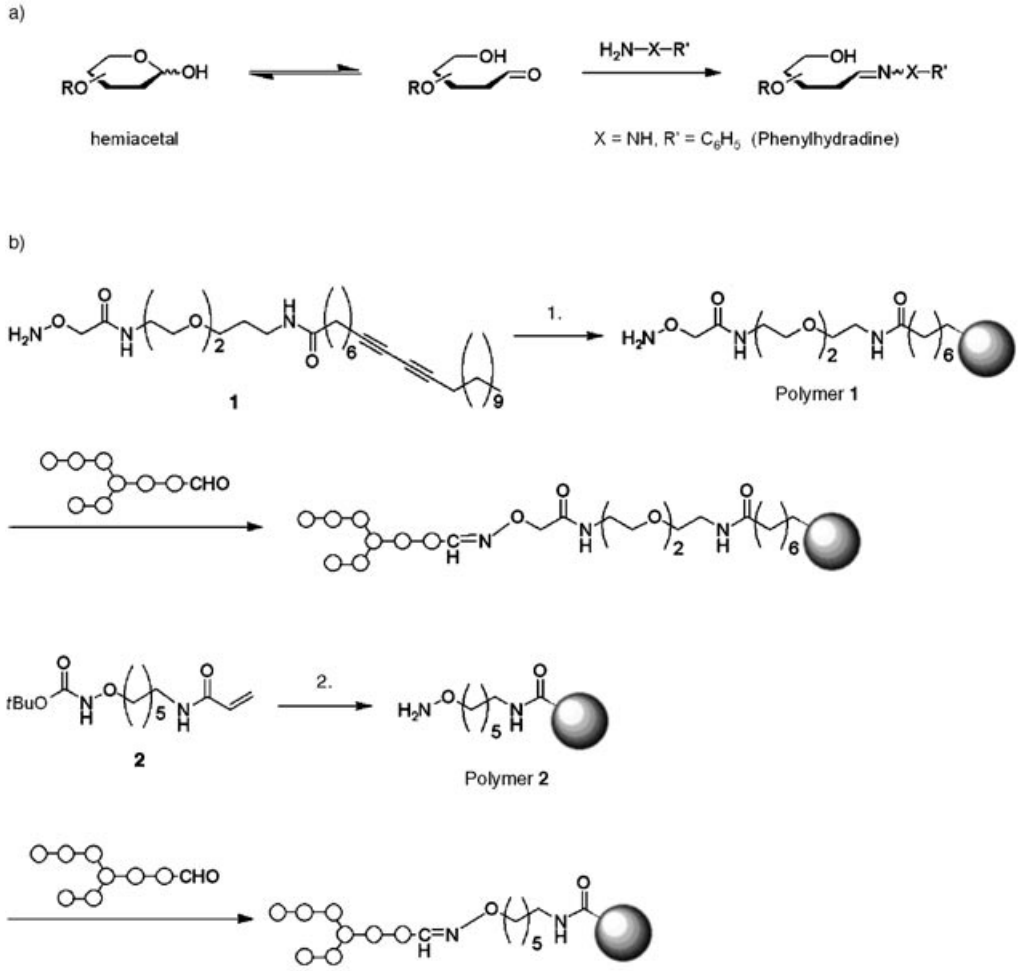
The unique chemical characteristics of the sugar family encouraged us to design novel polymers for capturing only carbohydrates from crude samples on the basis of the

[*] Prof. S.-I. Nishimura, Dr. K. Niikura, Dr. M. Kurogochi, T. Matsushita, R. Kamitani, Dr. H. Nakagawa, Dr. K. Deguchi, Dr. N. Miura, Dr. K. Monde
Division of Biological Sciences
Graduate School of Science
Frontier Research Center for Post-Genome Science and Technology
Hokkaido University
N21, W11, Sapporo 001-0021 (Japan)
Fax: (+81) 11-706-9042
E-mail: shin@glyco.sci.hokudai.ac.jp
Prof. S.-I. Nishimura, M. Fumoto, Dr. H. Hinou
Research Center for Glycoscience
National Institute of Advanced Industrial Science and Technology (AIST)
Sapporo 062-8517 (Japan)
Dr. H. Kondo
Discovery Research Laboratories
Shionogi & Co. Ltd.
Osaka 553-0002 (Japan)

[**] This work was partly supported by a grant for the National Project on "Functional Glycoconjugate Research Aimed at Developing for New Industry" from the Ministry of Education, Science, and Culture of Japan. Part of this study was presented at the Annual Conference of the Society for Glycobiology, December 3–6, 2003, San Diego, CA (USA).

Supporting information for this article is available on the WWW under http://www.angewandte.org or from the author.

 DOI: 10.1002/ange.200461685



Scheme 1. Fischer's concept and glycoblotting by Fischer-type polymer reagents. a) Reactions of carbohydrates with Fischer's reagents. b) Syntheses and reactions of Fischer-type polymer reagents: 1) liposomes prepared with phosphatidylcholine-related monomer, then UV (254 nm) irradiation, 2) radical copolymerization (ammonium persulfate) with acrylamide, followed by deprotection with 4 N HCl.

chemoselective ligation strategy. From considerations of the chemical stability, the handling of reagents, and the feasibility to the general synthesis of versatile polymer materials, we have developed several Fischer-type polymer reagents with reactive and stable oxylamino functional groups by polymerization of a diacetylene-containing lipid derivative **1** or an acrylamide-type derivative **2** (Scheme 1b). Liposomes composed of diacetylene-containing reagent **1** and phospholipid derivatives can be readily polymerized by UV irradiation to produce polymer-based nanoparticles having a mean diameter of 200–300 nm. Compound **2** gives a water-soluble polymer on radical copolymerization with acrylamide. We evaluated the feasibility of the polymers displaying oxylamino functional groups as a useful platform for the "trap and release" of the target oligosaccharides and glycopeptides.

The general protocol for the isolation and analysis of carbohydrates proposed here is simple and easy: a) Trapping carbohydrates by mixing polymers with unpurified proteolytic digest; b) collecting polymers by spin filtration from a crude mixture of peptides, salts, and other impurities; and c) releasing and subsequent MALDI-TOF MS analysis of the target carbohydrates (Figure 1). In this study we selected the *N*-glycans of human IgG to test the efficiency of protein glycomics in the strategy, since the importance of substantial structural alteration of IgG *N*-glycans has been extensively investigated,[3,8] The structural characterization of human IgG *N*-glycans was carried out using crude IgG samples obtained from human serum (100 μL, Figure 1). The merits of glycoblotting are evident since MALDI-TOF MS analysis of oligosaccharides trapped on the polymer has proven to be remarkably improved by using trapping and spin filtration (Figure 2b). When the crude digest was directly subjected to MALDI-TOF MS analysis we obtained an extremely complex spectrum (Figure 2a). The precursor ion peaks, which correspond to the *N*-glycans, cannot be distinguished from numerous other signals in the spectrum because of an excess amount of peptide impurities. As anticipated, highly sensitive and singly charged precursor ion peaks were generated in the presence of α-cyano-4-hydroxycinnamic acid (CHCA) and these greatly facilitated the subsequent sequencing and quantitative analysis of the oligosaccharides. In addition, it was found that the fragmentation of the precursor ions by the

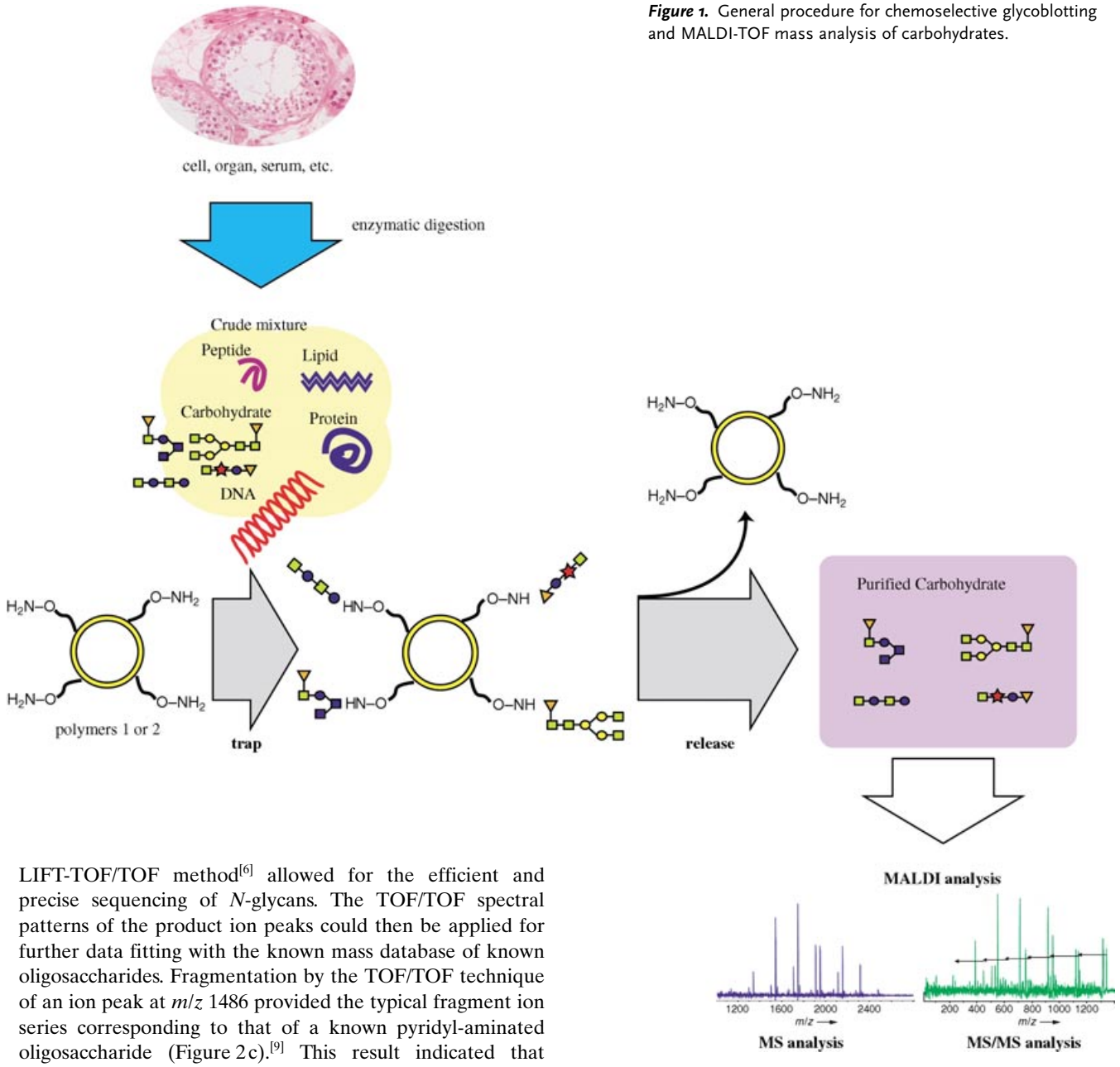


Figure 1. General procedure for chemoselective glycoblotting and MALDI-TOF mass analysis of carbohydrates.

LIFT-TOF/TOF method[6] allowed for the efficient and precise sequencing of *N*-glycans. The TOF/TOF spectral patterns of the product ion peaks could then be applied for further data fitting with the known mass database of known oligosaccharides. Fragmentation by the TOF/TOF technique of an ion peak at *m*/*z* 1486 provided the typical fragment ion series corresponding to that of a known pyridyl-aminated oligosaccharide (Figure 2c).[9] This result indicated that carbohydrate mass fingerprinting (CMF) analysis by the MALDI-TOF/TOF method allows for the precise structural identification of oligosaccharides captured from glycoproteins by glycoblotting.

It was clearly suggested that the structure of each oligosaccharide measured by the MALDI-TOF instrument was in good agreement with the data obtained by a conventional HPLC-based analysis of the corresponding pyridyl-aminated derivatives prepared from human IgG.[10] The relative intensity of each carbohydrate response given by the two methods exhibited quite similar results (see the Supporting Information), which suggests that the polymer captures carbohydrates with a reaction efficiency in proportion to the amount of each oligosaccharide existing in the sample mixtures. This result means that glycoblotting is a general and versatile method for trapping a variety of carbohydrates bound to proteins.

The CMF analysis based on TOF/TOF can also be used to determine the ratio of oligosaccharide isomers, as shown in the analysis of the ion peak at *m*/*z* 1648 (see Figure 3). The observed fragmentation of this precursor ion peak was found to be a mixture of two structural isomers and the analysis using the correlation coefficient calculated by using the known compounds enabled the ratio of the two isomers to be estimated as 62:38.

Our attention was next directed toward the application of the glycoblotting strategy for the identification of the *O*-GlcNAc site present in abundance in nucleocytoplasmic glycoproteins.[1b] Recently, we succeeded in the high-throughput identification of *O*-glycosylation (*O*-GalNAc) sites by high-speed analysis in the LIFT-TOF/TOF mode of the ion peaks of the glycosylated products in the presence of 2,5-

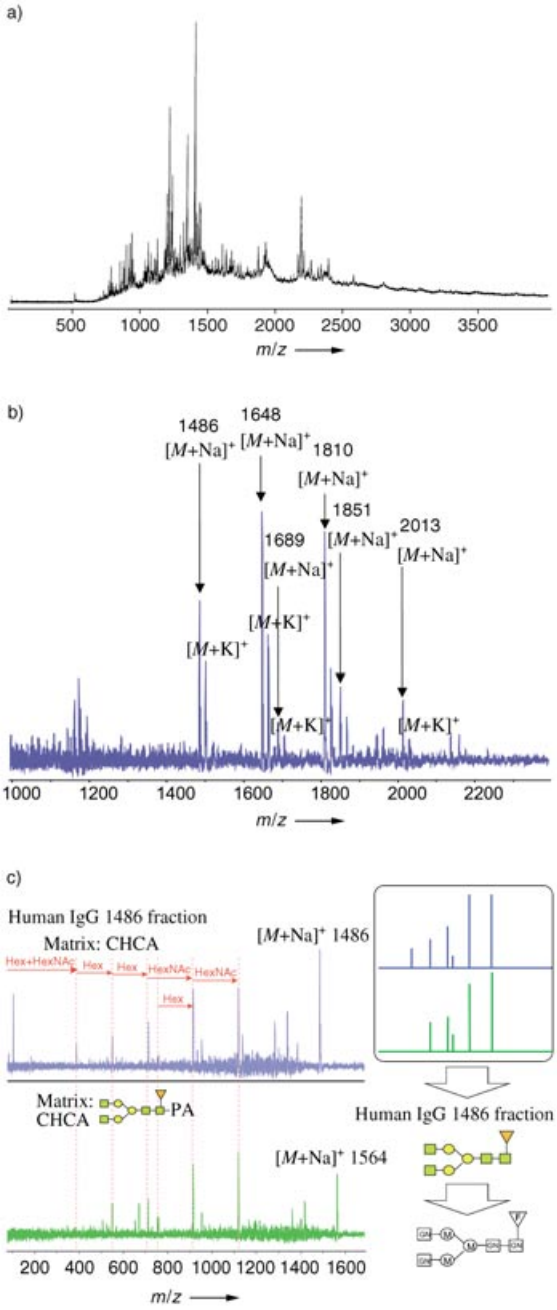
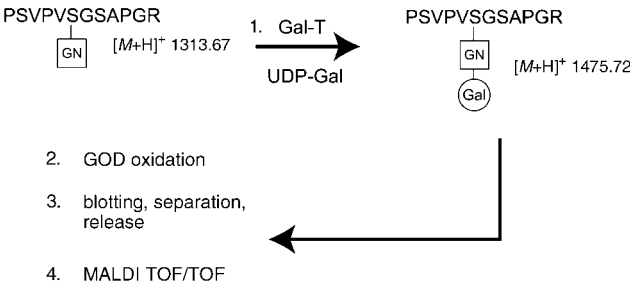


Figure 2. MALDI-TOF analysis of human IgG *N*-glycans. a) MALDI-TOF mass spectrum of crude *N*-glycans before glycoblotting. b) Major precursor ion peaks of *N*-glycans obtained by glycoblotting. c) Fragmentation by MALDI-LIFT-TOF/TOF tandem mass spectroscopy. Fragmentation of the precursor ion at m/z 1486 was performed according to the conditions reported previously in the presence of CHCA as a matrix.[11] Precise structural identification was carried out by comparing the fragmentation pattern of the known pyridyl-aminated oligosaccharide. All measurements were performed using an Ultraflex TOF/TOF mass spectrometer equipped with a reflector, and controlled by the Flexcontrol 1.2 software package (Bruker Daltonics GmbsH, Bremen, Germany). Ions generated by a pulsed UV laser beam (nitrogen laser, $\lambda = 337$ nm) were accelerated to a kinetic energy of 23.5 kV. Metastable ions generated by laser-induced decomposition of the selected precursor ions were analyzed without any additional collision gas. Precursor ions were accelerated to 8 kV and selected in a timed ion gate. The fragments were further accelerated by 19 kV in the LIFT cell, and their masses were analyzed after the ion reflector passage. Masses were automatically annotated by using XMASS 5.1.2 NT.

dihydroxybenzoic acid (DHB).[6a] To test the feasibility of glycoblotting and the subsequent TOF/TOF analysis of glycopeptides, we preliminarily investigated the enzymatic modification of a synthetic glycopeptide (PSVPVS(*O*-GlcNAc)GSAPGR) which was identified as the *O*-GlcNAc-modified peptide derived from basic phosphoprotein.[11] This glycopeptide was subjected to: a) Galactosylation with galactosyl transferase (GalT)[12] and b) oxidation with galactose oxidase (GOT) to generate a terminal aldehyde group at the C-6 position of the modified galactose residue (Scheme 2). A



Scheme 2. Strategy for the analysis of *O*-GlcNAc modified peptides.

crude glycopeptide mixture was directly used for further blotting with polymer reagent **2** and the subsequent analysis by MALDI-TOF mass spectrometry. Since peptide fragments bearing oligosaccharides prepared by tryptic digestion often exhibited poor solubility in water, water-soluble Fischer-type reagents may be a useful alternative to nanoparticles in terms of the efficiency. The TOF/TOF MS spectrum of the precursor ion at m/z 1473.71 measured with DHB gave highly sensitive and ideal fragmentation patterns because of the glycopeptide containing a (CHO)Galβ→4GlcNAcβ→Ser residue (Figure 4b). As anticipated, fragmentation in the presence of the DHB matrix occurred preferentially at peptide bonds without serious de-*O*-glycosylation (elimination) or fragmentation (cleavage) in the carbohydrate moiety. This characteristic greatly facilitated the identification of the O-glycosylation site in the target molecule (Figure 4c). This result suggests that the aldehyde attachment by enzymatic derivatization of the *O*-GlcNAc-modified peptide enables a high-throughput identification of O-glycosylation sites based on the glycoblotting method.[13] The merit of the present strategy is evident because the identification of new *O*-GlcNAc-modified proteins and the sites of modification would facilitate more global studies of the regulatory role of this posttranslational modification.

In conclusion, this study demonstrated that glycoblotting based on oxylamino-containing polymers (Fischer-type polymer reagents) could generally be used for capturing both common oligosaccharides and aldehyde-attached glycopeptides derived by enzymatic modifications. It should be noted that the efficiency and versatility in the capturing of carbohydrates from the crude proteolytic digests is crucial for ideal and practical protein glycomics. Therefore, the combined use of glycoblotting and high-performance MALDI-TOF/TOF analysis provides us with a new promising

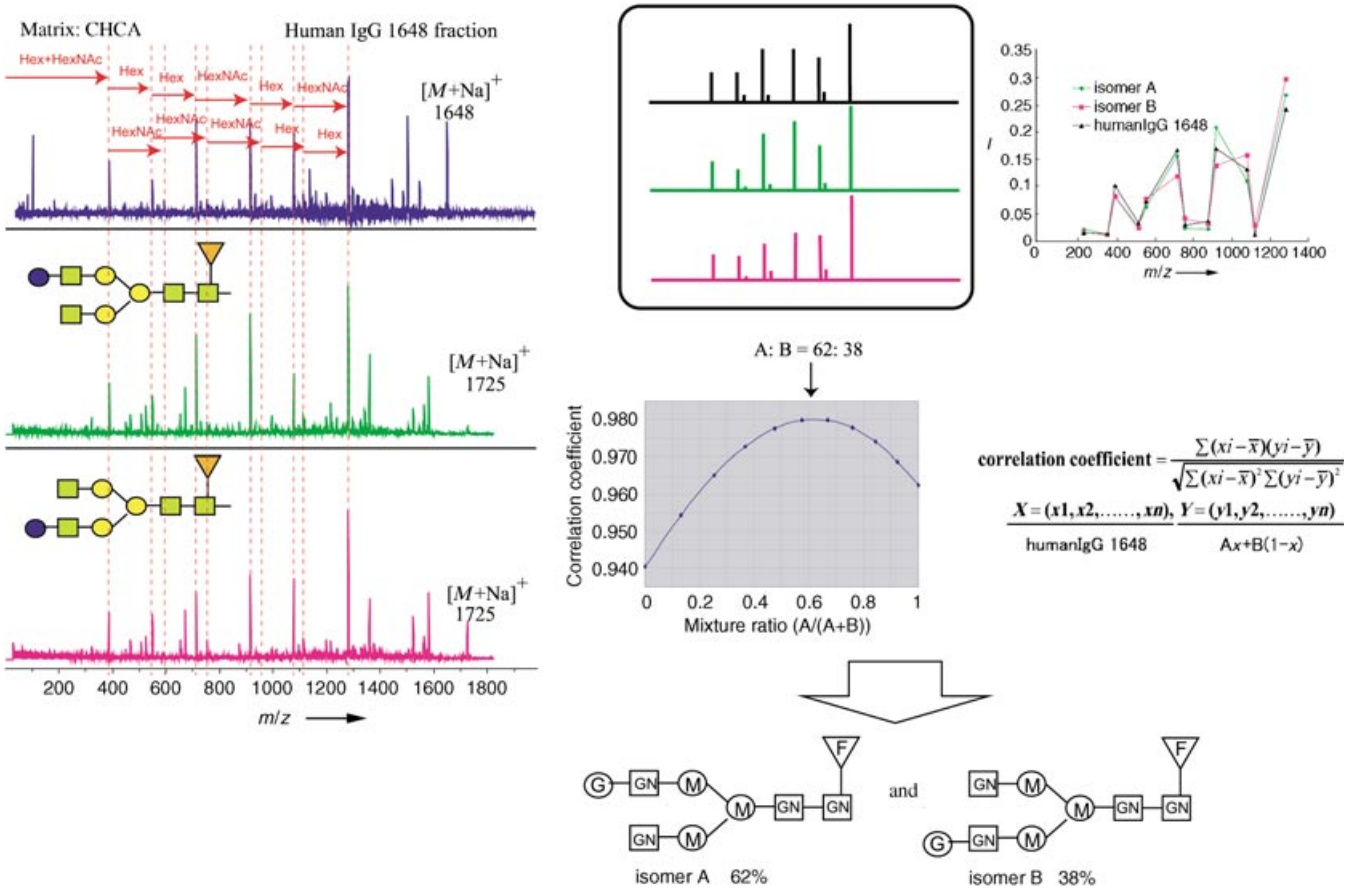


Figure 3. Identification of the ratio between isomers A and B found in the molecular ion peak at *m/z* 1648 by using carbohydrate mass fingerprinting (CMF) analysis recorded in LIFT-TOF/TOF mode. The relative intensities of LIFT-TOF/TOF product ion peaks of unknown carbohydrate structure were subjected to matching simulation with the correlation coefficient analysis. Two known oligosaccharide derivatives were used for achieving the standard CMF data.

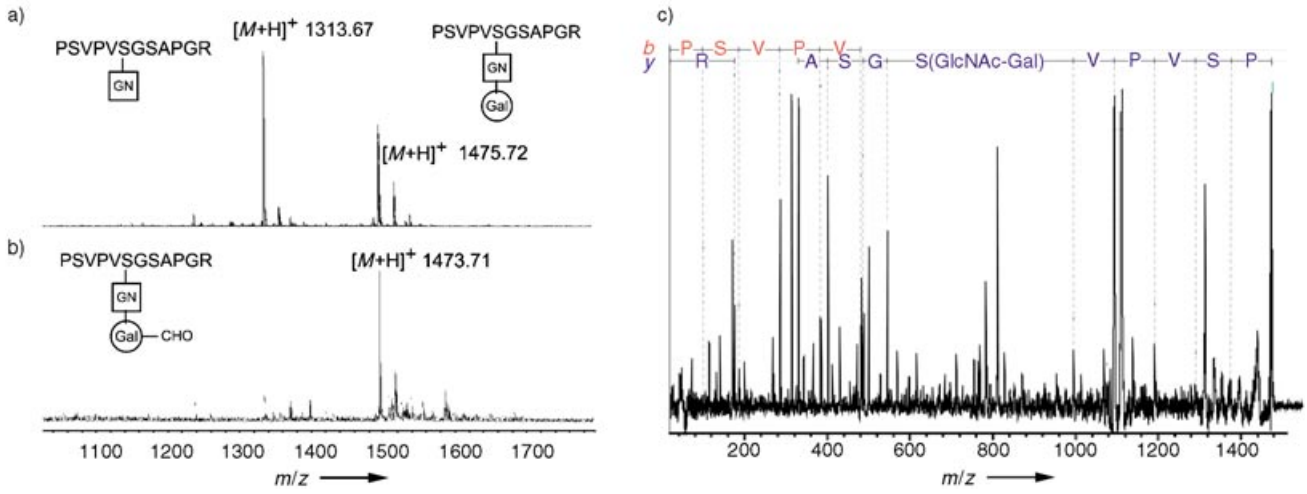


Figure 4. Identification of the *O*-GlcNAc modified peptides by MALDI-TOF mass spectrometry. a) MALDI-TOF MS of the reaction mixture after enzymatic galactosylation of glycopeptide, b) MALDI-TOF MS of an isolated glycopeptide derivative by glycoblotting method, and c) MALDI-LIFT-TOF/TOF tandem mass analysis of the precursor ion peak at *m/z* 1473.71.

strategy for early diagnoses and tailored treatments of a variety of diseases, as well as high-throughput protein glycomics.

Received: August 17, 2004

Keywords: carbohydrates · chemoselectivity · glycoproteins · high-throughput screening · mass spectrometry

[1] a) R. A. Dwek, *Chem. Rev.* **1996**, *96*, 683; b) N. E. Zachara, G. W. Hart, *Chem. Rev.* **2002**, *102*, 431.
[2] R. B. Parekh, R. A. Dwek, B. J. Sutton, D. L. Fernandes, A. Leung, D. Stanworth, B. T. W. Rademacher, *Nature* **1985**, *316*, 452.
[3] A. Varki in *Essentials of Glycobiology* (Eds.: A. Varki, R. Cummings, J. Esko, H. Freeze, G. Hart, J. Marth), Cold Spring Harbor Laboratory Press, New York, **1999**, pp. 537–549.
[4] R. L. Shields, J. Lai, R. Keck, L. Y. O'Connell, K. Hong, Y. G. Meng, S. H. A. Weikert, L. G. Presta, *J. Biol. Chem.* **2002**, *277*, 26733.
[5] F. I. Comer, G. W. Hart, *J. Biol. Chem.* **2000**, *275*, 29179.
[6] a) M. Kurogochi, T. Matsushita, S.-I. Nishimura, *Angew. Chem.* **2004**, *116*, 4163–4167; *Angew. Chem. Int. Ed.* **2004**, *43*, 4071–4075; b) M. Kurogochi, S.-I. Nishimura, *Anal. Chem.* **2004**, in press; c) M. Kurogochi, S.-I. Nishimura, Y. C. Lee, *J. Biol. Chem.* **2004**, in press.
[7] a) E. Fischer, *Ber. Dtsch. Chem. Ges.* **1884**, *17*, 579; b) E. Fischer, *Ber. Dtsch. Chem. Ges.* **1887**, *20*, 821.
[8] J. S. Axford, *Biochim. Biophys. Acta* **1999**, *1455*, 219.
[9] D. Suckau, A. Resemann, M. Schuerenberg, P. Hufnagel, J. Franzen, A. Holle, *Anal. Bioanal. Chem.* **2003**, *376*, 952.
[10] H. Nakagawa, Y. Kawamura, K. Kato, I. Shimada, Y. Arata, N. Takahashi, *Anal. Biochem.* **1995**, *226*, 130.
[11] K. D. Greis, W. Gibso, G. W. Hart, *J. Virol.* **1994**, *68*, 8339.
[12] K. Kamemura, B. K. Hayes, F. J. Comer, G. W. Hart, *J. Biol. Chem.* **2002**, *277*, 19229.
[13] An aldehyde group generated at the C-8 position of the terminal sialic acid of a glycoprotein can also be trapped by this method. The feasibility of this strategy was preliminarily demonstrated by trapping the tryptic digest of fetuin in which the diols at the C-8 and C-9 positions of sialic acids were selectively converted into aldehyde groups by treating with 1 mM $NaIO_4$ solution at 0 °C for 30 min (data not shown). The present method can not be used for the analysis of glycopeptides displaying high mannose-type oligosaccharides, since it seems that specific conversion of a terminal mannose residue into an appropriate aldehyde derivative is difficult.

