Proteomics Research

Post-Translational Modifications on Proteins: Facile and Efficient Procedure for the Identification of O-Glycosylation Sites by MALDI-LIFT-TOF/TOF Mass Spectrometry

*Masaki Kurogochi, Takahiko Matsushita, and Shin-Ichiro Nishimura**

Mass spectrometry has become one of the most important and powerful methods for the rapid identification of proteins in recent post-genome science and technology.[1] In addition to phosphorylation, a variety of post-translational modifications such as sulfation, acylations, methylation, and glycosylations were recently recognized to regulate the structures and functions of proteins.[2] Therefore, advent of an efficient and general strategy for identifying sites of post-translational modifications on proteins is highly awaited in proteomics research. Although glycosylations of proteins are basically classified into two major types, N-glycosylation and O-glycosylation, post-translational modification by glycosylation has been studied extensively in the case of *N*-linked-type glycans.[3] It has been known from studies of the *N*-glycans that *N*-acetyl-D-glucosamine links to an asparagine residue of the tripeptide sequence Asn-X-Ser(Thr) through formation of a stable amide bond. In contrast to N-glycosylation, there is no consensus sequence directing the initial incorporation of the sugar into mucin peptides by UDP-GalNAc, a polypeptide *N*-acetyl-D-galactosaminyltransferase (GalNAc-transferase). The difficulty in identifying O-glycosylation sites of mucins is a result of the instability of O-glycosidic bonds between GalNAc (GlcNAc) and Ser/Thr in these compounds. It has also been suggested that direct observation of O-glycosylation sites during collision-induced dissociation in mass spectrometry has proved difficult because the glycosidic linkages are labile and easily cleaved, thus resulting in very little fragmentation of the glycopeptides.[4] As a result, an alternative way for direct observation has been used: β-elimination followed by Michael addition of some nucleophiles to the dehydroalanine residue.[5] At present, it seems that NMR spectroscopy is virtually the only effective and reliable method for the structural elucidation of mucin-type glycopeptides with unstable O-glycosidic bonds, even though milligram-scale samples are required.[6]

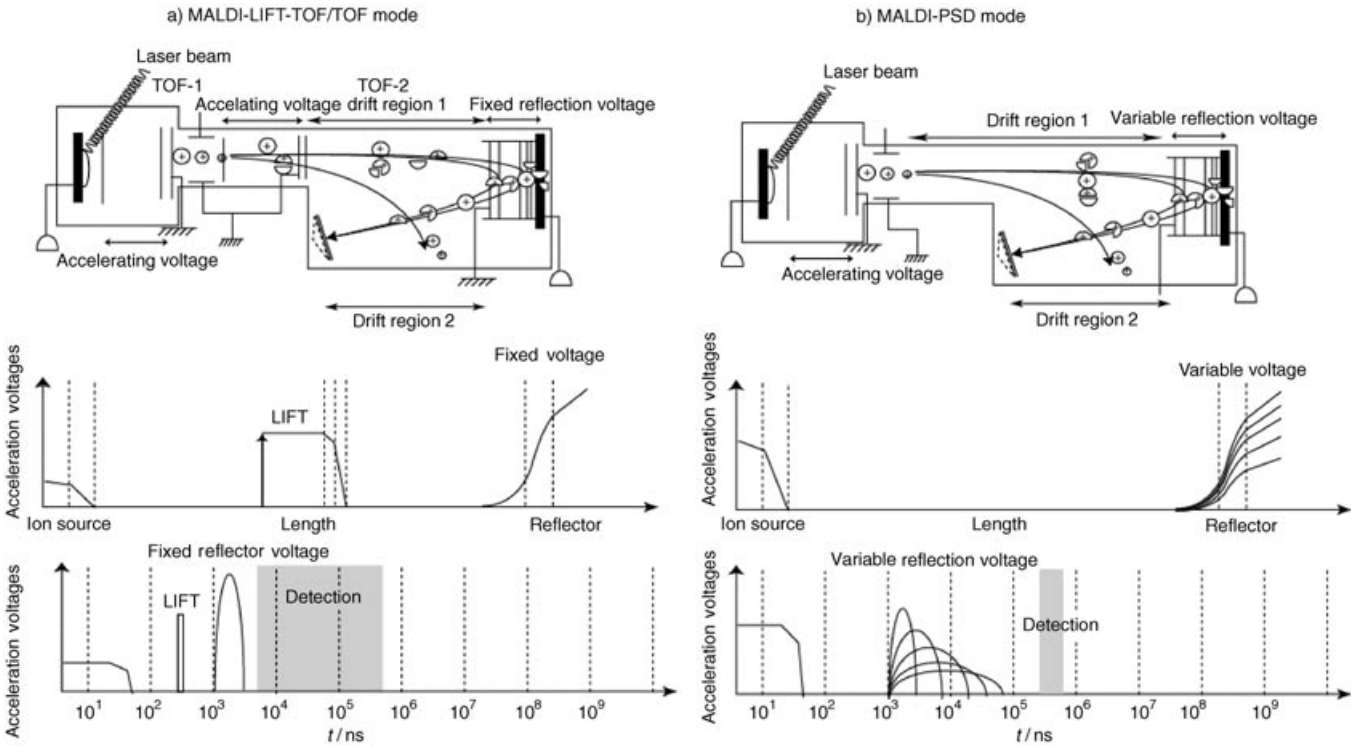


Figure 1. Schematic diagrams of LIFT-TOF/TOF (a) and PSD (b) in MALDI mass instruments.

In the course of our synthetic studies on functional glycopeptides and proteomics of modified β-galactosidases,[6b,7] we found that MALDI-TOF and MALDI-LIFT-TOF/TOF mass spectra measured in the presence of 2,5-dihydroxybenzoic acid (DHB) as the matrix often gave observable peaks corresponding to the target (glyco)peptides

[*] M. Kurogochi, T. Matsushita, Prof. S.-I. Nishimura
Division of Biological Sciences, Graduate School of Science
Frontier Research Center for the
Post-Genomic Science and Technology
Hokkaido University, N21, W11, Sapporo 001–0021 (Japan)
Fax: (+81) 11-706-9042
E-mail: shin@glyco.sci.hokudai.ac.jp

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

 DOI: 10.1002/ange.200460020

and glycoproteins. The artificial glyco-insulin carrying a sialooligosaccharide branch could also be detected as a single ion peak without any degradation of the sugar moiety,[7b] thus suggesting that the use of DHB might help in the ionization of unstable O-glycopeptides bearing α-sialosides. In addition, Schnaible et al. recently reported that peptide fragment ions having unstable S–S bonds can be clearly detected by mass spectrometry using a MALDI-LIFT-TOF/TOF instrument.[8] We thought that these interesting features of MALDI-TOF mass spectra must depend on the collision process used in the LIFT-TOF/TOF method having a significant affect on the intramolecular relaxation of the fragment ions.[9] Although a post source decay (PSD) as well as a collision-induced dissociation (CID) mode is used for MS/MS of glycopeptides,[10] precise analysis of fragment ion peaks often seems to be difficult because of preferential and fast deglycosylation.

As illustrated in Figure 1a, the LIFT-TOF/TOF instruments are characterized by a colinear arrangement of two TOF mass analyzers (TOF-1 and TOF-2), each equipped with an ion source that allows control over the acceleration and focusing of the ions. In TOF-1, the analyte ions are accelerated, and precursor ions are selected and fragmented. Then, fragment ions are selected to proceed to the source of TOF-2, where they are again accelerated and analyzed by the detector. As shown in the schematic diagram, interpretation of the LIFT-TOF/TOF method is straightforward as only one assignment must be verified,[9] while the PSD or CID mode requires several trials to control the reflectron voltage and enable separation and detection of each product ion (Figure 1b). As a result, the collision process in the MALDI-LIFT-TOF/TOF mass spectrometer under a mild and very fast experimental process seemed to offer much faster analysis and higher sensitivity than that of CID or PSD mode, even in cases of some unstable fragment ions. The merit of LIFT-TOF/TOF mass spectrometry in the peptide fragmentation studies prompted us to investigate the high-throughput identification of O-glycosylation sites of glycopeptides by MALDI-TOF mass spectrometry. Herein, we report that the MS/MS spectra of some synthetic O-glycopeptides measured on a MALDI-LIFT-TOF/TOF instrument in the presence of DHB as a matrix gave ideal fragmentation patterns and high sensitivity of structure-specific fragment ions that were suitable for identifying O-glycosylation sites of mucin-type glycoproteins.

Figure 2 shows MALDI-LIFT-TOF/TOF spectra of a peptide Ac-AHGVTSAPDTRPAPGSTA-NH_2 (**1**), and seven synthetic glycopeptide isomers having GalNAc residue at the position(s) T5, T10, or/and T17 (**2**)-(**8**). It was suggested that the DHB matrix produced much more reasonable and

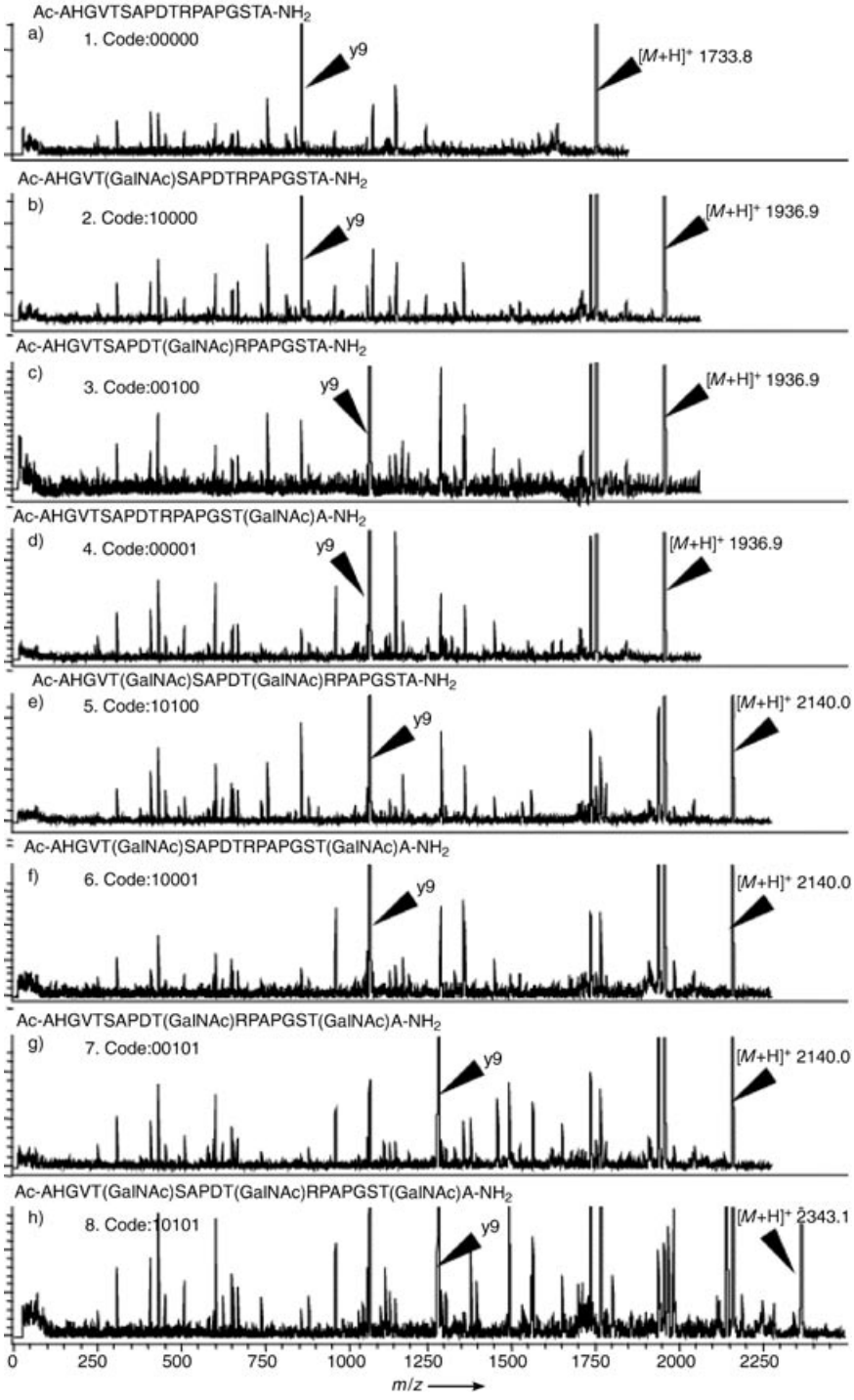


Figure 2. Fragmentations of peptide (**1**) and glycopeptides (**2**–**8**) by LIFT-TOF/TOF MS.

sensitive fragment ion peaks (b and y fragment ion series) bearing GalNAc residue(s) for the structural elucidation than α-cyano-4-hydroxycinnamic acid (CHCA) or sinapinic acid (SA; see the Supporting Information). In addition, we found that a major fragment ion peak (y9) of each glycopeptide generated by LIFT-TOF/TOF MS significantly sifted in accordance with the O-glycosylation site(s). These results greatly facilitated high-throughput identification of the O-glycosylation site by assignment of the fragment ions with Biotool 2.2 software (Bruker Daltonik GmbH). For example, a glycopeptide isomer bearing GalNAc residues at the three positions (**8**, code: 10101) gave important fragment peaks such as b4, b5, b8, b10, b16, y2, y3, y8, y9, y13, and y14 thus indicating the exact mass of the GalNAc residues attached to positions T5, T10, and T17 (Figure 3). In addition, significant fragment ions at *m*/*z* 1733, 1936, and 2140 corresponding to losses of neutral carbohydrate units also supported the reliability in the assignment of this compound. It was suggested that use of a MS/MS scoring method such as molecular weight search (MOWSE) with some selected query fragment ion peaks would greatly accelerate high-throughput identification of the O-glycosylation site(s) of mucin-type glycopeptides. Although MOWSE scoring predicted two plausible candidates (code: 10110 and code: 10101), further precise analysis of important ion peaks of the b and y series, such as b16, y2, and y3 gave a reasonable structure of compound **8** coded as 10101.

The versatility of the present method was preliminarily demonstrated by monitoring a transglycosylation reaction catalyzed by β-galactosidase as illustrated in Scheme 1. Transglycosylation by β-galactosidase was carried out according to the conditions reported by Vliegenthart et al.[4d] using a glycopeptide **8** (code: 10101) as an acceptor substrate and lactose as a donor substrate. It is known that this galactosylation reaction may proceed in moderate or relatively low yield and give Galβ(1→3)GalNAc derivatives as the main products. As anticipated, transglycosylation of **8** provided a new ion peak at *m*/*z* 2509.1 in addition to a peak corresponding to the starting material (*m*/*z* 2343.1; Figure 4). Although MOWSE scoring predicted three plausible candidates of the product, it was concluded from precise analysis of the y and b series of fragment ions that this peak involves two galactosylated isomers coded as 20101 and 10201.

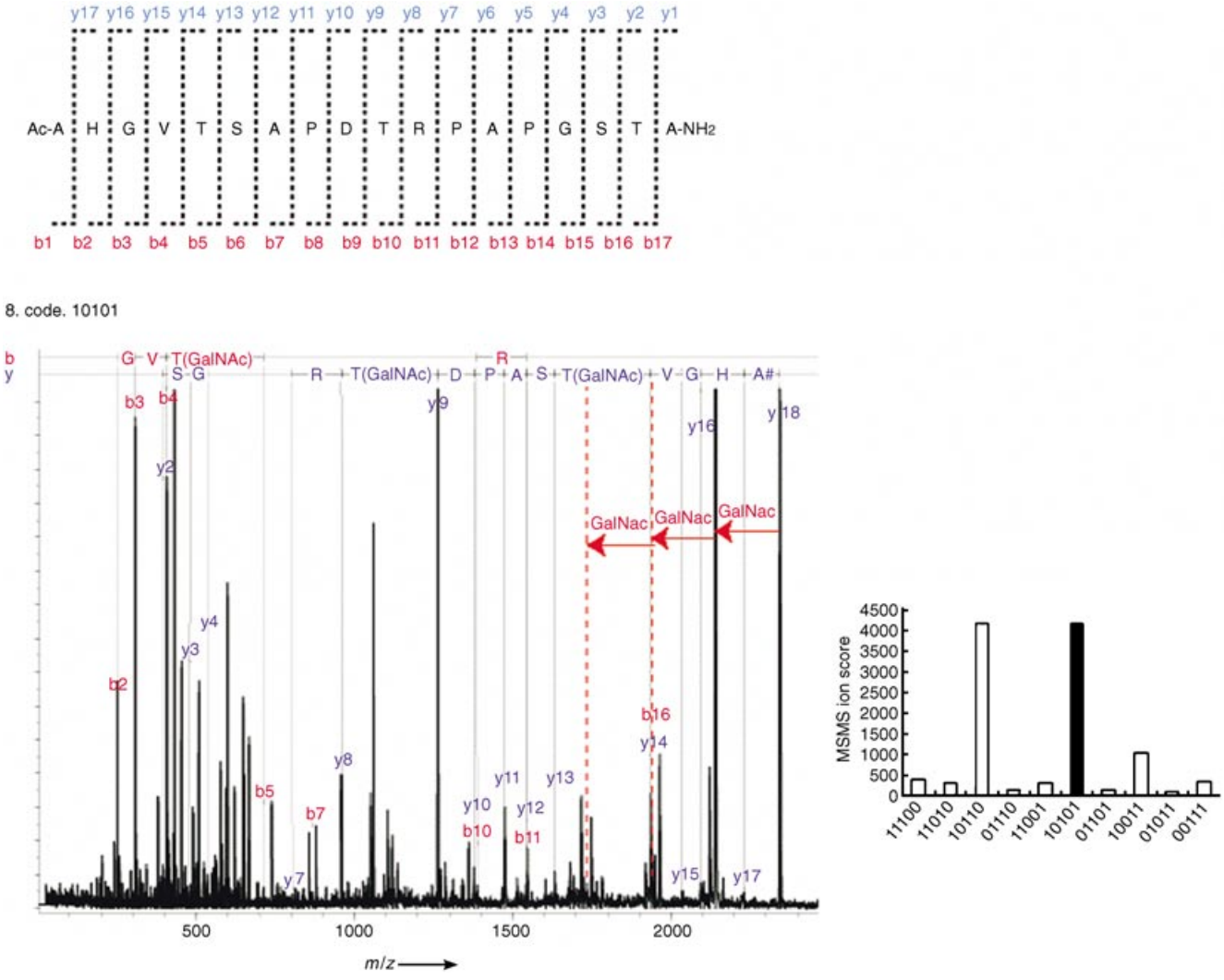
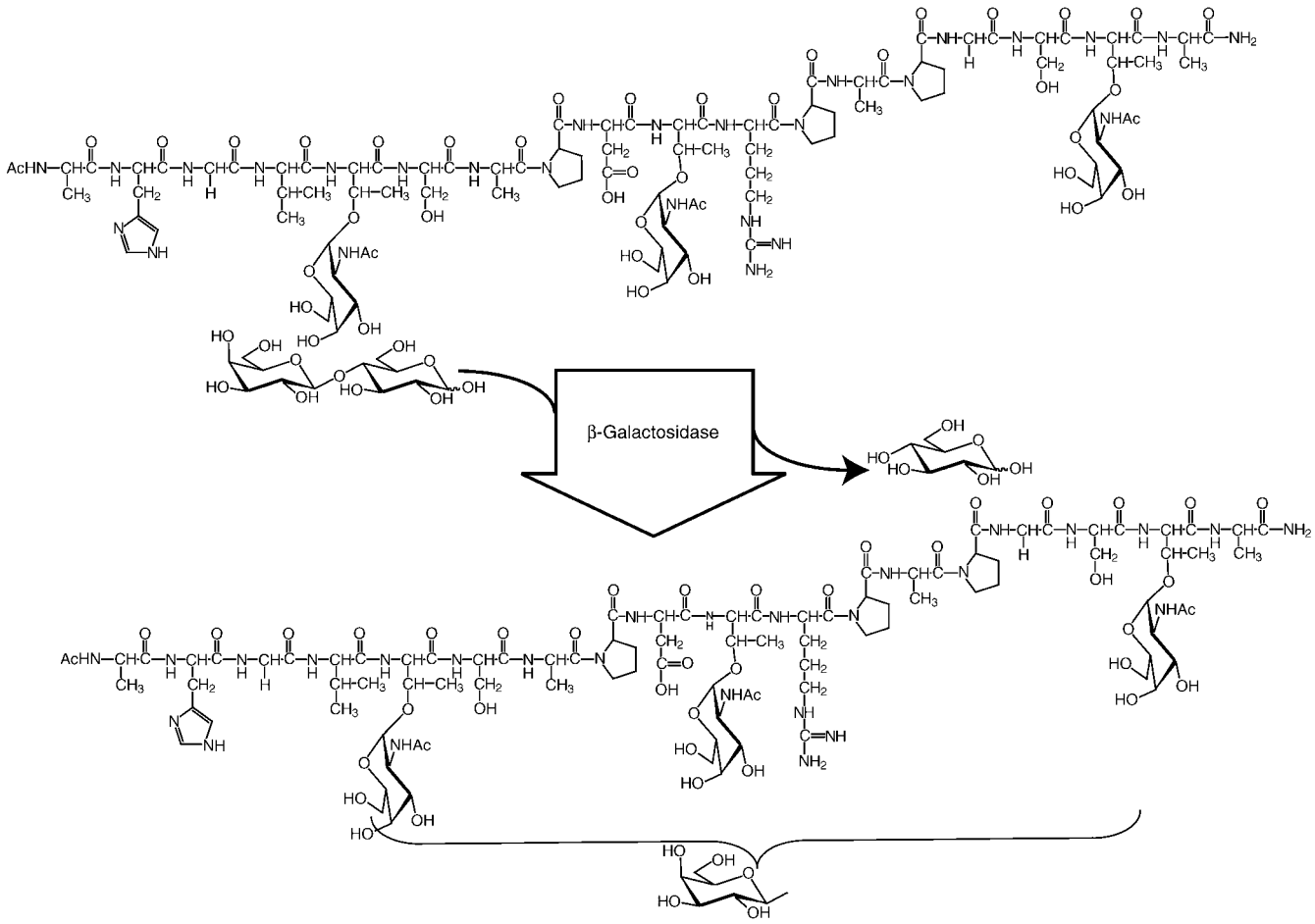


Figure 3. Precise analysis of LIFT-TOF/TOF spectra and MOWSE scoring for compound **8**.



Scheme 1. The transglycosylation reaction of compound **8** catalyzed by β-galactosidase in the presence of lactose as a glycosyl donor.

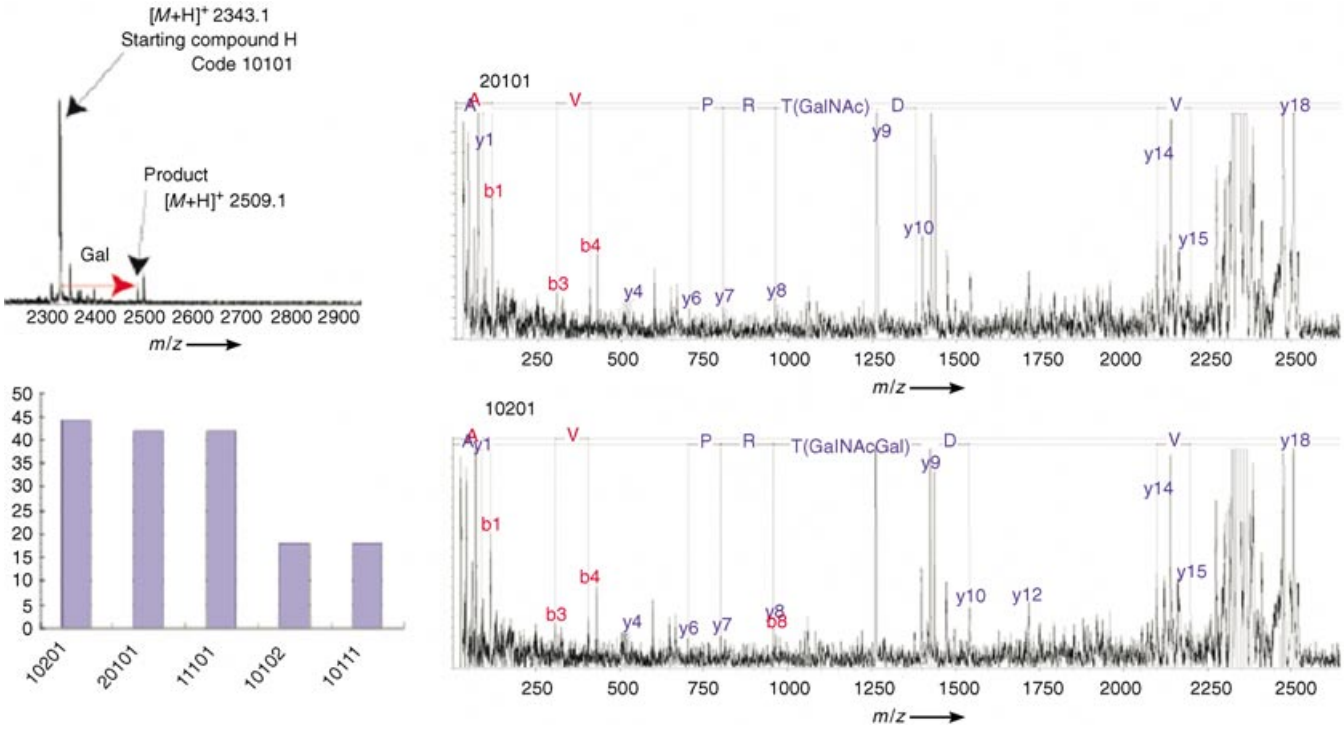


Figure 4. Transglycosylation reaction catalyzed by β-galactosidase with glycopeptide **8** and identification of sugar elongation site(s) by LIFT-TOF/TOF MS.

In conclusion, we have found that MALDI-LIFT-TOF/TOF spectrometry in the presence of a DHB matrix gives highly sensitive and ideal fragmentation patterns (b and y series of ions) generated from glycopeptide precursor ion peaks and they can be readily used for precise and reliable structure analyses including the identification of O-glycosylation sites of mucin-type glycopeptide derivatives. The merit of the present strategy is that the combined use of the fragmentation behavior of unknown glycopeptides and database searching of peptide mass fingerprints will greatly contribute to high-throughput proteomics research in terms of the post-translational modifications on functional proteins. In addition, this method could also be applied for screening systems of chemical and enzymatic synthesis of glycopeptides and glycoproteins in a combinatorial synthetic manner as preliminarily demonstrated in Figure 4.

Received: March 17, 2004 [Z460020]

Keywords: glycosylation · high-throughput screening · mass spectrometry · protein modifications · proteomics

[1] T. Keough, R. S. Youngquist, M. P. Lacey, *Anal. Chem.* **2003**, *75*, 156–165.
[2] a) T. Kouzarides, *EMBO. J.* **2000**, *19*, 1176–1179; b) A. E. McBridge, P. A. Silver, *Cell* **2001**, *106*, 5–8; c) L, Wells, K. Vosseller, G. W. Hart, *Science* **2001**, *291*, 2376–2378.
[3] a) K. Hankansson, H. J. Cooper, M. R. Emmett, C. E. Costello, A. G. Marshall, C. L. Nilsson, *Anal. Chem.* **2001**, *73*, 4530–4536; b) K. Hankansson, M. J. Chalmers, J. P. Quinn, M. A. McFarland, C. L. Hendrickson, A. G. Marshall, *Anal. Chem.* **2003**, *75*, 3256–3262.
[4] a) S. Muller, S. Goletz, N. Packer, A. Gooley, A. M. Lawson, F.-G. Hanisch, *J. Biol. Chem.* **1997**, *272*, 24780–24793; b) S. Goletz, B. Thiede, F.-G. Hanisch, M. Schultz, J. Peter-Katalinic, S. Muller, O. Seitz, U. Karsten, *Glycobiology* **1997**, *7*, 881–896; c) F.-G. Hanisch, B. N. Green, R. Bateman, J. Peter-Katalinic, *J. Mass Spectrom.* **1998**, *33*, 358–362; d) R. G. Gallego, G. Dudziak, U. Kragl, C. Wandrey, J. P. Kamerling, J. F. G. Vliegenthart, *Biochimie* **2003**, *85*, 275–286.
[5] a) K. D. Greis, B. K. Hayes, F. I. Comer, M. Kirk, S. Barnes, T. L. Lowary, G. W. Hart, *Anal. Biochem.* **1996**, *234*, 38–49; b) L. Wells, K. Vosseller, R. N. Cole, J. M. Cronshaw, M. J. Matunis, G. W. Hart, *Mol. Cell. Proteomics* **2002**, *1*, 791–804.
[6] a) D. M. Coltart, A. K. Royyuru, L. J. Williams, P. W. Glunz, D. Sames, S. D. Kuduk, J. B. Schwarz, X.-T. Chen, S. J. Danishefsky, D. H. Live, *J. Am. Chem. Soc.* **2002**, *124*, 9833–9844; b) Y. Tachibana, G. L. Fletcher, N. Fujitani, S. Tsuda, K. Monde, S.-I. Nishimura, *Angew. Chem.* **2004**, *116*, 874–880; *Angew. Chem. Int. Ed.* **2004**, *43*, 856–862.
[7] a) T. Ohta, N. Miura, N. Fujitani, F. Nakajima, K. Niikura, R. Sadamoto, C.-T. Guo, T. Suzuki, K. Monde, S.-I. Nishimura, *Angew. Chem.* **2003**, *115*, 5344–5347; *Angew. Chem. Int. Ed.* **2003**, *42*, 5186–5189; b) M. Sato, R. Sadamoto, K. Niikura, K. Monde, H. Kondo, S.-I. Nishimura, *Angew. Chem.* **2004**, *116*, 1542–1546; *Angew. Chem. Int. Ed.* **2004**, *43*, 1516–1520; c) M. Kurogochi, S.-I. Nishimura, Y. C. Lee, *J. Biol. Chem.* **2004**, submitted.
[8] V. Schnaible, S. Wefing, A. Resemann, D. Suckau, A. Bucker, S. Wolf-Kummeth, D. Hoffmann, *Anal. Chem.* **2002**, *74*, 4980–4988.
[9] D. Suckau, A. Resemann, M. Schuerenberg, P. Hufnagel, J. Franzen, A. Holle, *Anal. Bioanal. Chem.* **2003**, *376*, 952–965.
[10] a) N. Xu, Z.-H. Huang, B. L. M. de Jonge, D. A. Gage, *Anal. Chem.* **1997**, *248*, 7–14; b) A. Tsarbopoulos, U. Bahr, B. N. Pramanik, M. Karas, *Int. J. Mass Spectrom. Ion Processes* **1997**, *169*/170, 251–261; c) R. J. Chalkley, A. L. Burlingame, *J. Am. Soc. Mass Spectrom.* **2001**, *12*, 1106–1113; d) R. J. Chalkley, A. L. Burlingame, *Mol. Cell. Proteomics* **2003**, *2*, 182–190.

