Combinatorial Chemistry

Molecular-Weight-Tagged Glycopeptide Library: Efficient Construction and Applications**

*Hiromi Ito, Akihiko Kameyama, Takashi Sato, Katsue Kiyohara, Yoshiaki Nakahara, and Hisashi Narimatsu**

Combinatorial chemistry has become indispensable for drug discovery and biotechnological research with high-throughput screening. Recently, glycoproteomics or glycomics has become a major academic and clinical research priority. Oligosaccharides on glycoconjugates, including glycoproteins and glycolipids, play vital roles in biological processes by influencing the stability of proteins, protein conformation, intra- and intercell signaling, and binding specificity for other biomolecules.[1] However, the functions of glycoconjugates are little understood at the molecular level, mostly because of a lack of sensitive and high-throughput methods for analyzing the structure and interaction of oligosaccharides. To overcome these problems, technologies such as an observational multistage tandem mass-spectral (MS^n) library[2] and arrays for carbohydrate and lectin[3] have been developed. However, a large variety of structurally defined oligosaccharides, that is, an oligosaccharide library,[4] is prerequisite for these technologies. A convenient and effective method for the construction of an oligosaccharide library is desired. Herein, we describe a method for the efficient construction of a glycopeptide library which can be easily converted to an oligosaccharide library by using human recombinant glycosyltransferases.[5] Furthermore, applications for the screening of ligand specificity for lectins by mass spectrometry are demonstrated.

We have developed a method for the construction of molecular-weight-tagged libraries, in which the structure and molecular weight of the components were designed to correspond one-to-one by using glycosyltransferases. Our general strategy for the construction of the libraries is shown in Figure 1. To achieve a diversity of oligosaccharides, the starting peptide or glycopeptide **X**, enzyme **I**, and the donor substrate **A** were added to a reaction mixture, and the reaction was stopped at around a 50% yield of the product

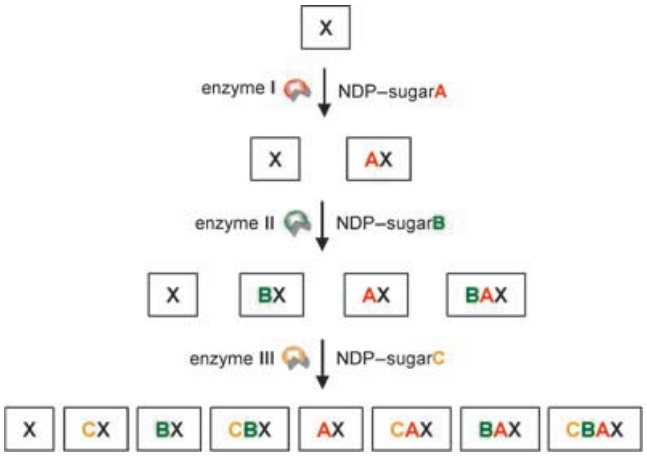

Figure 1. Basic strategy for the construction of oligosaccharide and glycopeptide libraries.

AX by monitoring the time course of the reaction using mass spectrometry. The reaction mixture was kept at 0°C during monitoring to stop the enzymatic reaction from proceeding. After enzyme **I** had been irreversibly denatured and inactivated by heat, enzyme **II** and the donor substrate **B** were added to the mixture of **X** and **AX**, and the mixture was again incubated. Then the transfer reaction of **B** was stopped at around a 50% yield, and enzyme **II** was inactivated. Sequentially, enzyme **III** and substrate **C** were added to the above solution containing **X**, **BX**, **AX**, and **BAX**, and the reaction was stopped at around a 50% yield. Finally, the glycopeptides were easily purified using a solid-phase extraction device, such as a reversed-phase microcolumn. After the three reactions, a mixed library, which theoretically contains eight different compounds (**X**, **CX**, **BX**, **CBX**, **AX**, **CAX**, **BAX**, and **CBAX**), should be generated.

The major advantages of our method are as follows: First, the formation of glycosidic linkages can be strictly controlled because of the excellent regio- and stereoselectivity of glycosyltransferases, which exhibit a clear specificity for substrates.[5] Second, by terminating each enzyme reaction halfway, it is possible to construct a library that has theoretically 2^n (n is the number of reaction steps) structures of oligosaccharides in a single tube. Third, the structure of each component in the library can be identified immediately by mass spectrometry alone, because these libraries are designed as molecular-weight-tagged libraries, that is, each product has a different molecular weight.

Our strategy has the limitation of not allowing the repetitive use of the same glycosyl donor or of diastereomeric donors (e.g., mannose and galactose donors) to achieve variation. Therefore, it is important for the systematic construction of the libraries to design reaction protocols that include the selection of enzymes, the reaction sequence, and the structure of the starting materials.

To demonstrate this strategy, we attempted to construct a library of O-linked glycopeptides, most of which are not commercially available. Extensions of the carbohydrate chain of glycopeptide **1** (β3Gal-Core 3-Muc1a) were achieved by sequential addition of each enzyme (ST6GalNAcI, fucosyl-

[*] Dr. H. Ito, Dr. A. Kameyama, Dr. T. Sato, K. Kiyohara, Prof. Dr. H. Narimatsu
Research Center for Glycoscience
National Institute of Advanced Industrial Science and Technology (AIST)
1-1-1 Umezono, Tsukuba, Ibaraki 305-8568 (Japan)
Fax: (+81) 29-861-3201
E-mail: h.narimatsu@aist.go.jp
Prof. Dr. Y. Nakahara
Department of Applied Biochemistry
Institute of Glycotechnology, Tokai University
Kitakaname 1117, Hiratsuka, Kanagawa 259-1292 (Japan)

[**] This work was performed as a part of the R&D Project of the Industrial Science and Technology Frontier Program supported by the New Energy and Industrial Technology Development Organization (NEDO).

 DOI: 10.1002/ange.200500802

transferase 2 (FUT2), and α3GalNAcT (A enzyme)) and its donor substrate (cytidine 5′-monophospho-*N*-acetylneuraminic acid (CMP-Neu5Ac), guanosine 5′-diphospho-β-L-fucose (GDP-Fuc), and uridine 5′-diphospho-*N*-acetylgalactosamine (UDP-GalNAc), respectively) in a single tube (Figure 2a). The time course of each reaction was monitored with a MALDI-TOF MS instrument (ReflexIV, Bruker-Daltonik) by directly applying the reaction solution to the target plate, and the reaction was stopped at around a 50% yield by inactivating the enzyme at 100°C for five minutes. Figure 2b shows the mass spectrum of the final reaction mixture. As α3GalNAcT transfers a GalNAc residue in the α1–3 linkage to the terminal Fucα1–2Gal structure, this library has six (over eight possible) compounds in this case. To each peak, the structure could be instantly assigned from the *m*/*z* value. In the same way, we synthesized a glycopeptide library having various *O*-linked oligosaccharides, that is, B antigen, Lewis^X and sialyl Lewis^X epitope, or Lewis^A and sialyl Lewis^A epitope elongated on the Core 2 and Core 3 structures, respectively (data not shown).

Reductive β elimination[6] by incubating the library at 50°C overnight in 20 μL of NaOH (50 mM) and $NaBH_4$ (500 mM) afforded the corresponding oligosaccharide library (Figure 3). Collision-induced dissociation (CID) spectra were acquired with a MALDI-QIT-TOF MS instrument (AXIMA-QIT, Shimadzu), and deposited in a CabosML database.[7]

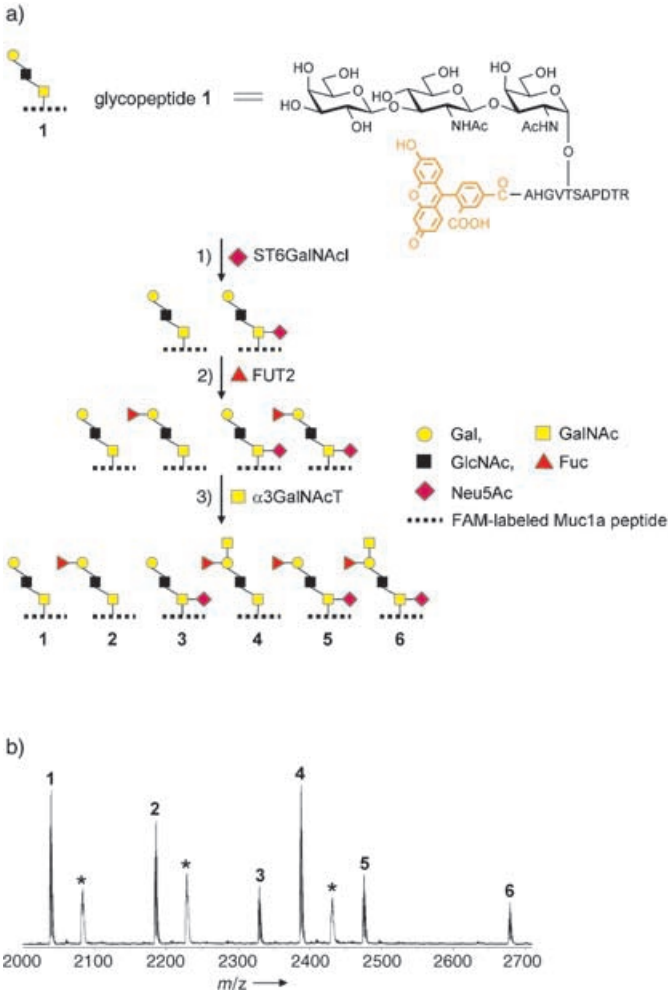

Figure 2. Demonstration of our construction strategy. a) Synthesis of an *O*-linked glycopeptide library including A antigen structures. b) MALDI-TOF mass spectrum of the prepared library. The asterisks indicate metastable ions from the sialylated compounds. Conditions for the enzymatic reactions: 1) ST6GalNAcI (5 μL), *N*-(2-hydroxylethyl)-piperazine-*N*′-2-ethanesulfonic acid (HEPES buffer) (25 mM, pH 7.0), $MnCl_2$ (10 mM), CMP-Neu5Ac (25 nmol), glycopeptide **1** (1 nmol), 37°C, 12 h; 2) FUT2 (2 μL), GDP-Fuc (5 nmol), 37°C, 2 h; 3) α3GalNAcT (2 μL), UDP-GalNAc (2 nmol), 37°C, 3 h. The recombinant glycosyltransferases are bound to anti-Flag M1 affinity gel (Sigma).

a)

Peak	Epitope	Structure of glycan moiety	*m*/*z* glycopeptide	*m*/*z* alditol
1	β4Gal-Core 3	Galβ1–4GlcNAcβ1–3GalNAc–	2037.66	611.32
2	blood group H	Fucα1–2Galβ1–4GlcNAcβ1–3GalNAc–	2183.72	757.40
3	blood group B	Galα1 (3) Fucα1–2Galβ1–4GlcNAcβ1–3GalNAc–	2345.77	919.50
4	blood group A	GalNAcα1 (3) Fucα1–2Galβ1–4GlcNAcβ1–3GalNAc–	2386.80	960.53

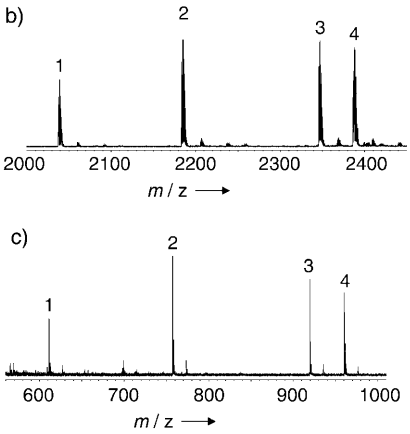

Figure 3. Conversion of a glycopeptide library to an oligosaccharide library. a) Structures and observed signals before and after β-elimination. b) MALDI-TOF mass spectrum of a glycopeptide library. c) MALDI-TOF mass spectrum of the released glycan from this glycopeptide library. The glycopeptides were observed as $[M+H]^+$ ions, the alditols as $[M+Na]^+$ ions.

The libraries described above have many uses. Herein, we demonstrate the screening of ligand specificity for lectins by using mass spectrometry. Ligand specificity for jacalin was determined quite easily by using the prepared library (Figure 4a). Jacalin covalently linked to agarose beads (0, 1, 2.5, and 5 μL) was added to library solutions. Bovine serum albumin (BSA)-conjugated agarose beads were used as the negative control. The solutions were incubated and centrifuged, and the glycopeptides remaining in the supernatants were identified by directly applying the supernatants with the matrix solution to the MS target plate. In this screening system, if the immobilized lectin has binding affinity for a

a)

Epitope	Structure	*m/z* $[M+H]^+$
Tn	GalNAcα1-peptide	1672.64
Core 1	Galβ1-3GalNAcα1-peptide	1834.72
Core 2	GlcNAcβ1 (6) Galβ1-3GalNAcα1-peptide	2037.85
Core 3	GlcNAcβ1-3GalNAcα1-peptide	1875.77
Core 4	GlcNAcβ1 (6) GlcNAcβ1-3GalNAcα1-peptide	2078.88

b)

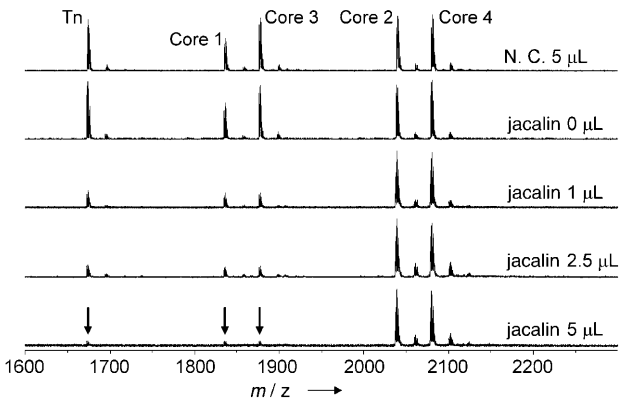

Figure 4. Screening of ligand specificity for jacalin by mass spectrometry. a) Components of the library used for screening. b) MALDI-TOF mass spectra of the library solution at various concentrations of jacalin. BSA covalently linked to agarose beads (5 μL) was used as the negative control (N.C.).

component of the library solution, the corresponding peak of this component becomes small or disappears on the mass spectrum. The intensities of three peaks corresponding to Tn, Core 1, and Core 3 structures in the mass spectra gradually decreased as the amount of added jacalin increased, whereas the intensity of two peaks, Core 2 and Core 4, did not decrease (Figure 4b). These results are consistent with the report that jacalin recognizes Tn and Core 1 residues,[8] and also include some additional findings: 1) jacalin binds to O-linked oligosaccharides, which are linearly elongated on the Gal residue of Core 1 and the GlcNAc residue of Core 3; 2) jacalin does not bind to *O*-glycans in which the O6 position of the GalNAc residue is occupied, that is, Core 2 and Core 4 structures.

In conclusion, we developed molecular-weight-tagged libraries thar were rapidly and readily prepared as mixtures, and the structure of each component in the libraries could be identified immediately by mass spectrometry alone. The libraries were useful for rapid and easy screening for ligand specificity of carbohydrate-binding proteins, such as lectins and antibodies. In the past few years, we have developed more than 160 human glycosyltransferases as recombinant enzymes.[5] It is possible to prepare libraries containing a large variety of natural oligosaccharide structures with these glycosyltransferases. We are increasing the variety of libraries as much as possible, based on different combinations of many enzymes and acceptor substrates, by carefully designing the reaction protocols, thus avoiding the complication attributable to components with the same molecular weight. In addition, this method will be applied to an automated system as a synthesizer for glycan and glycopeptide libraries. The libraries can also be used as carbohydrate microarrays in various fields of research, such as functional glycomics and drug discovery.

Received: March 4, 2005
Published online: June 27, 2005

Keywords: combinatorial chemistry · glycopeptides · high-throughput screening · mass spectrometry · transferases

[1] a) A. Varki, *Glycobiology* **1993**, *3*, 97–130; b) G. E. Ritchie, B. E. Moffatt, R. B. Sim, B. P. Morgan, R. A. Dwek, P. M. Rudd, *Chem. Rev.* **2002**, *102*, 305–320; c) A. Dove, *Nat. Biotechnol.* **2003**, *19*, 913–917.

[2] A. Kameyama, N. Kikuchi, S. Nakaya, H. Ito, T. Sato, T. Shikanai, Y. Takahashi, T. Takahashi, H. Narimatsu, *Anal. Chem.* **2005**, *77*, in press.

[3] a) T. Feizi, F. Fazio, W. Chai, C.-H. Wong, *Curr. Opin. Struct. Biol.* **2003**, *13*, 637–645; b) S. Park, I. Shin, *Angew. Chem.* **2002**, *114*, 3312–3314; *Angew. Chem. Int. Ed.* **2002**, *41*, 3180–3182; c) C. O. Mellet, J. M. Garcia Fernández, *ChemBioChem* **2002**, *3*, 819–822; d) D. Wang, *Proteomics* **2003**, *3*, 2167–2175; e) I. Shin, S. Park, M.-R. Lee, *Chem. Eur. J.* **2005**, *11*, 2894–2901; f) M. F. Templin, D. Stoll, M. Schrenk, P. C. Traub, C. F. Vöhringer, T. O. Joos, *Trends Biotechnol.* **2002**, *20*, 160–166.

[4] a) P. M. St. Hilaire, M. Meldal, *Angew. Chem.* **2000**, *112*, 1210–1228; *Angew. Chem. Int. Ed.* **2000**, *39*, 1162–1179; b) O. J. Plante, E. R. Palmacci, P. H. Seeberger, *Adv. Carbohydr. Chem. Biochem.* **2003**, *58*, 35–54; c) V. Wittmann, S. Seeberger, *Angew. Chem.* **2004**, *116*, 918–921; *Angew. Chem. Int. Ed.* **2004**, *43*, 900–903.

[5] a) *Handbook of Glycosyltransferases and Related Genes* (Eds.: N. Taniguchi, K. Honke, M. Fukuda), Springer, Tokyo, **2001**; b) H. Narimatsu, *Glycoconjugate J.* **2004**, *21*, 17–24.

[6] B. L. Schulz, N. H. Packer, N. G. Karlsson, *Anal. Chem.* **2002**, *74*, 6088–6097.

[7] N. Kikuchi, A. Kameyama, S. Nakaya, H. Ito, T. Sato, T. Shikanai, Y. Takahashi, H. Narimatsu, *Bioinformatics* **2005**, *21*, 1717–1718.

[8] a) M. V. Sastry, P. Banarjee, S. R. Patanjali, M. J. Swamy, G. V. Swarnalatha, A. Surolia, *J. Biol. Chem.* **1986**, *261*, 11726–11733; b) A. M. Wu, J. H. Wu, L.-H. Lin, S.-H. Lin, J.-H. Liu, *Life Sci.* **2003**, *72*, 2285–2302.

