Available online at www.sciencedirect.com
SCIENCE DIRECT®
Bioorganic & Medicinal Chemistry Letters 14 (2004) 2835–2838
Bioorganic & Medicinal Chemistry Letters

A facile route to dynamic glycopeptide libraries based on disulfide-linked sugar–peptide coupling

Shinsuke Sando,* Atsushi Narita and Yasuhiro Aoyama*

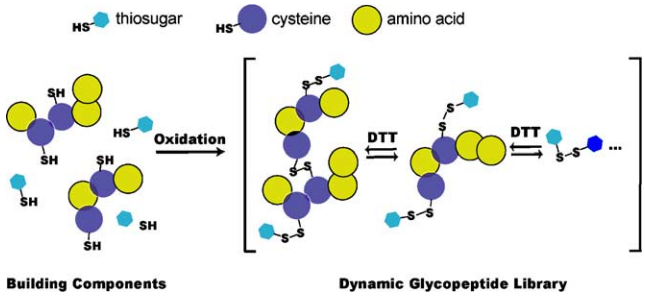
Department of Synthetic Chemistry and Biological Chemistry, Graduate School of Engineering, Kyoto University, Katsura, Nishikyo-Ku, Kyoto 615-8510, Japan

Received 22 January 2004; revised 17 March 2004; accepted 18 March 2004

Abstract—We report here that disulfide-linked dynamic glycopeptide libraries can be constructed from 1-thiosugar and cysteine-rich oligopeptide building blocks upon gentle air oxidation of a slightly basic (pH 7.8) aqueous solution thereof. A mixture of 1-thiogalactose and two oligopeptides H_2N–CysGlyCysGly–CO_2H and H_2N–GlyCycCysGlyGly–CO_2H, for example, affords a poorly HPLC-resolved disulfide library composed of various sugar–peptide conjugates and cyclic peptides, at least 10 of which can be identified by ESI mass spectrometry. The building components of disulfide members are exchangeable with each other in the presence of dithiothreitol as an initiator to allow dynamic equilibration. A preliminary SPR examination shows that the thiogalactose-derived library indeed contains active divalent galactoside species capable of cross-linking peanut lectin molecules.

Carbohydrates are the third class of informational biomolecules playing central roles in various biological events such as signal transduction, cell adhesion, and bacterial/viral infection.[1,2] There has been much current interest in the pharmaceutical application of carbohydrates and their mimics as cell-targeting agents, anti-infectious drugs, and so on.[3] Cell-surface oligosaccharides responsible for recognition events are often branched and exist as oligomeric forms as assembled on a protein scaffold (glycopeptides) so as to present a particular spatial arrangement of terminal saccharide moieties, which are cooperatively recognized in a multivalent fashion by relevant receptors. Such multivalent interactions enhance the three-dimensional sequence diversity of carbohydrates and strengthen otherwise weak carbohydrate–protein or carbohydrate–carbohydrate interactions.[4] A typical example is the three-point interaction between asialoglycoproteins (ASGP) and their receptors on the hepatic cell surfaces.[5] ASGP receptors contain three galactose-binding pockets and recognize spatially matched trivalent galactose clusters with inter-saccharide distances of 15, 20, and 25 Å more than 10,000-times more strongly than monovalent galactose ligands.[5a]

While the recent progress in automated synthetic methodology has indeed provided an easier access to complex carbohydrates and glycopeptides,[6] it is still a tough task to rationally design active glycocluster motifs having a proper spatial arrangement. A potential alternative would be combinatorial library approach.[7] Particularly intriguing is the so-called dynamic combinatorial method (Scheme 1),[8,9] where small building blocks are linked together via reversible and hence interchangeable bonds such as disulfide,[9] so that the desired target-binding species could be automatically



Scheme 1. Schematic illustration of how dynamic glycopeptide library is constructed out of cysteine-rich oligopeptide and thiosugar building blocks.

Keywords: Glycopeptide; Carbohydrate; Dynamic combinatorial library; Disulfide-linking chemistry.
* Corresponding authors. Tel.: +81-75-383-2766; fax: +81-75-383-2767; e-mail addresses: ssando@sbchem.kyoto-u.ac.jp; aoyamay@sbchem.kyoto-u.ac.jp

0960-894X/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmcl.2004.03.050

amplified as a result of shift in equilibria. Ramström and Lehn[9b] employed this disulfide method to get a library of dimers of alkyl glycoside starting with thioalkyl glycosides. For the diversity of library members, the use of oligomerizable, that is, at least difunctional, spacers would be essential. If branching is to be introduced, spacer must be at least three-way connectable, that is, trifunctional. In the present work, we use cysteine-rich peptides as spacers and 1-thiosugars as end-caps. There would be no particular difficulty in preparing various peptides having different numbers of cysteine residues with designed spacings. Thiosugars with enhanced anomer stabilities are also readily available.[10] S-glycosides derived therefrom usually remain to be good substrates of sugar-binding proteins and, more importantly, are much stabler against glycosidase attack than the corresponding O-glycosides.[10] We report here that dynamic glycopeptide libraries can be easily constructed in this way.

Disulfide bridges (–S–S–) play an essential role in the folding and assembly of proteins. The glycopeptide libraries here are based on reversible S–S bond formation from commercially available sugar monothiols (β-D-1-thioglucose (**1**) and β-D-1-thiogalactose (**2**)) and oligopeptide dithiols (CysGlyCysGly (**3**) and GlyCysCysGlyGly (**4**)) (Fig. 1). We first evaluated whether the present 1-thiosugar could end-cap the terminal SH groups of growing peptide cores. For this purpose, we used 2,3-dimercapto-1-propanesulfonic acid (**5**) as a simple alkanedithiol.[9a] Equal amounts of thioglucose **1** (10 mM) and dithiol **5** (10 mM) were mixed in a slightly basic aqueous buffer (10 mM TEAA, pH 7.8) and the mixture was stirred in an open vial at room temperature for ~3 days, during which the thiols were nearly completely air oxidized to disulfide species ($-SH + -SH + 1/2O_2 \rightarrow -S-S- + H_2O$) and the solution became negative to Ellman reagent. The product distribution was analyzed by electrospray ionization (ESI) mass spectrometry. The spectrum (Fig. 2, showing only the high molecular-weight region) shows at least eight major species in addition to trace amounts of starting monomers **1** ($m/z = 196.0$) and **5** (187.9). The new species with $m/z = 370.8$, 381.2, 389.1, 556.9, 575.0, 742.8, 761.0 (783.0, $[M-2H+Na]^-$), and 947.0 (969.0, $[M-2H+Na]^-$) were identified as homo- and heterodimers *c*(**5**–**5**) (*c* = cyclic), **1**–**5**, and **1**–**1**, trimers *c*(**5**–**5**–**5**) and **1**–**5**–**1**, tetramers *c*(**5**–**5**–**5**–**5**) and **1**–**5**–**5**–**1**, and pentamer **1**–**5**–**5**–**5**–**1**.

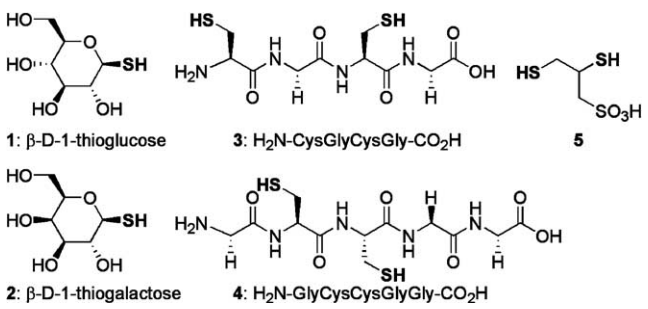


Figure 1. Chemical structures of building blocks used for the generation of dynamic glycopeptide library.

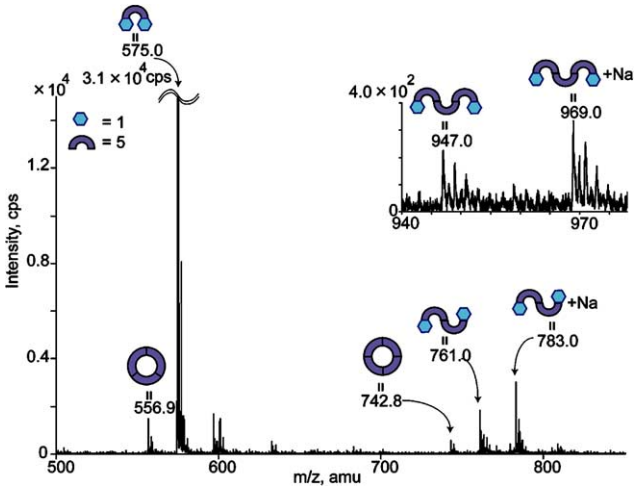


Figure 2. High molecular-weight region of negative ion QQQ ESI mass spectrum of the mixture of 1-thioglucose **1** (10 mM) and dithiol **5** (10 mM).

Reverse-phase HPLC trace of the mixture, on the other hand, showed two major slow-moving species in addition to a block of residual library members co-eluted in a low-retention region.[11] The two high-retention members were isolated and assigned, on the basis of mass and ^{1}H NMR spectra, as thioglucose dimers with or without intervening spacer unit, that is, **1**–**5**–**1** and **1**–**1** with **1**-based yields of ~45% and ~35%, respectively. The stereochemistry of thioglycosides was 85% β and 15% α, as revealed by NMR for the whole library. Anyway, sugar thiol **1** has a reactivity comparable to that of alkanedithiol **5** and thus easily end-caps the growing oligomeric spacer dithiols in competition with self-cyclization of the latter to give a library composed of homooligomers (**1**–**1**, *c*(**5**–**5**), *c*(**5**–**5**–**5**), and *c*(**5**–**5**–**5**–**5**), and heterooligomers (**1**–**5**–**1**, **1**–**5**–**5**–**1**, and **1**–**5**–**5**–**5**–**1**) under remarkably mild conditions, that is, gentle stirring of a slightly basic aqueous solution of monomeric building blocks at ambient temperature.

We then moved on the use of peptide spacers **3** and **4**. Gentle air oxidation of an aqueous solution (pH 7.8) of thiogalactose **2** (5 mM) and dithiol **3** (5 mM) at room temperature for 24 h afforded a poorly HPLC-resolved and hence difficult-to-separate disulfide mixture containing sugar dimer **2**–**2** ($m/z = 389.2$), peptide dimer in a cyclic form *c*(**3**–**3**) (671.3), and glycopeptides **2**–**3**–**2** (724.3) and **2**–**3**–**3**–**2** (1060.9) as identified by ESI mass spectrometry as above. The diversity could be enhanced by using two different peptide dithiols **3** and **4** in equal amounts. The number of constituent library members, even when limited to major ones, exceeded 10, including cross-coupling adduct **2**–**3**–**4**–**2** (1118.2, $[M-H]^-$; 1140.2, $[M-2H+Na]^-$) together with those derived from a single spacer such as **2**–**3**–**2** (725.1), **2**–**4**–**2** (782.2), and **2**–**3**–**3**–**2** (1061.2) and at least six cyclic peptides (Fig. 3, showing only the high molecular-weight region).

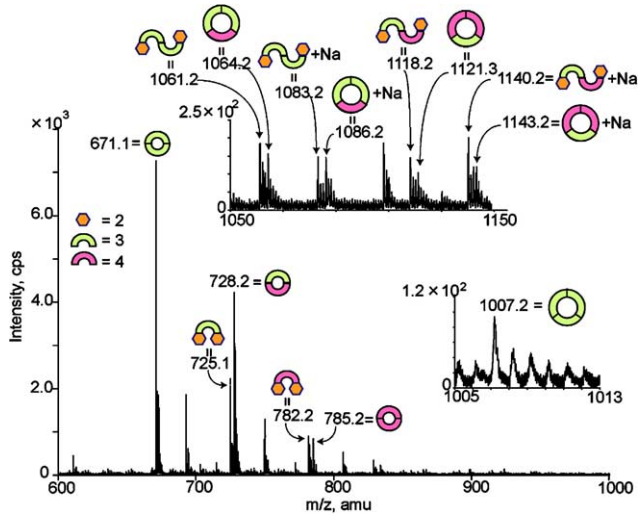


Figure 3. High molecular-weight region of negative ion QQQ ESI mass spectrum of the mixture of 1-thiogalactose **2** (1 mM) and oligopeptides **3** and **4** (1 mM).

A fascinating aspect of disulfide libraries is the dynamic nature thereof. In the presence of an appropriate disulfide exchange initiator such as dithiothreitol (DTT), a disulfide mixture undergoes shuffling. In ideal cases, the whole library could be converged into particular target/guest-binding form(s) due to complexation-induced shift in equilibria. Sugar thiol **2** (2 mM) and peptide dithiol **3** (2 mM) were separately preoxidized to give sugar dimer **2**–**2** ($m/z = 389.5$) and cyclic peptide *c*(**3**–**3**) (671.9), respectively. They were mixed in a slightly basic buffer (10 mM TEAA, pH 7.8) as above and the mixture stirred for 24 h. Nothing particular happened and the ESI mass spectrum of the mixture remained essentially the same as the sum of components **2**–**2** (389.5) and *c*(**3**–**3**) (671.9) (Fig. 4a). In the presence of DTT (0.1 mol%), however, disulfide exchange readily took place to give an 'equilibrium' mixture of **2**–**2**, *c*(**3**–**3**), and glycopeptide **2**–**3**–**2** (725.7) (Fig. 4b).

We also carried out a preliminary SPR (surface plasmon resonance) examination on the interaction with peanut lectin (PNL). PNL composed of four identical subunits is a galactoside-specific lectin (preferably for the β-isomer) and is cross-linked upon interaction with multivalent galactoside derivatives. A PNL-immobilized[12] SPR sensor chip was treated first with a **2**-derived library (**2**–**5**, **2**–**3**, **2**–**4**, or **2**–**3**–**4**) and then with PNL to result in well-behaved reversible association/dissociation of the latter, as shown in Figure 5 (blue line) for the case of pretreatment with library **2**–**5**. No association of PNL was observed without pretreatment (red line). These results clearly indicate that the library derived from β-thiogalactose contains active divalent galactoside mediator(s) to link PNL in bulk solution onto that immobilized on the chip.[13] The complexation in solution is also reversible and the isolation of sugar-complexed PNL to allow identification of the active β-galactoside species has so far been unsuccessful.

In this work, we studied disulfide-linked sugar–peptide coupling to provide a simple route to glycopeptides. A

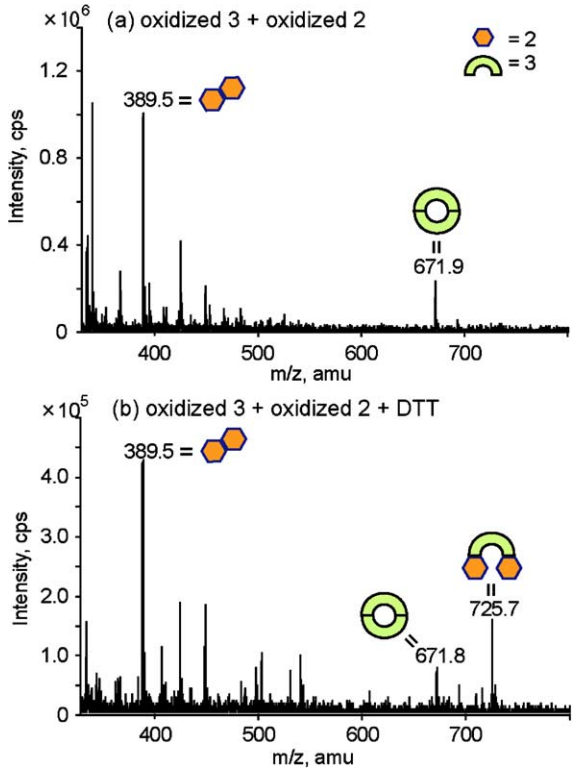


Figure 4. Negative ion QQQ ESI mass spectra of the mixture of pre-oxidized thiogalactose **2** (**2**–**2**) and pre-oxidized oligopeptide **3** (*c*(**3**-**3**)) in the (a) absence or (b) presence of disulfide exchange initiator DTT.

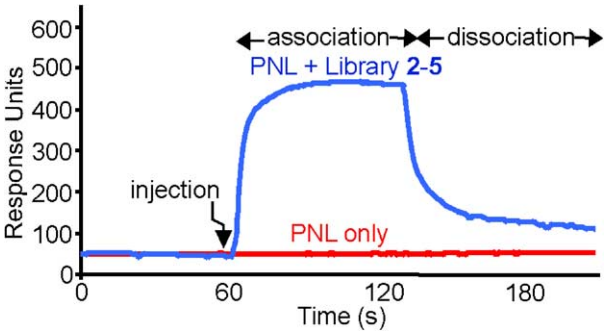


Figure 5. SPR resonance curves for the association/dissociation of PNL (20 μM, 10 μL) on a PNL-immobilized sensor chip with (blue line) or without (red line) pretreatment with library **2**–**5** (20 μM each of **2** and **5**, 30 μL) in the flow of DULBECCO'S phosphate buffer at a flow rate of 10 μL/min.

particular intention was to develop a one-pot, bottom-up method to construct glycopeptide libraries with a hope to expand the scope of glycobiology and glycopharmacology. Preliminary results obtained here are summarized as follows: (1) The reactivities of the glycosidic and peptide (cysteine) SH groups match with each other to give a balance in sugar–sugar, peptide–peptide, and sugar–peptide S–S bond formation, giving rise to a rich disulfide library, which is poorly HPLC-resolved but analyzed by means of ESI mass spectrometry. (2) The terminal SH groups in growing peptide oligomers are either sugar capped to give glycopeptides or self-cyclized into sugar-free cyclic peptides. The latter

process thus serves as a major competitor of the former. (3) The β-stereochemistry of starting thiosugars is 85% retained in the S–S bond formation, thus readily leading to divalent β,β-galactoside(s) as PNL connectors. The remaining 15%, on the other hand, undergoes inversion to α, thus further increasing the diversity of the libraries. (4) The disulfide libraries thus obtained are dynamic and undergo equilibration in the presence of DTT.

The basic requirements for glycopeptide dynamic combinatorial libraries are thus met. The size of a library depends on the diversity of available building blocks. As for sugars, any saccharides, mono- or oligo-, having a reducing terminus can be building blocks, since the anomeric OH-to-SH conversion is rather easy to carry out.[10] As for peptides, advantage is even more pronounced. A variety of oligopeptides with a rich sequence diversity are readily prepared using a solid-phase technique and even in a combinatorial manner. Referring to item 2 above, the present method may also be used to generate cyclic peptide libraries simply by using various oligopeptides containing two cysteine residues. The significance of this should not be overlooked. Cyclic peptides are promising pharmaceutical agents. There are a number of examples of cyclic peptide antibiotics.[14] Further work is now under way along these two lines.

Acknowledgements

We thank Profs. T. Imanaka, I. Saito, and K. Nakatani of Kyoto University for ESI mass measurements. This work was supported by Grant-in-Aid for Scientific Research on Priority Areas (C) 'Genome Science' and 21st COE on Kyoto University from the Japanese Government. This work was also supported by RFTF of JSPS.

References and notes

1. (a) Hakomori, S. *J. Biol. Chem.* **1990**, *256*, 18713; (b) Hakomori, S. *Pure Appl. Chem.* **1991**, *63*, 473.
2. For recent reviews, see: (a) Bertozzi, C. R.; Kiessling, L. L. *Science* **2001**, *291*, 2357; (b) Roth, J. *Chem. Rev.* **2002**, *102*, 285; (c) Ritchie, G. E.; Moffat, B. E.; Sim, R. B.; Morgan, B. P.; Dwek, R. A.; Rudd, P. M. *Chem. Rev.* **2002**, *102*, 305; (d) Zachara, N. E.; Hart, G. W. *Chem. Rev.* **2002**, *102*, 431.
3. Ritter, T. K.; Wong, C.-H. *Angew. Chem., Int. Ed.* **2001**, *40*, 3508.
4. (a) Kiessling, L. L.; Pohl, N. L. *Chem. Biol.* **1996**, *3*, 71; (b) Roy, R. *Trends Glycosci. Glycotechnol.* **1996**, *8*, 79; (c) Mammen, M.; Choi, S.-K.; Whitesides, G. M. *Angew. Chem., Int. Ed.* **1998**, *37*, 2754; (d) Lundquist, J. J.; Toone, E. J. *Chem. Rev.* **2002**, *102*, 555.
5. (a) Lee, R. T.; Lee, Y. C. *Glycoconj. J.* **2000**, *17*, 543; (b) Liljeblad, M.; Lundblad, A.; Pahlsson, P. *Biosens. Bioelectron.* **2002**, *17*, 883.
6. For recent reports on carbohydrate synthetic methodology, see for example: (a) Nishimura, S.-I. *Curr. Opin. Chem. Biol.* **2001**, *5*, 325; (b) Ritter, T. K.; Mong, K.-K. T.; Liu, H.; Nakatani, T.; Wong, C.-H. *Angew. Chem., Int. Ed.* **2003**, *42*, 4657; (c) Seeberger, P. H. *Chem. Commun.* **2003**, 1115.
7. (a) Arya, P.; Ben, R. N. *Angew. Chem., Int. Ed.* **1997**, *36*, 1280; (b) Schweizer, F.; Hindsgaul, O. *Curr. Opin. Chem. Biol.* **1999**, *2*, 291.
8. For recent reports on dynamic combinatorial chemistry (DCC), see for example: (a) Sakai, S.; Shigemasa, Y.; Sasaki, T. *Tetrahedron Lett.* **1997**, *38*, 8145; (b) Klekota, B.; Miller, B. L. *Trends Biotechnol.* **1999**, *17*, 205; (c) Cousins, G. R. L.; Poulsen, S.-A.; Sanders, J. K. M. *Curr. Opin. Chem. Biol.* **2000**, *4*, 270; (d) Karan, C.; Miller, B. L. *J. Am. Chem. Soc.* **2001**, *123*, 7455; (e) Lehn, J.-M.; Eliseev, A. V. *Science* **2001**, *291*, 2331; (f) Rowan, S. J.; Cantrill, S. J.; Cousins, G. R. L.; Sanders, J. K. M.; Stoddart, J. F. *Angew. Chem., Int. Ed.* **2002**, *41*, 898; (g) Otto, S.; Furlan, R. L. E.; Sanders, J. K. M. *Curr. Opin. Chem. Biol.* **2002**, *6*, 321; (h) Otto, S.; Furlan, R. L. E.; Sanders, J. K. M. *Drug Discovery Today* **2002**, *7*, 117.
9. For recent reports on DCC using disulfide-linking chemistry, see for example: (a) Otto, S.; Furlan, R. L. E.; Sander, J. K. M. *J. Am. Chem. Soc.* **2000**, *122*, 12063; (b) Ramström, O.; Lehn, J. M. *ChemBioChem.* **2000**, *1*, 41; (c) Otto, S.; Furlan, R. L. E.; Sanders, J. K. M. *Science* **2002**, *297*, 590; (d) Otto, S.; Kubik, S. *J. Am. Chem. Soc.* **2003**, *125*, 7804; (e) Brisig, B.; Sanders, J. K. M.; Otto, S. *Angew. Chem., Int. Ed.* **2003**, *42*, 1270.
10. (a) Driguez, H. *ChemBioChem.* **2001**, *2*, 311; (b) Zhu, X.; Schmidt, R. R. *Chem. Eur. J.* **2004**, *10*, 875.
11. Library distributions were determined by measuring the absorbance of disulfide –S–S– group at 250 nm.
12. PNL was immobilized on a carboxy-functionalized SPR chip upon coupling (taking advantage of free amino groups of the protein) using an amine-coupling kit.
13. Libraries derived from β-thioglucose (**1**) are hardly effective in cross-linking concanavalin A, which is specific to α-glucosides (and mannosides). The glycosidic stereochemistry in a **1**-derived library is 85% β and 15% α (vide supra) and the probability of active α,α-species responsible for cross-linking is expected to be ca. 2%.
14. (a) Kohli, R. M.; Walsh, C. T.; Burkart, M. D. *Nature* **2002**, *418*, 658; (b) Pfeiffer, B.; Peduzzi, E.; Moehle, K.; Zurbriggen, R.; Glück, R.; Pluschke, G.; Robinson, J. A. *Angew. Chem., Int. Ed.* **2003**, *42*, 2368.