Temporary attachment of carbohydrates to cyclopeptide templates: a new strategy for single-bead analysis of multivalent neoglycopeptides

Valentin Wittmann,[a,b,]* Sonja Seeberger[a] and Hermann Schägger[c]

[a]*Institut für Organische Chemie und Chemische Biologie, Goethe-Universität, Marie-Curie-Straße 11, 60439 Frankfurt, Germany*
[b]*Fachbereich Chemie, Universität Konstanz, 78457 Konstanz, Germany*
[c]*Zentrum der Biologischen Chemie, Klinikum der Goethe-Universität, Theodor-Stern-Kai 7, 60596 Frankfurt, Germany*

Abstract—Employing a cleavable carbohydrate–peptide linker, a new strategy for single-bead analysis of multivalent cyclic neoglycopeptides based on Edman degradation is described. Edman degradation of glycopeptides is hampered by the acid lability of the glycosidic bond and potential incompatibilities of phenylthiohydantoin (PTH) derivatives of glycosylated amino acids with PTH derivatives of the proteinogenic amino acids. To overcome this problem, carbohydrates are detached from the cyclopeptide templates before single-bead analysis, allowing for micro sequencing under routine conditions. With this strategy, application of multivalent cyclic neoglycopeptides in split-mix libraries with a subsequent screening process becomes possible for the first time. © 2003 Elsevier Ltd. All rights reserved.

Conformationally restricted cyclic peptides are frequently employed as templates for the spatially defined presentation of pharmacophores[1] or whole peptide domains.[2] Cyclic peptides carrying several copies of a carbohydrate ligand are of great interest as potential multivalent lectin ligands.[3,4] In particular, the influence of the spatial presentation of carbohydrates on lectin mediated signal transduction processes is subject of current research in the area of glycobiology.[5] The screening of one-bead one-compound libraries[6] of cyclic glycopeptides obtained by the split-mix protocol,[7] however, has not yet been described.

Recently, we reported the efficient convergent synthesis of split-mix libraries of cyclic neoglycopeptides of type **1**.[4] Here we present a strategy for the direct analysis of these compounds bound to single resin beads. Together, these techniques make it possible for split-mix libraries of type **1** to be applied in a screening process.[8]

Side-chain cyclized neoglycopeptide **5** (Scheme 1) which after deacetylation is a potential trivalent lectin ligand served as a model compound for our analytical strat-
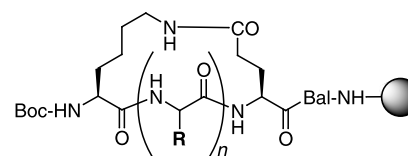
1:   R = amino acid side chain or
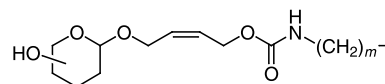
egy. It was obtained by convergent solid-phase synthesis[4] on aminoethyl TentaGel via intermediates **2** and **3**. The carbohydrates were attached to side-chain amino groups of **3** via urethane formation by treatment with active carbonate **4** in the presence of Hünig's base (DIEA). Quantitative derivatization was ensured by monitoring the coupling progress with bromophenol blue[9] after having thoroughly washed the resin with *N,N*-dimethylformamide (DMF).

Identification of active members of one-bead one-compound libraries is based on the analysis of the minute amounts of substances bound to single resin beads (typically <200 pmol).[10] Direct analysis of synthetic products is generally advantageous over indirect coding strategies because failed/side reactions can be moni-

tored. Whereas automated single-bead Edman sequencing of peptides is well established,[11] two problems arise during the analysis of glycopeptides. On one hand, for each glycosylated amino acid the corresponding phenylthiohydantoin (PTH) derivative has to be synthesized and its compatibility with the other PTH amino acids verified. On the other hand, the glycosidic bond is not stable under the strongly acidic conditions which leads to several PTH derivatives and, therefore, complicates the assignment and reduces the sensitivity.[12] The encoded ladder synthesis reported by Meldal et al.[13] to analyze libraries of linear glycopeptides is the only published method for the single-bead analysis of glycopeptides so far. However, this method is not suited for cyclic glycopeptides because the conformationally restricted target compounds would be 'diluted' by substantial amounts of linear fragments.

Our strategy for single-bead analysis of multivalent cyclic neoglycopeptides is based on the use of a cleavable linker between the carbohydrates and the cyclopeptide scaffold. With this setup it is possible to remove the carbohydrates after a screening process and sequencing is confined to the amino acids used for attachment of the sugars the

quantitative derivatization of which can be secured as described above. In compound **5** the temporary attachment is realized by the but-2-en-1-yloxycarbonyl linker, representing a modified Aloc protecting group. Treatment of **5** with [Pd(PPh$_3$)$_4$] and morpholine leads back to non-modified **3**.

As a reference, micro sequencing was first carried out with **3** before attachment of the sugars. To this end the N-terminus of **3** was deprotected with TFA/CH$_2$Cl$_2$ (Scheme 2). Subsequently, a single resin bead of **6** was analyzed with an ABI 473A protein sequencer. Figure 1A shows the HPLC diagrams of each degradation step. As expected, no peak characteristic for Lys was found in the first step since the PTH derivative of Lys is connected to the solid support via its side chain (**8**). Only in the case of an incomplete cyclization during library synthesis (**2**→**3**), a signal is expected in this step, representing a cyclization control. Also incomplete deprotection of the Aloc group in **2** as a reason for a failed cyclization is detectable as could be shown by micro sequencing of linear **2** after Boc deprotection (retention time of PTH-Lys(Aloc): 6.0 min, cf. Fig. 1B). Further sequencing of **8** (Fig. 1A, steps 2–7) then unambiguously gave the sequence of **3**.

In comparison, Figure 1C shows the result of micro sequencing of **5** after deprotectin with TFA/CH$_2$Cl$_2$. Clearly visible are each multiple peaks for the carbohydrate-modified amino acids D-Dab, Dpr, and D-Lys in step 3, 5, and 6, respectively, and at the same time complete absence of the non-modified diamino acids.

Scheme 1. Temporary attachment of GlcNAc residues to cyclic peptide **3** (All = allyl, Aloc = allyloxycarbonyl, Bal = β-alanine, Dab = 2,4-diaminobutyric acid, Ddv = 1-(4,4-dimethyl-2,6-dioxocyclohexylidene)isovaleryl, DIEA = diisopropylethylamine, Dpr = 2,3-diaminopropionic acid, HBTU = O-benzotriazol-1-yl-N,N,N′,N′-tetramethyluronium hexafluorophosphate, HOBt = 1-hydroxybenzotriazole).

Scheme 2. Edman degradation of side-chain cyclized peptide **3** (But = butyryl).
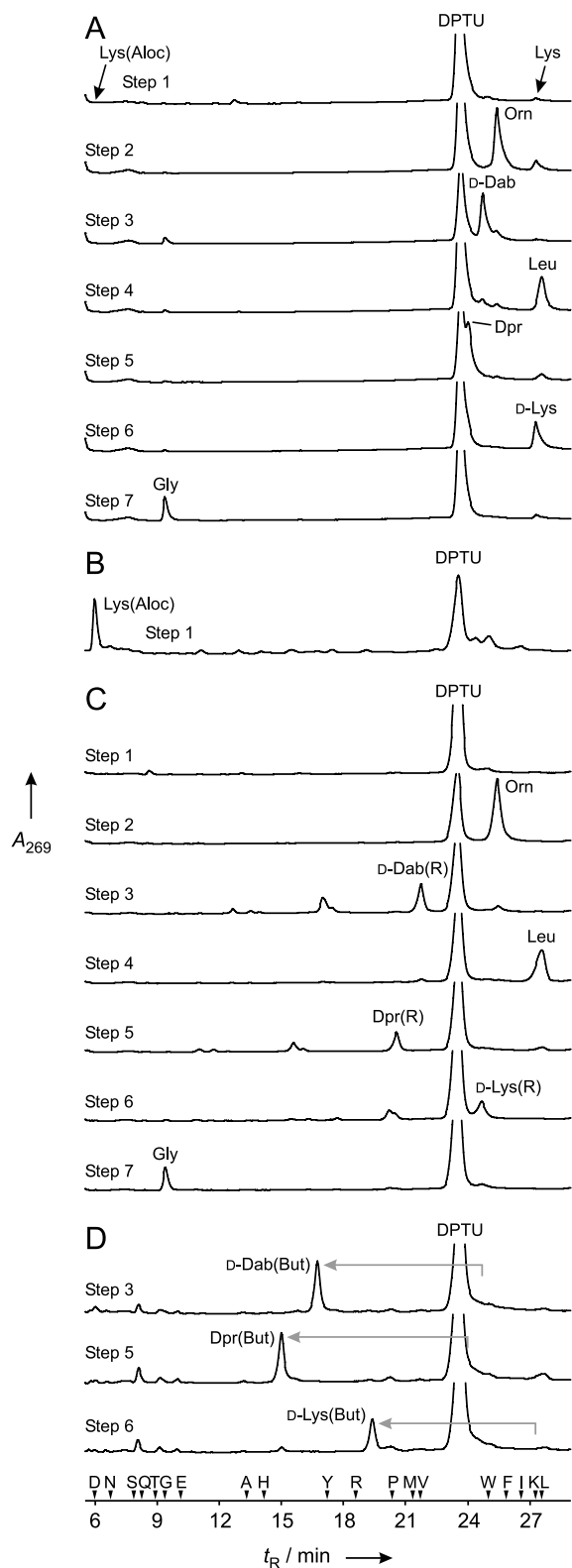
Figure 1. HPLC diagrams of the degradation steps of single-bead micro sequencing of (A) **6** (obtained by the reaction sequence **2**→**3**→**6**), (B) **2** (after treatment with TFA/CH$_2$Cl$_2$ 1:1), (C) **5** (after treatment with TFA/CH$_2$Cl$_2$ 1:1), and (D) **7**. On top of the time axis the retention times of the PTH derivatives of the proteinogenic amino acids are given.

After removal of the carbohydrates in **5** with [Pd(PPh$_3$)$_4$] and morpholine, the sequencing result of **3** thus obtained was virtually identical to the result shown in Figure 1A. This demonstrates that both the attach-ment and the removal of the carbohydrates proceeded completely.

Whereas the retention times of PTH-D-Dab and PTH-Dpr are clearly distinguishable from those of the PTH derivatives of the natural amino acids, the PTH-Dpr signal overlaps with that of diphenylthiourea (DPTU) making its detection in small amounts difficult. Separation of both peaks by varying the HPLC gradient was not possible without overlap of other amino acids. Therefore, **3** was treated with different capping reagents before sequencing in order to optimize the retention time of PTH-Dpr (and thus also of PTH-D-Dab and PTH-D-Lys). The best result was obtained with butyric acid anhydride (But$_2$O) (**3**→**7**, Scheme 2) leading to a shift of the PTH diamino acid signals in the less crowded region between 15 and 20 min (Fig. 1D). With this modification single-bead sequencing of the peptide is easily accomplished.[14]

In summary, we have described a new strategy for single-bead analysis of multivalent cyclic neoglycopeptides which allows application of these compounds in split-mix libraries with subsequent screening process for the first time. The use of a cleavable carbohydrate–peptide linker enables Edman sequencing under routine conditions regardless of the used carbohydrate derivative. In contrast to coding strategies, Edman sequencing allows control of the critical cyclization step as well as detection of failed couplings. In addition, the occurrence of linear peptide fragments is avoided. The concept of temporary attachment is not confined to carbohydrates but can be applied to any other amino-acid modification.

Acknowledgements

References

1. (a) Hruby, V. J. *Nat. Rev. Drug Discov.* **2002**, *1*, 847–858; (b) Senderowitz, H.; Rosenfeld, R. *J. Receptor Signal Transduction Res.* **2001**, *21*, 489–506; (c) Haubner, R.; Finsinger, D.; Kessler, H. *Angew. Chem.* **1997**, *109*, 1440–1456; *Angew. Chem. Int. Ed. Engl.* **1997**, *36*, 1374–1389; (d) Eichler, J.; Lucka, A. W.; Houghten, R. A. *Peptide Res.* **1994**, *7*, 300–307; (e) Rizo, J.; Gierasch, L. M. *Ann. Rev. Biochem.* **1992**, *61*, 387–418.

2. Tuchscherer, G.; Mutter, M. *Pure Appl. Chem.* **1996**, *68*, 2153–2162.

3. (a) Sprengard, U.; Schudok, M.; Schmidt, W.; Kret-zschmar, G.; Kunz, H. *Angew. Chem.* **1996**, *108*, 359–362; *Angew. Chem. Int. Ed. Engl.* **1996**, *35*, 321–324; (b) Franzyk, H.; Christensen, M. K.; Jørgensen, R. M.; Meldal, M.; Cordes, H.; Mouritsen, S.; Bock, K. *Bioorg. Med. Chem.* **1997**, *5*, 21–40.

4. Wittmann, V.; Seeberger, S. *Angew. Chem.* **2000**, *112*, 4508–4512; *Angew. Chem., Int. Ed.* **2000**, *39*, 4348–4352.

5. (a) Bertozzi, C. R.; Kiessling, L. L. *Science* **2001**, *291*, 2357–2364; (b) Lundquist, J. J.; Toone, E. J. *Chem. Rev.* **2002**, *102*, 555–578; (c) Brewer, C. F.; Miceli, M. C.; Baum, L. G. *Curr. Opin. Struct. Biol.* **2002**, *12*, 616–623.

6. Lam, K. S.; Lebl, M.; Krchňák, V. *Chem. Rev.* **1997**, *97*, 411–448.

7. Furka, Á.; Sebestyén, F.; Asgedom, M.; Dibó, G. *Int. J. Pept. Protein Res.* **1991**, *37*, 487–493.

8. Wittmann, V.; Seeberger, S. *Angew. Chem.* **2003**, *115*; *Angew. Chem., Int. Ed.* **2003**, *42*, in press.

9. Krchňák, V.; Vágner, J.; Šafář, P.; Lebl, M. *Collect. Czech. Chem. Commun.* **1988**, *53*, 2542–2548.

10. (a) Yan, B. *Curr. Opin. Chem. Biol.* **2002**, *6*, 328–332; (b) Affleck, R. L. *Curr. Opin. Chem. Biol.* **2001**, *5*, 257–263.

11. Lam, K. S.; Salmon, S. E.; Hersh, E. M.; Hruby, V. J.; Kazmeiersky, W. M.; Knapp, R. J. *Nature* **1991**, *354*, 82–84.

12. This can be used to identify glycosylation sites of proteins of known sequence. See: Zeng, R.; Xu, Q.; Shao, X.-X.; Wang, K.-Y.; Xia, Q.-C. *Eur. J. Biochem.* **1999**, *266*, 352–358.

13. (a) St. Hilaire, P. M.; Lowary, T. L.; Meldal, M.; Bock, K. *J. Am. Chem. Soc.* **1998**, *120*, 13312–13320; (b) Halkes, K. M.; St. Hilaire, P. M.; Crocker, P. R.; Meldal, M. *J. Comb. Chem.* **2003**, *5*, 18–27.

14. Solid-phase peptide synthesis was carried out on Tenta-Gel S-NH$_2$ (130 μm beads; Rapp-Polymere, Tübingen) without employing a linker. Detachment of 4-(glycosyl-oxy)but-2-en-1-yloxycarbonyl residues (**5**→**3**): In a micro glass reactor equipped with a G4 filter plate, a few beads of **5** were shaken with 100 μL of a solution of [Pd(PPh$_3$)$_4$] (11 mg) and morpholine (50 μL) in DMF/DMSO 1:1 (2 mL) for 15 h under exclusion of light and air. The beads were washed with DMF (5×), 0.5 % sodium diethyldithio-carbamate in DMF (2×), DMF (5×), and CH$_2$Cl$_2$ (3×). **7**: The beads of **3** obtained were treated with But$_2$O/pyri-dine 2:3 (2×15 min, each 200 μL) and washed with DMF (10×), and CH$_2$Cl$_2$ (3×). Boc groups were removed with TFA/CH$_2$Cl$_2$ 1:1 (2×15 min, each 200 μL) and the beads were thoroughly washed with CH$_2$Cl$_2$ and dried in vac-uum. Micro sequencing: A single bead of **7** was placed between two PVDF membranes (φ 25 mm) using a stereo microscope (Olympus SZ4045TR) and sequenced with an ABI 473A protein sequencer according to the manufac-turer's instructions. The HPLC gradient was slightly modified: 0 min: 10% B, 18 min: 36% B, 27 min: 36% B, 27.1 min: 90% B. Eluent A: 5% THF in H$_2$O with 18 mL buffer solution (prepared from 100 mL 3 M NaOAc buffer, pH 3.8, and 23.5 mL 3 M NaOAc buffer, pH 4.6) per liter. Eluent B: acetonitrile. Flow: 0.21 mL min^{-1}.