Pergamon

Tetrahedron Letters 39 (1998) 1815–1818

TETRAHEDRON LETTERS

C-linked Glycosyl Azido Acid in Novel Solid-Phase *C*-Glycopeptide Synthesis.

Ulf Tedebark[1], Morten Meldal[1*], Luigi Panza[2] and Klaus Bock[1]

[1]Carlsberg Laboratory, Department of Chemistry, Gamle Carlsberg Vej 10, DK-2500 Valby, Denmark.

[2]Department of Organic and Industrial Chemistry, University of Milano, via Venezian 21, I-20133 Milano Italy.

Received 29 October 1997; revised 23 December 1997; accepted 9 January 1998

Abstract: A *C*-linked glycosyl decapeptide was synthesised on solid-phase by a combination of the Fmoc-method and a novel method of using azido acids with reductive deprotection. The unprotected *C*-glycosyl azido acid building block was incorparated by a TBTU coupling, subsequently reduced using DTT in DMF and followed by further Fmoc-amino acid-OPfp ester couplings. Mouse hemoglobin (67-76) carrying the C-linked glycoside in position 72 was prepared as a metabolically stable T-cell glycopeptide antigen.

Interest in protein-bound carbohydrates have increased over the last two decades due to the growing understanding in different biological interactions.[1,2,3] The carbohydrate part is mostly responsible for different properties of solubility, molecular recognition, protease resistance and immunogenicity. These properties have stimulated an interest in glycopeptides as therapeutic drugs.[4] Solid-phase synthesis of *O*-glycopeptides is performed by either protected glycosylated serine or threonine amino acid building blocks[5] or direct glycosylation on the solid-phase.[6,7] Methodology for solid-phase synthesis of *N*- and *O*-linked glycopeptides has been developed. However, often the sensitivity of the glycosidic linkage towards both enzymatic and chemical degradation is problematic.[1] Therefore, it is of great interest to exchange the less stable *O*-glycosidic linkage between the amino acid and the carbohydrate to a more stable *C*-linked analogue.[8-12] Recently, the *C*-analogue **1** of *O*-β-D-glucopyranosyl serine[13] was prepared from benzylated gluconolactone. Compound **2** (Figure 1) was employed in the *C*-linked glycosyl decapeptide analogue of CBA/J mouse hemoglobin Hb (67-76) VITAFNEGLK. The *C*-linked glycosyl azido acid was introduced at the position of Asn-72, where introduction of glycan structures previously have been demonstrated to convert the non-immunogenic peptide into strong glycan specific immunogens.[14]

In the present work, the synthesis of a *C*-linked glycosyl decapeptide, using an unprotected *C*-β-D-glucopyranosyl serine analogue **2** with the amino group protected as an azide, is described. Azides have been employed as amino group precursors in organic solution chemistry for the last two decades. However, the conditions of reduction used were usually heterogeneous and therefore not compatible with the requirements for solid-phase reactions. The use of 1,4-dithio-DL-threitol (DTT) and other thiols in quantitative reduction of the azido group in carbohydrates and peptides bound to the solid-phase were recently described[15] as well as the synthesis of sterically demanding peptides using α-azido acids as building blocks.[16] The application of azides and the efficiency of DTT reduction ($-N_3$ to $-NH_2$) facilitated the synthesis of *C*-linked glycosyl azido acid as building blocks requiring less protecting group manipulation. Furthermore, the carboxylic acid of such building

0040-4039/98/$19.00
PII: S0040-4039(98)00095-1

blocks may be converted to the highly activated and chirally stable acid chlorides to be used, alternatively to the *in situ* activation with *e.g.* *O*-(benzotriazol-1-yl)-*N,N,N',N'*-tetramethyluronium tetrafluoroborate (TBTU), in difficult coupling reactions.

The acetyl protected *C*-linked serine analogue **1**[13] was synthesised by oxidation-vinylation of tetrabenzyl glucose followed by glycosylation with methyl glycolate to yield the 1-vinyl-α-glucoside. Stereoselective [2,3]-Wittig sigmatropic rearrangement using LDA and subsequent reduction of the double bond, yielded the 2-hydroxy-β-*C*-linked glucosyl Ser precursor. The α-hydroxyl group was substituted, with inversion of configuration, to the azide using diphenylphosphoryl azide and tetrabutylammonium azide. Finally, acetolysis of the benzyl groups gave compound **1**. The optical purity of **1** was confirmed with NMR-spectroscopy (^{13}C, ^{1}H, COSY) at 600 MHz.[17]

Compound **1** was deacetylated using NaOMe in MeOH at pH 11 (moist pH-paper) and then water was added to hydrolyse the methyl ester (over night). The reaction mixture was neutralised using Amberlite IR 120 H^+ ion exchange resin, concentrated and **2** was used without further purification. A small sample was purified on RP-HPLC and analysed by ES-MS and NMR-spectroscopy (^{13}C, ^{1}H, COSY).[18] Unfortunately, the NMR experiments revealed that racemization had occured at the α-carbon during hydrolysis.

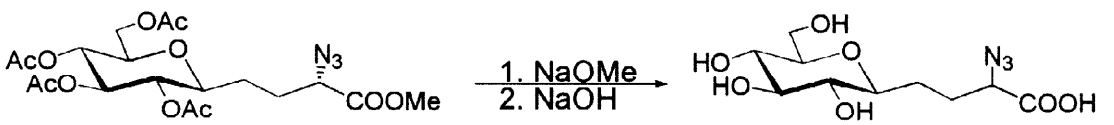

Figure 1 *The hydrolysis of the C-glycosyl azido acid building block afforded a mixture of the two α-stereoisomers*[18]

Polyethylene glycol polyacrylamide (PEGA) resin was derivatised with the base labile *p*-hydroxymethylbenzoic acid (HMBA) linker using TBTU/N-ethylmorpholine (NEM) activation.[19] The first amino acid (AA), Fmoc-Lys(Boc)-OH was coupled to the HMBA linker using 1-(2-mesitylenesulfonyl)-3-nitro-1H-1,2,4-triazol (MSNT)/1-methylimidazole (MeIm) activation.[20] The tetrapeptide (H-Glu(O*t*Bu)-Gly-Leu-Lys(Boc)-HMBA-PEGA) was synthesised using Fmoc-amino acid-pentafluorophenyl (OPfp) esters with 3,4-dihydro-3-hydroxy-4-oxo-1,2,3-benzotriazine (Dhbt-OH) catalysis and 20% piperidine/DMF for Fmoc cleavage. The unprotected *C*-glycosyl azido acid building block (1.5 equiv.) was introduced using TBTU/NEM activation. The product was cleaved from a few beads and analysed by ES-MS revealing unreacted tetrapeptide in addition to the expected product. Therefore, the coupling reaction was repeated (0.5 equiv.) followed by subsequent reduction of the azide using 2 M DTT and 1 M diisoprpylamine (DIPEA) in DMF at 50°C. The reaction was terminated after 3 hrs and a few beads were acetylated (Ac_2O/DMF), deprotected, cleaved off from the resin and analysed by ES-MS. Attempts to use MALDI-TOF-MS to follow the progress of reduction were unsuccessful. However, according to ES-MS complete reduction of the azide had been accomplished and a minor amount of unreacted tetrapeptide was also observed. The solid-phase synthesis of the *C*-linked glycosyl decapeptide (H-Val-Ile-Thr(*t*Bu)-Ala-Phe-Ala(*C*-β-D-Glc-CH_2-)-Glu(O*t*Bu)-Gly-Leu-Lys(Boc)-HMBA-PEGA) was completed using OPfp esters and Dhbt-OH catalysis. Cleavage, of the acid-labile protection groups, using 95% TFA followed by cleavage from the resin using 0.1 M NaOH, neutralisation using Amberlite IR 120 H^+ ion exchange resin and lyophilisation gave 52 mg of the crude peptide. RP-HPLC gave 23 mg (57% based on the loading) of the wanted *C*-linked glycosyl decapeptide mixture and 7mg (22%) of the nonapeptide, lacking the Ala(*C*-β-D-Glc-CH_2-) moiety. The two *a*-stereoisomers **3**[21] and **4**[22] were

separated by RP-HPLC and analysed by MALDI-TOF-MS, ES-MS and NMR-spectroscopy (^{1}H, COSY). The T-cell stimulation with compounds **3** and **4** is currently being investigated.

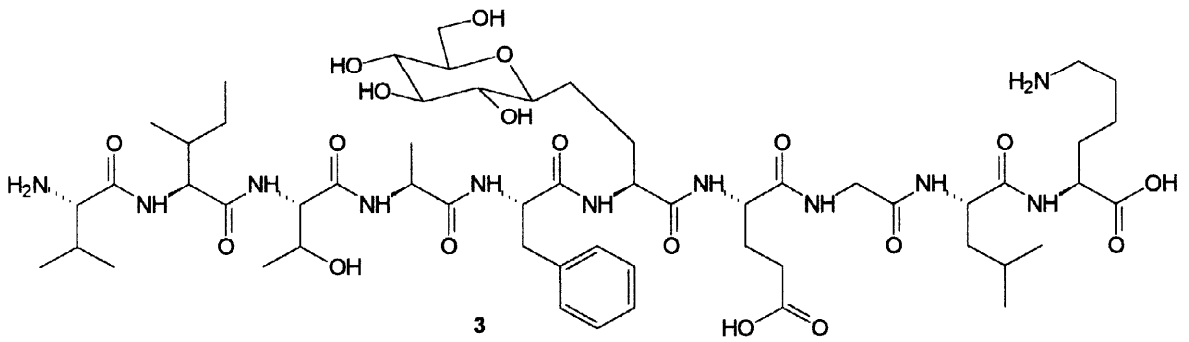

Figure 2 *Compound 3[21], the major of the two α-stereoisomers of the* C*-glycodecapeptide separated by RP-IIPLC.*

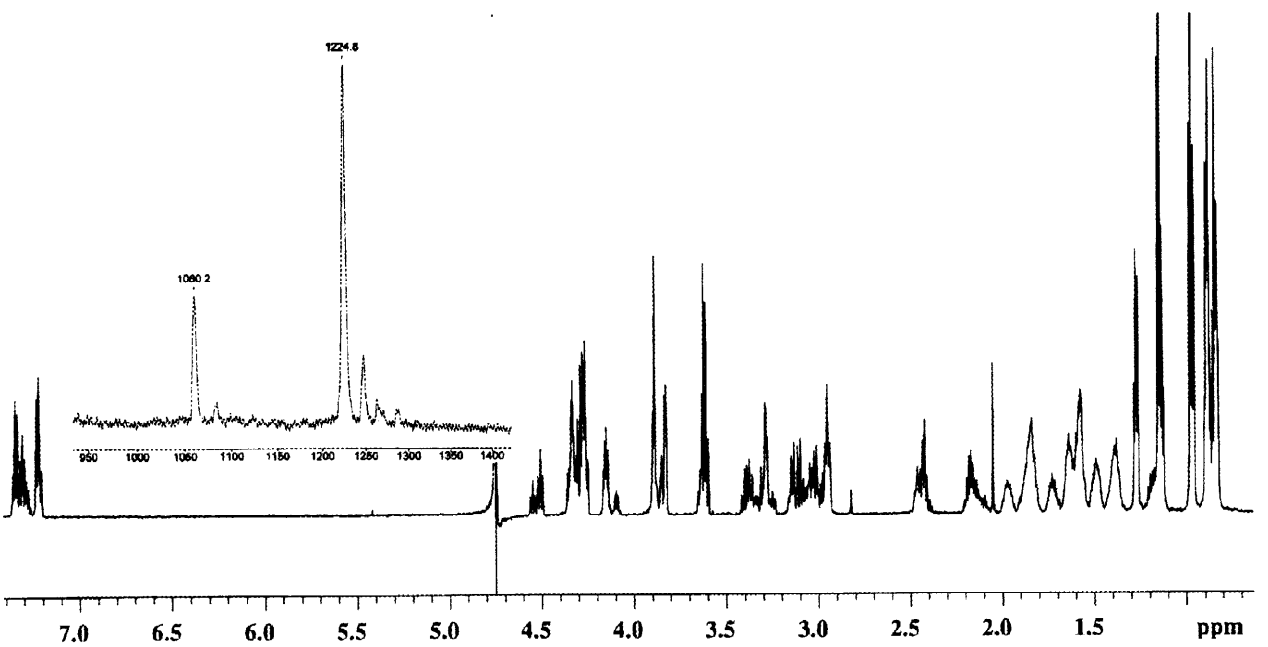

Figure 3 *NMR spectrum of the C-linked glycosyl decapeptides 3[21] and 4[22] together with the MALDI-TOF-MS (Bradykinin 1060.2 as ref.)*

Acknowledgement: Mr. Bent O. Petersen and Dr. Axel Meissner for help with the 600 MHz NMR, Mrs. Hanne Christiansen for help with the RP-HPLC, Dr Anita M. Jansson for help with the ES-MS, are greatfully acknowledged. This work has been supported by the EU-Carenet 2 program (ERB 4061 PL 95-0372).

References and Notes:

1. Dwek, R. A. *Chem. Rev.* **1996**, *96*, 683
2. Woods, R. J.; Edge, C. J.; Dwek, A. *Struc. biol.* **1995**, *1*, 499
3. Hart, G. W.; Greis, K. D.; Dong, D. L.-Y.; Blomberg, M. A.; Chou, T.-Y.; Jiang, M.-S.; Roquemore, E. P.; Snow, D. M.; Kreppel, L. K.; Cole, R. N.; Comer, F. I.; Arnold, C. S.; Hayes, B. K.In *Glycoimmunology*; Alavi, A., Axford, J. S. Eds.; Plenum Press: New York, 1995; pp 115-123.
4. Kihlberg, J.; Ahman, J.; Walse, B.; Drakenberg, T.; Nilsson, A.; Soderbergahlm, C.; Bengtsson, B.; Olsson, H. *J.Med.Chem.* **1995**, *38*, 161
5. Meinjohanns, E.; Meldal, M.; Bock, K. *Tetrahedron Lett.* **1995**, *36*, 9205

6. Meldal, M.; Renil, M.; Juliano, M. A.; Jansson, A. M.; Meinjohanns, E.; Buchardt, J.; Schleyer, A.In *Peptides 1996, Proc. Eur. Pept. Symp.* Epton, R. Ed.; Mayflower Scientific: Kingswinford, UK, 1997; pp in press
7. Schleyer, A.; Meldal, M.; Renil, M.; Paulsen, H.; Bock, K. *Angew. Chem. Int. Ed. Engl.* **1997**, *109*, 2064
8. Kessler, H.; Wittmann, V.; Köck, M.; Kottenhahn, M. *Angew. Chem.* **1992**, *104*, 874
9. Petrus, L.; BeMiller, J. N. *Carbohydr. Res.* **1992**, *230*, 197
10. Bertozzi, C. R.; Hoeprich, P. D.; Bednarski, M. D. *J. Org. Chem.* **1992**, *57*, 6092
11. Gurjar, M. K.; Mainkar, A. S.; Syamala, M. *Tetrahedron Assym.* **1993**, *4*, 2343
12. Burkhart, F.; Hoffmann, M.; Kessler, H. *Angew. Chem. Int. Ed. Engl.* **1997**, *36*, 1191
13. Lay, L.; Meldal, M.; Nicotra, F.; Panza, L.; Russo, G. *J. Chem. Soc., Chem. Commun.* **1997**, 1469
14. Galli-Stampino, L.; Meinjohanns, E.; Frische, K.; Meldal, M.; Jensen, T.; Werdelin, O.; Mouritsen, S. *J. Cancer. Res.* **1997**, *57*, 3214
15. Meinjohanns, E.; Meldal, M.; Jensen, T.; Werdelin, O.; Galli-Stampino, L.; Mouritsen, S.; Bock, K. *J. Chem. Soc., Perkin Trans. 1* **1997**, 871
16. Meldal, M.; Juliano, M. A.; Jansson, A. M. *Tetrahedron Lett.* **1997**, *38*, 2531
17. **Methyl 2-*S*-azido-4-*C*-(2,3,4,6-tetra-*O*-acetyl-β-D-glucopyranos-1-yl)-butanoate 1** (ES-MS $(M+Na)^+$ calcd 496.4 obsd 496.0) ^{13}C NMR (600 MHz, $CDCl_3$, 300K $CDCl_3$ 77.7 ppm) 21.3, 21.3, 21.4 ($\underline{C}H_3CO$), 28.0, 28.2 (C^{β}, C^{γ}), 53.3 (CH_3O), 62.6, 63.0 (C6, C^{α}), 69.3, 72.4, 74.9, 76.4, 78.2 (C1-C5), 170.2, 170.4, 171.0, 171.4, 171.4 (CO). ^1H NMR[8] (600 MHz, $CDCl_3$, 300K $CHCl_3$ 7.27 ppm) 1.56, 1.70 (m, 2H, H^{γ}), 1.76, 2.07 (m, 2H, H^{β}), 2.0, 2.02, 2.05, 2.08 (m, 12H, CH_3CO), 3.44 (dt, 1H, $J_{1,2}$ 9.4, $J_{1,\gamma}$ 2.5 Hz, H1), 3.63 (ddd, 1H, $J_{5,4}$ 9.8, $J_{5,6a}$ 2.2, $J_{5,6b}$ 5.1 Hz, H5), 3.80 (s, 3H, CH_3O), 3.92 (dd, 1H, $J_{\alpha,\beta}$ 4.5, 8.8 Hz, H^{α}), 4.10 (dd, 1H, $J_{6,6}$ 12.3 Hz, $J_{6,5}$ 2.2 Hz, $H6^a$), 4.21 (dd, 1H, $J_{6,6}$ 12.3, $J_{6,5}$ 5.0 Hz, $H6^b$), 4.88 (t, 1H, $J_{2,3}$ 9.5 Hz, H2), 5.04 (t, 1H, $J_{4,5}$ 9.6 Hz, H4), 5.17 (t, 1H, $J_{3,4}$ 9.3 Hz, H3).
18. **2-*R,S*-Azido-4-*C*-(2,3,4,6-tetra-*O*-acetyl-β-D-glucopyranos-1-yl)-butanoic acid 2** (ES-MS $(M+H)^+$ calcd 292.3 obsd 292.2) ^{13}C NMR (250 MHz, D_2O, 300K) 25.2x2, 25.4, 25.5 (C^{β}a,b; C^{γ}a,b), 59.3 (C6), 60.4, 60.7 (C^{α}a,b), 68.3x2, 71.6, 71.7, 75.7, 76.6 77.2, 77.9 (C1a,b-C5a,b), 173.1, 173.1 (COOH a,b). ^1H NMR (600 MHz, D_2O, 300 K, HDO 4.75 ppm) 1.52, 1.93 (m, 2H, H^{γ}a), 1.50, 1.94 (m, 2H, H^{γ}b), 1.82, 2.09 (m, 2H, H^{β}a) 1.96 (m, 2H, H^{β}b) 3.15 (dt, 1H, $J_{2,3}$ 9.0 Hz, H2), 3.29-3.33 (m, 3H, H1a,b; H4, H5), 3.41 (dt, 1H, $J_{3,4}$ 9.0 Hz, H3), 3.64 (dd, 1H, $J_{6,6}$ 11.9, $J_{6,5}$ 5.0 Hz, $H6^b$a,b), 3.85 (d, 1H, $J_{6,6}$ 11.7 Hz, $H6^a$a,b), 4.23 (dd, 1H, J 4.4, 5.1, 7.3, 8.0 Hz, H^{α}a,b).
19. Knorr, R.; Trzeciak, A.; Bannwarth, W.; Gillessen, D. *Tetrahedron Lett.* **1989**, *30*, 1927
20. Blankemeyer-Menge, B.; Nimtz, M.; Frank, R. *Tetrahedron Lett.* **1990**, *31*, 1701
21. **H-Val-Ile-Thr-Ala-Phe-Ala(*C*-β-D-Glc-CH_2-)-Glu-Gly-Leu-Lys-OH 3** (major component) (ES-MS $(M+H)^+$ calcd 1224.8 obsd 1224.7, MALDI-TOF-MS $(M+H)^+$ calcd 1224.8 obsd 1224.8). ^1H NMR (600 MHz, D_2O, 300 K, HDO 4.75 ppm) 0.83, 0.88 (L^{δ}, Leu-9), 0.85 (I^{δ}, Ile-2), 0.97 (V^{γ}, Val-1), 1.15 (T^{γ}, Thr-3), 1.17 (I^{γ}, Ile-2), 1.27 (A^{β}, Ala-4), 1.38 (K^{γ}, Lys-10), 1.55 (L^{γ}, Leu-9), 1.59, 1.87 (L^{β}, Leu-9), 1.64 (A^{γ}, Ala-6), 1.64 (K^{δ}, Lys-10), 1.71, 1.86 (K^{β}, Lys-10), 1.82 (I^{β}, Ile-2), 1.96, 2.14 (E^{β}, Glu-7), 2.17 (V^{β}, Val-1), 2.42 (E^{γ}, Glu-7), 2.94 (K^{ε}, Lys-10), 2.95 (A^{β}, Ala-6), 3.01, 3.05 (F^{β}, Phe-5), 3.12 (H2), 3.28 (H4), 3.34 (H5), 3.41 (H3), 3.64 3.86 ($H6^{a,b}$), 3.83 (V^{α}, Val-1), 3.89 (G^{α}, Gly-8), 4.14 (T^{β}, Thr-3), 4.15 (L^{α}, Leu-9), 4.16 (H1), 4.25 (K^{α}, Lys-10), 4.27 (A^{α}, Ala-4), 4.27 (A^{α}, Ala-6), 4.30 (I^{α}, Ile-2), 4.34 (T^{α}, Thr-3), 4.34 (E^{α}, Glu-7), 4.50 (F^{α}, Phe-5), 7.22, 7.31, 7.35 (aromatic H, Phe-5).
22. **H-Val-Ile-Thr-Ala-Phe-Ala(*C*-β-D-Glc-CH_2-)-Glu-Gly-Leu-Lys-OH 4** (minor component) (ES-MS $(M+H)^+$ calcd 1224.8 obsd 1224.7, MALDI-TOF-MS $(M+H)^+$ calcd 1224.8 obsd 1224.8). ^1H NMR (600 MHz, D_2O, 300 K, HDO 4.75 ppm) 0.84, 0.90 (L^{δ}, Leu-9), 0.85 (I^{δ}, Ile-2), 0.97 (V^{γ}, Val-1), 1.13 (T^{γ}, Thr-3), 1.28 (A^{β}, Ala-4), 1.38 (K^{γ}, Lys-10), 1.47 (I^{γ}, Ile-2), 1.58 (L^{γ}, Leu-9), (L^{β}, Leu-9, nd), 1.64 (A^{γ}, Ala-6), 1.64 (K^{δ}, Lys-10), 1.71, 1.86 (K^{β}, Lys-10), 1.82 (I^{β}, Ile-2), 1.97, 2.10 (E^{β}, Glu-7), 2.08 (E^{γ}, Glu-7), 2.17 (V^{β}, Val-1), 2.94 (K^{ε}, Lys-10), 2.95 (A^{β}, Ala-6), 3.00, 3.07 (F^{β}, Phe-5), 3.08 (H2), 3.26 (H4), 3.29 (H5), 3.38 (H3), 3.63, 3.83 ($H6^{a,b}$), 3.83 (V^{α}, Val-1), 3.89 (G^{α}, Gly-8), 4.08 (T^{β}, Thr-3), (L^{α}, Leu-9, nd), 4.16 (H1), 4.25 (K^{α}, Lys-10), 4.25 (E^{α}, Glu-7), 4.26 (A^{α}, Ala-4), 4.27 (A^{α}, Ala-6), 4.27 (T^{α}, Thr-3), 4.30 (I^{α}, Ile-2), 4.54 (F^{α}, Phe-5), 7.22, 7.31, 7.35 (aromatic H, Phe-5).