Synthesis of *N*-glycopeptides Applying Glycoamino Acid Building Blocks with a Combined Fmoc/Boc Strategy

Kinga Rákosi[1], Orsolya Szolomájer-Csikós[1], László Kalmár[1,2], Zoltán Szurmai[3], János Kerékgyártó[2] and Gábor K. Tóth[1,*]

[1]*Department of Medical Chemistry, Faculty of General Medicine, University of Szeged, 6720 Szeged Dóm tér 8., Hungary,* [2]*Department of Botany, Bio-organic Laboratory, Faculty of Sciences and Technology, University of Debrecen, 4032 Debrecen Egyetem tér 1., Hungary,* [3]*Department of Applied Ecology, Bio-organic Laboratory, Faculty of Sciences and Technology, University of Debrecen, 4032 Debrecen Egyetem tér 1., Hungary*

Abstract: Mono-, di- and trisaccharide representing the reducing terminal of the core structure of *N*-glycans were incorporated into Leu-Lys-Asn-Gly-Gly-Pro hexapeptide that is a partial structure of the Trp-cage mini-protein by linear assembly. These studies provide evidence that the used combination of Fmoc and Boc strategy and mild conditions result in glycopeptides in high purity and reasonable yield.

Keywords: Carbohydrates, Boc deprotection, *N*-glycopeptides, solid-phase synthesis, tin(IV) chloride.

INTRODUCTION

Great amount of new compounds are formed by partly post-translational modifications following the ribosomal protein synthesis, partly modifications of amino acids by mostly plants, microorganisms and fungi. Beside glycosylated, phosphorylated, etc. peptides and proteins, such compounds are amino acid- and peptide-based heterocyclic compounds (alkaloids and some antibiotics, etc.) widespread in the plant kingdom and used as medicines. These post-translational modifications have fundamental importance in biological recognition processes [1-6]. One of the most challenging task among them is the rational preparation of the glycosylated peptides especially those having complex oligosaccharide moieties. There are two main strategies for the synthesis of glycopeptides: the synthon (stepwise approach: normally proceeds through a glycosyl amino acid which usually serves as the building block for solid-phase construction of a peptide sequence) and convergent (the required carbohydrate chain and peptide are each built independently, and the amide linkage is created late in the synthesis) method [7-13]. Both of them can be implemented in liquid or solid-phase.

In the framework of our research project on the synthesis of *N*-glycopeptides [14-17] the preparation of mono- di- and trisaccharide containing hexapeptide conjugates was carried out by both convergent [17] and linear assembly in order to compare these two chemical methods directly. Here we report the synthesis of target conjugates by linear (stepwise) assembly.

*Address correspondence to this author at the Department of Medical Chemistry, Faculty of General Medicine, University of Szeged, 6720 Szeged Dóm tér 8., Hungary; Tel: +36-62-545139; Fax: +36-62-545971; E-mail: tgabor@mdche.szote.u-szeged.hu

RESULTS AND DISCUSSION

Mono-, di- and trisaccharide representing the reducing terminal of the core structure of *N*-glycans were incorporated into Leu-Lys-Asn*-Gly-Gly-Pro hexapeptide that is the 7-12 analogue of the Trp-cage mini-protein [18] by linear assembly, where * is the site of glycosylation.

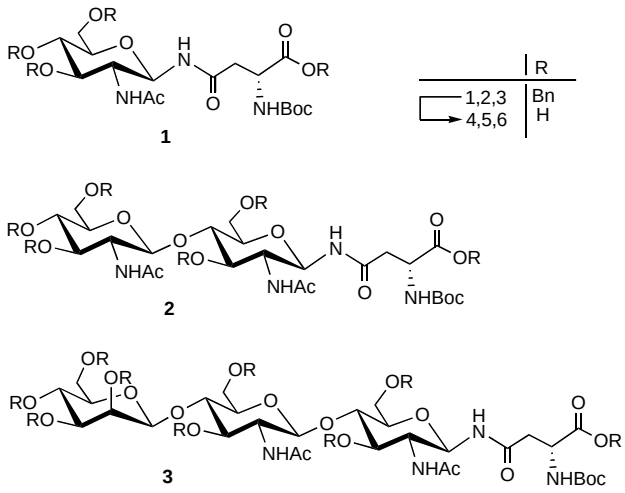
The Trp-cage miniprotein, TC5b (NLYIQWLKDGG PSSGRPPPS), is only a 20 residue long polypeptide, notable for its protein-like 3D-fold in quasi physiological conditions. In its 3D structure an α-helix, a 3_{10}-helix and a polyproline II secondary structural elements shield the central hydrophobic Trp6 residue. Tc5b is among the most investigated miniproteins, therefore it can be a good model for studying the influence of posttranslational modifications in the determination of 3D structures.

The preparation of the fully O-benzylated glycosyl azides GlcNAc(β1→N_3), GlcNAc(β1→4)GlcNAc(β1→N_3) and Man(β1→4)GlcNAc(β1→4)GlcNAc(β1→N_3) representing the reducing terminal of the core structure of *N*-glycans was carried out as described earlier [14]. Chemoselective reduction (PtO_2/H_2) of the azido function of glycosyl azides resulted in fully *O*-benzylated glycosylamines. The coupling reactions by *in situ* trapping of the amines with selectively protected activated (HBTU/*i*Pr_2NEt) aspartic acid resulted in carbohydrate-aspartic acid derivatives **1**, **2**, **3** in yields of 95%, 92% and 85%, respectively (Scheme 1.). Both the ^{1}H- and ^{13}C-NMR Spectra [19] revealed that the anomerization of the amines could be avoided under the applied reaction conditions. Catalytic hydrogenolysis ($Pd(OH)_2/H_2$) of compound **1**, **2**, **3** afforded key glycoamino acids **4**, **5**, **6**, respectively.

According to literature the Fmoc-protected glycosylated asparagine derivatives proved to be suitable building blocks for the solid-phase peptide synthesis, but in some cases the preparation of Boc-protected glycosylated amino acid build-

0929-8665/11 $58.00+.00

ing blocks is more convenient than the Fmoc-protected ones due to compatibility issues between carbohydrate- and peptide-chemistry [20]. One of the main barrier of the production of the appropriately protected glycosylated amino acid building blocks is the finding of the suitable orthogonal protecting groups, namely for the α-amino and α-carboxyl group and glyco part of the molecule. A number of carboxyl and hydroxyl protecting groups are selectively cleavable in the presence of Boc and not selectively cleavable in presence of Fmoc. Therefore the extension described in this paper could led to the increased set of applicable protecting groups for glycopeptide synthesis.



Scheme (1).

The glycopeptides - Leu-Lys-[GlcNAc(β1→N)]Asn-Gly-Gly-Pro-OH, Leu-Lys[GlcNAc(β1→4)GlcNAc(β1→N)] Asn-Gly-Gly-Pro-OH, Leu-Lys[Man(β1→4)GlcNAc(β1→4) GlcNAc(β1→N)]Asn-Gly-Gly-Pro-OH- were manually synthesised on Boc-Pro-Merrifield resin (loading: 0.64 mmol/g) and 3 fold molar excess of DCC-activated Boc-amino acids were used throughout. Standard Boc-amino acid side chain protecting groups were applied except the lysine, where we used Boc-Lys(Z). The scale of assembly was 0.2 mmol. Each coupling reaction was carried out for 2 h. Boc-deprotection was performed with 50% (v/v) TFA/DCM, followed by neutralization with 10% (v/v) TEA/DCM. Cleavage from the solid support and side chain deprotection was achieved by 120 min treatment with HBr (min. 33% in glacial acetic acid):TFA (1:9) (v/v) containing 2% dithiothreitol (m/v). The isolation of the glycosylated hexapeptide derivatives was made by RP-HPLC using a Phenomenex Jupiter C18 300 Angstrom 10μ column (250x10.0 mm) and eluents (A) 0.1% TFA and (B) 80% MeCN, 0.1% TFA. The flow rate was 2 ml/min and detection was performed at 220 nm. The glycopeptides were identified by electrospray ionization mass spectrometry and the purity was examined by analytical RP-HPLC [Eluent A: 0.1% TFA and eluent B: 80% MeCN, 0.1% TFA; flow: 1.0 ml/min; Phenomenex Luna C18(2) 100 Angstrom 5μ column (250x4.6 mm)].

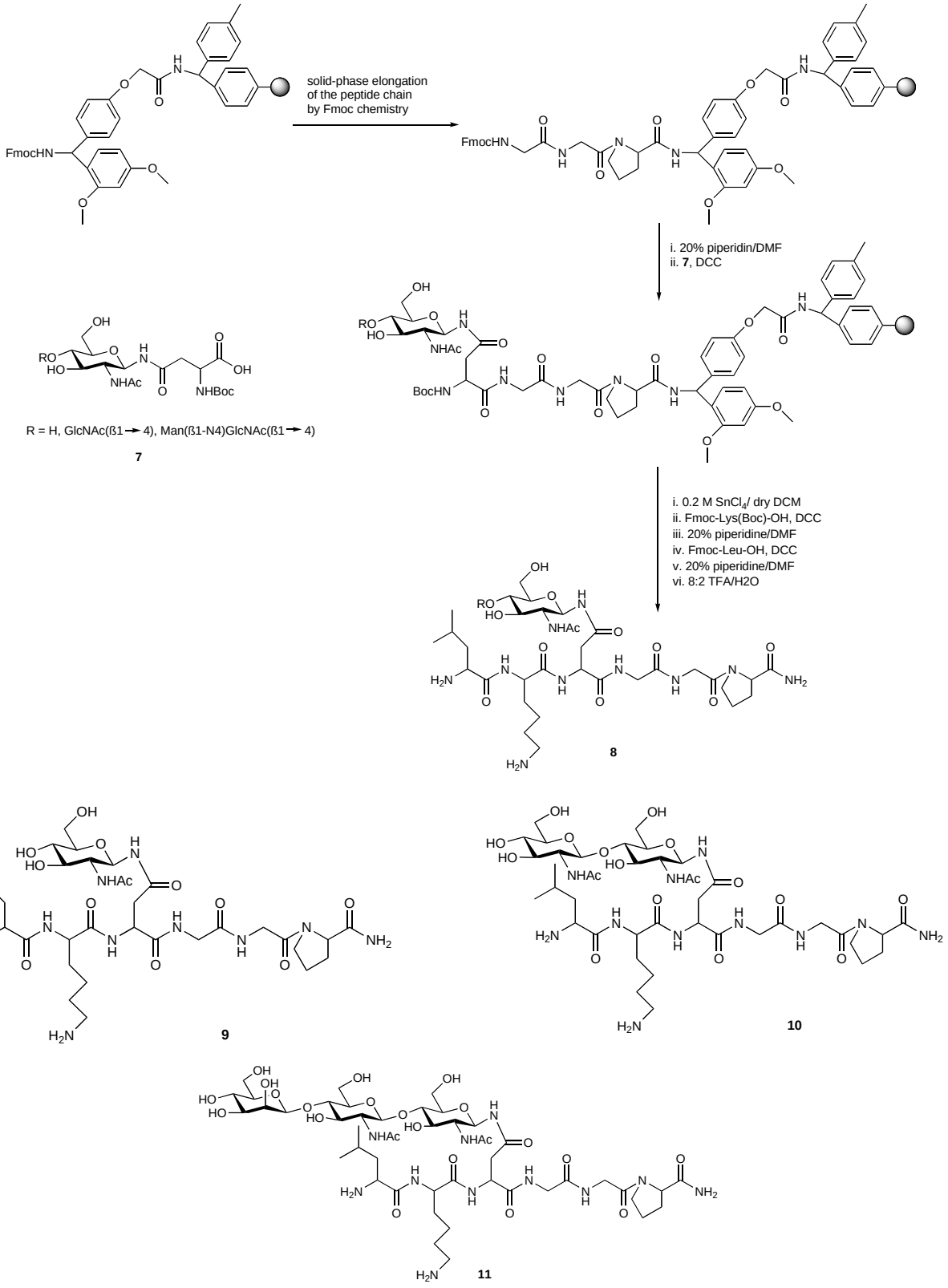
Although the first glycopeptide could be isolated in good yield and relatively high purity, in case of the disaccharide-hexapeptide derivative substantial loss of a monosaccharide unit could be observed during cleavage from the resin. We had tried to optimize the cleavage conditions by reducing on a hand the amount of HBr/acetic acid in the cleavage mixture and on the other hand the reaction time, but it had not resulted in better products. Isolation of the disaccharide-hexapeptide derivative from the hexapeptide containing only the monosaccharide unit was found to be very difficult. Changing the eluents to (A) 0.05M NH_4OAc buffer (pH 7.5) and (B) 60% MeOH/ 0.05M NH_4OAc buffer (pH 7.5) improved the separation, so the disaccharide-hexapeptide derivative could be isolated in a pure form.

Cleavage from the resin of the Leu-Lys-[Man-(β1→4) GlcNAc(β1→4)GlcNAc(β1→N)]Asn-Gly-Gly-Pro-OH peptide resulted in a low yield of the trisaccharide-hexapeptide derivative.

Therefore we decided to modify the strategy for the incorporation of the Boc-protected glycosylated amino acid derivatives in the model peptide by using a new, mild and selective Boc deprotecting agent – tin(IV) chloride. [21] The glycopeptides in this case were synthesised using the Fmoc chemistry on a TFA cleavable Rink amide MBHA resin. According to our investigations other resins common in Fmoc chemistry (Rink amide, 2-chlorotrityl resin, Wang resin) showed removal of substrate from the resin during cleavage with tin(IV) chloride, except the Rink amide MBHA.

The synthesis of Leu-Lys-[GlcNAc(β1→N)]Asn-Gly-Gly-Pro-NH_2, Leu-Lys-[GlcNAc(β1→4)GlcNAc(β1→N)] Asn-Gly-Gly-Pro-NH_2, Leu-Lys-[Man(β1→4)GlcNAc (β1→4)GlcNAc (β1→N)]Asn-Gly-Gly-Pro-NH_2 was carried out as follows: Rink amide-MBHA PS resin (loading: 0.81 mmol/g) was applied. The scale of assembly was 0.2 mmol. The first 3 amino acids were incorporated using the Fmoc/*t*Bu strategy with three equivalents of amino acid and three equivalents of *N,N'*-dicyclohexyl-carbodiimide in DMF for 3 hours. The coupling efficiency was monitored with the Kaiser test [22]. Fmoc-deprotection was made with 20% (v/v) piperidine in DMF. Before coupling of Boc-protected glycosylated aspartic acid derivatives **4**, **5**, **6** the resin was pre-washed with dry DCM and then the resin was treated twice with 0.2 M $SnCl_4$ in dry DCM for 10 minutes, giving the resin a red colour due to complex formation with $SnCl_4$.

Coupling of the last two amino acid was performed with the standard Fmoc-amino acids and *N,N'*-dicyclohexyl-carbodiimide. Cleavage of the peptides from the resin was performed by addition of a mixture containing TFA:H_2O (8:2) (v/v) for 3 hours (Scheme 2). After cleavage the peptides were precipitated onto the resin in ice cold diethyl ether and lyophilized after solubilization in water. Purification of the crude peptides was made by RP-HPLC on a Phenomenex Jupiter C18 300 Angstrom 10μ column (250x10.0 mm), using as eluents: (A) 0.1% TFA and (B) 80% MeCN, 0.1% TFA. Elution was made at a flow of 2 ml/min and detection was performed at 220 nm. The glycopeptides were identified by electrospray ionization mass spectrometry and the purity was examined by analytical RP-HPLC [Eluent A: 0.1% TFA and eluent B: 80% MeCN, 0.1% TFA; flow: 1.0 ml/min; Phenomenex Luna C18(2) 100 Angstrom 5μ column (250x4.6 mm)].



Scheme (2).

The upper procedure led to pure glycopeptides **9**, **10**, and **11** with yields of 56, 42, and 36%, respectively. The RP-HPLC chromatograms of the purified glycopeptides and the ESI-MS spectra are shown in Figures 1-3.

CONCLUSION

In summary, we have developed a strategy that combines the Fmoc and Boc SPPS approaches for the preparation of

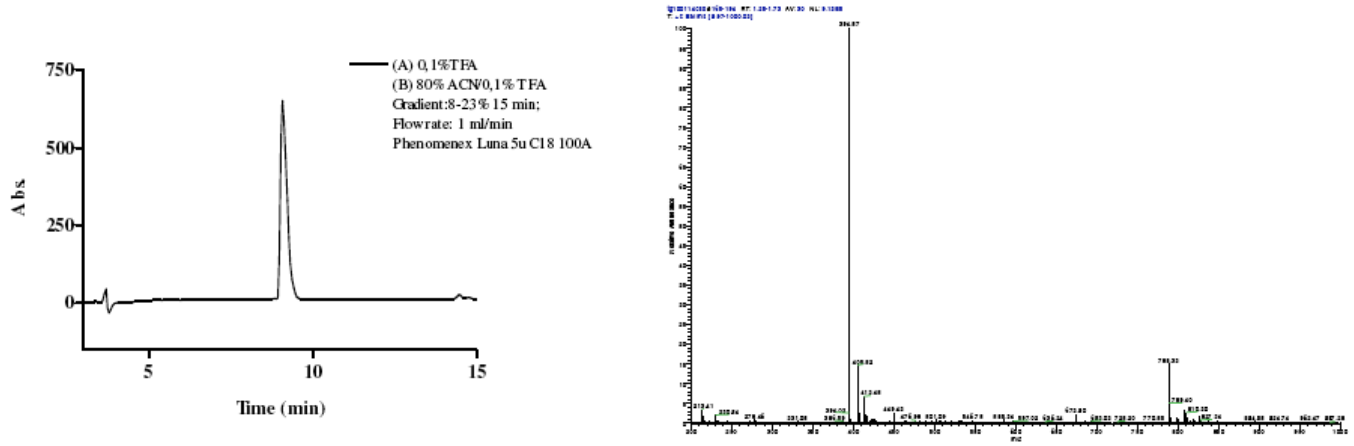


Figure 1. Analytical RP-HPLC (220nm) chromatogram and ESI-MS spectrum of the purified peptide **9**.

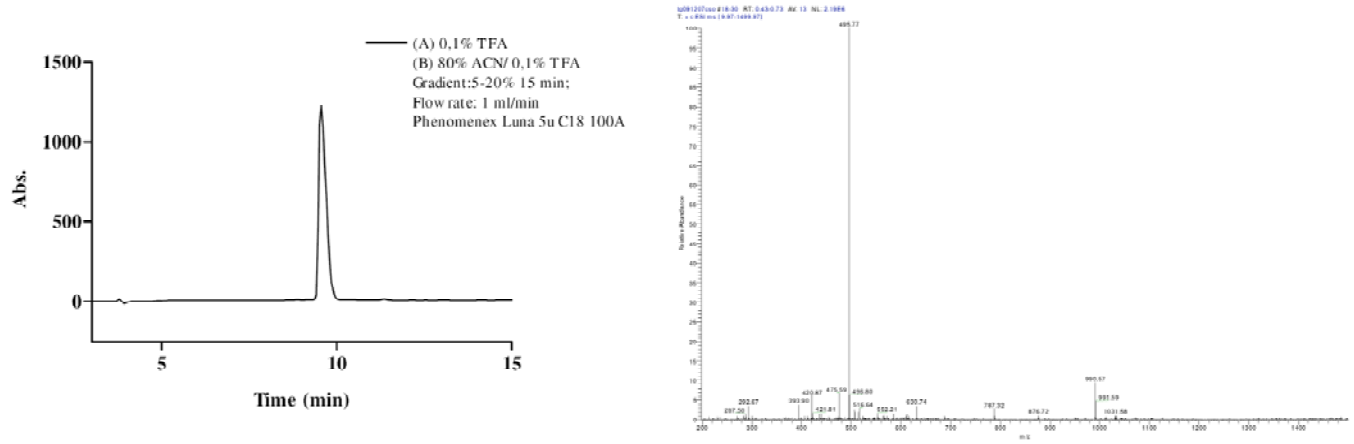


Figure 2. Analytical RP-HPLC (220nm) chromatogram and ESI-MS spectrum of the purified peptide **10**.

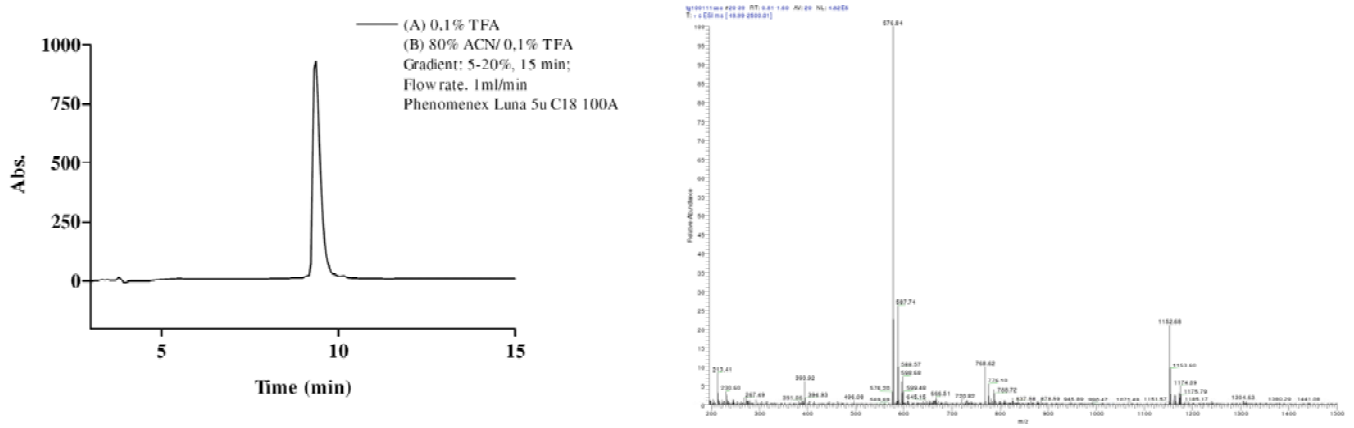


Figure 3. Analytical RP-HPLC (220nm) chromatogram and ESI-MS spectrum of the purified peptide **11**.

glycopeptides in high purity and resonable yield, by making use of $SnCl_4$ for Boc deprotection, which leaves the acid sensitive glycosidic bonds intact.

This combined technique offers an excellent cooperation between the Fmoc and Boc strategy protecting the building block with the medium acid sensitive Boc group and applying the mild environment removable Fmoc compatible resins and side-chain protecting groups.

ACKNOWLEDGMENTS

We are grateful for the grant from the Hungarian National Science Foundation (OTKA-71753) and TÁMOP-4.2.1/B-09/1/KONV-2010-0007 for the financial support.

REFERENCES AND NOTES

[1] Varki, A. Biological roles of oligosaccharides: all of the theories are correct. *Glycobiology*, **1993**, *3*, 97-130.

[2] Dwek, R. Glycobiology: Toward understanding the function of sugars. *Chem. Rev.*, **1996**, *96*, 683-720.

[3] Helenius, A.; Aebi, M. Intracellular functions of N-linked glycans. *Science*, **2001**, *291*, 2364-2369.

[4] Imperiali, B.; O'Connor, S.E. Effect of *N*-linked glycosylation on glycopeptide and glycoprotein structure. *Curr. Opin. Chem. Biol.*, **1999**, *3*, 643-649.

[5] Knight, P. The Carbohydrate frontier. *Nat. Biotechnol*., **1989**, *7*, 35-40.

[6] Mandal, M.; Dudkin, V.Y.; Geng, X.; Danishefsky, S.J. In pursuit of carbohydrate based HIV vaccines. Part 1: The total synthesis of hybrid-type gp120 fragments. *Angew. Chem., Int. Ed.,* **2004**, *43*, 2557-2561.

[7] Gamblin, D.P.; Scanlan, E.M.; Davis, B.G. In: *Synthesis of Glycopeptides and Glycoproteins in Comprehensive Glycoscience I-IV*, Elsevier Ltd: Oxford, **2007**; Vol. *I.*, 605-644, and references (17-30) therein.

[8] Gamblin D.P.; Scanlan E.M.; Davis B.G. Glycoprotein synthesis: an update. *Chem. Rev.*, **2009**, *109*(1), 131-163.

[9] Mizuno, M.; Muramoto, I.; Kobayashi, K.; Yaginuma, H.; Inazu, T. A simple method for the synthesis of N^{β}-glycosylated-asparagine and -glutamine derivatives. *Synthesis*, **1999**, (1), 162-165.

[10] Kajihara, Y.; Suzuki, Y.; Yamamoto, N.; Sasaki, K.; Sakakibara, T.; Juneja, L. R. Prompt chemoenzymatic synthesis of diverse complex-type oligosaccharides and its application to the solid-phase synthesis of a glycopeptide with Asn-linked sialyl-undeca- and asialo-nonasaccharides. *Chem. Eur. J.*, 2**004**, *10*, 971-985.

[11] Meinjohanns, E.; Meldal, M.; Paulsen, H.; Dwek, R.A.; Bock, K. Versatile solid-phase thiolytic reduction of azido and N-Dts groups in the synthesis of haemoglobin (67-76) O-glycopeptides and photoaffinity labelled analogs to study glycan T-cell specificity. *J. Chem. Soc., Perkin Trans. 1.*, **1998**, 549-560.

[12] Michael, K. In: *Frontiers in Modern Carbohydrate Chemistry, ACS Symposium Series*, **2007**, Vol. *960*, pp. 328-353.

[13] Cohen-Anisfeld, S.T.; Lansbury, P.T. A practical, convergent method for glycopeptide synthesis. *J. Am. Chem. Soc*., **1993**, *115*, 10531-10537.

[14] Kerékgyártó, J.; Ágoston, K.; Batta, Gy.; Kamerling, J.P.; Vliegenthart, J.F.G. Synthesis of fully and partially benzylated glycosyl azides via thioalkyl glycosides as precursors for the preparation of N-glycopeptides. *Tetrahedron Lett.,* **1998**, *39*, 7189-7192.

Received: October 8, 2010 **Revised: February 4, 2011** **Accepted: February 9, 2011**

[15] Tóth, G.K.; Kádár, K.; Hegyi, O.; Csikós, O.; Kalmár, L.; Kerékgyártó, J. *J. Pept. Res.*, **2008**, *14*(8), 71-72.

[16] (a) Tóth, G.K.; Kádár, K.; Hegyi, O.; Csikós, O.; Kalmár, L.; Kerékgyártó, J. (b) Kerékgyártó, J.; Kalmár, L.; Szurmai, Z.; Harangi, J.; Hegyi, O.; Tóth, G.K. In: Proceedings of the 4th Central European Conference Chemistry towards Biology, Dobogókő, Hungary, September 8-11, 2008, (a) 41, (b) 42.

[17] Kerékgyártó, J.; Kalmár, L.; Szurmai, Z.; Hegyi, O.; Tóth, G.K. submitted to *Org. Lett.*, **2011.**

[18] Neidigh, J.W.; Fesinmeyer R.M.; Andersen, N.H. Designing a 20-residue protein. *Nat. Struct. Biol.*, **2002**, *9*, 425-430.

[19] Selected spectroscopic and physical data are the following: compound **1**: (have been prepared earlier along other route [14], but the stereochemistry of the anomeric center was not verified by spectroscopic data) $[\alpha]_D$ +41.3 (c 0.13, $CHCl_3$); δ_H ($CDCl_3$) 4.71 (dd, 1H, H-1, $J_{1,2}$ 10.2 Hz); δ_C ($CDCl_3$) 80.42 (C-1 and C-3, JC_1-H_1 156 Hz); MALDI-TOF: $C_{45}H_{53}O_{10}N_3$ (795.37): *m/z* 818.35 $[M+Na]^+$, 838.34 $[M+K]^+$. Compound **2**: $[\alpha]_D$ +11.1 (c 0.11, $CHCl_3$); δ_H ($CDCl_3$) 4.65 (dd, 1H, H-1, $J_{1,2}$ 8.0 Hz), 4.36 (d, 1H, H-1', $J_{1',2'}$ 8.0 Hz); δ_C ($CDCl_3$) 100.07 (C-1', $JC_{1'}$-$H_{1'}$ 160 Hz), 80.42 (C-1, JC_1-H_1 156 Hz); MALDI-TOF: $C_{67}H_{78}O_{15}N_4$ (1178.55): *m/z* 1201.67 $[M+Na]^+$, 1217.63 $[M+K]^+$. Compound **3**: $[\alpha]_D$ -13.8 (c 0.13, $CHCl_3$); δ_H ($CDCl_3$) 4.63 (dd, 1H, H-1, $J_{1,2}$ 8.8 Hz), 4.48 (bs, 1H, H-1", $J_{1",2"}$ < 1.0 Hz), 4.35 (d, 1H, H-1' $J_{1',2'}$ 9.7 Hz); δ_C ($CDCl_3$) 101.10 (C-1", $JC_{1"}$-$H_{1"}$ 156 Hz), 100.23 (C-1', $JC_{1'}$-$H_{1'}$ 162 Hz), 80.14 (C-1, JC_1-H_1 156 Hz); MALDI-TOF: $C_{94}H_{106}O_{20}N_4$ (1610.74): *m/z* 1633.83 $[M+Na]^+$, 1649.75 $[M+K]^+$. Compound **4**: $[\alpha]_D$ +13.0 (c 0.15, H_2O); δ_H (D_2O) 5.06 (d, 1H, H-1, $J_{1,2}$ 9.8 Hz); δ_C (D_2O) 79.15 (C-1, JC_1-H_1 156 Hz); MALDI-TOF: $C_{17}H_{29}O_{10}N_3$ (435.43): *m/z* 458.35 $[M+Na]^+$, 474.32 $[M+K]^+$. Compound **5**: $[\alpha]_D$ +2.9 (c 0.10, H_2O); δ_H (D_2O) 5.05 (d, 1H, H-1, $J_{1,2}$ 9.6 Hz), 4.59 (d, 1H, H-1', $J_{1',2'}$ 8.2 Hz); δ_C (D_2O) 102.04 (C-1', $JC_{1'}$-$H_{1'}$ 161 Hz), 78.87 (C-1, JC_1-H_1 156 Hz); MALDI-TOF: $C_{25}H_{42}O_{15}N_4$ (638.27): *m/z* 661.35 $[M+Na]^+$, 677.34 $[M+K]^+$. Compound **6**: $[\alpha]_D$ -6.3 (c 0.11, H_2O); δ_H (D_2O) 5.05 (d, 1H, H-1, $J_{1,2}$ 9.5 Hz), 4.75 (bs, 1H, H-1", $J_{1",2"}$ < 1.0 Hz), 4.60 (d, 1H, H-1', $J_{1',2'}$ 7.5 Hz); δ_C (D_2O) 102.01 (C-1', $JC_{1'}$-$H_{1'}$ 161 Hz), 100.72 (C-1", $JC_{1"}$-$H_{1"}$ 161 Hz), 78.95 (C-1, JC_1-H_1 156 Hz); MALDI-TOF: $C_{31}H_{52}O_{20}N_4$ (800.32): *m/z* 823.38 $[M+Na]^+$, 839.37 $[M+K]^+$.

[20] Atherton, E.; Bury, C.; Sheppard, R.C.; Williams, B.J. Stability of fluorenylmethoxycarbonylamino groups in peptide synthesis. Cleavage by hydrogenolysis and by dipolar aprotic solvents. *Tetrahedron Lett.*, **1979**, *32*, 3041-3042.

[21] Freeman, N.S.; Gilon, C. The use of tin(iv) chloride for mild and selective Boc deprotection on Tfa cleavable Rink-amide MBHA resin. *Synlett*, **2009**, *13*, 2097-2100.

[22] Kaiser, E.; Colescott, R.L.; Bossinger, C.D.; Cook, P.I. Color test for detection of free terminal amino groups in the solid-phase synthesis of peptides. *Anal. Biochem.*, **1970**, *34*, 595.