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

Kinga Rákosi¹, Orsolya Szolomájer-Csikós¹, László Kalmár¹,², Zoltán Szurmai³, János Kerékgyártó² and Gábor K. Tóth¹,⁴

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

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-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 (NYIQWLKDGGPSSGRPPPS), 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 minipeptides, 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→4)GlcNAc(β1→N3), GlcNAc(β1→4)GlcNAc(β1→N3) and Man(β1→4)GlcNAc(β1→4)GlcNAc(β1→N3) representing the reducing terminal of the core structure of N-glycans was carried out as described earlier [14]. Chemoselective reduction (PtO2/H2) 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/iPr2NEt) aspartic acid resulted in carbohydrate-aspartic acid derivatives 1, 2, 3 in yields of 95%, 92% and 85%, respectively (Scheme 1.). Both the 1H- and 13C-NMR Spectra [19] revealed that the anomerization of the amines could be avoided under the applied reaction conditions. Catalytic hydrogenolysis (Pd(OH)2/H2) 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-
ing blocks is more convenient than the Fmoc-protected ones due to compatibility issues between carbohydrate- and peptide-chemistry [20]. One of the main barriers 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.

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→3)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 three 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 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-NH₂, Leu-Lys-[GlcNAc(β1→4)GlcNAc(β1→N)] Asn-Gly-Gly-Pro-NH₂, Leu-Lys-[Man(β1→4)GlcNAc (β1→4)GlcNAc (β1→N)]Asn-Gly-Gly-Pro-NH₂ 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/tBu 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 SnCl₂ in dry DCM for 10 minutes, giving the resin a red colour due to complex formation with SnCl₂.

Cleavage of the first 3 amino acids 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:H₂O (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 (250x1.0 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)].
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
glycopeptides in high purity and resonable yield, by making use of SnCl₄ 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

[19] Selected spectroscopic and physical data are the following: compound I (have been prepared earlier along other route [14], but the stereochemistry of the anomeric center was not verified by spectroscopic data) [m/z] +13.8 (c 0.13, CHCl3); δN (CDCl3) 4.71 (dd, 1H, H-1, J1,2 10.2 Hz); δC (CDCl3) 50.42 (C-1 and C-3, JC=H, 156 Hz); MALDI-TOF: C39H40O15N4 (795.37): m/z 818.35 [M+Na]+, 838.34 [M+K]+. Compound 2: [m/z] +1.1 (c 0.11, CHCl3); δN (CDCl3) 4.65 (dd, 1H, H-1, J1,2 8.0 Hz), 4.36 (d, 1H, H-1', J1',2' 8.0 Hz); δC (CDCl3) 100.07 (C-1', JC=H, 160 Hz), 80.42 (C-1, JC=H, 156 Hz); MALDI-TOF: C25H27O10N4 (638.27): m/z 647.45 [M+Na]+, 661.35 [M+K]+. Compound 3: [m/z] -13.8 (c 0.13, CHCl3); δN (CDCl3) 4.63 (dd, 1H, H-1, J1,2 8.8 Hz), 4.48 (bs, 1H, H-1', J1',2' < 1.0 Hz), 4.35 (d, 1H, H-1' J1',2' 9.7 Hz); δC (CDCl3) 101.10 (C-1', JC=H, 156 Hz), 100.23 (C-1', JC=H, 162 Hz), 80.14 (C-1, JC=H, 156 Hz); MALDI-TOF: C23H23O10N3 (555.26): m/z 563.83 [M+Na]+, 577.85 [M+K]+. Compound 4: [m/z] +13.0 (c 0.15, H2O); δN (D2O) 5.06 (d, 1H, H-1, J1,2 9.8 Hz); δC (D2O) 79.15 (C-1, JC=H, 156 Hz); MALDI-TOF: C21H19O9N3 (435.43): m/z 458.35 [M+Na]+, 474.32 [M+K]+. Compound 5: [m/z] +2.9 (c 0.10, H2O); δC (D2O) 5.05 (d, 1H, H-1, J1,2 9.6 Hz), 4.59 (d, 1H, H-1', J1',2' 8.2 Hz); δC (D2O) 102.04 (C-1', JC=H, 161 Hz), 78.87 (C-1, JC=H, 156 Hz); MALDI-TOF: C19H17O8N3 (353.29): m/z 367.40 [M+Na]+. Compound 6: [m/z] -6.3 (c 0.11, H2O); δN (D2O) 5.05 (d, 1H, H-1, J1,2 9.5 Hz), 4.75 (bs, 1H, H-1', J1',2' < 1.0 Hz), 4.60 (d, 1H, H-1', J1',2' 7.5 Hz); δC (D2O) 102.01 (C-1', JC=H, 156 Hz), 100.72 (C-1', JC=H, 161 Hz), 78.95 (C-1, JC=H, 156 Hz); MALDI-TOF: C17H15O7N3 (250.22): m/z 263.38 [M+Na]+, 283.37 [M+K]+.