Tetrahedron Letters, Vol.31, No.41, pp 5889-5892, 1990
Printed in Great Britain

0040-4039/90 $3.00 + .00
Pergamon Press plc

AUTOMATED SOLID-PHASE SYNTHESIS OF GLYCOPEPTIDES. INCORPORATION OF UNPROTECTED MONO- AND DISACCHARIDE UNITS OF N-GLYCOPROTEIN ANTENNAE INTO T CELL EPITOPIC PEPTIDES

Laszlo Otvos, Jr.[a*], Laszlo Urge[b], Miklos Hollosi[b], Krzysztof Wroblewski[c], Grazyna Graczyk[c], Gerald D. Fasman[d], and Jan Thurin[a]

[a]The Wistar Institute of Anatomy and Biology, 3601 Spruce Street, Philadelphia, PA 19104, U.S.A.; [b]Department of Organic Chemistry, L. Eotvos University, Budapest, Hungary; [c]Department of Biochemistry and Biophysics, University of Pennsylvania, Philadelphia, PA, U.S.A.; [d]Department of Biochemistry, Brandeis University, Waltham, MA, U.S.A.

Abstract: A new reagent for custom synthesis of N-glycosylated peptides, Fmoc-Asn[GlcNAc-(1→4)-GlcNAc]-OH, has been prepared. The O-glycosidic bond survived the trifluoroacetic acid cleavage. High coupling efficiencies were obtained when this glycoamino acid derivative was used for solid-phase synthesis of a test peptide and a T cell epitopic dodecapeptide containing an Asn residue in mid-chain position. An automated peptide synthesizer was used to produce the parent Asn peptide and its N-acetyl-glycosyl mono- and disaccharide derivatives.

Many secretory and membrane proteins carry asparagine-linked oligosaccharides.[1] It is also well known that viral envelope glycoproteins predominantly are N-glycosylated. For example, gp 120 of HIV-1 is extremely heavily glycosylated (31-36 N-linked oligosaccharides per molecule[2]), and the full pathogenic potential of HIV-1 *in vitro* is manifested only if its viral proteins are N-glycosylated.[3] Chitobiose is the starting sugar structure for all N-glycosylation;[4] the removal of the rest of the carbohydrate residues seems not to result in dramatic conformational[5] and immunological[6] change. It is possible that the chitobiose core alone can provide sufficient conformational-steering effect in many cases, where carbohydrate chains are known to stabilize biologically active protein structures. Solid-phase synthesis is the method of choice to produce a large number of peptides[7] suitable for screening epitopic regions of viral proteins. However, an appropriate reagent for N-glycosylated peptide synthesis on solid support has not been available until recently, so the presence and importance of sugar moieties on epitopic regions of glycoproteins (for example rabies G[8]) were only documented indirectly. A "glycosylating reagent" candidate should (i) be an amino-terminal protected Asn residue carrying at least two O-glycosidically linked N-acetyl-glucosamine (GlcNAc) moieties (chitobiose); (ii) couple with high efficiency to keep the final purification simple; (iii) be unaffected by the following repetitive deprotection and coupling steps; and (iv) survive the final cleavage from the resin.

We demonstrated earlier[9] that Fmoc-Asn(GlcNAc)-OH fulfills the second requirement only if the carbohydrate hydroxyls are unprotected. As illustrated by successful experiments with minimal side-chain protection (leaving the hydroxyl group of serine free) in the synthesis of fragments of ribonuclease T_1 in solution[10] and of solid-phase synthesis of phosphopeptides,[11] we have also shown that neither the pentafluorophenyl esters nor the symmetrical anhydrides of amino acids acylate the unprotected hydroxyl groups of sugar moieties during solid-phase peptide synthesis. Fmoc amino terminal protection was used earlier in synthesis of O-glycopeptides on solid support.[12] Trifluoroacetic acid (TFA) is used for final deprotection with concomitant release

of the peptide from the resin, but the acid lability of certain O-glycosidic bonds (especially fucosidic) limits the applicability of the procedure, particularly if no sugar-protection is employed.[13] Unfortunately the alternative Pd sensitive resins[14] are not easy to work with and variable yields of detached peptides are obtained.

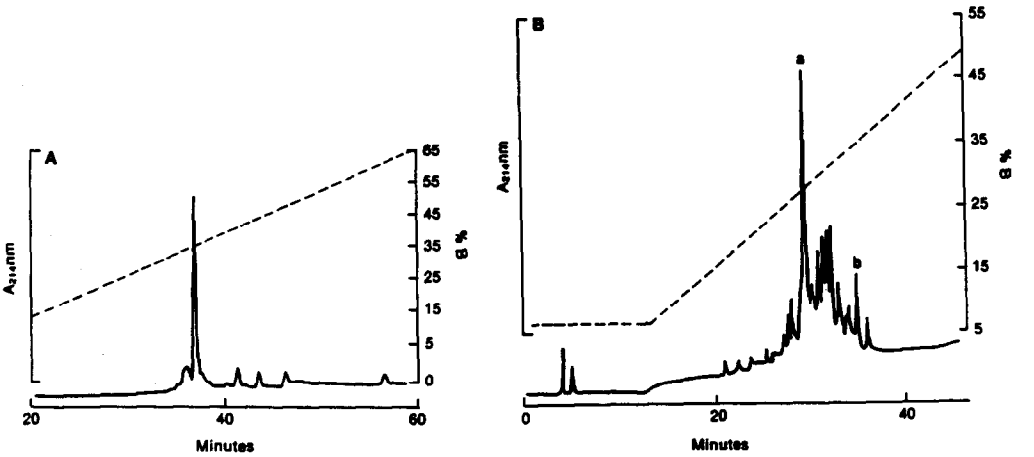
Because of limitations of the previously established procedures, we developed a method of synthesizing N^{α}-Fmoc-N^{β}-[2-acetamido-2-deoxy-4-O-(2-acetamido-2-deoxy-β-D-glucopyranosyl)-β-D-glucopyranosyl]-asparagine [Fmoc-Asn(GlcNAc-1→4-GlcNAc)-OH]. We tested its stability during TFA treatment, determined the coupling efficiency with this residue on a test peptide and synthesized T cell epitopic dodecapeptides corresponding to residues 312-323 of rabies virus glycoprotein, without glycosylation and glycosylated with GlcNAc and the chitobiose derivative for later comparative immunological and conformational studies.

Fmoc-Asn(GlcNAc-1→4-GlcNAc)-O^{t}Bu: 250 mg (0.61 mmol) Fmoc-Asp-O^{t}Bu and 112 mg (0.61 mmol) pentafluorophenol was dissolved in 2 ml of N,N-dimethylformamide (DMF) and 126 mg (0.61 mmol) N,N'-dicyclohexylcarbodiimide was added. After 30 min, 200 mg crude 1-amino-chitobiose[15] in 1.5 ml H_2O-DMF (1:2; v/v) mixture was added. The reaction mixture was stirred overnight and the insoluble material was filtered off. After removal of the solvent *in vacuo*, the product was washed with cold ether and cold water. +FAB-MS: M+Na = 839. RP-HPLC[16]: 43.6 min. TLC: R_F[17] = 0.3 (solvent a). Then 5 ml TFA was added to the ^{t}Bu ester for 15 min and evaporated. The resulting Fmoc-Asn(GlcNAc-1→4-GlcNAc)-OH was triturated with ether. Yield: 232 mg (50%). FAB-MS: M+H= 761, M+Na= 783. Notably, no M=558 [Fmoc-Asn(GlcNAc)-OH] peak was detected. The absence of this peak is an indication of the stability of the chitobiosyl amino acid in TFA. TLC: R_F[17] = 0.1 (solvent a), 0.4 (solvent b). ^{1}H-NMR (ppm, in D_2O at 310 K): Asn Cα: δ=4.01, t1H, J=9.16; Glc_A H_1: δ=4.22, d1H, J=5.49; Glc_B H_1: δ=4.12, d1H, J=5.49; Fmoc H_1: δ=5.52, s1H; Fmoc H_{ring}: δ=7.4-6.4, m8H; Ac_A: δ=1.674, s3H; Ac_B: δ=1.671, s3H.

Coupling efficiency of chitobiosyl amino acid to resin-bound glycine[9] was determined on the basis of HPLC peak area integration values.[9] Single coupling overnight with a 6 molar excess of preformed pentafluorophenyl ester (Pfp; for general preparation see[18]) in DMF resulted in 78% of chitobiosyl amino acid incorporation compared to 84% coupling efficiency obtained with the appropriate GlcNAc derivative. The slight decrease in coupling efficiency reflects the increased steric hindrance (this steric hindrance would be more pronounced if protected carbohydrate was used).

Double couplings with Fmoc-Asn-OPfp, Fmoc-Asn(GlcNAc)-OPfp and Fmoc-Asn(GlcNAc-1→4-GlcNAc)-OPfp (synthesis 1, 2 and 3, respectively, in the table below) were used to produce a synthetically difficult set of T cell epitopic dodecapeptides, GM 12, H-Gly-Lys-Ala-Tyr-Thr-Ile-Phe-Asn*-Lys-Thr-Leu-Met-NH_2, containing glycosylated amino acid at the proposed site (*). Earlier studies indicated that the biological activity of the peptide depends upon the presence of the carbohydrate side-chain.[8] Standard Fmoc synthetic protocol[19] was used on a SAM 2 automated peptide synthesizer. The rest of the amino acids were coupled for 2 hrs with a 6 molar excess of symmetrical anhydrides in a dichloromethane-DMF mixture and repetitive amino-terminal deprotection was achieved with 30% piperidine in DMF-toluene. Cleavage of peptides from the solid support was accomplished with 95% TFA, 5% thioanisole for 1 hr. Peptides were purified on RP-HPLC.[16] Peptides were analyzed by amino acid analysis, +FAB-MS and ^{1}H-NMR.[20]

Synthesis	Peptide	Retention time (min)[16]	FAB-MS
1	GM 12	31.2	M=1386
2	GM 12-GlcNAc	30.2	M=1589
3	GM 12-Chitobiose	29.8	M=1792



Panel A shows the RP-HPLC[16] profile of Fmoc-Asn(GlcNAc-1→4-GlcNAc)-OH after an additional hr in TFA. **Panel B** shows the RP-HPLC profile of the crude reaction mixture of synthesis 3. Peak **a** is the product, peak **b** is the Asn deleted peptide. Remarkably, double couplings resulted in 80% chitobiosylamino acid incorporation and no significant peak is found at 30.2 min, suggesting the stability of the O-glycosidic bond during the conditions of the solid-phase synthesis and cleavage.

We have presented an efficient way of synthesizing N-glycopeptides that contain the first two sugar moieties of oligosaccharide structure connected with an O-glycosidic bond which is found as a common starting structure in N-glycoproteins. The chitobiosyl amino acid can be prepared easily in two simple steps. The synthetic strategy allows the screening of large numbers of glycopeptide sequences.

Acknowledgements: The authors wish to thank Dr. Sheldon L. Miller and Shirley Peterson (editor) for their critical review of the manuscript. This work was supported by grants GM 45011 (to L.O.) and AI 28679 (to J.T.) from the National Institutes of Health, a grant from the W. W. Smith Charitable Trust, Philadelphia (to J.T.) and a grant from the National Scientific Research Foundation, Hungary (to M.H.).

References and Footnotes:

1. U. Fuhrmann, E. Bause and H. Ploegh, *Biochim. Biophys. Acta, 825*, 95 (1985).
2. R.A. Gruters, J.J. Neefjes, M. Tersmette, R.E. DeGoede, A. Tulp, H.G. Huismann, F. Miedema and H.L. Ploegh, *Nature, 330*, 74 (1987).
3. D.C. Montefiori, W.E. Robinson and W.M. Mitchell, *Proc. Natl. Acad. Sci. U.S.A., 85*, 9248 (1988).

4. N. Kornfeld and S. Kornfeld, *Annu. Rev. Biochem., 54*, 631 (1985).
5. M.C. Rose, W.A. Woter, H. Sage, C.F. Brown and B. Kaufman, *J. Biol. Chem., 259*, 3167 (1984).
6. E. A. Kabat, In: Structural Concepts in Immunology and Immunochemistry, Holt, Rinehart and Winston, New York, 1968, p. 820.
7. G. Barany, N. Kneib-Cordonier and D.G. Mullen, *Int. J. Pept. Protein Res., 30*, 705 (1987).
8. R.I. Macfarlan, B. Dietzschold, T. Wiktor, M. Kiel, R. Houghten, R.A. Lerner, J.R. Sutcliffe and H. Koprowski, *J. Immunol., 133*, 2748 (1984).
9. L. Otvos, Jr., K. Wroblewski, E. Kollat, A. Perczel, M. Hollosi, G.D. Fasman, H.C.J. Ertl and J. Thurin, *Pept. Res., 2*, 362 (1989).
10. H.T. Storey, J. Beacham, S.F. Cernosek, F.M. Finn, C. Yanaihara and K. Hofmann, *J. Amer. Chem. Soc., 94*, 6170 (1972).
11. L. Otvos, Jr., I. Elekes and V.M.-Y. Lee, *Int. J. Pept. Protein Res., 34*, 129 (1989).
12. V.H. Paulsen, K. Adermann, G. Merz, M. Schultz and U. Weichert, *Starke/Starch, 40*, 465 (1988).
13. H. Kunz, H. Waldmann and J. Marz., *Liebigs Ann. Chem., 1989*, 45.
14. H. Kunz and B. Dombo, *Angew. Chem. Int. Ed. Engl., 27*, 711 (1988).
15. Peracetyl chitobiose was dissolved in a methanol-water mixture and was saturated with NH_4HCO_3. After standing 3 days at 30°C the solvent was evaporated, and the residue was dissolved in saturated aqueous NH_4HCO_3 and was kept for 10 days. Water was evaporated from the reaction mixture 10 times followed by evaporation of ethanol. The product was used for glycoamino acid synthesis without further purification. (Free chitobiose and acetamide were detected on TLC as non-interacting contaminants.)
16. Reversed-phase chromatographic (RP-HPLC) conditions: Column: Altex Ultrasphere ODS, flow rate: 3 ml/min, detection: 214 nm; 0.1 AUFS, solvent A: 0.1% TFA; solvent B: 0.1% TFA in CH_3CN, gradient: 1.33%/min B increase starting from 5% at 13 min.
17. Thin-layer chromatography: Ethyl acetate:(pyridine:acetic acid:water = 20:6:11; v/v/v) = 3:2 (a) and 2:3 (b).
18. D. Hudson, *J. Org. Chem., 53*, 617 (1988).
19. G.B. Fields and R.L. Noble, *Int. J. Pept. Protein Res., 35*, 161 (1990).
20. Two-dimensional DQF-COSY ^{1}H-NMR in D_2O at 310K (GM12:1; GM12-GlcNAc:2; GM12-chitobiose:3):

Amino acid	α			β			γ			δ			ε			ζ		
	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
Gly	3.70	3.77	3.69															
Lys(2)	4.22	4.22	4.22	1.62	1.63	1.63	1.30	1.30	1.30	1.60	1.62	1.62	2.89	2.89	2.89			
Ala	4.25	4.25	4.25	1.28	1.28	1.28												
Tyr	4.56	4.56	4.56	2.57 2.78	2.57 2.78	2.57 2.78				7.02	7.02	7.02	6.72	6.71	6.71			
Thr(2)	4.20 4.18	4.22 4.18	4.21 4.18	3.95 4.12	3.95 4.09	3.95 4.08	0.98 1.10	0.99 1.15	0.99 1.15									
Ile	4.07	4.04	4.03	1.70	1.71	1.71	0.76	0.77	0.77	1.03 1.29	1.03 1.28	1.03 1.28	0.77	0.78	0.78			
Phe	4.50	4.50	4.50	2.92 3.03	2.92 3.03	2.92 3.03				7.15	7.15	7.15	7.20	7.20	7.20	7.28	7.28	7.28
Asn	4.17	4.16	4.15	1.68 1.72	1.70 1.80	1.70 1.80												
Leu	4.26	4.27	4.27	1.51	1.52	1.52	1.51	1.52	1.52	0.78 0.82	0.80 0.85	0.79 0.84						
Met	4.35	4.37	4.37	1.92 2.05	1.92 2.05	1.92 2.05	2.45 2.55	2.45 2.54	2.45 2.55	1.99	2.03	2.03						

Carbohydrate signals (δ) in GM 12-GlcNAc (1) and in GM 12-Chitobiose (A,B):

	1	2	3	4	5	6a	6b	Ac
1:	4.97; J=7.69	3.74	3.53	3.42	3.40	3.68	3.77	1.93
A:	4.95	3.77	3.68	3.48	3.41	3.83	3.69	1.98
B:	4.50	3.67	3.57	3.44	3.59	3.73	3.57	2.03

(Received in USA 2 August 1990)