Tetrahedron Letters, Vol.31, No.48, pp 6991-6994, 1990
Printed in Great Britain

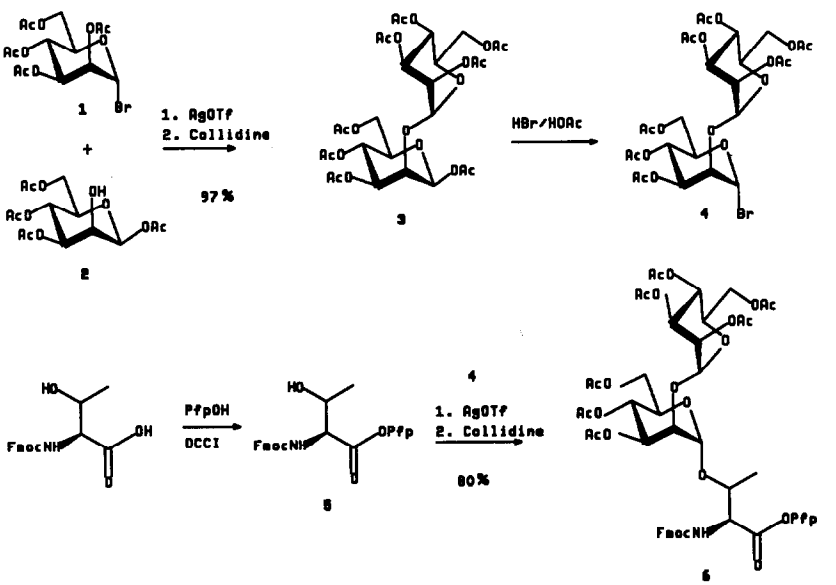
0040-4039/90 $3.00 + .00
Pergamon Press plc

THE ACTIVE ESTER N-Fmoc-3-O-[Ac_4-α-D-Manp-(1→2)-Ac_3-α-D-Manp-1-]-THREONINE-O-Pfp AS A BUILDING BLOCK IN SOLID-PHASE SYNTHESIS OF AN O-LINKED DIMANNOSYL GLYCOPEPTIDE.

Anita M. Jansson, Morten Meldal* and Klaus Bock
Department of Chemistry, Carlsberg Laboratory, Gamle Carlsberg Vej 10, DK-2500 Valby, Copenhagen, Denmark.

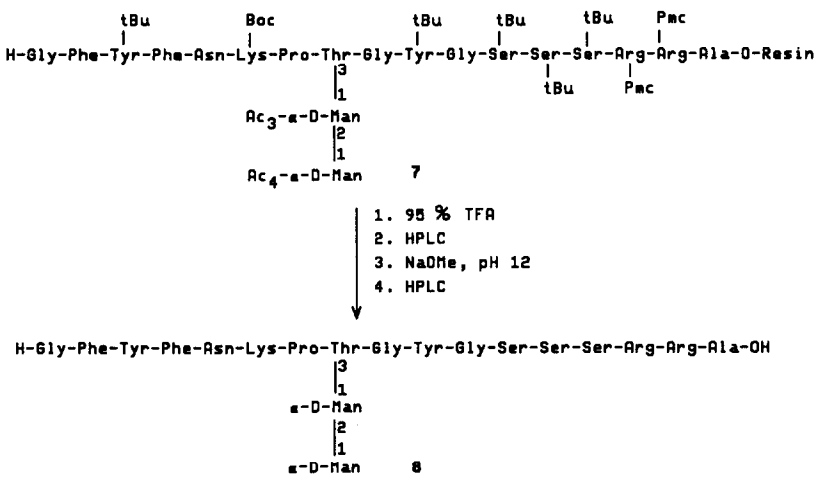
Abstract: The active ester N^{α}-Fmoc-3-O-[Ac_4-α-D-Manp-(1→2)-Ac_3-α-D-Manp-1-]-Thr-O-Pfp (6) was synthesized by direct condensation of Ac_4-α-D-Manp-(1→2)-Ac_3-α-D-Manp-Br (4) with Fmoc-Thr-O-Pfp (5) and used as a building block in an automated continuous-flow solid-phase synthesis of an O-glycosylated heptadeca amino acid fragment of the insulin-like growth factor 1 (IGF-1).

Glycopeptides are biopolymers which contain one or more carbohydrate chains linked covalently to a peptide backbone. In O-glycosylated peptides the oligosaccharides play an important role for the biological activity and development of efficient techniques for the chemical synthesis of model compounds and analogues of glycopeptides is of considerable interest. O-glycosylated peptides may be synthesized by use of suitable protected and glycosylated hydroxy amino acids (e.g. threonine or serine) as building blocks in a solid-phase peptide synthesis[1-5]. However, since O-glycosyl-β-hydroxy carboxyl acid derivatives are sensitive to both strong acid (bond cleavage and anomerization) and strong base[6] (β-elimination and racemization) suitable



acid- and base labile protecting groups are required. Protection of the α-amino group with the base labile group fluoren-9-ylmethoxycarbonyl (Fmoc) has been reported to be attractive in this respect due to its sensitivity to mild organic bases such as morpholine[4,7]. For temporary protection of the α-carboxyl group the pentafluorophenyl (Pfp) esters have several advantages since they serve the dual purpose of protecting the carboxylic acid during glycosylation and at the same activating the carboxyl group for the subsequent amide bond formation[1,8] thus avoiding elaborate orthogonal protection schemes. A suitable choice of protecting groups for the hydroxyl groups of the disaccharide has proved to be acetates, which easily can be removed by treatment with sodium methoxide in methanol[2,9]. This strategy has been successfully employed in the solid-phase synthesis of an O-glycosylated heptadeca amino acid peptide fragment of the insulin-like growth factor 1 (IGF-1), which, in addition to the native non-glycosylated form, recently has been isolated from yeast (*Saccharomyces cerevisiae*) and characterized[10]. The glycosylation site of IGF-1 has been determined to be threonine-29[10] and the synthesized fragment corresponds to amino acids 22 to 38.

Silver triflate-promoted reaction of 2,3,4,6-tetra-O-acetyl-α-D-mannopyranosyl bromide (1) with 1,3,4,6-tetra-O-acetyl-β-D-mannopyranose[11] (2) in dry dichloromethane at -40°C for 40 min followed by addition of 2,4,6-collidine afforded peracetylated α-D-Man*p*-(1→2)-β-D-Man*p* (3) in 97% isolated yield. Compound 3 was transformed into the corresponding α-bromide 4 by treatment with 10 eqv. hydrogen bromide in acetic acid and dichloromethane. The active ester Fmoc-Thr-O-Pfp[12] (5) was prepared by diisopropyl carbodiimide promoted esterification of Fmoc-Thr-OH with pentafluorophenol (Pfp-OH) in THF followed by purification on silica gel under dry conditions. The dimannosyl bromide 4 was reacted with 5 for 8 hours at -30 to -40°C



in dry dichloromethane promoted by silver triflate. After neutralization with 2,4,6-collidine, warming to room temperature and chromatography on silica gel under dry conditions Fmoc-3-O-[Ac_4-α-Man*p*-(1→2)-Ac_3-α-D-Man*p*-1]-Thr-O-Pfp[13] (**6**) was obtained in 80% yield. The long reaction time may be due to influence of the Pfp group, and a similar observation was made in the glycosylation reaction of the primary group in Fmoc-Ser-O-Pfp[1]. Compound **6** was subsequently used as a building block in the automated continuous flow solid-phase synthesis[14] of the heptadeca amino acid fragment of IGF-1. The first amino acid, alanine, was coupled to the resin via the symmetrical anhydride, $(FmocAla)_2O$, by catalysis of 4-N,N-dimethylaminopyridine (DMAP), and the protected threonine derivative **6** was coupled to the peptide chain with 3,4-dihydro-3-hydroxy-4-oxo-1,2,3-benzotriazine (Dhbt-OH) as an auxiliary nucleophile. All other amino acids were added as suitable protected Dhbt-esters[15]. Deprotection of the α-amino groups during synthesis was effected by treatment with 50% morpholine in DMF for 20 min and the acylation times were determined with a solid-phase spectrophotometer by measuring of the yellow colour due to formation of an ion pair between Dhbt-OH and unreacted amino groups[14b].

By treatment of the protected resin bound glycopeptide **7** with 95% trifluoroacetic acid the O-glycosylated peptide was cleaved from the resin with simultaneous removal of the acid labile protecting groups. The crude peptide was purified by preparative reversed-phase HPLC. O-Deacetylation was performed with sodium methoxide in methanol at pH 12 for 30 min at room temperature (the reaction time was less than 10 min according to HPLC) followed by neutralization with carbon dioxide. (No β-elimination or racemization could be observed by NMR). The deprotected glycopeptide, **8** (25 mg), was obtained after preparative HPLC. Characterisation of **8**[16] was performed by 1D- and 2D-NMR spectroscopy and amino acid analysis confirmed the amino acid composition. Investigations regarding the conformation of compound **8** are in progress. The corresponding non-glycosylated peptide fragment was also prepared and characterized for comparative conformational studies regarding the influence of the sugar moiety on the conformation of the peptide chain.

References and notes.

1. M. Meldal and K. J. Jensen, J.Chem.Soc.,Chem.Commun. 483 (1990)
2. B. Lüning, T. Norberg and J. Tejbrant, J.Chem.Soc.,Chem.Commun, 1267 (1989)
3. S. Lavielle, N. C. Ling, R. Saltman and R. C. Guillemin, Carbohydr.Res. **89**, 229 (1981)
4. H. Paulsen, G. Merz and U. Weichert, Angew.Chem.Int.Ed.Engl. **27**, 1365 (1988)
5. a) B. Scolaro, L. Biondi, F. Filira, E. Peggion and R. Rocchi, Proc.20th Eur.Pept.Symp.,Tübingen, FRG (1988), ed. G. Jung and E. Bayer, de Gruyter, Berlin, 322 (1989) b) F. Filira, L. Biondi, F. Cavaggion, B. Scolaro and R. Rocchi, Int.J.Peptide Protein Res. **36**, 86 (1990)
6. K. Wakabayashi and W. Pigman, Carbohydr.Res. **35**, 3 (1974)
7. a) H. Kunz, S. Birnbach and P. Wernig, Carbohydr.Res. **202**, 207 (1990) b) H. Kunz and B. Dombo, Angew.Chem.Int.Ed.Engl. **27**, 711 (1988)
8. E. Atherton, L. R. Cameron and R. C. Sheppard, Tetrahedron, **44**, 843 (1988)
9. H. Paulsen and J.-P. Hölck, Carbohydr.Res. **109**, 89 (1982)

10. P. Gellerfors, K. Axelsson, A. Helander, S. Johansson, L. Kenne, S. Lindqvist, B. Pavlu, A. Skottner and L. Fryklund, J.Biol.Chem. **264**, 11444 (1989)

11. J. O. Deferrari, E. G. Gros and I. O. Mastronardi, Carbohydr.Res. **4**, 432 (1967)

12. I. Schön and L. Kisfaludy, Synthesis, 303 (1986)

13. Compound **6**. 500 MHz ^{1}H-NMR; δ,ppm (in $CDCl_3$, ref $CHCl_3$=7.30 at 300 K) J,Hz: Ac_4-α-D-Man*p*-(1→2) 4.95 (br s,H-1); 5.29 (d,3 Hz,H-2); 5.41 (dd,3 and 10 Hz,H-3); 5.35 (H-4); 4.05 (H-5); 4.28 (dd,5 and 12 Hz,H-6); 4.20 (H-6'); Ac_3-α-D-Man*p*-(1→Thr) 5.15 (br s,H-1); 3.99 (H-2); 5.31 (H-3); 5.33 (H-4); 4.10 (H-5); 4.01 (dd,3 and 12 Hz,H-6); 4.21 (H-6'); 2.0-2.2 (Ac); Thr 4.87 (dd,2 and 10 Hz,H-α); 4.57 (H-β); 1.45 (d,6 Hz,H-γ); 5.75 (d,10 Hz,NH); Fmoc 4.32 (t,7 Hz,CH); 4.53 (CH_2); 7.35-7.83 (8 H-arom.)
125.77 MHz ^{13}C-NMR; δ,ppm (in $CDCl_3$=77.00 at 300 K): Ac_4-α-D-Man*p*-(1→2) 99.11 (C-1); 69.74 (C-2); 68.39 (C-3); 66.28 (C-4); 69.23 (C-5); 62.26 (C-6); Ac_3-α-D-Man*p*-(1→Thr) 100.00 (C-1); 76.74 (C-2); 69.68 (C-3); 66.38 (C-4); 69.43 (C-5); 62.26 (C-6); Ac 20.51- 20.76 (CH_3); 169.34-169.68 (CO); Thr 58.60 (C-α); 76.88 (C-β); 18.18 (C-γ); Fmoc 47.15 (CH); 67.67 (CH_2); 120.04,125.07,127.12,127.81,141.35 (C-arom.);(signals from OCONH and C-Pfp not detected due to low S/N ratio.)

14. a) A. Dryland and R. C. Sheppard, J.Chem.Soc.Perkin Trans 1, 125 (1986) b) L. R. Cameron, J. L. Holder, M. Meldal and R. C. Sheppard, J.Chem.Soc.Perkin Trans 1, 2895 (1988)

15. E. Atherton, J. L. Holder, M. Meldal, R. C. Sheppard and R. M. Valerio, J.Chem.Soc.Perkin Trans.1, 2887 (1988)

16. Compound **8**. 500 MHz ^{1}H-NMR; δ,ppm (in D_2O, ext ref dioxane=3.76 at 310 K and pD=7.35) J,Hz: α-D-Man*p*-(1→2) 4.94 (d,2 Hz,H-1); 4.02 (dd,2 and 3.5 Hz,H-2); 3.79 (dd,3.5 and 9.5 Hz,H-3); 3.63 (dd,9.5 and 9.5 Hz,H-4); 3.71 (H-5); 3.70-3.75 (H-6); 3.82-3.90 (H-6'); α-D-Man*p*-(1→Thr) 5.10 (d,1 Hz,H-1); 3.83 (H-2); 3.89 (H-3); 3.63 (dd,9.5 and 9.5 Hz,H-4); 3.71 (H-5); 3.70-3.75 (H-6); 3.82-3.90 (H-6'); H-Gly 3.53 (d,16.5 Hz,H-α); 3.57 (d,16.5 Hz,H-α'); Phe.s 4.50 (H-α1); 4.47 (H-α2); 2.92 and 3.05 (dd,6.5 and 14 Hz,H-β1); 2.78 and 2.90 (dd,8 and 14 Hz,H-β2); 7.10,7.19,7.26,7.32 (H-arom.); Tyr.s 4.53 (2H-α); 2.82,2.89, 2.91 and 3.05 (H-β); 6.77,6.82,7.01,7.09 (H-arom.); Asn 4.58 (dd,6 and 7.5 Hz,H-α); 2.62 (dd,7.5 and 15.5 Hz,H-β); 2.71 (dd,6 and 15.5 Hz,H-β'); Lys 4.32 (H-α); 1.73 (2H-β); 1.38 (2H-γ); 1.61 (2H-δ); 2.91 (2H-ε); Pro 4.51 (H-α); 1.91 (H-β); 2.25 (H-β'); 1.98 (2H-γ); 3.59 (H-δ); 3.71 (H-δ'); Thr 4.42 (d,3.5 Hz,H-α); 4.23 (dd,3.5 and 6.5 Hz,H-β); 1.24 (d,6.5 Hz,H-γ); Gly.s 3.94 (4H-α); Ser.s 4.49 (H-α1); 4.45 (H-α2); 4.43 (H-α3); 3.92 and 3.86 (H-β1); 3.89 and 3.84 (H-β2); 3.85 (2H-β3); Arg.s 4.49 (H-α1); 4.32 (H-α2); 1.60 (H-β1); 1.72 (H-β1'); 1.83 (2H-β2); 1.61 (H-γ1); 1.57 (H-γ2); 3.16 (4H-δ); Ala 4.09 (q,7.3 Hz,H-α); 1.30 (d,7.3 Hz,H-β).
125.77 MHz ^{13}C-NMR; δ,ppm (in D_2O,ext ref dioxane=67.40 at 310K and pD=7.35): α-D-Man*p*-(1→2) 103.30 (C-1); 70.99 (C-2); 71.39 (C-3); 67.79[a] (C-4); 74.15 (C-5); 61.94[b] (C-6); D-Man*p*-α-(1→Thr) 100.49 (C-1); 80.17 (C-2); 70.91 (C-3); 68.03[a] (C-4); 74.15 (C-5); 62.02[b] (C-6); Gly.s 42.55,43.25,43.53 (C-α); Phe.s 52.37,55.64[c] (C-α); 38.20,40.18[d] (C-β); 128.13,128.91,129.00,129.69,129.73,130.07,130.15,137.12,137.12(C-arom.); Tyr.s 55.78,56.36[c] (C-α); 37.19[d] (2C-β); 116.39,116.47,131.45,131.50,155.46,155.53(C-arom.); Asn 51.15 (C-α); 37.90[d] (C-β); Lys 54.21 (C-α); 31.03[e] (C-β); 22.75 (C-γ); 27.32[e] (C-δ); 41.49[d] (C-ε); Pro 61.26 (C-α); 30.21[e] (C-β); 25.63 (C-γ); 48.90 (C-δ); Thr 58.96 (C-α); 76.73 (C-β); 18.59 (C-γ); Ser.s 56.61,56.67[f] (C-α); 61.94,62.02[b] (C-β); Arg.s 54.21,55.64[c] (C-α); 29.10[e] (2C-β); 25.16,25.26 (C-γ); 41.49 (2C-δ); Ala 51.98 (C-α); 18.48 (C-β); Carbonyls 171.76, 172.38, 172.49, 172,70, 173.06, 173.28, 174.00, 174.74, 175.34.
a,b,c,d,e,f: The assignments can be interchanged

(Received in Germany 31 August 1990)