Pergamon

Tetrahedron Letters, Vol. 36, No. 50, pp. 9205-9208, 1995
Elsevier Science Ltd
Printed in Great Britain
0040-4039/95 $9.50+0.00

0040-4039(95)01941-3

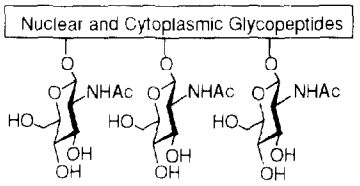
Efficient Synthesis of *O*-(2-Acetamido-2-deoxy-β-D-glucopyranosyl)-Ser/Thr Building Blocks for SPPS of *O*-GlcNAc Glycopeptides

Ernst Meinjohanns, Morten Meldal and Klaus Bock*

Department of Chemistry, Carlsberg Laboratory, Gamle Carlsberg Vej 10, DK-2500 Valby, Copenhagen, Denmark

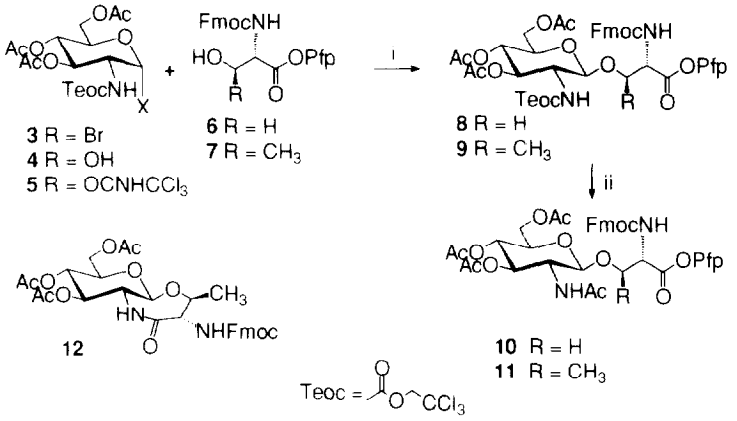
Abstract: *Suitably protected building blocks for solid-phase synthesis of β-O-GlcNAc glycopeptides. N^{α}-Fmoc-Ser(Ac$_3$-β-D-GlcNAc)-OPfp* ***10*** *and N^{α}-Fmoc-Thr(Ac$_3$-β-D-GlcNAc)-OPfp* ***11*** *have been synthesized by stereoselective glycosylation of N^{α}-Fmoc-Ser-OPfp* ***6*** *and N^{α}-Fmoc-Thr-OPfp* ***7****, respectively, with the 2-trichloroethoxycarbonylamino (Teoc) glycosyl donors* ***3*** *and* ***5****, followed by in situ reduction of the Teoc-group and simultaneous N-acetylation using zinc dust in tetrahydrofuran/acetic anhydride/acetic acid (3:2:1).*

Recently, much attention has been focused on a new form of protein glycosylation.[1] A wide variety of nuclear and cytoplasmic proteins, including several structural proteins and transcription factors, are modified by the addition of single *N*-acetylglucosamine residues β-glycosidically linked to the hydroxy side chains of serine and threonine. This novel type of post-translational glycosylation is both an abundant and transient modification. The degree of glycosylation of individual proteins is often modulated during the cell cycle or in response to specific physiological stimuli. The addition of *O*-GlcNAc appears to be highly dynamic and it has been postulated that this modification plays a regulatory role in many ways analogous to protein phosphorylation.[2] The increasing evidence of the importance of this post-translational glycosylation has stimulated substantial efforts towards efficient synthesis of *O*-GlcNAc glycopeptides to delineate their biological functions.



The most efficient and reliable approach to the synthesis of *O*-glycopeptides is the use of suitably protected *O*-glycosylated serine and threonine amino acids as building blocks in the stepwise assembly of glycopeptides.[3,4] Most previous published approaches towards glycosylated amino acid building blocks required multistep procedures, e.g. exchange of protecting groups used for the α-amino group and/or the carbohydrate hydroxyl groups and activation of the carboxylic group after selective removal of the protecting group both lowering the overall yield. An alternative strategy involving direct glycosylation of the active esters derivatives N^{α}-Fmoc-Ser-OPfp **6** and N^{α}-Fmoc-Thr-OPfp **7** has been developed which alleviates the manipulation of protecting groups in the glycosyl amino acids and allows direct use in peptide synthesis.[5] The pentafluorophenyl(Pfp) ester serves as a protective group during glycosylation and as an activating group during the peptide-bond formation. The synthesis of 2-acetamido-2-deoxy-β-D-glucopyranosides of serine or threonine requires glycosyl donors containing participating protective groups in the 2-position. Depending on the 2-amino protecting group, particular problems have been observed in the synthesis of the desired building blocks. Glycosylation with donors derived from *N*-acetylglucosamine or the oxazoline method require multistep procedures. Furthermore, the reactions require high temperatures, strong acidic conditions and the yields are often very poor. These conditions are incompatible with the temperature sensitive Pfp esters. Therefore, various glucosamine donors possessing modified amino

functionalities have been investigated. In three previous papers[6-8] we have described alternative strategies which simplifies the direct glycosylation of the active esters derivatives **6** and **7** by employing the *N*-allyloxycarbonyl-(Aloc)[9] and the *N*-dithiasuccinoyl (Dts)[7] amino protective groups. The desired *N*-acetyl group can than be generated alternatively after glycopeptide assembly on solid phase (*N*-Dts) or prior to its incorporation onto a peptide (*N*-Aloc).[8] The synthesis of *O*-GlcNAc building blocks has been simplified by performing direct glycosylation of the N^{α}-Fmoc-Ser and N^{α}-Fmoc-Thr amino acids.[7]



Scheme: Synthesis of the building blocks **10** and **11**. *Reagents and conditions*: i. AgOTf, molecular sieves, CH_2Cl_2, (-40°C). ii. Zn in THF/Ac_2O/HOAc (3:2:1).

In the present work an even more efficient synthesis of the *O*-linked GlcNAc serine and threonine building blocks **10** and **11** is described. The procedure relies upon the use of the *N*-trichloroethoxycarbonyl (Teoc)[10] for the protection of the amino group. The *N*-Teoc group has a number of advantages that make it a useful protecting group in the assembly of these building blocks: i). the neighboring group participation of *N*-Teoc leads to the desired β-glycosides ii). the undesired oxazoline formation is prevented during glycosylation and iii). the Teoc group can be easily removed under mild conditions in a chemospecific manner compatible with the sensitive Pfp ester.

The synthesis of the *N*-Teoc glycosyl donors **3** and **5** was accomplished by described procedures.[11,12]

Glycosyl donor **3** is readily available from glucosamine hydrochloride without chromatographic purification. Thus, the product of the reaction of glucosamine hydrochloride with Teoc-Cl is crystallized and after quantitative acetylation it is converted into **3** with HBr in acetic acid. The subsequent silver trifluoromethansulfonate (AgOTf) mediated glycosylation of N^{α}-Fmoc-Ser-OPfp **6** and N^{α}-Fmoc-Thr-OPfp **7** with the glycosyl bromide **3** was carried out in dichloromethane at -40°C. The desired β-glycosides **8** and **9** were obtained stereoselectively in isolated yields of 85 and 75 %, respectively, after silica gel purification.[13] The easy synthesis of **8** and **9** makes this procedure a valuable alternative to previous procedures. In our previous studies[7,8] of direct glycosylations of the amino acids **6** and **7** we found the use of the trichloroacetimidate glycosylation method superior to the glycosyl bromide approach. Therefore, the imidate **5**[12] was treated with the serine and threonine Pfp esters **6** and **7** in the presence of AgOTf at room temperature and the desired β-glycosides **8** and **9** were furnished stereoselectively in 87 and 90 % yield, respectively, following silica gel purification.[13] It has already been established that the Teoc group can easily be removed with zinc in acetic acid.[14] On the assumption that acetic anhydride is a better acylating reagent than the Pfp ester, we tested the conversion of the *N*-Teoc protected glycosyl amino acids **8** and **9** into N^{α}-Fmoc-Ser(Ac_3-β-D-GlcNAc)-OPfp **10** and N^{α}-Fmoc-Thr(Ac_3-β-D-GlcNAc)-OPfp **11** by reductive cleavage of the Teoc-group using zinc dust in tetrahydrofuran/acetic acid in the presence of acetic anhydride (3:1:2).[15] The reaction is complete within 30 min. The *O*-GlcNAc building blocks **10** and **11** were purified by chromatography on silica gel and subsequently crystallized from diethyl ether to afford compounds **10** and **11** in 75 and 78 % yield, respectively. During the conversion of **9** to **11** the formation of the cyclic lactam **12**[16] as a minor side product (less than 5%) was observed, however no equivalent lactam formation during the synthesis of **10** was detected.

In conclusion, the *N*-Teoc glucosamine donors **3** and **5** in combination with AgOTf activation affords excellent yields of the desired β-glycosides **8** and **9**. The "one step" *N*-Teoc deprotection and *N*-acetylation can be readily accomplished via reduction with zinc dust in tetrahydrofuran/acetic anhydride/acetic acid (3:2:1), minimizing side

reactions (lactam formation). The simplicity of this method allows the synthesis of the two building blocks **10** and **11** on a gram scale with only four synthetic steps yielding products that are active esters which can be used directly for the synthesis of *O*-GlcNAc glycopeptides on solid phase.

Acknowledgements: This work was supported by the EC Program (M.M., grant SCI*-CT92-0765).

References and Notes:

1 For papers on *O*-GlcNAc glycopeptides see: a). C.-R. Torres and G. W. Hart, *J. Biol. Chem.*, 1984, **259**, 3308 b). R. S. Haltiwanger, W. G. Kelly, E. P. Roquemore, M. A. Blomberg, L.-Y. D. Dong, L. Kreppel, T.-Y. Chou and G. W. Hart, *Biochem. Soc. Trans.*, 1992, **20**, 264 c). G. W. Hart, R. S. Haltiwanger, G. D. Holt and W. G. Kelly, *Ann. Rev. Biochem.*, 1989, **58**, 841 d). K. P. Kearse and G. W. Hart, *Arch. Biochem. Biophys.*, 1991, **290**, 543 e). W. G. Kelly, M. E. Dahmus, and G. W. Hart, *J. Biol. Chem.*, 1993, **268**, 10416 f). G. W. Hart, W. G. Kelly, M. A. Blomberg, E. P. Roquemore, D. L.-Y. Dong, L. Kreppel, T.-Y. Chou, D. Snow and K. Greis, in *Complex Carbohydrates in Drug Research. Structural and Functional Aspects*, ed. K. Bock and H. Clausen, Mundsgaard, Copenhagen, 1993, pp. 280. g). W. Kelly, E. Dahmus and G. W. Hart, *J. Biol. Chem.*, 1993, **268**, 10416 h). D. L.-Y. Dong, Z.-S. Xy, M.R. Cheuvier, R. J. Cotter, D. W. Cleveland and G. W. Hart, *J. Biol. Chem.*, 1993, **268**, 16679

2 B. K. Hayes, K. D. Greis and G. W. Hart, *Anal. Biochem.*, 1995, **228**, 115

3 S. Peters, T. Bielfeldt, M. Meldal, K. Bock and H. Paulsen, *J. Chem. Soc. Perkin Trans. 1*, 1992, 1163

4 a). M. Meldal, *Curr. Opin. Struct. Biol.*, 1994, **4**, 710-718 b). M. Meldal Glycopeptide Synthesis. In Neoglycoconjugates: Preparation and Applications; Y. C. Lee and R. T. Lee Eds.; Academic Press, Inc: San Diego, 1994; pp. 145 c). M. Meldal and K. Bock, *Glycoconjugate J.*, 1994, **11**, 59

5 M. Meldal and K. J. Jensen, *J. Chem. Soc., Chem Commun.*, 1990, 483

6 A. Vargas-Berenguel, M. Meldal, H. Paulsen and K. Bock, *J. Chem. Soc. Perkin Trans. 1*, 1994, 2615

7 E. Meinjohanns, M. Meldal, H. Paulsen and K. Bock, *J. Chem. Soc. Perkin Trans. 1*, 1995, 405

8 E. Meinjohanns, A. Vargas-Berenguel, M. Meldal, H. Paulsen and K. Bock, *J. Chem. Soc. Perkin Trans. 1*, 1995, in press

9 a). G. Arsequell, L. Krippner, R.A. Dwek, S. Y. C. Wong, *J. Chem. Soc., Chem. Commun.*, 1994, 2383 b). L. A. Salvador, M. Elofsson, J. Kihlberg, *Tetrahedron*, 1995, **51**, 5643

10 T. M. Windholz and B. R. Jonston, *Tetrahedron Lett.*, 1967, 2555

11 M. Imoto, H. Yoshimura, T. Shimamoto, N. Sakaguchi, S. Kusumoto and T. Shiba, *Bull. Chem. Soc. Jpn.*, 1987, **60**, 2205-2214

12 H. Paulsen and B. Helpap, *Carbohydr. Res.*, 1991, **216**, 289-313

13 Synthesis of **8** and **9**: *Method A:* The bromide **3** (3g, 5.52 mmol) in dichloromethane (25 ml) was added dropwise over 15 min to a solution of N^{α}-Fmoc-Ser-OPfp **6** (2.45 g, 4.97 mmol) or N^{α}-Fmoc-Thr-OPfp **7** (2.52 g, 4.97 mmol), AgOTf (1.56 g, 6.07 mmol) and molecular sieves (3Å) in dichloromethane at - 40° C in the dark. The mixture was stirred for 2h at - 40° C and than allowed to warm up to room temperature. The mixture was then without further neutralisation filtered through Celite and concentrated. VLC [light petroleum-ethyl acetate (3:1)] yielded the title compounds **8** (4.04 g, 85%) and **9** (3.61 g, 75%).

Method B: Imidate **5** (1 g, 1.60 mmol), compound **6** (715 mg, 1.45 mmol) or **7** (735 mg, 1.45 mmol), AgOTf (452 mg,1.76 mmol) and molecular sieves (3Å) were placed in pre-dried flask in the dark and dry dichloromethane was injected. The solution was stirred overnight, filtered through Celite and concentrated. VLC [light petroleum - ethyl acetate (3:1)] yielded the title compounds **8** (1.21 g, 87%) and **9** (1.26 g, 90%), respectively. *Compound 8*: $[\alpha]_D^{25}$ - 4.4 (c 1.0, $CDCl_3$) $C_{39}H_{34}Cl_3F_5N_2O_{14}$ [MALDI-MS $(M+H)^+$ calcd 957.06,obsd 957.14]. ^1H NMR (250 MHz, $CDCl_3$): 6.08 (1H, FmocNH), 5.54 (1H, NH), 5.30 (1H, H-3), 5.09 (1H, H-4), 4.93 (1H, α-H Ser), 4.74 and 4.64 (2H, CH_2CCl_3), 4.73 (1H, $J_{1,2}$=8.18 Hz H-1), 4.57 and 4.36 (2H, $FmocCH_2$), 4.44 and 4.01 (2H, β- and β'-H Ser), 4.28 (1H, FmocCH), 4.25 and 4.19 (2H, H-6a and H-6b), 3.71 (1H, H-5), 3.68 (1H, H-2), 2.07 (9H, 3 x OAc). ^{13}C NMR (62.9 Mhz, $CDCl_3$): 100.92 (C-1), 74.92 (CH_2CCl_3), 72.38 (C-5), 72.10 (C-3), 68.89 (C-4), 68.88 (β-C Ser), 67.94 (Fmoc CH_2), 62.28 (C-6), 56.50 (C-2), 54.54 (C-α Ser), 47.47 (Fmoc CH), 20.98 (3x OAc). *Compound 9*: $[\alpha]_D^{27}$ - 17.7 (c 0.95, $CDCl_3$) $C_{40}H_{36}Cl_3F_5N_2O_{14}$ [MALDI-MS $(M+H)^+$ calcd 971.09, obsd 971.23]. ^1H NMR (250 MHz, $CDCl_3$): 6.01 (Fmoc NH), 5.33(NH), 5.29 (H-3), 5.12 (H-4), 4.79 and 4.72 (CH_2CCl_3), 4.77 (α-H Thr), 4.72 ($J_{1,2}$=8.07 Hz, H-1), 4.67 (β-H Thr), 4.53 and 4.35 ($FmocCH_2$), 4.31 (FmocCH), 4.26 and 4.10 (H-6a and H-6b), 3.72 (H-5), 3.69 (H-2), 2.09, 2.07 and

2.05 (3 x OAc), 1.36 (γ-H, Thr), ^{13}C NMR (62.9 MHz, $CDCl_3$): 98.44 (C-1), 74.98 (CH_2CCl_3), 73.22 (C-β, Thr), 72.24 (C-5), 71.91 (C-3), 68.74 (C-4), 68.02 ($FmocCH_2$), 62.19 (C-6), 59.05 (C-α Thr), 56.79 (C-2), 47.51 (FmocCH), 21.05, 20.99, 20.83 (OAc), 16.74 (C-γ Thr)

14 R. B. Woodward, K. Heusler, J. Gosteli, P. Naegeli, W. Oppolzer, R. Ramage, S. Ranganathan, H. Vorbrüggen, *J. Am. Chem. Soc.*, 1966, **88**, 852

15 *Typical experimental procedure:* To a stirred solution of compound **8** (2 g, 2.09 mmol) or **9** (2 g, 2.06 mmol) in tetrahydrofuran/acetic anhydride/acetic acid (60 ml, 3:2:1) was added zinc dust (5 g, 325 mesh). After stirring for 2 h, the mixture was filtered through Celite, rinsed several times with freshly distilled THF and concentrated. Purification by chromatography on dried silica gel (eluent: ethyl acetate/light petroleum 2:1) and subsequently crystallisation afforded the crystalline compounds **10** (1.29 g, 75%) and **11** (1.34 g, 78 %), respectively. NMR data, melting points and optical rotations $[\alpha]_D$ are in agreement with those previously reported[6]. A small fraction (less than 5%) of lactam **12** was isolated from the purification of **11** by silica gel chromatography.

16 *Compound **12**:* $[\alpha]_D^{27}$ + 17.5 (c 1.7, $CDCl_3$) $C_{31}H_{34}N_2O_{11}$; [MALDI-MS $(M+H)^+$ calcd 611.62, obsd 611.57) ^{1}H NMR (250 MHz, $CDCl_3$): 6.25 (1H, NH), 5.86 (1H, 5.2 Hz, NHFmoc), 5.14 (1H, H-4), 4.99 (1H, H-3), 4.74 (1H, α-H Thr), 4.69 (1H, 7.73 Hz, H-1), 4.43 and 4.37 (2H, $FmocCH_2$), 4.35 (1H, β-H Thr), 4.21 and 4.11 (2H, H-6a and H-6b), 4.14 (1H, FmocCH), 3.67 (1H, H-5), 3.62 (1H, H-2), 1.31 (3H, 6.02 Hz, γ-H Thr), ^{13}C NMR (62.9 MHz, $CDCl_3$): 96.05 (C-1), 72.97 (C-5), 72.89 (C-3), 71.92 ($FmocCH_2$), 67.59 (C-β Thr), 67.51 (C-4), 62.05 (C-6), 59.15 (C-2), 58.41 (C-α, Thr), 47.55 (FmocCH), 21.14, 20.94, 20.93 (3 x OAc), 12.83 (C-γ, Thr).

(*Received in UK* 5 *September* 1995; *revised* 12 *October* 1995; *accepted* 13 *October* 1995)