Pergamon

*Tetrahedron Letters*, Vol. 37, No. 5, pp. 595-598, 1996

0040-4039/96 $15.00 + 0.00

0040-4039(95)02229-5

# βGal(1-3)GalNAc Block Donor for the Synthesis of TF and α Sialyl(2-6)TF as Glycopeptide Building Blocks

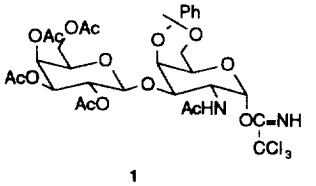
**Dongxu Qiu, Sham S. Gandhi and R. Rao Koganty***

Biomira Inc., 2011-94 Street, Edmonton, Alberta, Canada T6N 1H1

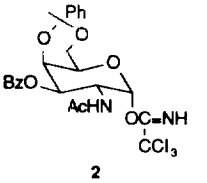
***Abstract:*** 4,6-Benzylidene protected N-acetylgalactosamine β-glycosylated at the 3-position with peracetylated galactose is found to be an excellent donor for the synthesis of Thomsen-Freidenreich (TF) family of antigens. These may serve further as building blocks for the synthesis of mucin derived glycopeptides and as intermediates for further extension to trisaccharides such as sialyl-TF and 6-O-βGlcNAc-TF(core 2). TF and Sialylated TF are widely regarded as tumor associated and are being investigated as antigens for the immunotherapy of cancers of epithelial origin.

The disaccharide, βGal(1-3)αGalNAc, O-linked to serine and threonine in mucins and glycoproteins, is classified as biosynthetic core 1 glycan[1] and is popularly known as carcinoma associated Thomsen-Friedenreich (TF) antigen. Cancer associated mucins from gastric[2] and rectal[3] epithelium were found to express this disaccharide in abnormally increased quantities. This disaccharide is also found on the cell surface of chronic and acute myelogenous leukemic leukocytes.[4] Sialylated versions of TF, particularly α2-6 sialyl-TF, have been found on the mucins expressed by human breast cancer cell lines.[5] Synthetic glycopeptides derived from cancer associated mucins, bearing this structure, may be used as immunogens to stimulate immune responses against cancer cells expressing this disaccharide[6]. Synthesis of serine and threonine with large α-O-linked carbohydrate structures as building blocks for the glycopeptide synthesis has been a challenge for many years.[7] Chemical synthesis of α-O-linked N-acetylgalactosamine based mucin type glycopeptides, depends on the accessibility to large amounts of O-glycosylated amino acids. Sequential glycosylations of serine and threonine for Fmoc based glycopeptide synthesis involves a complex manipulation of the selectivities of base sensitive protecting groups. Fmoc protected serine and threonine are among highly sensitive and hindered aglycons used in glycosylation reactions.

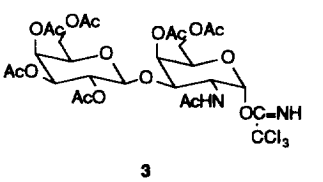
Recently, we reported[8] the influence of fused cyclic acetal at 4 and 6 hydroxyls of N-acetylgalactosamine in promoting the formation of α-linked glycosides of serine, threonine and a variety of other acceptors. With large scale process development as a target, we found that a similar disaccharide block Gal(1-3)GalNAc trichloroacetimidate (**1**), derived from 4,6-benzylidenyl GalNAc, serves as an excellent donor for the synthesis of αO-glycosyl serine/threonine. The yields of exclusively α-glycosides are higher compared to the monosaccharide donor (**2**) reported earlier.[8] In contrast, the peracetylated block donor (**3**) yielded a 1:1 mixture of α and β glycosides in poor yields while an oxazoline derivative is formed as the major product, proving the extensive participation of the 2-acetamido group at the anomeric carbon. In galactose analogues, 4,5 (axial and equatorial, respectively) fused 1,3 dioxane ring system appears to prevent the participating group at 2-position (e.g. 2-O-acyl, 2-acetamido) from interacting with the anomeric carbon.



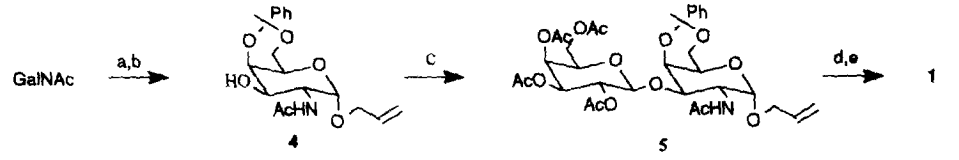

**1**




**2**




**3**

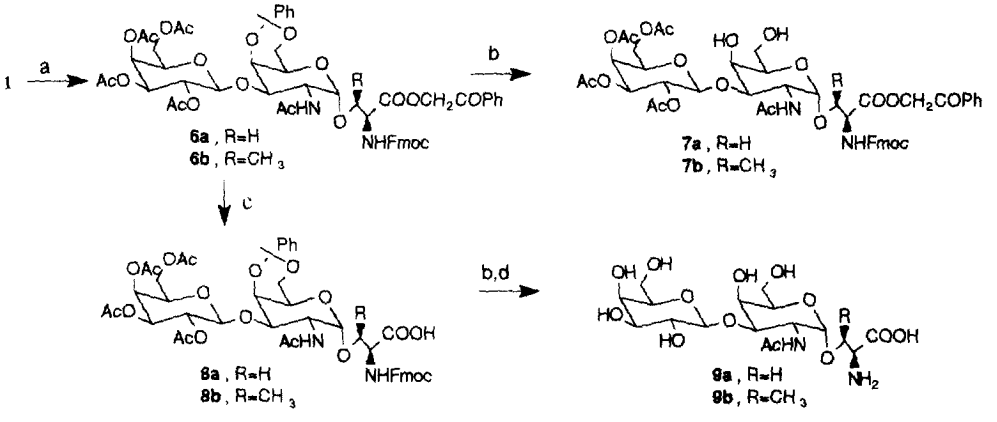



a.Allyl alcohol, HCl, 60°C, 55%; b. Acetonitrile, p-toluene sulfonic acid, benzaldehyde dimethyl acetal, 55°C, 62%; c. Acetobromogalactose, $Hg(CN)_2$, benzene, nitromethane, 50°C, 65%; d.[Bis(methyldiphenylphosphene) (1,5-cyclo octadiene)] iridium (I) hexaflourophosphate, THF, r.t., 58%; e. $CH_2Cl_2$, $CCl_3CN$, DBU, 70%.

Scheme I

The disaccharide block is synthesized (Scheme I) by heating N-acetylgalactosamine at 60°C in dry allyl alcohol as a solvent/reactant and catalytic amount (2%) of dry HCl gas to obtain allyl glycoside in about 55% yield. 4,6-Hydroxyls are protected with benzaldehyde dimethylacetal to form 3-OH analog 4. Glycosylation with acetobromogalactose using $Hg(CN)_2$ as catalyst in dry benzene and nitromethane (1:1) gives the allyl disaccharide **5** in about 65% yield. Allyl group is deblocked using [Bis(methyldiphenylphosphene)(1,5-cyclo octadiene)] iridium (I) hexaflourophosphate as catalyst to obtain 1-OH of the disaccharide from which the donor **1** is formed in about 70% yield. A mixture of the donor (0.9 mmol), the acceptor (0.61 mmol) and 3Å molecular sieves (0.5 g) in 5 mL of dry THF was stirred at room temperature for 10 minutes under argon. After cooling the reaction mixture to -20 ± 5°C, 0.5 mL of 0.1 mol $BF_3$-$Et_2O$ in THF was added dropwise in 10 minutes. The mixture was stirred at -20 ± 5°C for 30 minutes and warmed to room temperature. The solvent was removed in vacuo and the residue was purified by silicagel column using hexane/ethylacetate/methanol (10:10:1).

The significance of using the block donor for glycosylation, is the excellent yields of the α-glycosides **6**, of serine and threonine (Scheme II). 4,6-Benzylidene may be removed using 80% aqueous acetic acid to obtain 4,6-diol **7** which may be used for the synthetic extentions, while the removal of phenacyl group gives **8** which can be used as a glycopeptide building block. $C^{13}$ and $H^1$ NMR data of deblocked structures **9** are in agreement with reported data.[9]

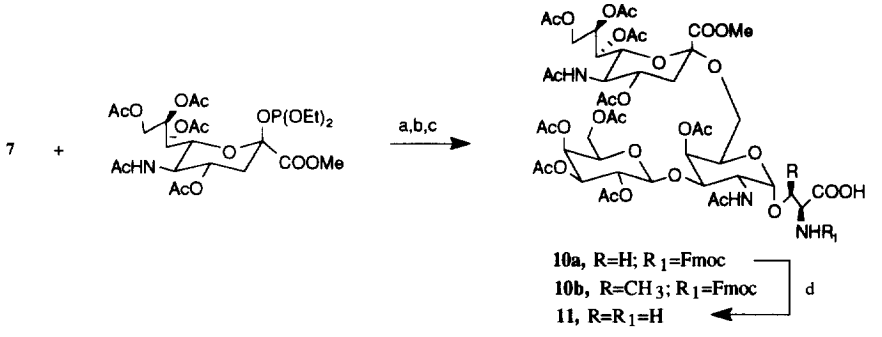



a. N-Fmoc-L serine or threonine phenacyl ester, $BF_3$-$OEt_2$ (0.05m), -20°C, THF; b. 80% $CH_3COOH$ in $H_2O$, 80°C, 73%; c. 80% $CH_3COOH$ in EtOAc, Zinc, 90%; d. 0.1N NaOH

Scheme II

Further extention of the synthesis (Scheme III), using 4,6-diol (**7**) and a sialyl donor, gives sialyl-TF (**10**) as a building block for glycopeptide synthesis. NMR of a partially deblocked structure is in agreement with reported data.[10] The reactions leading to **10** are simple and the yields are moderate to high. This disaccharide

block gives quick access to several glycopeptide building blocks including the core-2 trisaccharide (not reported here), βGlcNAc-TF, in large scale.




**a.** TMS-OTf, -20°C, $CH_3CN$, ~75% (α:β=1:1.3); **b.** $Ac_2O$, Pyridine, 80%; **c.** 80% $CH_3COOH$ in EtOAc, Zinc, 82%; **d.** Morpholine, 81%

Scheme III

4,6-Benzylidene, in addition to being a protecting group, plays an important role in altering the overall course of the N-acetylgalactosamine reactivity exerting a significant steric control on the glycosidic bond formation. α-Glycoside is by far the dominant product while the expected side products, mainly due to the influence of 2-acetamido group, such as oxazoline and the β-glycoside are only formed in trace amounts. The table compares the yields of α- and β-glycosides from various donors and different molar ratios of serine and threonine.

Table. % Yield of serine/threonine glycosides with anomer or anomeric ratio in parentheses.

| **Donor** / **Acceptor** | **1** | **2** | **3** |
|---|---|---|---|
| Equivalents of Serine (N-Fmoc) Phenacyl ester | | | |
| 1.5 | 63 (α) | 45 (α) | 20 (1:1) |
| 0.67 | 68 (α) | 62 (α) | — |
| 0.33 | — | 72 (α) | — |
| Equivalents of Threonine (N-Fmoc) Phenacyl ester | | | |
| 1.5 | 31 (α) | 20 (α) | 5-10 (1:4) |
| 0.67 | 36 (α) | 29 (α) | — |
| 0.33 | — | 39 (α) | — |

Donors such as **1** facilitate a fast access to the related α-glycosides in large quantities, required if the industrial scale synthesis of glycopeptides is contemplated. Research into further applications of N-acetylgalactosamine based donors is on going. Block donors solve the problem of dealing with several base sensitive protecting groups on carbohydrate as well as on the serine/threonine. For example, stepwise synthesis

of a trisaccharide such as **11** becomes complicated due to the base sensitive protecting groups if monosaccharide donor **2** were to be used.

## REFERENCES


1. Schachter, H.; Brockhausen, I. In: *Glycoconjugates: Composition , Structure and Function;* Allen, H. J.; Kisailus, E. C. Eds.; Marcel Dekker, Inc.: New York. **1992**, 263-332.

2. Yamashita, Y.; Chung, Y. S.; Horie, R.; Kannagi, R.; Sowa, M. *J. Natl. Can. Inst.* **1995**, *87*, 441-446.

3. Kurosaka, A.; Nakajima, H.; Funakoshi, J.; Matsuyama, M.; Nagayo, T.; Yamashina, I. *J. Biol. Chem.* **1983**, *258*, 11594-11598.

4. Fukuda, M.; Carlsson, S. R.; Klock, J. C.; Dell, A. *J. Biol. Chem.* **1986**, *261*, 12796-12806.

5. Hull, S. R.; Bright, A.; Carraway, K. L.; Abe, M.; Hayes, D. F.; Kufe, D.W. *Cancer Commun.* **1989**, *1*, 261-267.

6. (a) Fung, P. Y. S.; Madej, M.; Koganty, R. and Longenecker, B. M. *Cancer Res.* **1990**, *50*, 4308-4314.
   (b) MacLean, G. D.; Bowen-Yacashin, M. B.; Samuel, J.; Meikle, A.; Stuart, G.; Nation, J.; Poppema, S.; Jerry, M.; Koganty, R.; Wong, T. and Longenecker, B. M. *J. Immunother.* **1992**, *11*, 292-305.
   (c) MacLean, G. D.; et al. Cancer Immunol. Immunother. **1993**, *36*, 215-222.
   (d) Longenecker, B. M.; MacLean, G. D. *The Immunologist.* **1993**, *1*, 89-93.

7. (a) Kunz, H. *Angew. Chem. Int. Ed. Engl.* **1987**, *26*, 294-308.
   (b) Kunz, H. *Adv. Carbohydr. Chem. Biochem.* **1994**, *50*, 277-310.

8. Yule, J. E.; Wong, T. C.; Gandhi, S. S.; Qiu, D.; Riopel, M. A.; Koganty, R. R. *Tetrahedron Lett.* **1995**, *38*, 6839-6842.

9. NMR data for the deblocked TF-ser/thr. (Lit. ref. Paulson, H.; Holck, J. *Carbohydr. Res.* **1982**, *109*, 89-107.)
   **9a**, $[\alpha]D^{20}$ + 90.0 (c0.5, $H_2O$); $^1$H-NMR (300 MHz, $D_2O$), δ=4.82 (d, 1H, J=3.5 Hz, H-1), 4.35 (d, 1H, J=8.0 Hz, H-1'), 4.24 (2d, 1H, J=3.5 and 11.5 Hz, Serβ-H), 3.39 (2d, 1H, J=8.0 and 10.5 Hz, H-2'), 1.90 (s, 3H, NAc); $^{13}$C-NMR (300 MHz, $D_2O$), δ=104.7 (C-1'), 98.3 (C-1).
   **9b**, $[\alpha]D^{20}$ + 92.0 (c0.5, $H_2O$); $^1$H-NMR (300 MHz, $D_2O$), δ=4.85 (d, 1H, J=3.5 Hz, H-1), 4.34 (d, 1H, J=8.0 Hz, H-1'), 4.18 (2d, 1H, J=3.5 and 11.5 Hz, Thrβ-H), 3.37 (2d, 1H, J=8.0 and 10.5 Hz, H-2'), 1.92 (s, 3H, NAc), 1.29 (d, 3H, J=6.5 Hz, $ThrCH_3$); $^{13}$C-NMR (300 MHz, $D_2O$), δ=105.5 (C-1'), 100.1 (C-1).

10. NMR data for the sialyl TF-ser. (Lit. ref. Iijima, H.; Ogawa, T. *Carbohydr. Res.* **1989**, *186*, 95-106.)
    **11**, $[\alpha]D^{20}$ + 40.0 (c0.5, MeOH); $^1$H-NMR (500 MHz, $CD_3OD$) δ=5.34-5.40 (m, 4H, H-4a, H-4c, H-8b), 5.32 (2d, 1H, J=2.2 and 8.5Hz, H-7b), 5.05 (2d, 1H, J=3.5 and 10.5Hz, H-3c), 4.94 (2d, 1H, J=7.5 and 10.5Hz, H-2c), 4.82 (d, 1H, J=3.5 Hz, H-1a), 4.72 (d, 1H, J=7.5 Hz, H-1c), 4.42 (2d, 1H, J=3.5 and 11.0Hz, H-2a), 3.82 (s, 3H, $COOCH_3$), 2.64 (2d, 1H, J=4.0 and 12.5Hz, H-3beq), 2.14, 2.13, 2.11, 2.10, 2.03 (6H), 2.02, 2.00 1.97, 1.93, 1.82 (10s, 33H, 9 $OCOCH_3$ and $2NCOCH_3$).


(*Received in USA* 23 *October* 1995; *revised* 16 *November* 1995; *accepted* 20 *November* 1995)