Pergamon

Tetrahedron Letters, Vol. 38, No. 1, pp. 45-48, 1997

0040-4039/97 $17.00 + 0.00

PII: S0040-4039(96)02252-6

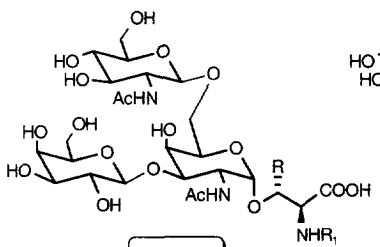
Mucin Type Glycopeptides: Synthesis of Core 2, Core 6 and F1-α Building Blocks and Some Unexpected Reactions

Dongxu Qiu and R. Rao Koganty*

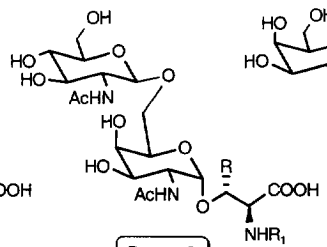
Biomira Inc., 2011-94 Street, Edmonton, Alberta, Canada T6N 1H

Abstract: Protected Core 2 and 6 and F1-α analogs were synthesized using the 4,6-diol of Tn-Ser/Thr and TF-Ser/Thr. Acetylation of glucosamine in acetic anhydride/pyridine resulted in an additional and unusual N-acetylation at the Fmoc protected amine of serine and threonine followed by β-elimination of the aglycons. Normal glucosamine acetylation was carried out at -20°C for 20 minutes to accomplish in synthesis of Core 2, Core 6 and F1-α as glycopeptide building blocks.

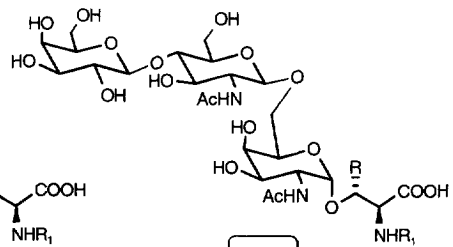
O-Linked oligosaccharides found on cancer associated mucins are heterogeneous in structure and size, built primarily on α-N-acetylgalactosamine linked to serines and threonines of protein core.[1] O-Linked structures are broadly classified into six core types from which all oligosaccharides extend in mucins.[2] The core extensions on cancer associated mucins are severely restricted and often altered resulting in prematurely terminated abnormal structures which have been well studied.[3] These abnormal carbohydrates are widely regarded as targets for the immunotherapy of cancers.[4] Core 2 (**1-4**) and core 6 (**5-8**) are two interesting structures that seem to lead to normal and abnormal paths of carbohydrate biosynthesis, in spite of the appearance that core 6 may be a precursor to core 2. However, normal biosynthetic path follows the 3 before 6 rule [2] in glycosylating the GalNAc precluding core 6 as a normal structure. This is further supported by the fact that core 6 leads to another cancer associated structure called F1-α which has been found to be widely expressed on gastric cancer mucins.[5] Yamashita et al[5] have reported the synthesis of F1-α as a lipid analog and generated a monoclonal antibody, MAb F1-α-75 which strongly stains tissue from gastric, pancreatic and colon cancers but not the corresponding normal adult tissue. Our interest in mucin glycopeptides and their use for the cancer immunotherapy lead us to the synthesis of core 2 and core 6 as well as F1-α with serine and threonine as aglycons which may be used in the automated peptide synthesizers.



Core 2



Core 6



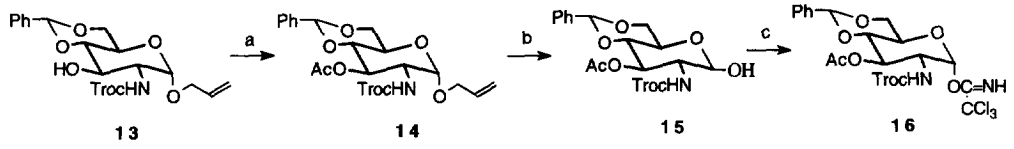
F1-α

1. R=H, R_1=Ac
2. R=R_1=H
3. R=CH_3, R_1=Ac
4. R=CH_3, R_1=H

5. R=H, R_1=Ac
6. R=R_1=H
7. R=CH_3, R_1=Ac
8. R=CH_3, R_1=H

9. R=H, R_1=Ac
10. R=R_1=H
11. R=CH_3, R_1=Ac
12. R=CH_3, R_1=H

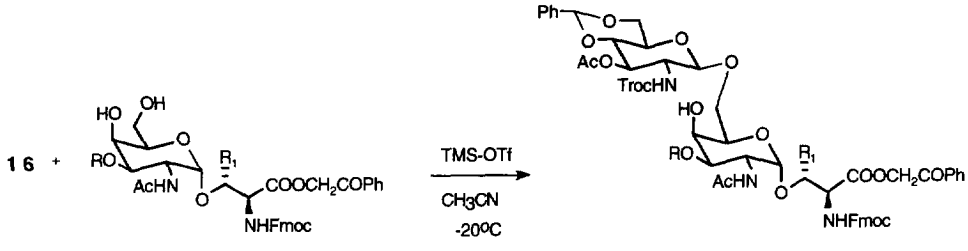
3-O-Acetyl-4,6-benzylidine-N-trichloroethoxycarbonyl (troc)-glucosaminyl trichloroacetimidate **16** was synthesized (scheme 1) from the known intermediate **13**.[6]



a. Ac_2O/Pyridine, r.t., overnight, 54%; **b**. [Bis(methyldiphenylphosphene)(1,5-cyclo-octadiene)]iridium (I) hexafluorophosphate, THF, r.t. 65%; **c**. CH_2Cl_2, CCl_3CN, DBU, r.t., 57%

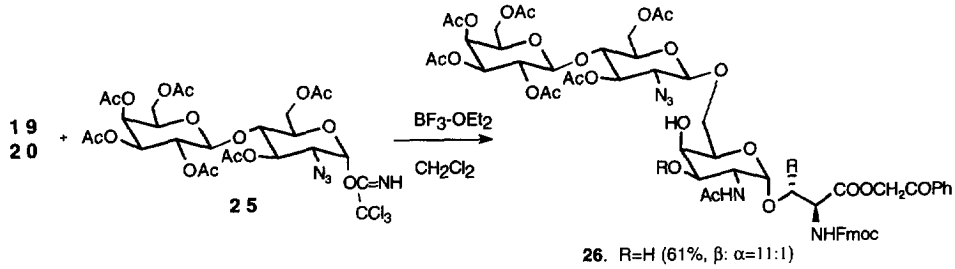
Scheme 1

α-Glycosylation of serine and threonine (Fmoc and phenancyl ester protected) using 4,6-benzylidine, 3-O-benzoyl (or 3-O-tetraacetylgalactosyl) trichloroacetoimitate has been carried out as described earlier.[7] Subsequent deblocking of 4,6-benzylidine in 80% acetic acid at 80°C gave precursors **17-20**. Reaction of donor **16** with **17** and **18** gave **21** and **22** (core 2 structure) with **19** and **20** gave **23** and **24** (core 6) respectively. F1-α serine (**26**) and threonine (**27**) were obtained from the reaction of disaccharide donor **25**[8] with **19** and **20**, respectively (scheme 2).



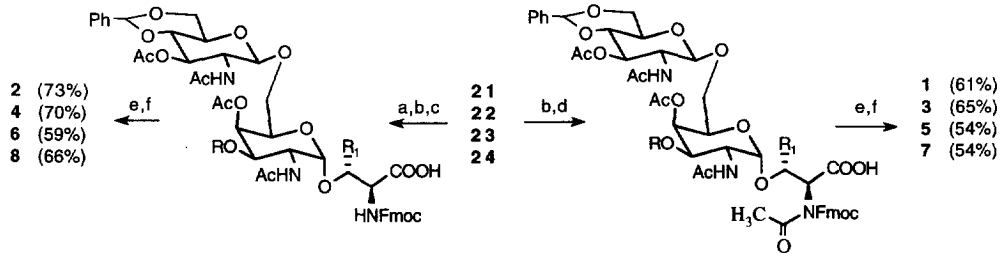
17. R=tetra-O-acetyl-β-D-galactoside, R1=H
18. R=tetra-O-acetyl-β-D-galactoside, R_1=CH_3
19. R=Bz, R_1=H
20. R=Bz, R_1=CH_3

21. R=tetra-O-acetyl-β-D-galactoside, R_1=H (69%, only β)
22. R=tetra-O-acetyl-β-D-galactoside, R_1=CH_3 (65%, only β)
23. R=Bz, R_1=H (74%, only β)
24. R=Bz, R_1=CH3 (77%, only β)



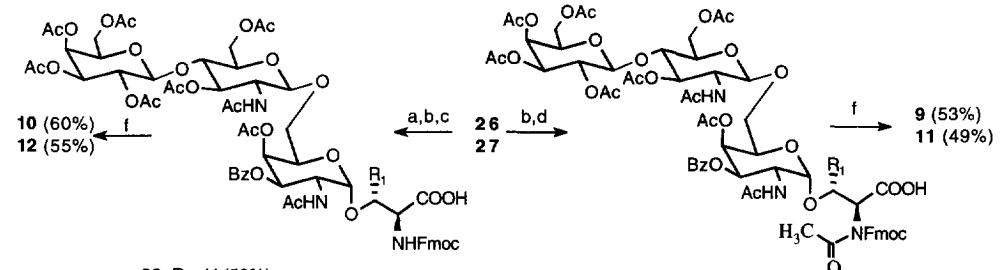
26. R=H (61%, β: α=11:1)
27. R=CH_3 (55%, β: α=14:1)

Scheme 2



28. R=tetra-O-acetyl-β-D-galactoside, R_1=H (73%)
29. R=tetra-O-acetyl-β-D-galactoside, R_1=CH_3 (77%)
30. R=Bz, R_1=H(60%)
31. R=Bz, R_1=CH3(67%)

32. R=tetra-O-acetyl-β-D-galactoside, R_1=H (69%)
33. R=tetra-O-acetyl-β-D-galactoside, R_1=CH_3 (65%)
34. R=Bz, R_1=H(74%)
35. R=Bz, R_1=CH3(77%)



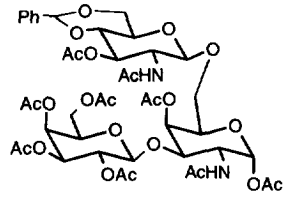
36. R_1=H (50%)
37. R_1=CH_3 (80%)

38. R_1=H (48%)
39. R_1=CH_3 (45%)

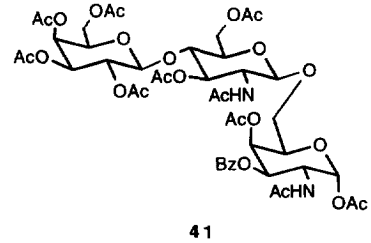
a. Acetic anhydride/pyridine (1:2), r.t., 12 hours; **b**. Zinc, 80% Acetic acid/ethylacetate, r.t.; **c**. Acetic anhydride/pyridine (1:2), -20°C, 20 min.; **d**. Acetic anhydride/pyridine (1:2), r.t., 5 hours; **e**.80% aq acetic acid, 80°C, 2 hours; **f**. 0.1 N NaOH/methanol (2:1), overnight;

Scheme 3

The N-troc group of **21-24** and the azido group of **26** and **27** are easily converted to free amine using activated zinc and acetic acid reduction. Under these conditions the phenacylester on the serine and threonine is also removed generating free acid. But, in an attempt to N-acetylate the glucosamine (scheme 3) in pyridine/acetic anhydride(2:1) at room temperature for 5 hours, Fmoc protected amine group of serine/threonine was further acetylated along with N-acetylation of glucosamine. This was further confirmed by total deblocking of **32-35**, **38** and **39** (i, 80% aq acetic acid at 80°C for 2 hours, ii. 0.1 N NaOH/methanol at 2:1 ratio) which gave **1**, **3**, **5**, **7**, **9** and **11** respectively, all with three N-acetyl groups both before and after deblocking. Prolonged acetylation (60 hours) in pyridine/acetic anhydride resulted in β-elimination of serine/threonine, forming 1-O-acetylated structures **40**, **41**. Normal acetylation of glucosamine was accomplished by stirring in pyridine/acetic anhydride at -20°C for 20 minutes to obtain core 2 (**28**, **29**), core 6 (**30**, **31**) and F1-α (**36**, **37**) building blocks and total deblocking afforded carbohydrate antigens **2**, **4**, **6**, **8**, **10** and **12**, respectively.



40



41

During the preparation of this manuscript, Core 2 synthesis through a separate route was described by another group.[10]

REFERENCES

1. (a) Strous, G. J.; Dekker, J. *Crit. Rev. Biochem. Mol. Bol.* **1992**, *27*, 57-92.
(b) Carraway, K. L. ans Hull, S. *Glycobiology* **1991**, *1*, 131-138.

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

3. (a) Bhavanandan, V. P. *Glycobiology* **1991**, *1*, 493-503.
(b) Yamashita, Y.; Chung, Y. S.; Horie, R.; Sowa, M. *J. Natl. Can. Inst.* **1995**, *87*, 441-446
(c) Kurosaka, A. et al. *J. Biol. Chem.* **1983**, *258*, 11594-11598
(d) Hull, S. R. et al. *Cancer Commun.* **1989**, *1*, 261-267

4. (a) Fung, P. Y. S. et al. *Cancer Res.* **1990**, *50*, 4308-4314
(b) MacLean, G. D. et al. *J. Immunther.* **1992**, *11*, 292-305.
(c) MacLean, G. D.et al. *Cancer Immunol. Immunother.* **1993**, *36*, 215-222.
(d) Longenecker, B. M. and MacLean, G. D. *The Immunologist* **1993**, *1*, 89-93.

5. (a) Horie, R.; Hara, K.; Nakano, K. *Carbohydr. Res.* **1992**, *230*, C11-C15
(b)Yamashita, Y. et al. *Int. J. Cancer* **1994**, *58*, 349-355.

6. Fukase, K. et al. *Tetrahedron Lett.* **1995**, *36*, 7455-7458.

7. (a) Yule, J. E. et al. *Tetrahedron Lett.* **1995**, *36*, 6839-6842
(a) Qiu, D.; Gandhi, S. S.; Koganty, R. R. *Tetrahedron Lett.* **1996**, *37*, 595-598.

8. Grundler, G. and Schmidt, R. R. *Liebigs Ann. Chem.* **1984**, 1826-1847

9 NMR data for the selected compounds.

1. δ=4.86 (d, 1H, J=3.5Hz, H-1a), 4.57 (d, 1H, J=8.5Hz, H-1b), 4.48(d, 1H, J=7.5Hz, H-1c), 4.45 (m, 1H, Ser-α-H), 4.34 (2d, 1H, J=3.5, 11.5Hz, H-2), 2.06, 2.03, 2.01 (3s, 9H, 3NAc);
2. δ=4.90 (d, 1H, J=3.5Hz, H-1a), 4.54 (d, 1H, J=8.0 Hz, H-1b), 4.46(d, 1H, J=7.5Hz, H-1c), 4.35 (2d, 1H, J=3.5, 11.5Hz, H-2a), 2.05, 2.03 (2s, 6H, 2NAc);
3. δ=4.99 (d, 1H, J=3.5Hz, H-1a), 4.56 (d, 1H, J=8.5 Hz, H-1b), 4.51 (d, 1 H, J=4.0 Hz, Thr-α-H), 4.43(d, 1H, J=7.5 Hz, H-1c), 2.06, 2.03, 2.01 (3s, 9H, 3NAc), 1.19 (d, 3H, J=6.5 Hz, CH_3).
4. δ=4.95 (d, 1 H, J=3.5Hz, H-1a), 4.55 (d, 1H, J=8.0 Hz, H-1b), 4.46(d, 1H, J=7.5Hz, H-1c), 4.39 (m, 1 H, Thr-α-H), 2.06, 2.04 (2s, 6H, 2NAc), 1.37 (d, 3H, J=6.5Hz, CH_3);
5, δ=4.94 (d, 1H, J=3.5Hz, H-1a), 4.67 (d, 1H, J=8.5Hz, H-1b), 4.50 (2d, 1H, J=3.5, 5.0Hz, Ser-α-H), 4.16 (2d, 1H, J=3.5, 11.5Hz, H-2), 2.16 (s, 3H, NAc), 2.14 (s, 6H, 2 NAc);
6. δ=4.97 (d, 1H, J=3.5Hz, H-1a), 4.63 (d, 1H, J=8.5Hz, H-1b), 4.26 (2d, 1H, J=3.5, 11.0Hz, H-2), 2.14, 2.13 (2s, 6H, 2NAc);
7, δ=5.08 (d, 1H, J=3.5Hz, H-1a), 4.66 (d, 1H, J=8.5Hz, H-1b), 4.58 (d, 1H, J=3.5Hz, Thr-α-H), 2.17, 2.15, 2.12 (3s, 9H, 3NAc), 1.30(d, 3H, J=6.5Hz, CH_3);
8. δ=5.05 (d, 1H, J=4.0Hz, H-1a), 4.65 (d, 1 H, J=8.5Hz, H-1b), 4.44 (m, 1H, Thr-α-H), 2.13, 2.12 (2s, 6H, 2NAc), 1.40(d, 3H, J=6.5Hz, CH_3);
9. δ=4.95(d, 1H, J=3.5Hz, H-1a), 4.71(d, 1H, J=7.5Hz, H-1b), 4.58(d, 1H, J=8.0Hz, H-1c), 4.52(dd, 1H, J=3.5, 5.0Hz, Ser-α-H), 4.28(dd, 1H, J=3.5, 11.0Hz, H-2a), 3.65(dd, 1H, J=8.0, 10.0Hz, H-3b), 2.18, 2.16, 2.15(3s, 9H, 3NAc);
10. δ=4.99(d, 1H, J=3.5Hz, H-1a), 4.69(d, 1H, J=7.5Hz, H-1b), 4.55(d, 1H, J=8.0Hz, H-1c), 4.49(m, 1H, Ser-α-H), 3.62(dd, 1H, J=8.0, 10.0Hz, H-3b), 2.15, 2.13(2s, 6H, 2NAc);
11. δ=5.02(d, 1H, J=3.5Hz, H-1a), 4.63(d, 1H, J=7.5Hz, H-1b), 4.51(d, 1H, J=4.5Hz, Thr-α-H), 4.50(d, 1H, J=8.0Hz, H-1c), 3.57(dd, 1H, J=8.0, 10.0Hz, H-3b), 2.11, 2.10, 2.05(3s, 9H, 3NAc), 1.24(d, 3H, J=6.5Hz, CH_3);
12. δ=4.96(d, 1H, J=3.5Hz, H-1a), 4.62(d, 1H, J=7.5Hz, H-1b), 4.55(d, 1H, J=8.0Hz, H-1c), 4.42(m, 1H, Thr-α-H), 357(dd, 1H, J=8.0, 10.0Hz, H-3b), 2.09, 2.08(2s, 6H, 2NAc), 1.40(d, 3H, J=6.5Hz, CH_3);
40. δ=6.17 (d, 1H, J=3.5 Hz, H-1a), 5.86 (d, 1H, J=9.0Hz, NH), 5.64 (d, 1H, J=9.5Hz, NH), 5.48(s, 1H, CHPh), 5.37 (m, 2H, H-4a, H-4b), 5.25 (t, 1 H, J=10.0Hz, H-3c), 5.13 (dd, 1H, J=8.0, 10.5Hz, H-2b), 4.98 (dd, 1H, J=3.5, 10.5 Hz, H-3b), 4.64 (d, 1H, J=8.0Hz, H-1b), 4.60 (d, 1H, J= 8.5Hz, H-1c), 2.18, 2.17, 2.14, 2.08, 2.07, 2.06, 2.01, 1.99, 1.96 (9s, 27H, 9Ac).
41. δ=6.18 (d, 1H, J=3.5Hz, H-1a), 5.72 (d, 1H, J=9.5 Hz, NH), 5.64 (d, 1H, J=9.0Hz, NH), 5.50 (dd, 1H, J=1.0, 3.0Hz, H-4), 5.40 (dd, 1H, J=3.0, 11.0Hz, H-3a), 5.35 (dd, 1H, J=1.0, 3.0Hz, H-4), 5.06-5.15(m, 2H), 4.96 (dd, 1H, J=3.5, 10.0Hz, H-3c), 4.88 (m, 1H, H-2a), 4.57 (d, 1H, J=7.5Hz, H-1b), 4.48 (d, 1H, J=7.5Hz, H-1c), 2.22, 2.18, 2.16, 2.12, 2.07, 2.05, 2.04, 1.97, 1.89 (9s, 30H, 10Ac).

10. Meinjohanns, E. et al. *J. Chem. Soc., Perkin Trans. 1* **1996**, 985-993.

(*Received in USA* 22 *October* 1996; *revised* 8 *November* 1996; *accepted* 11 *November* 1996)