*Mendeleev Commun.*, 2002, 12(5), 183–184

# Synthesis of N-acetyllactosamine-terminated oligosaccharides — fragments of glycoprotein O-chains

**Galina V. Pazynina, Vyacheslav V. Severov and Nicolai V. Bovin***

*M. M. Shemyakin–Yu. A. Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, 117997 Moscow, Russian Federation. Fax: +7 095 330 5592; e-mail: bovin@carb.siobc.ras.ru*

10.1070/MC2002v012n05ABEH001653

The pentasaccharide Galβ1-4GlcNAcβ1-6(Galβ1-4GlcNAcβ1-3)GalNAcαOR, its partial trisaccharides Galβ1-4GlcNAcβ1-3GalNAcαOR and Galβ1-4GlcNAcβ1-6GalNAcαOR and the tetrasaccharide Galβ1-4GlcNAcβ1-6(Galβ1-3)GalNAcαOR were synthesised as spacered derivatives (R = $CH_2CH_2CH_2NH_2$) using a Troc-protected lactosamine glycosyl donor at the key stage.

Eight different cores, *i.e.*, regions connected with Ser or Thr of glycoprotein O-chains, are known.[1] Recently,[2] we reported the synthesis of oligosaccharides repeating four core structures, *i.e.*, GlcNAcβ1-3GalNAcαOR (core 3), GlcNAcβ1-3-(GlcNAcβ1-6)GalNAcαOR (core 4), GlcNAcβ1-6GalNAcαOR (core 6) and Galβ1-3(GlcNAcβ1-6)GalNAcαOR (core 2); R = = $CH_2CH_2CH_2NH_2$.

GlcNAcβ1-3GalNAcαO-Ser/Thr (core 3)

GlcNAcβ1-6
GalNAcαO-Ser/Thr (core 4)
GlcNAcβ1-3

GlcNAcβ1-6GalNAcαO-Ser/Thr (core 6)

GlcNAcβ1-6
GalNAcαO-Ser/Thr (core 2)
Galβ1-3

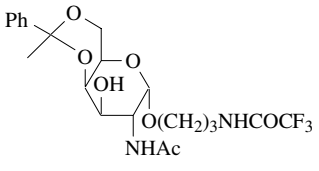
Galβ1-4GlcNAcβ1-3GalNAcαO–$CH_2CH_2CH_2NH_2$ **1**

Galβ1-4GlcNAcβ1-6
GalNAcαO–$CH_2CH_2CH_2NH_2$ **2**
Galβ1-4GlcNAcβ1-3

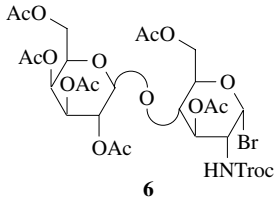
Galβ1-4GlcNAcβ1-6GalNAcαO–$CH_2CH_2CH_2NH_2$ **3**

Galβ1-4GlcNAcβ1-6
GalNAcαO–$CH_2CH_2CH_2NH_2$ **4**
Galβ1-3

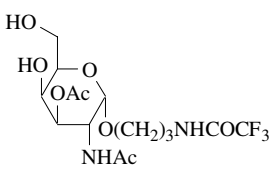
Here, we describe the synthesis of more complex oligosaccharides, namely, containing an additional terminal β-Gal moiety, as spacered compounds **1**–**4** suitable for further transformations into various glycoprobes.[3] The above core glycopeptides have been identified in mucin-type glycoproteins; *e.g.*, core 6 structure is a cancer-associated antigen[4] and core 4



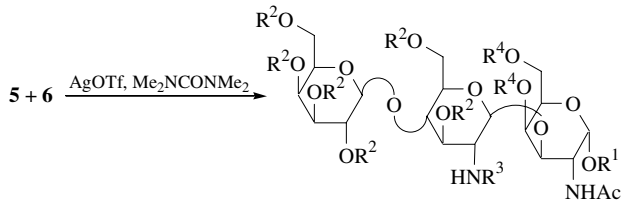
**5**



**6**



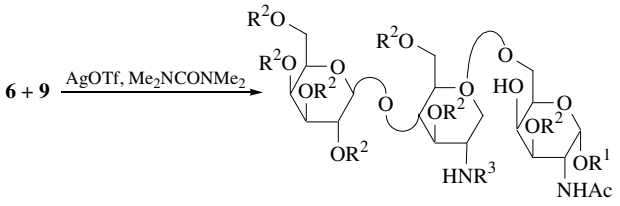
**9**



i → **7** $R^1 = (CH_2)_3NHCOCF_3$, $R^2$ = Ac, $R^3$ = Troc, $R^4$ = PhCH, 84%

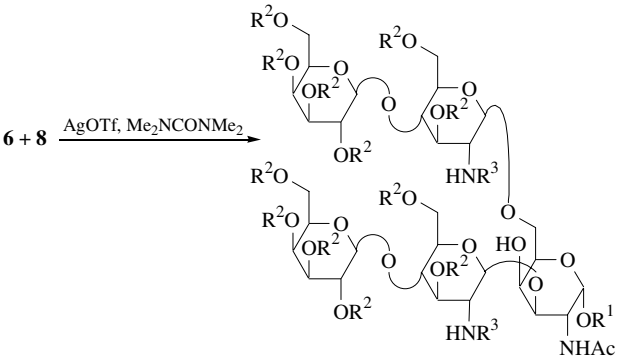
ii, iii, iv, v → **8** $R^1 = (CH_2)_3NHCOCF_3$, $R^2$ = Ac, $R^3$ = Troc, $R^4$ = H, 72%

ii, iii, iv, v → **1** $R^1 = (CH_2)_3NH_2$, $R^2 = R^4$ = H, $R^3$ = Ac, 75%



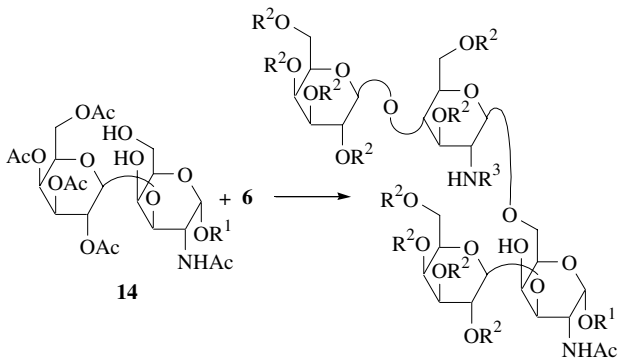
iii, iv, v → **12** $R^1 = (CH_2)_3NHCOCF_3$, $R^2$ = Ac, $R^3$ = Troc, 78%

iii, iv, v → **3** $R^1 = (CH_2)_3NH_2$, $R^2$ = H, $R^3$ = Ac, 74%



iii, iv, v → **10** $R^1 = (CH_2)_3NHCOCF_3$, $R^2$ = Ac, $R^3$ = Troc, 71%

iii, iv, v → **2** $R^1 = (CH_2)_3NH_2$, $R^2$ = H, $R^3$ = Ac, 50%



**14**

iii, iv, v → **15** $R^1 = (CH_2)_3NHCOCF_3$, $R^2$ = Ac, $R^3$ = Troc, 65%

iii, iv, v → **4** $R^1 = (CH_2)_3NH_2$, $R^2$ = H, $R^3$ = Ac, 48%

**Scheme 1** *Reagents and conditions*: i, AcOH; ii, $Ac_2O$, Py; iii, Zn, AcOH, $Ac_2O$; iv, MeONa, MeOH; v, NaOH, $H_2O$.

trisaccharide is probably a crypted carcinoma-associated Tk antigen.[5] Core 2 oligosaccharide is the specific motif, which defines P-selectin binding properties[6] of SiaLe$^x$-containing glycoprotein PSGL-1. Gal-terminated oligosaccharides **1**–**4** are potential ligands for galectins, the family of β-galactoside binding lectins, whose natural cellular ligands are not yet clarified.

We used 3′-trifluoroacetamidopropyl 2-acetamido-4,6-*O*-benzylidene-2-deoxy-α-D-galactopyranoside **5**, which was described previously,[7] as a starting building block. To synthesise trisaccharide **1**, compound **5** reacted with bromide **6** (Scheme 1). Resulting trisaccharide **7** was used as the starting material in the synthesis of pentasaccharide **2**: it was 4,6-O-debenzylidenated, and obtained diol **8** was glycosylated regioselectively at the 6-position with bromide **6**. To synthesise trisaccharide **3**, the 3-OH group in compound **5** was protected by acetylation, the benzylidene group was removed and the obtained diol **9** was 6-O-glycosylated as described above. Tetrasaccharide **4** was synthesised by 6-O-regioselective glycosylation of disaccharide **14**[8] with bromide **6**.

In all cases, N-Troc-protected lactosaminyl bromide **6** (Troc is 1,1,1-trichloroethyloxycarbonyl),[9] obtained by acetylation and subsequent bromine treatment of the corresponding commercial ethyl thioglycoside, was used for stereocontrolled β-glycosylation. This choice is stipulated by good yields, high β-stereoselectivity of glycosylation,[2,9] and compatibility of this protection with other protecting groups used in the synthesis. Preparative yields in glycosylation with this reagent are noted in Scheme 1, and they are comparable with published data.[9] β-Stereoselectivity of glycosylation with donor **6** was confirmed by the corresponding $J_{1,2}$ (8.0–9.5 Hz). The structures of synthesised oligosaccharides were confirmed by $^1$H NMR data using homonuclear correlation spectroscopy and conventional analysis of coupling constants.†

Deprotection of sugar moieties and the spacer-arm by conventional methods (Scheme 1) gave rise to oligosaccharides **1**–**4** as 3-aminopropyl glycosides; the synthesis of these oligosaccharides in the convenient spacered form has not been described earlier.

Compounds **1**–**4** coupled with fluorescein-labelled polyacrylamide were used for the study of malignant cell galectins, the data will be published elsewhere.

This work was supported in part by NIH (grant no. 1U54GM62116).

† *Galβ1-4GlcNAcβ1-3GalNAcα1-O(CH$_2$)$_3$NH$_2$* **1**: $^1$H NMR (500 MHz, $D_2O$) δ: 4.778 (d, 1H, H-1 GalNAc, $J_{1,2}$ 3.9 Hz), 4.540 (d, 1H, H-1 GlcNAc, $J_{1,2}$ 7.6 Hz), 4.410 (d, 1H, H-1 Gal, $J_{1,2}$ 7.8 Hz), 4.196 (dd, 1H, H-2 GalNAc, $J_{1,2}$ 3.9 Hz, $J_{2,3}$ 11.3 Hz), 4.151 (dd, 1H, H-4 GalNAc, $J_{3,4}$ 2.7 Hz, $J_{4,5} \leq 1$ Hz), 3.85–3.92 (m, 4H, H-3 and H-5 GalNAc, H-4 Gal, H-6$^a$ GlcNAc), 3.781 (dd, 1H, H-6$^b$ GlcNAc, $J_{5,6}$ 4.7 Hz, $J_{6^a,6^b}$ 12.2 Hz), 3.63–3.76 (m, 9H, H-2, H-3, H-4 GlcNAc, H-5, H-6$^a$, H-6$^b$ Gal, H-6$^a$, H-6$^b$ GalNAc, OCH$_2$ sp), 3.607 (dd, 1H, H-3 Gal, $J_{2,3}$ 10.0 Hz, $J_{3,4}$ 3.5 Hz), 3.47–3.53 (m, 2H, H-5 GlcNAc, OCH$_2$ sp), 3.475 (dd, 1H, H-2 Gal, $J_{2,3}$ 10.0 Hz, $J_{1,2}$ 7.8 Hz), 3.055 (m, 2H, NCH$_2$ sp), 1.983 (s, 3H, NHCO*Me*), 1.953 (s, 3H, NHCO*Me*), 1.940 (m, 2H, CH$_2$ sp). MS, *m/z*: 666 (643 + 23) ($M^+ + Na^+$). $[\alpha]_D$ +49 (*c*1, $H_2O$).

*Galβ1-4GlcNAcβ1-6(Galβ1-4GlcNAcβ1-3)GalNAcα1-O(CH$_2$)$_3$NH$_2$* **2**: $^1$H NMR (500 MHz, $D_2O$) δ: 4.515 (d, 1H, H-1 GlcNAc, $J_{1,2}$ 7.8 Hz), 4.453 (d, 1H, H-1 GlcNAc, $J_{1,2}$ 8.3 Hz), 4.391 (d, 2H, 2H-1 Gal, $J_{1,2}$ 7.8 Hz), 4.172 (dd, 1H, H-2 GalNAc, $J_{1,2}$ 3.7 Hz, $J_{2,3}$ 11.0 Hz), 4.124 (dd, 1H, H-4 GalNAc, $J_{3,4}$ 2.9 Hz, $J_{4,5} \leq 1$ Hz), 4.077 (dd, 1H, H-6$^a$ GalNAc, $J_{5,6^a}$ 2.2 Hz, $J_{6^a,6^b}$ 10.5 Hz), 3.85–3.96 (m, 6H, 2H-6$^a$ GlcNAc, H-6$^a$ and H-3 GalNAc, 2H-4 Gal), 3.764 (dd, 2H, 2H-6$^b$ GlcNAc, $J_{5,6^b}$ 4.9 Hz, $J_{6^a,6^b}$ 12.2 Hz), 3.60–3.73 (m, 14H, 2H-5, 2H-6$^a$, 2H-6$^b$ Gal, 2H-2, 2H-3, 2H-4 GlcNAc, H-5 GalNAc, OCH sp), 3.586 (dd, 2H, 2H-3 Gal, $J_{3,4}$ 3.4 Hz, $J_{2,3}$ 10.0 Hz), 3.532 (m, 1H, H-5 GlcNAc), 3.40–3.50 (m, 4H, H-5 GlcNAc, 2H-2 Gal, OCH sp), 3.045 (m, 2H, NCH$_2$ sp), 1.963 (s, 3H, NHCO*Me*), 1.936 (s, 3H, NHCO*Me*), 1.934 (s, 3H, NHCO*Me*), 1.85–1.98 (m, 2H, CH$_2$ sp). MS, *m/z*: 1031 (1008 + 23) ($M^+ + Na^+$). $[\alpha]_D$ +18 (*c*1, $H_2O$).


## References

1 E. F. Hounsell, M. J. Davies and D. V. Renouf, *Glycoconjugate J*., 1996, **13**, 19.
2 G. V. Pazynina and N. V. Bovin, *Mendeleev Commun*., 2000, 132.
3 N. V. Bovin, *Glycoconjugate J*., 1998, **15**, 431.
4 Y. Yamashita, Y. S. Chung, R. Horie, R. Kannagi and M. Sowa, *J. Natl. Cancer. Inst*., 1995, **87**, 441.
5 M. Meichenin, J. Rocher, O. Galanina, N. Bovin, N. Nifant'ev, A. Sherman, E. Cassagnau, M. F. Heymann, J. Bara, R. H. Fraser and J. le Pendu, *Cancer Res*., 2000, **60**, 5499.
6 P. Mehta, R. D. Cummings and R. P. McEver. *J. Biol. Chem*., 1998, **273**, 32506.
7 N. V. Bovin, T. V. Zemlyanukhina and A. Ya. Khorlin, *Bioorg. Khim*., 1986, **12**, 533 [*Sov. J. Bioorg. Chem.* (*Engl. Transl.*), 1986, **12**, 282].
8 N. V. Bovin, T. V. Zemlyanukhina and A. Ya. Khorlin, *Bioorg. Khim*., 1985, **11**, 1256 (in Russian).
9 V. Ellervik and G. Magnusson, *Carbohydr. Res*., 1996, **280**, 251.


*Received: 31st July 2002; Com. 02/1980*

*Galβ1-4GlcNAcβ1-6GalNAcα1-O(CH$_2$)$_3$NH$_2$* **3**: $^1$H NMR (500 MHz, $D_2O$) δ: 4.865, (d, 1H, H-1 GalNAc, $J_{1,2}$ 3.7 Hz), 4.526 (d, 1H, H-1 GlcNAc, $J_{1,2}$ 8.1 Hz), 4.453 (d, 1H, H-1 Gal, $J_{1,2}$ 7.9 Hz), 4.139 (dd, 1H, H-2 GalNAc, $J_{1,2}$ 3.7 Hz, $J_{2,3}$ 11.0 Hz), 4.069 (dd, 1H, H-6$^a$ GalNAc, $J_{5,6^a}$ 3.2 Hz, $J_{6^a6^b}$ 10.8 Hz), 3.96–4.01 (m, 2H, H-6$^a$ GlcNAc, H-6$^b$ GalNAc), 3.953 (dd, 1H, H-4 GalNAc, $J_{3,4}$ 3.2 Hz, $J_{4,5} \leq 1$ Hz), 3.909 (dd, 1H, H-4 Gal, $J_{3,4}$ 3.4 Hz, $J_{4,5} \leq 1$ Hz), 3.884 (dd, 1H, H-3 GalNAc, $J_{3,4}$ 3.2 Hz, $J_{2,3}$ 11.0 Hz), 3.825 (dd, 1H, H-6$^b$ GlcNAc, $J_{5,6^b}$ 4.9 Hz, $J_{6^a6^b}$ 12.5 Hz), 3.66–3.79 (m, 8H, H-2, H-3, H-4 GlcNAc, H-5, H-6$^a$, H-6$^b$ Gal, H-5 GalNAc, OCH sp), 3.648 (dd, 1H, H-3 Gal, $J_{3,4}$ 3.4 Hz, $J_{2,3}$ 9.8 Hz), 3.593 (m, 1H, H-5 GlcNAc), 3.527 (dd, 1H, H-2 Gal, $J_{1,2}$ 7.9 Hz, $J_{2,3}$ 9.8 Hz), 3.46–3.53 (m, 1H, OCH$_2$ sp), 3.104 (m, 2H, NCH$_2$ sp), 2.022 (s, 3H, NHCO*Me*), 2.002 (s, 3H, NHCO*Me*), 1.91–2.01 (m, 2H, CH$_2$ sp). MS, *m/z*: 666 (643 + 23) ($M^+ + Na^+$). $[\alpha]_D$ +42 (*c*1, $H_2O$).

*Galβ1-4GlcNAcβ1-6(Galβ1-3)GalNAcα1-O(CH$_2$)$_3$NH$_2$* **4**: $^1$H NMR (500 MHz, $D_2O$) δ: 4.900 (d, 1H, H-1 GalNAc, $J_{1,2}$ 3.7 Hz), 4.557 (d, 1H, H-1 GlcNAc, $J_{1,2}$ 8.2 Hz), 4.491 (d, 1H, H-1 Galβ1-3, $J_{1,2}$ 8.1 Hz), 4.474 (d, 1H, H-1 Galβ1-4, $J_{1,2}$ 9.0 Hz), 4.350 (dd, 1H, H-2 GalNAc, $J_{1,2}$ 3.7 Hz, $J_{2,3}$ 11.0 Hz), 4.245 (dd, 1H, H-4 GalNAc, $J_{4,5} \leq 1$ Hz, $J_{3,4}$ 2.3 Hz), 4.114 (dd, 1H, H-6$^a$ GalNAc, $J_{5,6^a}$ 2 Hz, $J_{6^a6^b}$ 10.5 Hz), 3.99–4.08 (m, 3H, H-3 and H-6$^b$ GalNAc, H-6$^a$ GlcNAc), 3.947 (dd, 1H, H-4 Galβ1-4, $J_{3,4}$ 3.2 Hz, $J_{4,5} < 1$ Hz), 3.927 (dd, 1H, H-4 Galβ1-3, $J_{3,4}$ 3.1 Hz, $J_{4,5} \leq 1$ Hz), 3.862 (dd, 1H, H-6$^b$ GlcNAc, $J_{5,6^b}$ 4.9 Hz, $J_{6^a6^a}$ 12.3 Hz), 3.70–3.83 (m, 11H, H-5 GalNAc, H-2, H-3, H-4 GlcNAc, 2H-5, 2H-6$^a$, 2H-6$^b$, Gal, OCH sp), 3.685 (dd, 1H, H-3 Galβ1-4, $J_{2,3}$ 10.0 Hz, $J_{3,4}$ 3.2 Hz), 3.634 (dd, 1H, H-3 Galβ1-3, $J_{3,4}$ 3.2 Hz, $J_{2,3}$ 9.8 Hz), 3.60–3.66 (m, 1H, H-5 GlcNAc), 3.51–3.59 (m, 3H, 2H-2 Gal, OCH sp), 3.152 (m, 2H, NCH$_2$ sp), 2.040 (s, 3H, NHCO*Me*), 1.90–2.10 (m, 2H, CH$_2$ sp). MS, *m/z*: 828 (805 + 23) ($M^+ + Na^+$). $[\alpha]_D$ +40 (*c*1, $H_2O$).