Synthesis of GlcNAc-terminated oligosaccharides — fragments of glycoprotein O-chains

Galina V. Pazynina and Nicolai V. Bovin*

M. M. Shemyakin–Yu. A. Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, 117871 Moscow, Russian Federation. Fax: +7 095 330 5592; e-mail: bovin@carb.siobc.ras.ru

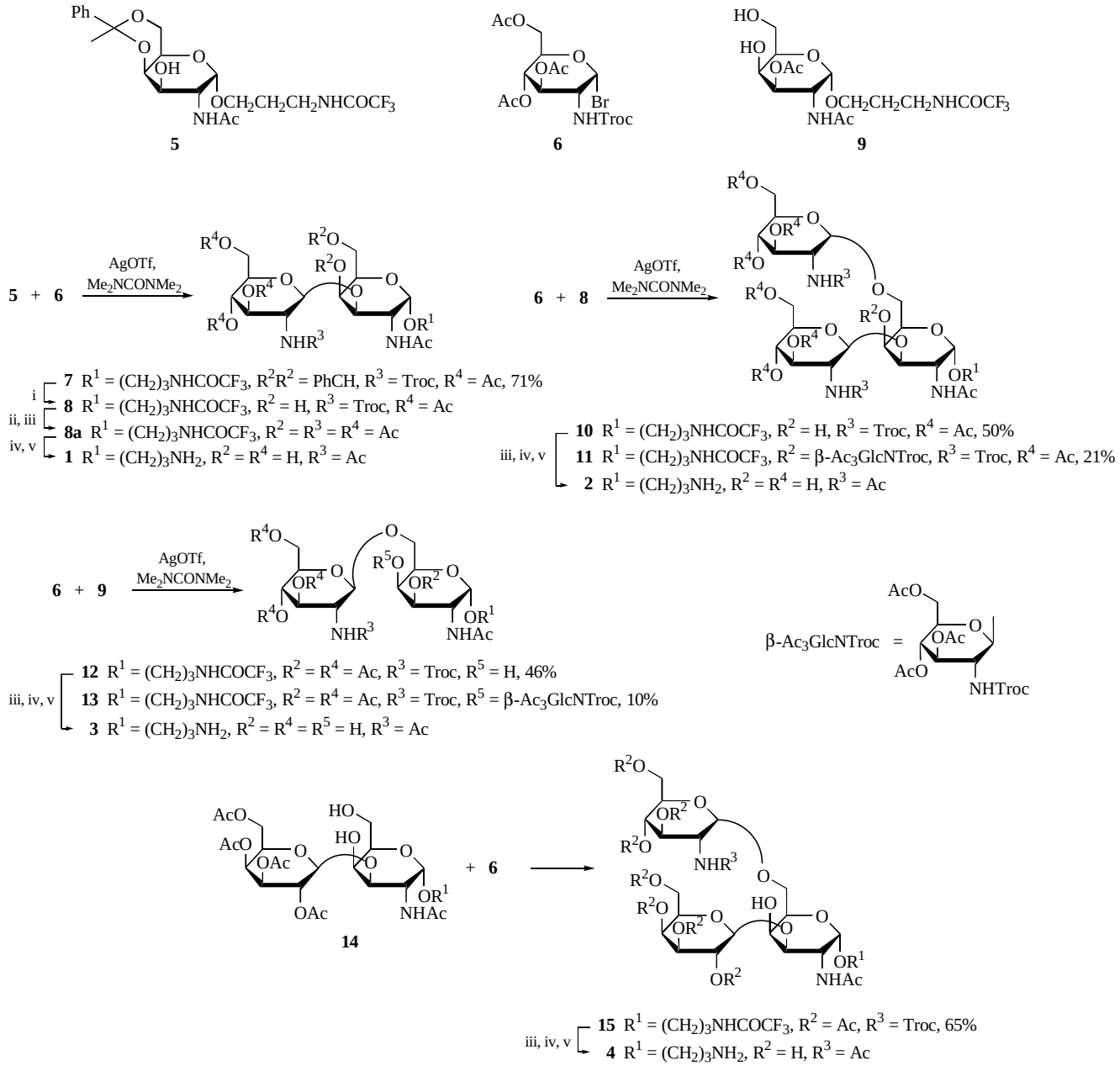
10.1070/MC2000v010n04ABEH001277

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

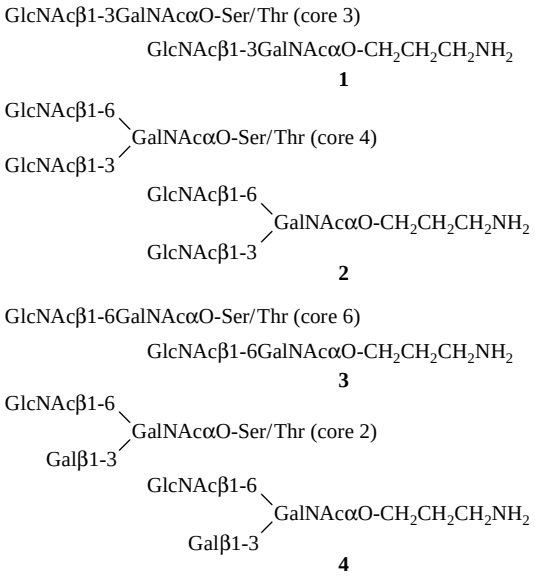
Eight different cores, *i.e.*, regions connected with Ser or Thr of mammalian glycoprotein O-chains are known; four of them are terminated by a GlcNAcβ moiety.[1]

The above GlcNAcβ-terminated oligosaccharides have been identified in mucin-type glycoproteins; at least one of them, core 6 structure, being a substrate for enzyme β-galactosyltransferase, plays a role in oncotransformation.[2] Here we describe the synthesis of core 2, 3, 4, and 6 oligosaccharides as spacered compounds **1**–**4** suitable for further transformation to different glycoprobes,[3] which are indispensable tools for glycobiology.

We used 3′-trifluoroacetamidopropyl 2-acetamido-4,6-*O*-benzylidene-2-deoxy-α-D-galactopyranoside **5** as a starting building block.[4] To synthesise disaccharide **1**, compound **5** was glycosylated at the single unprotected 3-OH position (Scheme 1).



Scheme 1 *Reagents and conditions*: i, AcOH; ii, Ac_2O, Py; iii, Zn, AcOH, Ac_2O; iv, MeONa, MeOH; v, Et_3N, H_2O.



Obtained disaccharide **7** was the starting material in the synthesis of trisaccharide **2**; **7** was 4,6-O-debenzylidenated followed by regioselective glycosylation at the position 6-OH of obtained diol **8**. In order to synthesise disaccharide **3**, the 3-OH group in compound **5** was protected by acetylation followed by removal of the benzylidene group; obtained diol **9** was 6-O-glycosylated as described above. Trisaccharide **4** was synthesised by 6-O-regioselective glycosylation of disaccharide **14**[5] with bromide **6**.

In all cases, *N*-Troc-protected glucosyl bromide **6** (Troc is 1,1,1-trichloroethyloxycarbonyl)[6] was used for stereocontrolled β-glycosylation, this choice is stipulated by good yields, high β-stereoselectivity of glycosylation,[6] and compatibility of this protection with other protecting groups used in the synthesis. Preparative yields of glycosylation with this reagent are shown in Scheme 1, the yields are comparable with published data.[6] Because glycosyl donor **6** was taken in some excess, glycosylation of diols **8** and **9** gave rise not only to the aimed monoglycosylation products, but also to corresponding 4,6-di-*O*-glycosyl derivatives **11** and **13** (Scheme 1). The β-stereochemistry of glycosylation with donor **6** is confirmed by corresponding $J_{1,2}$ values of 8.0–9.5 Hz, the regioselectivity of the glycosylation reaction in cases of diol aglycons **8**, **9** and **14** was confirmed by acetylation of the unprotected hydroxyl group in glycosylation product and an NMR study of the acetylated derivatives, using homonuclear correlation spectroscopy (COSY) and conventional analysis of coupling patterns. Mass spectra (MALDI, Vision-2000, Thermo Bio Corp., England) of deprotected compounds: **1**, 505 ($M + Na^+$); **2**, 707 ($M + Na^+$); **3**, 505 ($M + Na^+$); **4**, 666 ($M + Na^+$).

Deprotection of sugar moieties and spacer-arm with conventional methods (Scheme 1) gave rise to oligosaccharides **1**–**4** as 3-aminopropyl glycosides in overall yields of 80–90%.† The synthesis of these oligosaccharides in a convenient spacered form was not described earlier.

Compounds **1**–**4** coupled with polyacrylamide[3] were used for characterization of anti-T_k antibodies, the data will be published elsewhere.

References

1 E. F. Hounsell, M. J. Davies and D. V. Renouf, *Glycoconjugate J*., 1996, **13**, 19.
2 Y. Yamashita, Y. S. Chung, R. Horie, R. Kannagi and M. Sowa, *J. Natl. Cancer. Inst*., 1995, **87**, 441.
3 N. V. Bovin, *Glycoconjugate J*., 1998, **15**, 431.
4 N. V. Bovin, T. V. Zemlyanukhina and A. Ya. Khorlin, *Bioorg. Khim*., 1986, **12**, 533 [*Sov. J. Bioorg. Chem.* (*Engl. Transl.*), 1986, **12**, 282].
5 N. V. Bovin, T. V. Zemlyanukhina and A. Ya. Khorlin, *Bioorg. Khim*., 1985, **11**, 1256 (in Russian).
6 V. Ellervik and G. Magnusson, *Carbohydr. Res*., 1996, **280**, 251.

Received: 8th February 2000; Com. 00/1603

† ^{1}H NMR data (D_2O, 500 MHz, δ/ppm); solutions contained 5% of trifluoroacetic acid. Signals of the $OCH_2CH_2CH_2N$ group: 3.6–3.7, 3.0–3.1, 1.8–2.0.

For **1**, GlcNAcβ1-3 unit: 4.516 (d, H-1, $J_{1,2}$ 8.5 Hz), 3.631 (dd, H-2, $J_{2,3}$ 9.9 Hz), 3.481 (dd, H-3, $J_{3,4}$ 9.3 Hz), 3.397 (dd, H-4, $J_{4,5}$ 9.3 Hz), 3.494 (m, H-5), 3.833 (dd, H-6a, $J_{5,6a}$ 1.5 Hz, $J_{6a,6b}$ 12.8 Hz), 3.351 (unresolved m, H-6b), 1.955 (s, Ac); GalNAcα1-Osp unit: 4.790 (d, H-1, $J_{1,2}$ 3.8 Hz), 4.197 (dd, H-2, $J_{2,3}$ 11.8 Hz), 3.898 (dd, H-3, $J_{3,4}$ 3.0 Hz), 4.150 (d, H-4, $J_{4,5} < 1.0$ Hz), 3.658–3.778 (m, H-6a and H-6b), 1.983 (s, Ac).

For **2**, GlcNAcβ1-3 unit: 4.487 (d, H-1, $J_{1,2}$ 8.7 Hz), 3.620 (dd, H-2, $J_{2,3}$ 10.2 Hz), 3.385–3.490 (m, H-3), 3.370 (dd, H-4, $J_{4,5}$ 10.2 Hz), 3.855 (dd, H-6a, $J_{5,6a}$ 1.5 Hz, $J_{6a,6b}$ 12.8 Hz), 1.937 (s, Ac); GalNAcβ1-6 unit: 4.424 (d, H-1, $J_{1,2}$ 8.8 Hz), 3.601 (dd, H-2, $J_{2,3}$ 10.2 Hz), 3.353 (dd, H-4, $J_{4,5}$ 10.2 Hz), 3.809 (dd, H-6a, $J_{5,6a}$ 1.5 Hz, $J_{6a,6b}$ 12.8 Hz), 1.961 (s, Ac); GalNAcα1-Osp unit: 4.732 (d, H-1, $J_{1,2}$ 3.8 Hz), 4.164 (dd, H-2, $J_{2,3}$ 11.5 Hz), 3.869 (dd, H-3, $J_{3,4}$ 3.0 Hz), 4.114 (d, H-4, $J_{4,5} \leqslant 1$ Hz), 3.942 (m, H-5), 4.008 (dd, H-6a, $J_{5,6a}$ 3.0 Hz, $J_{6a,6b}$ 11.5 Hz), 1.937 (s, Ac).

For **3**, GlcNAcβ1-6 unit: 4.449 (d, H-1, $J_{1,2}$ 9.0 Hz), 3.640 (dd, H-2, $J_{2,3}$ 10.2 Hz), 3.468 (dd, H-3, $J_{3,4}$ 8.3 Hz), 3.367 (dd, H-4, $J_{4,5}$ 9.6 Hz), 3.868 (dd, H-6a, $J_{5,6a}$ 1.5 Hz, $J_{6a,6b}$ 12.5 Hz), 1.958 (s, Ac); GalNAcα-Osp unit: 4.814 (d, H-1, $J_{1,2}$ 3.8 Hz), 4.088 (dd, H-2, $J_{2,3}$ 11.5 Hz), 3.836 (dd, H-3, $J_{3,4}$ 3.2 Hz), 3.905 (d, H-4, $J_{4,5} \leqslant 1$ Hz), 3.935 (m, H-5), 4.018 (dd, H-6a, $J_{5,6a}$ 3.5 Hz, $J_{6a,6b}$ 11.5 Hz), 1.972 (s, Ac).

For **4**, GlcNAcβ1-6 unit: 4.489 (d, H-1, $J_{1,2}$ 8.5 Hz), 3.682 (dd, H-2, $J_{2,3}$ 10.2 Hz), 3.513 (dd, H-3, $J_{3,4}$ 8.2 Hz), 3.426 (dd, H-4, $J_{4,5}$ 8.2 Hz), 3.906 (dd, H-6a, $J_{5,6a}$ 1.5 Hz, $J_{6a,6b}$ 11 Hz), 1.995 (s, Ac); Galβ1-3 unit: 4.432 (d, H-1, $J_{1,2}$ 7.7 Hz), 3.499 (dd, H-2, $J_{2,3}$ 11.1 Hz), 3.588 (dd, H-3, $J_{3,4}$ 3.5 Hz), 3.884 (dd, H-4, $J_{4,5} \leqslant 1$ Hz); GalNAcα-Osp unit: 4.852 (d, H-1, $J_{1,2}$ 3.8 Hz), 4.301 (dd, H-2, $J_{2,3}$ 11.1 Hz), 3.982 (dd, H-3, $J_{3,4}$ 3.0 Hz), 4.193 (d, H-4, $J_{4,5} \leqslant 1$ Hz), 4.003 (m, H-5), 4.058 (dd, H-6a, $J_{5,6a}$ 2.5 Hz, $J_{6a,6b}$ 10.8 Hz), 1.995 (s, Ac).