Pergamon

Tetrahedron Letters 41 (2000) 6279–6284

TETRAHEDRON LETTERS

First total synthesis of sialylated and sulfated Lewis^x mucin Core 2 structures as potential tumor associated antigens

Bao-Guo Huang, Rakesh K. Jain, Robert D. Locke, James L. Alderfer, Walter A. Tabaczynski and Khushi L. Matta*

Molecular & Cellular Biophysics, Roswell Park Cancer Institute, Elm & Carlton Streets, Buffalo, NY 14263, USA

Received 7 March 2000; accepted 2 June 2000

Abstract

Two branched Core 2 structures, (3-*O*-SO_3Na)Galβ1→4(Fucα1→3)GlcNAcβ1→6(NeuAcα2→3Galβ1→3)GalNAcαOMe (**1**) and its positional isomer NeuAcα2→3Galβ1→4(Fucα1→3)GlcNAcβ1→6(3-OSO_3Na-Galβ1→3)GalNAcαOMe (**2**), were chemically synthesized for the first time as potential tumor associated antigens. © 2000 Elsevier Science Ltd. All rights reserved.

Keywords: antigens; glycosylation; Core 2 structure; oligosaccharides.

In practice, knowledge of well defined carbohydrate antigens has enabled chemists to undertake their synthesis.[1–4] During our studies of enzymes involved in the assembly of tumor associated antigens, we realized that knowledge acquired on their specificity could provide valuable information regarding the structures of unknown antigens. We hypothesize that a study of sulfotransferases combined with knowledge of enzyme specificities, such as those of α-fucosyltransferase and sialyltransferase, can indicate the potential existence of novel sulfated glycoconjugates in, e.g. breast, colon and ovarian tumor tissue sources. For example, our laboratory was the first to examine the specificity of α(1→3/4)-L-fucosyltransferase from ovarian tumor tissues and cell lines. We postulated that some sulfated glycoproteins could contain sulfated Lewis^a type structures.[5] Feizi et al.[6] subsequently reported the existence of such 3-*O*-sulfated Lewis^x and Lewis^a structures as part of ovarian adenocarcinoma glycoproteins. We have examined sulfotransferase specificity using Core 2 [Galβ1→4GlcNAcβ1→6(Galβ1→3)GalNAcα→OR] acceptors. It is striking that in colon tissue and cell lines sulfotransferase acts upon the C-3 position of Gal in the Galβ1→4GlcNAc arm of this Core 2 acceptor, whereas in breast tumor tissue C-3 of Gal in the Galβ1→3GalNAcα1→ sequence is sulfated.[7,8] We have now synthesized the branched Core 2 oligosaccharide structures, (3-*O*-SO_3Na)Galβ1→4(Fucα1→3)GlcNAcβ1→6 (NeuAcα2→3Galβ1→3)GalNAcαOMe (**1**) and its positional isomer

* Corresponding author. Tel: (716) 845-2397; fax: (716) 845-3458; e-mail: klmatta@yahoo.com

PII: S0040-4039(00)00975-8

NeuAcα2→3Galβ1→4(Fucα1→3)GlcNAcβ1→6(3-OSO_3Na-Galβ1→3)GalNAcαOMe (**2**), as novel epitopes for colon and breast carcinoma, respectively. During our studies, Kim and co-workers reported that a sulfated Lewis^x determinant is a major structural moiety of O-linked glycans obtained from human colon carcinoma mucin.[9] These findings provide support for our above-mentioned hypothesis.

Compounds **1** and **2** were synthesized from three key building blocks (**3**–**5**, Fig. 1).[10,11] Our synthetic strategy for **1** (Scheme 1) was based on the employment of a novel glycosyl donor (**6**), having the 3,4-isopropylidene galactopyranosyl moiety, which was easily obtained from **3** in two steps. The core oligosaccharide **8** was chosen as the key glycosyl acceptor for the synthesis of **1**. Compound **8** has the C-2, C-3 and C-4 hydroxy groups of galactose in Galβ1,3GalNAc free for regioselective sialylation at the C-3 position and 3,4-isopropylidene protection on Gal in the Lewis^x moiety, which upon acetylation followed by deisopropylidenation provided the site for selective sulfation at C-3. On the other hand, compound **15** was employed as a key acceptor for the construction of the target molecule **2** (Scheme 2). Compound **15** has C-2, C-3 and C-4 hydroxy groups on Gal of Lewis^x free for selective sialylation at the C-3 position and 3,4-isopropylidene protection on Gal of Galβ1,3GalNAc for further selective sulfation under a similar sequence of reactions.

Selective deacetylation of Lewis^x donor **3** followed by isopropylidene formation produced the key glycosyl donor **6**. Glycosylation of **6** with disaccharide acceptor **4** in the presence of NIS–triflic acid[12] led to the formation of exclusively 1,6-β-linked pentasaccharide **7** in 50% yield. The ^{1}H NMR spectrum of **7** displayed characteristic signals at δ 1.15 and 1.22 (2s, 18H, 2×Piv) and δ 1.86, 1.93, 1.99 and 2.10 (4s, 12H, 4×CH_3CO). Treatment of **7** with hydrazine hydrate in ethanol (v/v, 1:9) at 80°C for 2 h resulted in the selective removal of the phthalimido group and acetates to produce the free amine, which was converted to compound **8** in 75% yield. Only two *N*-acetyl groups are seen in the ^{1}H NMR spectrum (δ 1.72, 1.98), which is consistent with the structure. Under these conditions for conversion of **7** to **8**, the 6-*O*-pivaloyl group remained intact (δ 1.16, s, 9H and δ 1.19, s, 9H). Reaction of **8** with sialic acid donor **5** promoted by NIS–triflic acid in propionitrile at –75°C gave the sialylated oligosaccharide **9** in 48% yield. The NeuAcα2-3Galβ linkage was evidenced by ^{13}C NMR (δ 98.7, C-2 for NeuAc and δ 74.6, C-3 for Gal) and ^{1}H NMR (δ 1.74–2.12, 7s, 21H, 7×$COCH_3$). The low reaction temperature was very important in order to avoid the loss of fucosyl residue during glycosylation.[13] Acetylation of **9** afforded **10**, which was hydrolyzed with 60% acetic acid to produce 3,4-diol **11**. Selective sulfation of **11** with SO_3–pyridine complex in pyridine gave 3-*O*-sulfo compound **12**. The sulfation at C-3 of Gal was confirmed by ^{13}C NMR (δ 79.2, C-3 for Gal, with δ 68.1 for C-2 and 65.7 for C-4). Removal of the benzyl groups in **12** by hydrogenation followed by deacetylation and hydrolysis of methyl ester produced the target molecule **1** (48% from **12**).

For the synthesis of compound **2**, disaccharide **13** was chosen as a key glycosyl acceptor. Compound **13** was prepared in 39% overall yield from compound **4**. Silylation of **4** with diphenyl-*t*-butylsilyl chloride in DMF followed by selective deacetylation, 3,4-isopropylidene formation and desilylation gave **13**. Condensation of the resulting glycosyl acceptor **13** with Lewis^x donor **3** in the presence of NIS–triflic acid afforded the pentasaccharide **14**. Selective deacetylation of **14** with sodium methoxide in methylene chloride and methanol (v/v, 1:1) gave compound **15**. This was further sialylated to produce compound **16**. The ^{1}H NMR spectrum of compound **16** gave signals at δ 1.84–2.12 (6s, 18H, 6×$COCH_3$) and δ 1.22, 1.23 (2s, 18H, 2×Piv). The site of sialylation was evidenced by ^{13}C NMR (δ 76.5, C-3, Gal). Acetylation of **16** with acetic anhydride afforded compound **17**. Removal of 3,4-isopropylidene with 60% acetic acid gave the

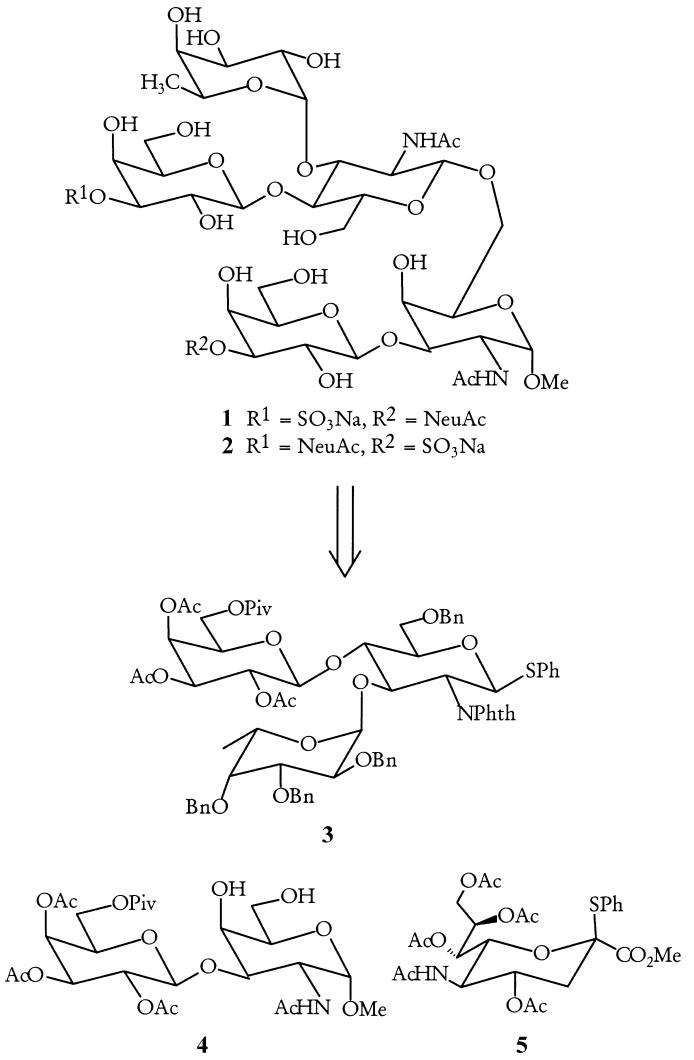
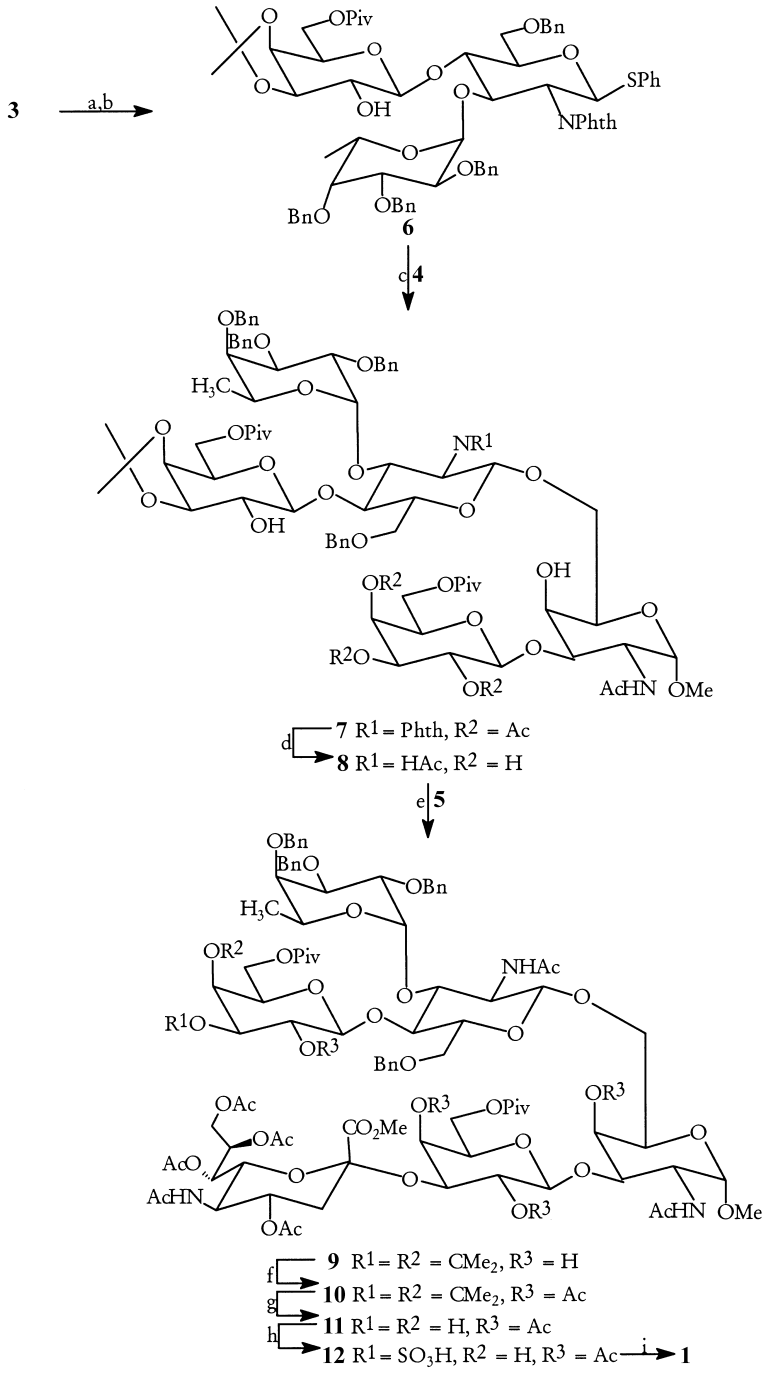


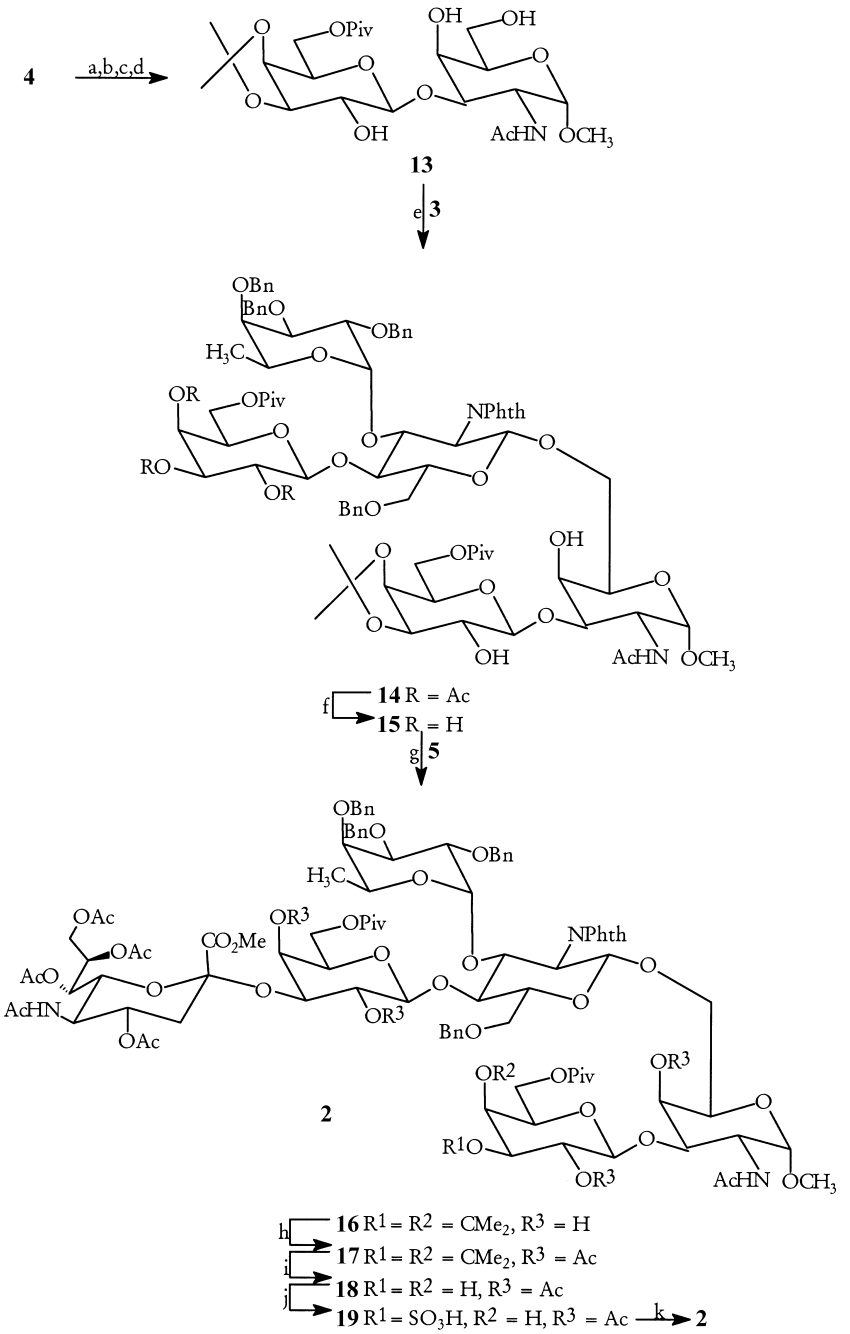
Figure 1. The building blocks **3**, **4** and **5** used in the synthesis of compounds **1** and **2**

diol **18**. Selective sulfation of **18** with SO_3–pyridine complex gave compound **19**. Deprotection of **19** with LiI in pyridine[14] followed by hydrazine hydrate produced a free amine intermediate, which was treated with acetic anhydride in methanol and triethylamine, deacetylation and sodium resin sequentially to furnish the final product **2** (40% from **19**). The structures of the newly synthesized molecules **1**, **2** and some intermediates were characterized by NMR and mass spectra.[15]

Application of 1D and 2D NMR techniques allowed unambiguous assignment of all ^1H and ^{13}C resonances of compounds **1** and **2**. ^1H COSY spectra were used to sort ^1H resonances into spin systems. These were assigned to particular residues using structural reporter groups. ^{13}C assignments were made using a combination of HMQC and HMBC spectra. Glycosidic linkages and substitution sites were confirmed using the HMBC data. Details of the NMR assignments will be published elsewhere.



Scheme 1. Reagents and conditions: (a) NaOMe, MeOH:CH_2Cl_2 (1:1, v/v), 0°C, 5 h, 80%; (b) CSA in DMP, rt, 2 h, 82%; (c) 1.05 equiv. **6**, NIS–TfOH, CH_2Cl_2, –75°C, 2 h, 50%; (d) hydrazine hydrate:EtOH (1:9, v/v), 80°C, 2 h; MeOH:Et_3N:Ac_2O (4:2:1, v/v/v), rt, 2 h, 75%; (e) 2.5 equiv. **5**, NIS–TfOH, CH_3CH_2CN, –75°C, 2 h, 48%; (f) Ac_2O, pyr., rt, 12 h; (g) 60% HOAc, 55°C, 3 h, 54% for steps f and g; (h) 10 equiv. SO_3–pyridine in pyr., 0°C, 12 h, 75%; (i) 10% Pd/C, H_2, MeOH, 12 h; NaOMe, MeOH, rt, 48 h; H_2O, 5 h, Na^+ resin, 48% from **12**



Scheme 2. Reagents and conditions: (a) Ph_2Si^tBuCl, imidazole, DMF, 0°C, 2 h; (b) NaOMe, MeOH:CH_2Cl_2 (1:1, v/v), 0°C, 2 h; (c) CSA in DMP, rt, 2 h; (d) TBAF in THF, rt, 30 min, 39% for four steps; (e) 1.05 equiv. **3**, NIS–TfOH, CH_2Cl_2, –75°C, 2 h, 32%; (f) NaOMe, MeOH:CH_2Cl_2 (1:1, v/v), 0°C, 3 h, 70%; (g) 2.5 equiv. **5**, NIS–TfOH, CH_3CH_2CN, –75°C, 3 h; (h) Ac_2O, pyr., rt, 16 h, 32% for two steps; (i) 60% HOAc, 65°C, 1.5 h, 84%; (j) SO_3–pyridine complex, pyr., 0°C, 16 h, 90%; (k) LiI, pyr., 120°C, 4 h; hydrazine hydrate:MeOH (5:1, v/v), 80°C, 3 h, MeOH Et_3N:Ac_2O (4:2:1, v/v/v), rt, 2 h, NaOMe, MeOH, rt, 48 h; Na^+ resin, 40% from **19**

In summary, novel procedures for the total synthesis of the sulfated hexasaccharides **1** and **2** have been developed. These procedures will also provide a practical approach toward the synthesis of a series of branched Core 2 type linked oligosaccharides, which contain allyl or other functional groups at the anomeric position, suitable for attachment to protein for immunological studies.

Acknowledgements

These investigations were supported by Grant Nos. CA-63218 and CA-35329 (K.L.M.), and in part by CA-16056 (RPCI-NMR facility) awarded by the National Cancer Institute.

References

1. Stults, C. L. M.; Sweeley, C. C.; Macher, B. A. *Methods Enzymol*. **1989**, *179*(14), 167.
2. Livingston, P. O.; Zhang, S.; Lloyd, K. O. *Cancer Immunol. Immunother*. **1997**, *45*, 1.
3. Toyokumi, T.; Singhal, A. K. *Chem. Soc. Rev*. **1995**, 231.
4. Seeberger, P. H.; Bilodeau, M. T.; Danishefsky, S. J. *Aldrichim. Acta*, **1997**, *30*, 75.
5. Chandrasekaran, E. V.; Jain, R. K.; Matta, K. L. *J. Biol. Chem*. **1992**, *267*, 23806.
6. Yuen, C.; Lawson, A. M.; Chai, W.; Larking, M.; Stoll, M. S.; Stuart, A. C.; Sullivan, F. X.; Ahern, T. J.; Feizi, T. *Biochemistry* **1992**, *31*, 9126.
7. Chandrasekaran, E. J.; Jain, R. K.; Matta, K. L. *Glycobiology* **1997**, *7*, 753.
8. Chandrasekaran, E. V.; Jain, R. K.; Rhodes, J. M.; Chawda, R.; Piskorz, C.; Matta, K. L. *Glycoconjugate J*. **1999**, *16*, 523–536.
9. Capon, C.; Wieruszeski, J. M.; Lemoine, J.; Byrd, J. C.; Leffler, H.; Kim, Y. S. *J. Biol. Chem*. **1997**, *272*, 31957.
10. Huang, B.-G.; Locke, R. D.; Jain, R. K.; Matta, K. L. *Bioorg. Med. Chem. Lett*. **1997**, *7*(9), 1157.
11. Koenig, A.; Jain, R.; Vig, R.; Norgard-Sumnicht, K. E.; Matta, K. L.; Varki, A. *Glycobiology* **1997**, *7*, 79.
12. Konradason, P.; Mootoo, D. R.; McDevitt, R. E.; Fraser-Reis, B. *J. Chem. Soc., Chem. Commun*. **1990**, 270.
13. Jain, R. K.; Vig, R.; Rampal, R.; Chandrasekaran, E. V.; Matta, K. L. *J. Am. Chem. Soc*. **1994**, *116*, 12123.
14. Nicolaou, K. C.; Hummel, C. W.; Iwabuchi, Y. *J. Am. Chem. Soc*. **1992**, *114*, 3126.
15. NMR spectra were recorded at 30°C with Bruker AM-400 and AMX-600 spectrometers (^{1}H frequencies are 400 and 600 MHz respectively). *Selected data*. For **1**: $[\alpha]_D$ +48 (*c* 0.70; H_2O); ^{1}H NMR (D_2O): δ 4.80 (d, $J=3.6$ Hz, H-1, GalNAcα-OMe), 4.55 (d, 1H, $J=7.8$ Hz, H-1′, Galβ1-3GalNAc), 4.59(d, 1H, $J=8.3$ Hz, H-1‴, GlcNAc), 5.14 (d, 1H, $J=4.0$ Hz, H-1⁗, Fuc), 4.60 (d, 1H, $J=7.8$ Hz, H-1⁗′, Galβ1-4GlcNAc), 3.38 (s, 3H, OMe), 2.78 (dd, 1H, $J=4.7$, 12.4 Hz, H-3″e), 2.03, 2.04, 2.06 (3s, 9H, 3×NAc), 1.81 (dd, 1H, $J=11.8$, 12.4 Hz, H-3″a), 1.20 (d, 3H, $J=6.7$ Hz, H-6⁗); ^{13}C NMR (D_2O): δ 173.6, 173.1, 174.0 (3×NHCOCH$_3$), 172.8 (C-1″, NeuAc), 97.1 (C-1, Galα-OMe), 103.4(C-1′, Galβ1-3GalNAc), 98.7 (C-2″, NeuAcα), 100.4 (C-1‴, GlcNAc), 97.6 (C-1⁗, Fuc), 100.4 (C-1⁗′, Galβ1-4GlcNAc), 76.2 (C-3), 69.2 (C-6), 74.6 (C-3′, site of sialylation), 73.8 (C-3‴), 72.4 (C-4‴), 79.2 (C-3⁗′, site of sulfation on Galβ1-4GlcNAc), 53.9 (OMe), 54.7 (C-2‴), 21.0, 21.1, 21.3 (3×NHCOCH$_3$), 14.2(C-6⁗, Fuc); MS *m*/*z*: 1300.4 (M–Na)$^-$. For **2**: $[\alpha]_D$ +75 (*c* 0.50; H_2O); ^{1}H NMR (D_2O): δ 4.79 (d, 1H, H-1, $J=3.7$ Hz, GalNAcα-OMe), 4.58 (d, 1H, H-1′, $J=7.9$ Hz, Galβ1-3GalNAc), 4.59 (d, 1H, H-1″, $J=8.2$ Hz, GlcNAcβ1-6GalNAc), 5.13 (d, 1H, $J=4.0$ Hz, H-1‴, Fuc), 4.55 (d, 1H, $J=7.9$ Hz, H-1⁗′, Galβ1-4GlcNAc), 3.38 (s, 3H, OMe), 2.79 (dd, 1H, $J=4.7$, 12.4 Hz, H-3⁗′e), 2.03, 2.04, 2.06 (3s, 9H, 3×NHCOCH$_3$), 1.82 (dd, 1H, $J=12.1$, 12.4 Hz, H-3⁗′a), 1.20 (d, 3H, H-6‴, Fuc); ^{13}C NMR (D_2O): δ 173.1, 173.6, 174.0 (3×NHCOCH$_3$), 172.8 (C-1⁗′, NeuAc), 97.2 (C-1, GalNAcα-OMe), 103.3 (C-1′, Galβ1-3GalNAc), 100.4 (C-1″, GlcNAcβ1-6GalNAc), 97.6 (C-1‴, Fuc), 100.6 (C-1⁗′, Galβ1-4GlcNAc), 98.7 (C-2⁗′, NeuAc), 76.5 (C-3), 69.2 (C-6), 79.2 (C-3′, site of sulfation), 73.8 (C-3″), 72.4 (C-4″), 74.7 (C-3⁗′, site of sialylation), 53.9 (OMe), 21.0, 21.0, 21.3 (3×NHCOCH$_3$), 14.3 (C-6‴); MS *m*/*z*: 1300.0 (M–Na)$^-$.