Supplemental information belonging to:

Sialylation of Galactosyl-lactoses Using *Trypanosoma cruzi Trans*-Sialidase as Biocatalyst and Bovine κ-Casein-Derived Glycomacropeptide as Donor Substrate

Maarten H. Wilbrink,[a] Geert A. ten Kate,[a] Sander S. van Leeuwen,[a] Peter Sanders,[a] Erik Sallomons,[b] Johannes A. Hage,[c] Lubbert Dijkhuizen,[a] Johannis P. Kamerling[a]

Microbial Physiology, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, Groningen, The Netherlands[a]; Royal FrieslandCampina Innovation Center, Wageningen, The Netherlands[b]; Darlingii, Son, The Netherlands[c]

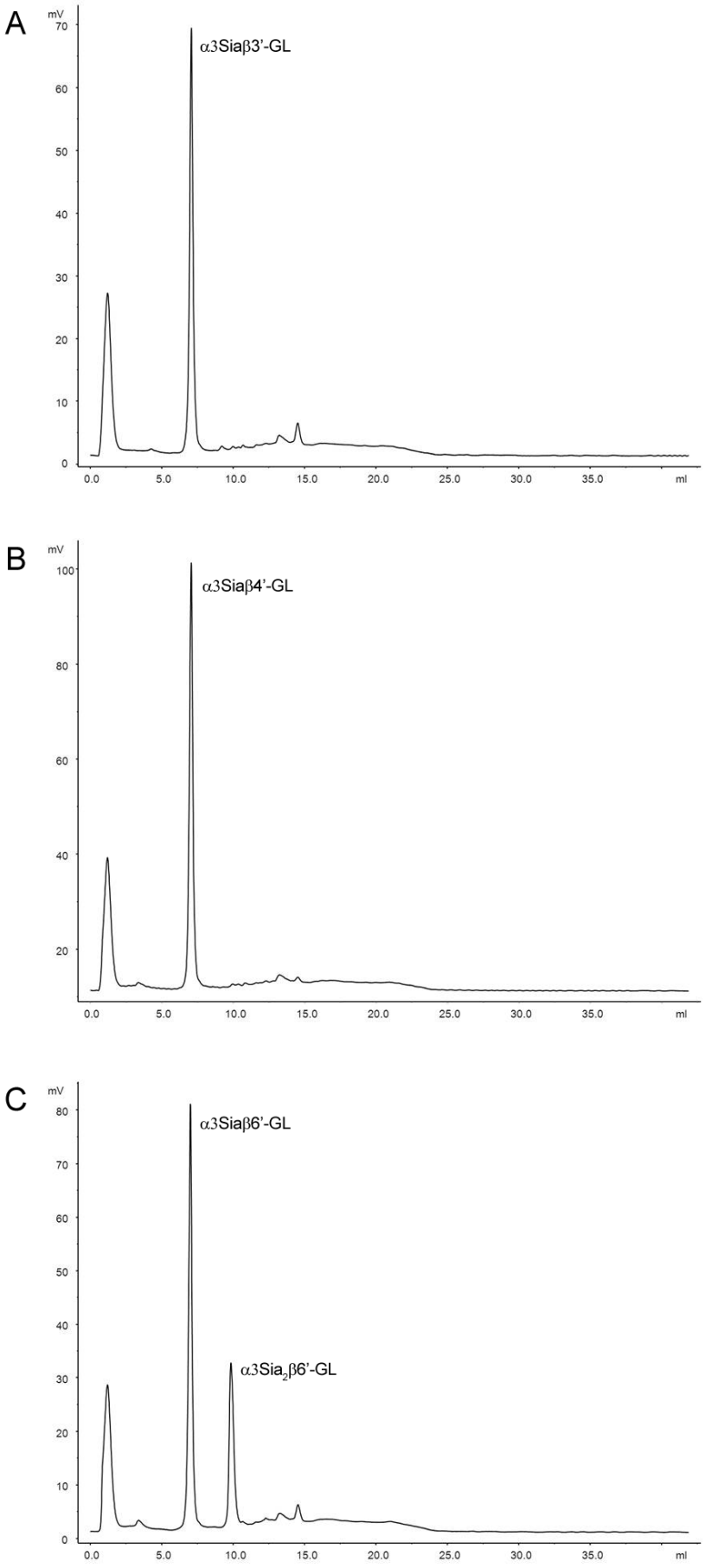

FIG. S1. Anion-exchange chromatograms on Resource Q of TcTS-catalyzed incubations of (A) β3′-GL, yielding α3Siaβ3′-GL, (B) β4′-GL, yielding α3Siaβ4′-GL, and (C) β6′-GL, yielding α3Siaβ6′-GL and $\alpha3Sia_2\beta6'$-GL, using GMP as sialic acid donor (24 h, 25°C, pH 5.0; UV detection).

NMR analysis of sialylated 3′-galactosyl-lactose (β3′-GL), 4′-galactosyl-lactose (β4′-GL), and 6′-galactosyl-lactose (β6′-GL).

In the following discussion of the NMR data, the sequence of each GL is represented by **C**-**B**-**A** with **C** as terminal Gal unit, **B** as internal Gal unit, and **A** as reducing Glc unit; **D** stands for Neu5Ac(α2-3). The ^{1}H NMR spectra of the acceptor substrates β3′-GL, β4′-GL, and β6′-GL are included in Figs. S2, S3, and S4, respectively.

The ^{1}H NMR spectrum of mono-sialylated β3′-GL (Fig. S2, Table 1, MALDI-TOF-MS: $[M-H]^-$ *m/z* 795.26) showed anomeric signals at δ 5.223 (**Aα** H-1), δ 4.664 (**Aβ** H-1), δ 4.513 (**B** H-1), and δ 4.689 (**C** H-1). The **D** H-3a and H-3e signals at δ 1.802 and δ 2.764, respectively, are indicative of a Neu5Ac(α2-3) residue (1). The ^{1}H NMR spectrum is identical to that earlier reported for Neu5Ac(α2-3)Gal(β1-3)Gal(β1-4) Glc (α3Siaβ3′-GL) (2).

In the ^{1}H NMR spectrum of mono-sialylated β4′-GL (Fig. S3, Table 1, MALDI-TOF-MS: $[M-H]^-$ *m/z* 795.26) anomeric signals are observed at δ 5.220 (**Aα** H-1), δ 4.663 (**Aβ** H-1), δ 4.473 (**B** H-1), and δ 4.660 (**C** H-1). The Neu5Ac(α2-3) residue is reflected by the **D** H-3a and H-3e signals at δ 1.796 and δ 2.762, respectively (1). Going from β4′-GL to mono-sialylated β4′-GL (Fig. S3), the sialylation of O-3 of the terminal Gal residue in β4′-GL is evident from the strong downfield shift of **C** H-3 (δ 3.66 → δ 4.107). Summarizing the analytical data, the structure of mono-sialylated β4′-GL is Neu5Ac(α2-3)Gal(β1-4)Gal(β1-4)Glc (α3Siaβ4′-GL).

The ^{1}H NMR spectrum of mono-sialylated β6′-GLa (Fig. S4B, Table S1, MALDI-TOF-MS: $[M-H]^-$ *m/z* 795.26) showed anomeric signals at δ 5.217 (**Aα** H-1), δ 4.662 (**Aβ** H-1), δ 4.535 (**B** H-1), and δ 4.470 (**C** H-1). The Neu5Ac H-3a and H-3e signals at δ 1.810 and δ 2.756, respectively, confirmed the presence of a Neu5Ac(α2-3) residue (1). Using two-dimensional ^{1}H-^{1}H COSY, ^{1}H-^{1}H TOCSY, and ^{1}H-^{1}H ROESY spectra, in combination with two-dimensional ^{13}C-^{1}H HSQC spectra (data not shown), all ^{1}H and ^{13}C chemical shifts could be assigned (Table S1). Similar to residue **A** in β6′-GL (Table S1), residue **A** showed a δ-value pattern fitting with a 4-substituted Glc residue, i.e. **Aα** H-4 at δ 3.64, **Aβ** H-4 at δ 3.65, **Aα** and **Aβ** C-4 at δ 80.0. Residue **B** turned out to be a di-substituted Gal(β1-4) residue. An O-6 substitution is indicated by clear downfield shifts of both ^{1}H and ^{13}C δ-values, i.e. **B** H-6a at δ 4.037, **B** H-6b at δ 3.92, and **B** C-6 at δ 70.1 (compare with residues **B** and **C** in β6′-GL) (3). An O-3 substitution is reflected by the downfield chemical shift values of **B** H-3 at δ 4.125 and **B** C-3 at δ 76.3 (compare with **B** H-3 at δ 3.66 and **B** C-3 at δ 73.8 in β6′-GL), combined with slight upfield shifts for **B** C-2 and **B** C-4 (compare with **B** C-2 and **B** C-4 in β6′-GL) (3) (Table S1). Residue **C** shows the δ-value pattern fitting a terminal Gal(β1-6) residue. The ROESY

spectrum showed cross-peaks between **C** H-1 and **B** H-6a,6b and between **B** H-1 and **A** H-4, supporting the **C**1-6**B**1-4**A** sequence. In view of the results, the Neu5Ac residue should be located at O-3 of residue **B**. Summarizing the NMR and MS data, the structure of mono-sialylated β6′-GLa is Gal(β1-6)[Neu5Ac(α2-3)] Gal(β1-4)Glc (α3Siaβ6′-GLa).

In the ^{1}H NMR spectrum of mono-sialylated β6′-GLb (Fig. S4C, Table S1, MALDI-TOF-MS: $[M-H]^-$ *m/z* 795.26) anomeric signals are detected at δ 5.228 (**Aα** H-1), δ 4.660 (**Aβ** H-1), δ 4.528 (**C** H-1), and δ 4.462 (**B** H-1). The Neu5Ac H-3a and H-3e signals at δ 1.815 and δ 2.755, respectively, confirmed the presence of a Neu5Ac(α2-3) residue (1). Using two-dimensional NMR spectroscopy as carried out for α3Siaβ6′-GLa, all ^{1}H and ^{13}C chemical shifts could be assigned (Table S1). Residue **A** showed for both the α- and β-anomeric configuration a similar δ-value pattern as found for α3Siaβ6′-GLa, indicating a 4-substituted reducing Glc residue. Internal residue **B** was found to be a 6-O-substituted Gal(β1-4) residue with significant downfield shifts of **B** H-6a at δ 4.073 and **B** C-6 at δ 69.8 (compare with residues **B** and **C** in β6′-GL) (3). Residue **C** showed evidence for an O-3 substitution as revealed by the downfield shifts of **C** H-3 at δ 4.100 and **C** C-3 at δ 76.8 (compare with **C** H-3 at δ 3.67 and **C** C-3 at δ 73.8 in β6′-GL), combined with slight upfield shifts for **C** C-2 and **C** C-4 (compare with **C** C-2 and **C** C-4 in β6′-GL) (3). In the ROESY spectrum cross-peaks are detected between **C** H-1 and **B** H-6a,6b and between **B** H-1 and **A** H-4, supporting the **C**1-6**B**1-4**A** sequence. In view of the results, the Neu5Ac residue should be located at O-3 of residue **C**. Summarizing the NMR and MS data, the structure of mono-sialylated β6′-GLb is Neu5Ac(α2-3)Gal(β1-6)Gal(β1-4)Glc (α3Siaβ6′-GLb).

The ^{1}H NMR spectrum of di-sialylated β6′-GL (Fig. S4D, Table S1, MALDI-TOF-MS: $[M-H]^-$ *m/z* 1086.36) showed anomeric signals at δ 5.222 (**Aα** H-1), δ 4.656 (**Aβ** H-1), δ 4.514 (**C** H-1), and δ 4.545 (**B** H-1). The Neu5Ac H-3a and H-3e signals at δ 1.812 and δ 2.752, respectively, with twice the intensity compared with α3Siaβ6′-GLa and α3Siaβ6′-GLb, confirmed the presence of two Neu5Ac(α2-3) residues (1). Using two-dimensional NMR spectroscopy, as carried out for the mono-sialylated β6′-GL components, all ^{1}H and ^{13}C chemical shifts were determined (Table S1). The δ-value pattern of residue **A**, both in the α- and β-configuration, matches that of β6′-GL, indicating a 4-substituted reducing Glc residue. Residue **B** showed a δ-value pattern comparable with that of residue **B** in α3Siaβ6′-GLa, with evidence for both O-3 (**B** H-3, δ 4.126; **B** C-3, δ 76.3) and O-6 (**B** H-6a, δ 4.044; **B** C-6, δ 70.1) substitution of the Gal(β1-4) residue (3). Residue **C** showed a δ-value pattern comparable with that of residue **C** in α3Siaβ6′-GLb, showing evidence for O-3 substitution (**C** H-3, δ 4.093; **C** C-3, δ 76.5) of the terminal Gal residue (3). The ROESY spectrum showed correlations between **B** H-1 and **A** H-4 and between **C** H-1 and **B** H-6a,6b, in agreement

with the **C**1-6**B**1-4**A** sequence. Summarizing the NMR and MS data, the structure of di-sialylated β6′-GL is Neu5Ac(α2-3)Gal(β1-6)[Neu5Ac(α2-3)]Gal(β1-4)Glc ($\alpha 3Sia_2\beta 6'$-GL).

References

1. **Vliegenthart JFG, Kamerling JP.** 2007. ^{1}H NMR Structural-reporter-group concepts in carbohydrate analysis, p. 133-191. *In* Kamerling JP, Boons GJ, Lee YC, Suzuki A, Taniguchi N, Voragen AGJ (ed), Comprehensive Glycoscience – From Chemistry to Systems Biology, vol. 2. Elsevier, Amsterdam, The Netherlands.

2. **Fukuda K, Yamamoto A, Ganzorig K, Khuukhenbaatar J, Senda A, Saito T, Urashima T.** 2010. Chemical characterization of the oligosaccharides in Bactrian camel (*Camelus bactrianus*) milk and colostrum. J. Dairy Sci. **93:**5572-5587.

3. **Bock K, Thøgersen H.** 1982. Nuclear magnetic resonance spectroscopy in the study of mono- and oligosaccharides. Annu. Rep. NMR Spectrosc. **13**:1-57.

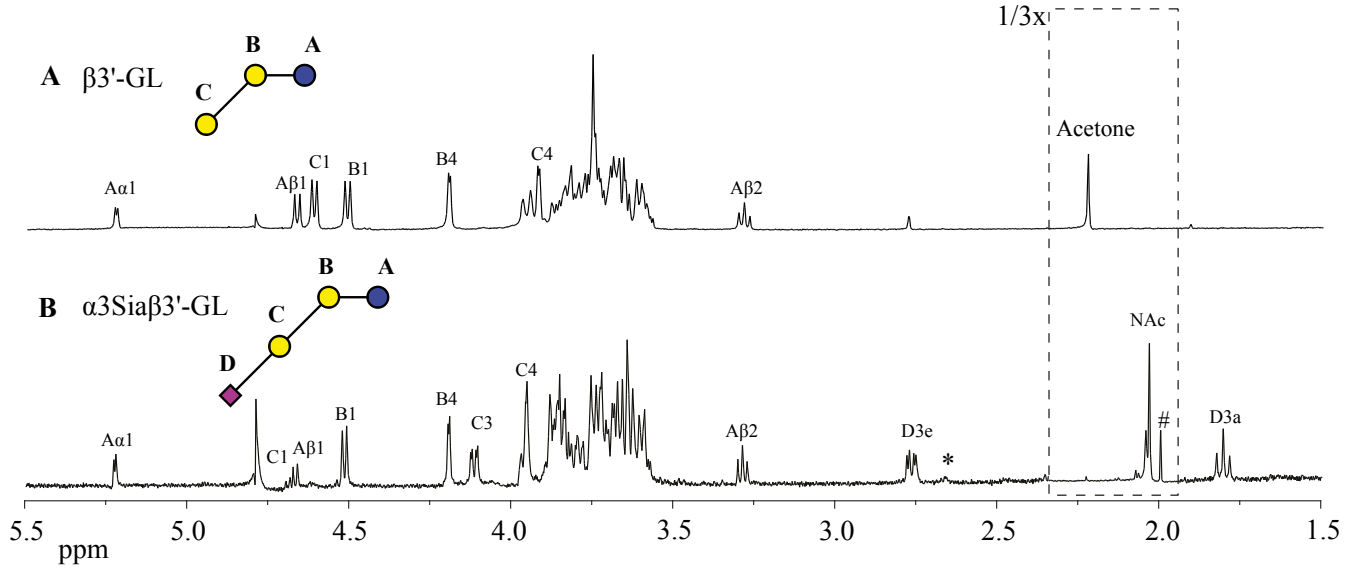

FIG. S2. One-dimensional 600-MHz ^{1}H NMR (D_2O, 298 K) spectra of (A) Gal(β1-3)Gal(β1-4)Glc (β′3-GL) and (B) Neu5Ac(α2-3)Gal(β1-3)Gal(β1-4)Glc (α3Siaβ3′-GL). Coding system: Glc, blue circle; Gal, yellow circle; Neu5Ac, pink square; * and #, contaminants.

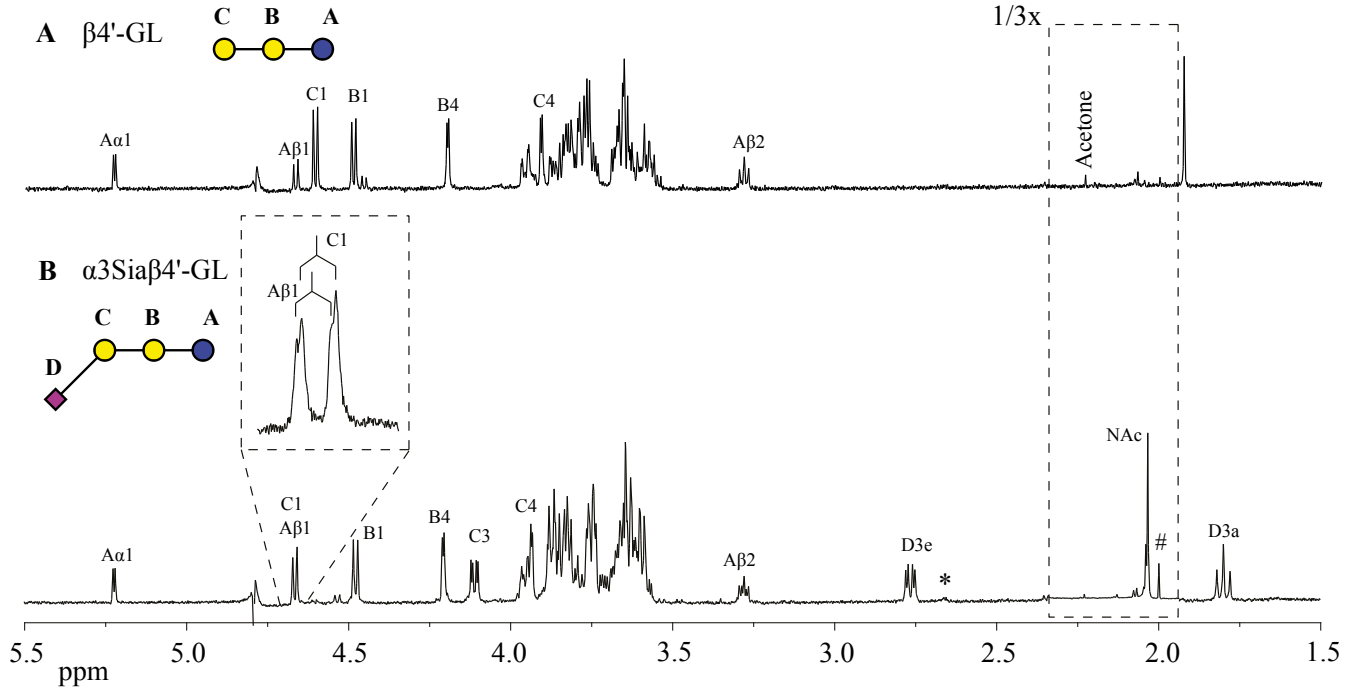

FIG. S3. One-dimensional 600-MHz ^{1}H NMR (D_2O, 298 K) spectra of (A) Gal(β1-4)Gal(β1-4)Glc (β′4-GL) and (B) Neu5Ac(α2-3)Gal(β1-4)Gal(β1-4)Glc (α3Siaβ4′-GL). Coding system, see Fig. S2; * and #, contaminants.

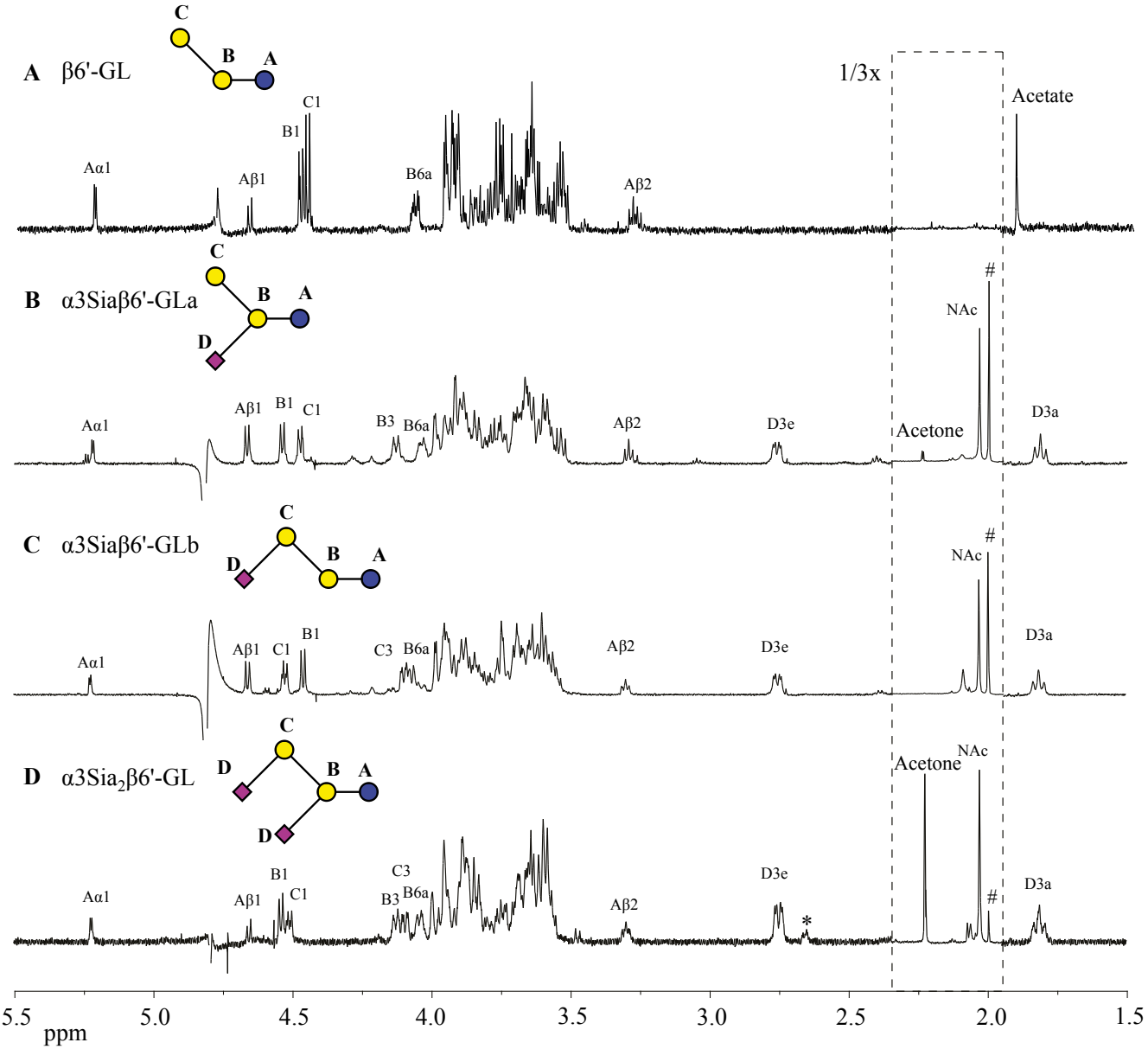

FIG. S4. One-dimensional 600-MHz ^{1}H NMR (D_2O, 298 K) spectra of (A) Gal(β1-6)Gal(β1-4)Glc (β′6-GL), (B) Gal(β1-6)[Neu5Ac(α2-3)]Gal(β1-4)Glc (α3Siaβ6′-GLa), (C) Neu5Ac(α2-3)Gal(β1-6)Gal(β1-4)Glc (α3Siaβ6′-GLb), and (D) Neu5Ac(α2-3)Gal(β1-6)[Neu5Ac(α2-3)]Gal(β1-4)Glc (α3Sia_2β6′-GL). Coding system, see Fig. S2; * and #, contaminants.

TABLE S1. ^{1}H and ^{13}C chemical shifts[a] (D_2O, 298 K) of β6′-GL, α3Siaβ6′-GLa, α3Siaβ6′-GLb, and α3Sia$_2$β6′-GL.

	β6′-GL		α3Siaβ6′-GLa		α3Siaβ6′-GLb		α3Sia$_2$β6′-GL	
	^{1}H	^{13}C	^{1}H	^{13}C	^{1}H	^{13}C	^{1}H	^{13}C
Aα-1	5.223	93.1	5.217	92.6	5.228	92.7	5.222	92.9
Aα-2	3.57	72.1	3.59	72.1	3.60	72.0	3.60	72.0
Aα-3	3.83	72.8	3.84	72.6	3.83	72.6	3.83	72.6
Aα-4	3.63	80.2	3.64	80.0	3.62	80.2	3.65	80.2
Aα-5	3.94	71.1	3.94	71.0	3.96	71.0	3.95	71.2
Aα-6a	3.88	61.1	3.88	60.9	3.88	61.0	3.88	61.1
Aα-6b	3.84		3.83		3.82		3.84	
Aβ-1	4.667	96.9	4.662	96.6	4.660	96.7	4.656	96.7
Aβ-2	3.294	74.9	3.293	74.5	3.306	74.8	3.299	74.8
Aβ-3	3.63	75.8	3.67	76.0	3.64	75.8	3.65	75.9
Aβ-4	3.65	80.2	3.65	80.0	3.63	80.2	3.61	80.2
Aβ-5	3.60	75.9	3.60	75.4	3.60	75.2	3.60	75.6
Aβ-6a	3.94	61.4	3.96	60.9	3.96	61.0	3.95	61.1
Aβ-6b	3.80		3.81		3.81		3.80	
B-1	4.483	104.4	4.535	104.0	4.462	104.3	4.545	103.9
B-2	3.53	72.0	3.58	70.2	3.54	71.8	3.58	70.2
B-3	3.66	73.8	4.125	76.3	3.67	73.9	4.126	76.3
B-4	3.940	69.7	3.986	68.4	3.98	69.4	3.997	68.5
B-5	3.96	75.0	3.91	74.9	3.91	74.8	3.93	74.8
B-6a	4.079	70.3	4.037	70.1	4.073	69.8	4.044	70.1
B-6b	3.93		3.92		3.93		3.90	
C-1	4.460	104.4	4.470	104.4	4.528	104.1	4.514	104.1
C-2	3.54	72.0	3.53	72.8	3.57	70.2	3.56	70.2
C-3	3.67	73.8	3.66	73.8	4.100	76.8	4.093	76.5
C-4	3.974	69.7	3.914	69.4	3.95	68.5	3.942	68.6
C-5	3.68	76.3	3.68	75.9	3.69	76.1	3.68	75.3
C-6a	3.81	62.2	3.80	62.0	3.80	62.0	3.80	62.0
C-6b	3.76		3.75		3.75		3.75	
D-3a			1.810	40.6	1.815	40.6	1.812	40.5
D-3e			2.756		2.755		2.752	
D-4			3.68	69.5	3.68	69.4	3.69	69.5
D-5			3.83	52.8	3.84	52.6	3.83	52.7
D-6			3.66	73.7	3.66	73.8	3.67	73.8
D-7			3.60	69.2	3.60	69.2	3.61	69.2
D-8			3.86	72.6	3.86	72.9	3.85	72.6
D-9a			3.86	63.5	3.86	63.7	3.85	63.5
D-9b			3.64		3.64		3.65	
D-NAc			2.030	23.1	2.031	23.0	2.028	23.1

[a]In ppm relative to the signal of internal acetone ($\delta^{1}H$ 2.225, $\delta^{13}C$ 31.08). β6′-GL = Gal(β1-6)Gal(β1-4)Glc; α3Siaβ6′-GLa = Gal(β1-6)[Neu5Ac(α2-3)]Gal(β1-4)Glc; α3Siaβ6′-GLb = Neu5Ac(α2-3)Gal(β1-6)Gal(β1-4)Glc; α3Sia$_2$β6′-GL = Neu5Ac(α2-3)Gal(β1-6)[Neu5Ac(α2-3)]Gal(β1-4)Glc.