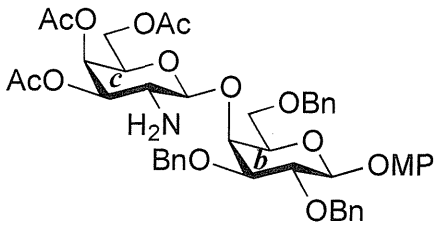
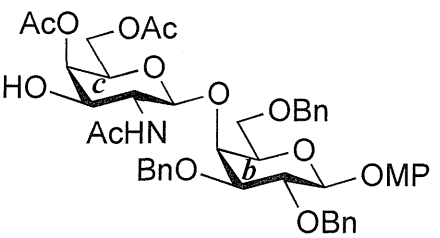
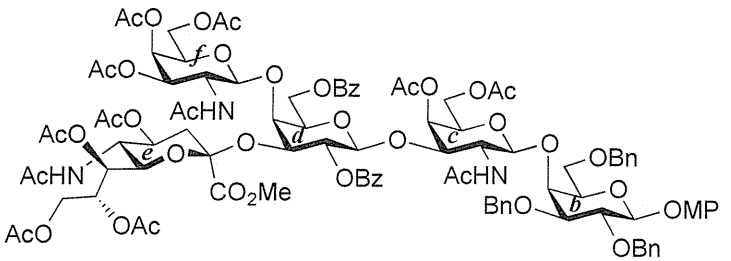
PhCH_2), 4.96 (d, 1H, J_{gem} = 11.0 Hz, PhCH_2), 4.91 (d, 1H, J_{gem} = 12.4 Hz, PhCH_2), 4.84 (d, 2H, $J_{1,2}$ = 7.6 Hz, H-1b, PhCH_2), 4.70 (m, 2H, H-3c, H-1c), 4.64 (d, 1H, PhCH_2), 4.55 (q, 2H, J_{gem} = 11.7 Hz, OCH_2CCl_3), 4.41 (d, 1H, PhCH_2), 4.10 (dd, 1H, $J_{5,6}$ = 7.6 Hz, J_{gem} = 11.0 Hz, H-6c), 4.06 (d, 1H, $J_{3,4}$ = 2.8 Hz, H-4b), 4.04 (dd, 1H, $J_{5,6'}$ = 6.2 Hz, H-6'c), 3.92 (m, 2H, H-2b, H-2c), 3.80 (dd, 1H, $J_{5,6}$ = 5.8 Hz, J_{gem} = 10.0 Hz, H-6b), 3.77 (s, 3H, OCH_3), 3.76–3.71 (m, 2H, H-6'b, H-5c), 3.66 (m, 2H, H-3b, H-5b), 2.14–1.94 (3 s, 9H, 3Ac); ^{13}C-NMR (150 MHz, $CDCl_3$) δ 170.3, 170.2, 155.3, 154.3, 151.4, 138.2, 138.1, 137.2, 129.0, 128.7, 128.5, 128.4, 128.2, 128.0, 127.8, 127.7, 127.6, 125.3, 118.6, 114.5, 102.9, 101.9, 95.8, 81.6, 79.6, 75.7, 75.3, 74.6, 74.3, 73.6, 73.5, 71.7, 70.9, 69.1, 66.5, 61.1, 55.6, 52.7, 20.6, 20.5. HRMS (ESI) *m/z*: found $[M+Na]^+$ 1040.2402, $C_{49}H_{54}Cl_3NO_{16}$ calcd for $[M+Na]^+$ 1040.2400.



4-Methoxyphenyl 3,4,6-tri-O-acetyl-2-amino-2-deoxy-β-D-galactopyranosyl-(1→4)-2,3,6-tri-O-benzyl-β-D-galactopyranoside (**25**). To a solution of **24** (100 mg, 98.3 μmol) in acetonitrile/AcOH (4:1, 3.3 mL) was added Zn (500 mg) at rt. After stirring for 30 min at rt as the reaction was monitored by TLC (1:1 toluene–EtOAc), the solution was diluted with EtOAc and filtered through Celite. The filtrate was then washed with satd aq Na_2CO_3 and brine. The organic layer was subsequently dried over Na_2SO_4, and concentrated. The residue was purified by silica gel column chromatography (1.5:1 toluene–EtOAc) to give **25** (81 mg, 98%). $[\alpha]_D$ −12.3° (c 0.4, $CHCl_3$); ^{1}H-NMR (600 MHz, $CDCl_3$) δ 7.36–7.24 (m, 15H, 3Ph), 7.06–6.78 (m, 4H, Ar), 5.29 (d, 1H, $J_{3,4}$ = 2.1 Hz, H-4c), 5.01 (d, 1H, J_{gem} = 11.0 Hz, PhCH_2), 4.87 (d, 1H, $J_{1,2}$ = 7.6 Hz, H-1b), 4.83 (d, 1H, J_{gem} = 11.7 Hz, PhCH_2), 4.82 (d, 1H, J_{gem} = 11.6 Hz, PhCH_2), 4.69 (d, 1H, PhCH_2), 4.66 (dd, 1H, $J_{2,3}$ = 10.7 Hz, H-3c), 4.55 (s, 2H, PhCH_2), 4.49 (d, 1H, $J_{1,2}$ = 8.2 Hz, H-1c), 4.08 (dd, 1H, $J_{5,6}$ = 7.6 Hz, J_{gem} = 11.0 Hz, H-6c), 4.04 (d, 1H, $J_{3,4}$ = 2.8 Hz, H-4b), 4.02 (dd, 1H, $J_{5,6'}$ = 6.2 Hz, H-6'c), 3.94 (dd, 1H, $J_{1,2}$ = 7.6 Hz, $J_{2,3}$ = 9.6 Hz, H-2b), 3.80 (dd, 1H, $J_{5,6}$ = 4.8 Hz, J_{gem} = 10.3 Hz, H-6b), 3.76 (s, 3H, OCH_3), 3.76–3.73 (m, 2H, H-6'b, H-5c), 3.66 (m, 1H, H-5b), 3.59 (dd, 1H, H-3b), 3.15 (dd, 1H, H-2c), 2.09–2.01 (3 s, 9H, 3Ac); ^{13}C-NMR (150 MHz, $CDCl_3$) δ 170.4, 170.3, 170.2, 155.2, 151.5, 138.2, 138.1, 137.9, 129.0, 128.6, 128.4, 128.3, 128.2, 128.0, 127.8, 127.6, 127.5, 125.3, 118.3, 114.5, 105.1, 102.9, 81.0, 78.8, 75.2, 74.8, 74.1, 73.9, 73.6, 73.6, 70.5, 69.8, 66.3, 61.5, 55.6, 51.9, 21.4, 20.8, 20.7, 20.6. HRMS (ESI) *m/z*: found $[M+Na]^+$ 866.3358, $C_{46}H_{53}NO_{14}$ calcd for $[M+Na]^+$ 866.3358.

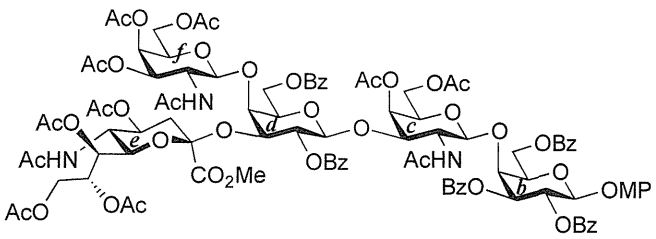


4-Methoxyphenyl 2-acetamido-4,6-di-O-acetyl-2-deoxy-β-D-galactopyranosyl-(1→4)-2,3,6-tri-O-benzyl-β-D-galactopyranoside (**26**). Compound **25** (270 mg, 320 μmol) was dissolved in 1,4-dioxane/AcOH (4:1, 32 mL) and the solution was stirred for 50 h at 60 °C as the reaction was monitored by TLC (1:1.5 toluene–EtOAc). Dilution of the mixture with $CHCl_3$ provided a solution, which was then washed with satd aq Na_2CO_3. The organic layer was subsequently dried over Na_2SO_4, and concentrated. The residue was purified by silica gel column chromatography (2:1 toluene–EtOAc) to give **26** (218 mg, 81%). $[\alpha]_D$ −3.5° (c 0.6, $CHCl_3$); ^{1}H-NMR (600 MHz, $CDCl_3$) δ 7.38–7.24 (m, 15H, 3Ph), 7.17 (d, 1H, $J_{2,NH}$ = 3.4 Hz, NHc), 7.05–6.78 (m, 4H, Ar), 5.91 (d, 1H, $J_{3,OH}$ = 1.4 Hz, OHc), 5.27 (d, 1H, $J_{3,4}$ = 3.4 Hz, H-4c), 5.12 (d, 1H, J_{gem} = 11.0 Hz, PhCH_2), 4.88 (d, 1H, $J_{1,2}$ = 7.6 Hz, H-1b), 4.87 (d, 1H, J_{gem} = 9.6 Hz, PhCH_2), 4.73 (d, 2H, PhCH_2), 4.56 (q, 2H, J_{gem} = 11.7 Hz, PhCH_2), 4.46 (d, 1H, $J_{1,2}$ = 8.2 Hz, H-1c), 4.15 (dd, 1H, $J_{5,6}$ = 6.5 Hz, J_{gem} = 11.3 Hz, H-6c), 4.05 (m, 2H, H-6'c, H-4b), 3.89–3.83 (m, 3H, H-2b, H-2c, H-6b), 3.77 (s, 3H, OCH_3), 3.76–3.69 (m, 4H, H-6'b, H-3b, H-5b, H-5c), 3.53 (br d, 1H, H-3c), 2.14–1.61 (3 s, 9H, 3Ac); ^{13}C-NMR (150 MHz, $CDCl_3$) δ 173.9, 170.5, 170.3, 155.4, 151.2, 138.1, 138.0, 136.4, 129.1, 129.0, 128.8, 128.5, 128.4, 128.2, 128.0, 127.9, 127.7, 127.5, 118.5, 114.5, 102.9, 102.7, 81.6, 79.9, 75.7, 75.3, 74.3, 73.6, 71.3, 69.3, 67.9, 61.9, 55.8, 55.6, 29.7, 22.3, 20.8, 20.7. HRMS (ESI) *m/z*: found $[M+Na]^+$ 866.3358, $C_{46}H_{53}NO_{14}$ calcd for $[M+Na]^+$ 866.3358.



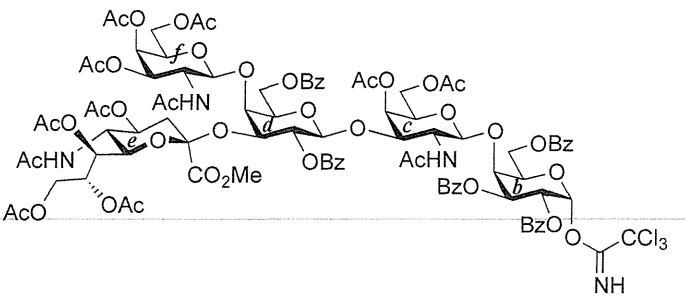
4-Methoxyphenyl 2-acetamido-3,4,6-tri-O-acetyl-2-deoxy-β-D-galactopyranosyl-(1→4)-{[methyl 5-acetamido-4,7,8,9-tetra-O-acetyl-3,5-dideoxy-D-glycero-α-D-galacto-2-nonulopyranosylonate]-(2→3)}-2,6-di-O-benzoyl-β-D-galactopyranosyl-(1→3)-2-acetamido-4,6-di-O-acetyl-2-deoxy-β-D-galactopyranosyl-(1→4)-2,3,6-tri-O-benzyl-β-D-galactopyranoside (**29**). To a mixture of **28** (180 mg, 135 μmol) and **26** (114 mg, 135 μmol) in CH_2Cl_2 (4.5 mL) was added 4 Å molecular sieves (300 mg) at rt. After stirring for 1 h, the mixture was cooled to 0 °C. TMSOTf (2.4 μL, 13.5 μmol) was then added to the mixture at 0 °C. After stirring for 2 h at 0 °C as the reaction was monitored by TLC (1:1 $CHCl_3$–acetone), the reaction was quenched by the addition of satd aq $NaHCO_3$. The solution was diluted with $CHCl_3$ and filtered through Celite. The filtrate was then washed with satd aq $NaHCO_3$ and brine. The organic layer was subsequently dried over Na_2SO_4, and concentrated. The residue was purified by silica gel column chromatography (40:1 $CHCl_3$–MeOH) to give **29** (203 mg, 75%). $[\alpha]_D$ −7.7° (c 0.2, $CHCl_3$); ^{1}H-NMR (600 MHz, DMSO-d_6) δ 8.00–7.19 (m, 25H, 5Ph), 7.57 (d, 1H, $J_{5,NH}$ = 8.9 Hz, NHe), 7.20 (d, 1H, $J_{2,NH}$ = 8.2 Hz, NHc), 6.96–6.78 (m, 4H, Ar), 6.78 (d, 1H, $J_{2,NH}$ = 6.9 Hz, NHf), 5.32 (d, 1H, $J_{3,4}$ = 3.5 Hz, H-4c), 5.29–5.25 (m, 2H, H-8e, H-3f), 5.23 (d, 1H, $J_{3,4}$ = 2.7 Hz, H-4f), 5.14–5.09 (m, 2H, H-2d, H-7e), 4.90 (d, 1H, $J_{1,2}$ = 7.6 Hz, H-1d), 4.85 (d, 2H, $J_{1,2}$ = 7.6 Hz, H-1b, H-1f), 4.78 (m, 1H, H-4e), 4.74 (d, 1H, J_{gem} = 11.7 Hz, PhCH_2), 4.68 (m, 2H, H-1c, PhCH_2), 4.59 (m, 2H, H-3d, PhCH_2), 4.52 (d, 1H, J_{gem} = 12.4 Hz, PhCH_2), 4.47 (m, 2H, H-6d,

PhCH_2), 4.41 (d, 1H, J_{gem} = 12.4 Hz, PhCH_2), 4.28 (m, 1H, H-3c), 4.25 (dd, 1H, $J_{5,6'}$ = 5.5 Hz, J_{gem} = 11.0 Hz, H-6'd), 4.08–4.00 (m, 3H, H-6c, H-6f, H-9e), 3.98–3.91 (m, 4H, H-6'c, H-6'f, H-9'e, H-5d), 3.86–3.72 (m, 8H, H-4d, H-5c, H-5f, H-2c, H-2f, H-6e, H-6b, H-5e), 3.75–3.68 (2 s, 6H, 2OCH_3), 3.66 (near t, 1H, H-2b), 3.59–3.53 (m, 4H, H-3b, H-4b, H-5b, H-6'b), 2.30 (dd, 1H, $J_{3eq,4}$ = 4.8 Hz, J_{gem} = 13.1 Hz, H-3e*eq*), 1.80 (near t, 1H, H-3e*ax*), 2.08–1.64 (12 s, 36H, 12 Ac); ^{13}C-NMR (150 MHz, $CDCl_3$) δ 171.3, 170.7, 170.6, 170.5, 170.4, 170.4, 170.2, 169.9, 169.8, 168.1, 166.0, 164.3, 155.1, 151.5, 138.4, 137.9, 133.2, 133.1, 130.0, 129.5, 128.6, 128.5, 128.4, 128.3, 128.0, 127.7, 127.5, 118.3, 114.4, 102.8, 100.7, 100.5, 100.0, 98.3, 80.7, 79.2, 75.3, 74.9, 74.1, 74.0, 73.8, 73.5, 73.2, 72.1, 72.0, 70.8, 70.3, 70.1, 70.0, 69.3, 68.7, 67.4, 66.9, 66.5, 63.7, 62.7, 62.1, 61.4, 55.6, 54.2, 53.1, 51.8, 49.2, 36.1, 31.9, 29.7, 29.3, 23.4, 23.1, 22.7, 21.1, 20.8, 20.7, 20.7, 20.7, 20.5, 20.4, 14.1. HRMS (ESI) *m/z*: found $[M+Na]^+$ 2038.7055, $C_{100}H_{117}N_3O_{41}$ calcd for $[M+Na]^+$ 2038.7055.

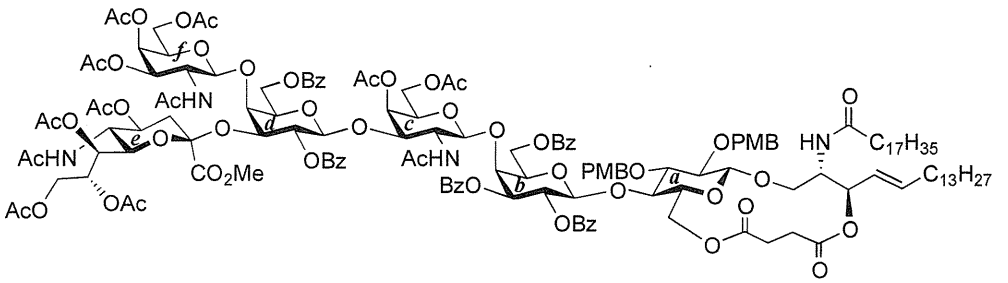


4-Methoxyphenyl 2-acetamido-3,4,6-tri-O-acetyl-2-deoxy-β-D-galactopyranosyl-(1→4)-{[methyl 5-acetamido-4,7,8,9-tetra-O-acetyl-3,5-dideoxy-D-glycero-α-D-galacto-2-nonulopyranosylonate]-(2→3)}-2,6-di-O-benzoyl-β-D-galactopyranosyl-(1→3)-2-acetamido-4,6-di-O-acetyl-2-deoxy-β-D-galactopyranosyl-(1→4)-2,3,6-tri-O-benzoyl-β-D-galactopyranoside (**30**). To a solution of **29** (215 mg, 107 μmol) in EtOH (10.7 mL) was added $Pd(OH)_2$/C (20%, 215 mg). After stirring for 16 h at rt under a hydrogen atmosphere as the reaction was monitored by TLC (10:1 $CHCl_3$–MeOH), the mixture was filtered through Celite. The filtrate was concentrated and the crude residue obtained was exposed to high vacuum for 3 h. The resulting residue was then dissolved in pyridine (535 μL). Benzoic anhydride (145 mg, 642 μmol) and DMAP (7.8 mg, 64.2 μmol) were added to the mixture at 0 °C. After stirring for 40 min at rt as the reaction was monitored by TLC (15:1 $CHCl_3$–MeOH), the reaction was quenched by the addition of MeOH at 0 °C. The mixture was co-evaporated with toluene and the residue was then diluted with $CHCl_3$, and washed with 2 M HCl, H_2O and satd aq $NaHCO_3$. The organic layer was subsequently dried over Na_2SO_4, and concentrated. The resulting residue was purified by silica gel column chromatography (25:1 $CHCl_3$–MeOH) to give **30** (180 mg, 82%). $[\alpha]_D$ +6.3° (c 0.4, $CHCl_3$); ^{1}H-NMR (600 MHz, DMSO-d_6) δ 8.01–7.35 (m, 26H, 5Ph, NHe), 7.11 (br d, 1H, $J_{2,NH}$ = 5.5 Hz, NHc), 7.01 (d, 1H, $J_{2,NH}$ = 8.2 Hz, NHf), 6.85–6.62 (m, 4H, Ar), 5.60 (t, 1H, $J_{1,2}$ = $J_{2,3}$ = 7.6 Hz, H-2b), 5.52 (dd, 1H, $J_{3,4}$ = 2.8 Hz, H-3b), 5.39 (d, 1H, H-1b), 5.30 (d, 1H, $J_{3,4}$ = 2.7 Hz, H-4c), 5.25 (m, 3H, H-3f, H-8e, H-4f), 5.12 (m, 2H, H-2d, H-7e), 4.87 (m, 2H, H-1d, H-1f), 4.81 (m, 1H, H-4e), 4.74 (d, 1H, $J_{1,2}$ = 8.3 Hz, H-1c), 4.59 (d, 1H, $J_{2,3}$ = 10.3 Hz, H-3d), 4.49–4.41 (m, 5H, H-6b, H-4b, H-6'b, H-6d, H-5b), 4.29 (br dd, 1H, $J_{2,3}$ = 10.3 Hz, H-3c), 4.24 (dd, 1H, $J_{5,6'}$ = 5.5 Hz, J_{gem} = 11.0 Hz, H-6'd), 4.06 (m, 2H, H-6f, H-9e), 3.98–3.93 (m, 3H, H-6'f, H-9'e, H-6c), 3.90–3.72 (m, 7H, H-4d, H-5d, H-5f, H-2f, H-6e, H-5e, H-6'c), 3.75 (s, 3H, OCH_3), 3.67 (m, 1H, H-5c), 3.62 (s, 3H, OCH_3), 3.60 (m, 1H, H-2c), 2.26 (near dd, 1H, H-3e*eq*), 2.09–1.75 (m, 37H, H-3e*ax*, 12Ac); ^{13}C-NMR (150 MHz, $CDCl_3$) δ 171.1, 170.6, 170.5, 170.5, 170.5, 170.3, 170.2, 169.9,

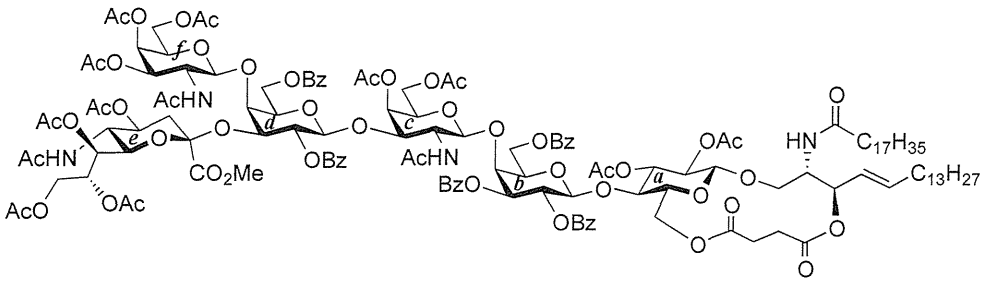
169.7, 168.1, 166.3, 166.1, 166.0, 165.3, 164.3, 155.5, 151.2, 133.6, 133.4, 133.1, 132.9, 130.1, 130.0, 130.0, 129.8, 129.7, 129.6, 129.5, 129.4, 128.6, 128.5, 128.4, 128.3, 118.7, 114.3, 100.9, 100.8, 100.4, 98.2, 97.8, 74.7, 74.0, 73.4, 72.7, 72.0, 71.9, 71.2, 70.6, 70.1, 70.0, 69.7, 69.4, 68.8, 67.3, 67.0, 66.5, 63.7, 63.5, 62.8, 62.3, 61.4, 55.5, 53.1, 52.1, 52.0, 49.3, 36.3, 29.7, 23.4, 23.2, 21.1, 20.9, 20.8, 20.8, 20.7, 20.5, 20.4. HRMS (ESI) *m/z*: found $[M+Na]^+$ 2080.6434, $C_{100}H_{111}N_3O_{44}$ calcd for $[M+Na]^+$ 2080.6433.



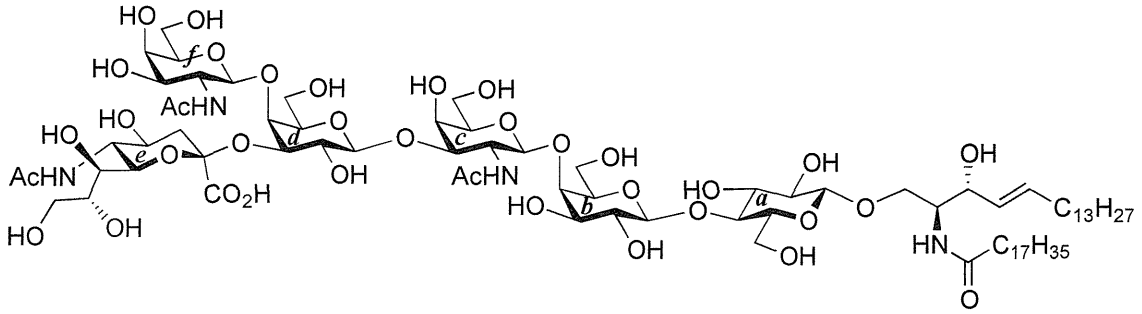
2-Acetamido-3,4,6-tri-O-acetyl-2-deoxy-β-D-galactopyranosyl-(1→4)-{[methyl 5-acetamido-4,7,8,9-tetra-O-acetyl-3,5-dideoxy-D-glycero-α-D-galacto-2-nonulopyranosylonate]-(2→3)}-2,6-di-O-benzoyl-β-D-galactopyranosyl-(1→3)-2-acetamido-4,6-di-O-acetyl-2-deoxy-β-D-galactopyranosyl-(1→4)-2,3,6-tri-O-benzoyl-α-D-galactopyranosyl trichloroacetimidate (**31**). To a solution of **30** (115 mg, 55.9 μmol) in acetonitrile/toluene/H_2O (6:5:3, 1.1 mL) was added CAN (245 mg, 447 μmol) at 0 °C. After stirring for 2 h at 0 °C as the reaction was monitored by TLC (10:1 $CHCl_3$–MeOH), the mixture was diluted with $CHCl_3$. The solution was then washed with H_2O, satd aq $NaHCO_3$ and brine. The organic layer was subsequently dried over Na_2SO_4, and concentrated. The residue was roughly purified by silica gel column chromatography (30:1 $CHCl_3$–MeOH). The product obtained was exposed to high vacuum for 20 h and then dissolved in CH_2Cl_2 (502 μL). CCl_3CN (101 μL, 1.00 mmol) and DBU (9.0 μL, 60.2 μmol) were added to the mixture at 0 °C. After stirring for 30 min at rt as the reaction was monitored by TLC (15:1 $CHCl_3$–MeOH), the reaction mixture was evaporated. The crude residue obtained was purified by silica gel column chromatography (40:1 $CHCl_3$–MeOH) to give **31** (94 mg, 81%). $[\alpha]_D$ +14.0° (c 0.6, $CHCl_3$); ^{1}H-NMR (600 MHz, DMSO-d_6) δ 9.65 (br s, 1H, C(=NH)), 8.01–7.37 (m, 26H, 5Ph, NHe), 7.21 (br s, 1H, NHc), 6.98 (d, 1H, $J_{2,NH}$ = 8.2 Hz, NHf), 6.57 (br s, 1H, H-1b), 5.77 (br d, 1H, H-2b), 5.64 (near dd, 1H, H-3b), 5.31 (s, 1H, H-4c), 5.23 (m, 3H, H-3f, H-8e, H-4f), 5.10 (m, 2H, H-2d, H-7e), 4.89 (d, 1H, $J_{1,2}$ = 7.6 Hz, H-1d), 4.86–4.81 (m, 3H, H-1f, H-1c, H-4e), 4.62 (m, 2H, H-4b, H-5b), 4.56 (br d, 1H, H-3d), 4.48 (m, 2H, H-6b, H-6d), 4.38 (near dd, 1H, H-6'b), 4.31 (br d, 1H, H-3c), 4.25 (near dd, 1H, H-6'd), 4.06 (m, 2H, H-6f, H-9e), 3.98–3.84 (m, 8H, H-6'f, H-9'e, H-6c, H-4d, H-5d, H-5f, H-2f, H-6'c), 3.81–3.69 (m, 6H, H-6e, OCH_3, H-5e, H-5c), 3.62 (m, 1H, H-2c), 2.26 (near dd, 1H, H-3e*eq*), 2.09–1.66 (m, 37H, H-3e*ax*, 12Ac); ^{13}C-NMR (150 MHz, $CDCl_3$) δ 171.0, 170.8, 170.4, 170.4, 170.4, 170.3, 170.2, 169.8, 169.7, 166.1, 165.9, 165.3, 164.2, 160.4, 133.5, 133.3, 133.1, 133.0, 129.9, 129.8, 129.7, 129.6, 129.5, 128.7, 128.4, 128.4, 128.3, 128.2, 100.7, 100.5, 98.2, 98.0, 93.4, 90.8, 74.8, 73.5, 73.1, 71.9, 71.9, 71.4, 71.3, 70.9, 70.4, 70.1, 70.0, 69.5, 68.7, 67.2, 66.9, 66.4, 63.8, 63.5, 62.8, 62.2, 61.4, 55.3, 53.0, 51.2, 49.1, 36.1, 29.6, 23.3, 23.1, 22.9, 22.6, 21.0, 20.8, 20.8, 20.7, 20.6, 20.4, 20.3, 14.1. HRMS (ESI) *m/z*: found $[M+Na]^+$ 2117.5111, $C_{95}H_{105}Cl_3N_4O_{43}$ calcd for $[M+Na]^+$ 2117.5110.



4-O-{2-Acetamido-3,4,6-tri-O-acetyl-2-deoxy-β-D-galactopyranosyl-(1→4)-[(methyl 5-acetamido-4,7,8,9-tetra-O-acetyl-3,5-dideoxy-D-glycero-α-D-galacto-2-nonulopyranosylonate)-(2→3)]-2,6-di-O-benzoyl-β-D-galactopyranosyl-(1→3)-2-acetamido-4,6-di-O-acetyl-2-deoxy-β-D-galactopyranosyl-(1→4)-2,3,6-tri-O-benzoyl-β-D-galactopyranosyl-(1→4)}-2,3-di-O-p-methoxybenzyl-β-D-glucopyranosyl-(1'→1)-(2S,3R,4E)-2-octadecanamido-4-octadecene-1,3-diol-3,6'-succinate (**32**). To a mixture of **31** (59 mg, 28.2 μmol) and **18β** (30 mg, 28.2 μmol) in $CHCl_3$ (940 μL) was added 4 Å molecular sieves (100 mg) at rt. After stirring for 1 h, the mixture was cooled to 0 °C. TMSOTf (0.5 μL, 2.82 μmol) was then added to the mixture at 0 °C. After stirring for 1.5 h at rt as the reaction was monitored by TLC (1.5:1 acetone–*n*-hexane), the reaction was quenched by the addition of satd aq $NaHCO_3$. The solution was diluted with $CHCl_3$ and filtered through Celite. The filtrate was then washed with satd aq $NaHCO_3$ and brine. The organic layer was subsequently dried over Na_2SO_4, and concentrated. The residue was purified by silica gel column chromatography (1:1 acetone–*n*-hexane) to give **32** (22 mg, 26%). $[\alpha]_D$ +7.8° (c 0.4, $CHCl_3$); ^{1}H-NMR (500 MHz, $CDCl_3$) δ 8.08–6.80 (m, 33H, 7Ar), 6.77 (d, 1H, $J_{2,NH}$ = 7.9 Hz, NHc), 5.73 (m, 2H, $J_{2,NH}$ = 8.9 Hz, H-5^{Cer}, NH^{Cer}), 5.58–5.53 (m, 2H, H-3b, H-2d), 5.47–5.35 (m, 3H, NHf, H-3^{Cer}, H-8e), 5.33 (d, 1H, $J_{3,4}$ = 3.1 Hz, H-4c), 5.31–5.20 (m, 4H, H-4f, H-2b, H-4^{Cer}, H-7e), 5.14 (d, 1H, $J_{5,NH}$ = 9.9 Hz, NHe), 5.09 (d, 1H, $J_{1,2}$ = 8.1 Hz, H-1f), 5.05 (d, 1H, $J_{1,2}$ = 8.4 Hz, H-1c), 4.99 (m, 1H, H-4e), 4.96 (d, 1H, $J_{1,2}$ = 7.8 Hz, H-1d), 4.84–4.73 (m, 3H, 2 ArCH_2, H-3f), 4.71 (d, 1H, $J_{1,2}$ = 7.7 Hz, H-1b), 4.66 (d, 1H, J_{gem} = 10.8 Hz, ArCH_2), 4.61 (m, 2H, H-6b, H-6d), 4.52 (d, 1H, ArCH_2), 4.47 (d, 1H, $J_{3,4}$ = 1.6 Hz, H-4d), 4.33–4.19 (m, 5H, H-6'b, H-2^{Cer}, H-1a, H-6f, H-6c), 4.14–4.03 (m, 5H, H-6'd, H-6'f, H-3c, H-9e, H-9'e), 4.00–3.87 (m, 6H, H-4b, H-6'c, H-6a, H-2c, H-3d, H-5e), 3.85–3.72 (m, 8H, H-6'a, OCH_3, H-5b, H-6e, H-5d, H-1^{Cer}), 3.66–3.58 (m, 3H, H-1'^{Cer}, H-3a, H-4a), 3.54 (m, 1H, H-5a), 3.48 (m, 2H, H-5c, H-5f), 3.33 (t, 1H, $J_{1,2}$ = $J_{2,3}$ = 7.2 Hz, H-2a), 3.16 (m, 1H, H-2f), 2.58–2.45 (m, 4H, 2C(=O)CH_2), 2.29 (dd, 1H, $J_{3eq,4}$ = 4.8 Hz, J_{gem} = 13.4 Hz, H-3e*eq*), 2.18–1.55 (m, 43H, C(=O)CH_2CH_2^{Cer}, H-3e*ax*, H-6^{Cer}, H-6'^{Cer}, 12Ac), 1.26 (m, 50H, 25-CH_2-), 0.88 (m, 6H, 2-CH_3^{Cer}); ^{13}C-NMR (125 MHz, $CDCl_3$) δ 172.8, 171.8, 171.2, 171.0, 170.6, 170.5, 170.4, 170.4, 170.2, 169.9, 169.8, 168.1, 166.2, 166.0, 165.6, 164.3, 159.3, 159.0, 137.1, 133.6, 133.3, 133.1, 130.5, 130.0, 130.0, 129.8, 129.8, 129.7, 129.6, 129.6, 129.6, 129.0, 128.8, 128.5, 128.4, 128.4, 128.3, 128.2, 125.6, 124.3, 120.2, 113.8, 113.7, 101.8, 100.9, 100.6, 98.3, 97.8, 81.8, 80.7, 78.5, 77.6, 75.0, 73.9, 73.7, 73.5, 73.3, 72.7, 72.2, 72.0, 71.2, 70.6, 70.4, 70.4, 70.1, 70.0, 69.5, 68.7, 67.3, 66.9, 66.5, 63.5, 62.8, 62.7, 62.2, 61.4, 55.3, 55.2, 55.1, 53.8, 53.1, 51.9, 49.2, 43.0, 38.7, 36.6, 36.2, 32.3, 31.9, 31.7, 30.3, 29.7, 29.7, 29.7, 29.6, 29.6, 29.5, 29.4, 29.3, 29.3, 29.1, 29.0, 28.9, 28.8, 25.9, 25.6, 23.9, 23.8, 23.4, 23.1, 23.0, 22.9, 22.8, 22.7, 22.6, 21.1, 20.8, 20.8, 20.8, 20.6, 20.4, 20.4, 14.1, 14.1, 14.0, 11.0, 10.9. HRMS (ESI) *m/z*: found [1/2M+Na]$^+$ 1514.6482, $C_{155}H_{202}N_4O_{54}$ calcd for [1/2M+Na]$^+$ 1514.6484.



4-O-{2-Acetamido-3,4,6-tri-O-acetyl-2-deoxy-β-D-galactopyranosyl-(1→4)-[(methyl 5-acetamido-4,7,8,9-tetra-O-acetyl-3,5-dideoxy-D-glycero-α-D-galacto-2-nonulopyranosylonate)-(2→3)]-2,6-di-O-benzoyl-β-D-galactopyranosyl-(1→3)-2-acetamido-4,6-di-O-acetyl-2-deoxy-β-D-galactopyranosyl-(1→4)-2,3,6-tri-O-benzoyl-β-D-galactopyranosyl-(1→4)}-2,3-di-O-acetyl-β-D-glucopyranosyl-(1'→1)-(2S,3R,4E)-2-octadecanamido-4-octadecene-1,3-diol-3,6'-succinate (**33**). To a mixture of **31** (53 mg, 25.3 μmol) and **1** (23 mg, 25.3 μmol) in $CHCl_3$ (843 μL) was added 4 Å molecular sieves (120 mg) at rt. After stirring for 1 h, the mixture was cooled to 0 °C. TMSOTf (0.5 μL, 2.53 μmol) was then added to the mixture at 0 °C. After stirring for 2.5 h at rt as the reaction was monitored by TLC (4:3 acetone–*n*-hexane), the reaction was quenched by the addition of satd aq $NaHCO_3$. The solution was diluted with $CHCl_3$ and filtered through Celite. The filtrate was then washed with satd aq $NaHCO_3$ and brine. The organic layer was subsequently dried over Na_2SO_4, and concentrated. The residue was purified by silica gel column chromatography (1:1 acetone–*n*-hexane) to give **33** (22 mg, 31%). The yields of **33** based on the use of 2.0 eq. and 3.0 eq. of **1** were 48% and 60%, respectively. $[\alpha]_D$ +8.0° (c 0.5, $CHCl_3$); ^{1}H-NMR (600 MHz, $CDCl_3$) δ 8.11–7.31 (m, 25H, 5Ph), 6.31 (d, 1H, $J_{2,NH}$ = 8.2 Hz, NHc), 5.98 (d, 1H, $J_{2,NH}$ = 6.2 Hz, NHf), 5.77 (m, 1H, H-5^{Cer}), 5.59 (d, 1H, $J_{2,NH}$ = 9.6 Hz, NH^{Cer}), 5.52 (m, 2H, H-3b, H-2d), 5.38 (m, 1H, H-8e), 5.34–5.29 (m, 4H, H-4c, H-4f, H-2b, H-3^{Cer}), 5.26–5.21 (m, 2H, H-4^{Cer}, H-7e), 5.18–5.14 (m, 2H, H-3a, H-1f), 5.10 (d, 1H, $J_{5,NH}$ = 9.7 Hz, NHe), 5.04 (d, 1H, $J_{1,2}$ = 8.3 Hz, H-1c), 5.00 (m, 1H, H-4e), 4.88 (dd, 1H, $J_{3,4}$ = 3.4 Hz, $J_{2,3}$ = 11.0 Hz, H-3f), 4.83 (t, 1H, $J_{1,2}$ = $J_{2,3}$ = 7.2 Hz, H-2a), 4.75 (d, 1H, $J_{1,2}$ = 7.6 Hz, H-1b), 4.69 (d, 1H, $J_{1,2}$ = 7.6 Hz, H-1d), 4.62 (m, 2H, H-6b, H-6d), 4.52 (d, 1H, $J_{3,4}$ = 2.8 Hz, H-4d), 4.34 (m, 3H, H-1a, H-6'b, H-6'd), 4.25 (m, 2H, H-2^{Cer}, H-6f), 4.19 (near dd, 1H, H-6c), 4.20–4.05 (m, 4H, H-6′f, H-3c, H-6′c, H-9e), 4.02–3.88 (m, 6H, H-9'e, H-4b, H-6a, H-2c, H-3d, H-5e), 3.85–3.75 (m, 8H, H-1^{Cer}, OCH_3, H-1'^{Cer}, H-5b, H-5d, H-6e), 3.70 (m, 2H, H-4a, H-6'a), 3.60 (m, 2H, H-5c, H-5f), 3.50 (near t, 1H, H-5a), 3.12 (m, 1H, H-2f), 2.59–2.40 (m, 4H, 2C(=O)CH_2), 2.29 (dd, 1H, J_{gem} = 13.0 Hz, $J_{3eq,4}$ = 4.8 Hz, H-3e*eq*), 2.19–1.55 (m, 49H, C(=O)$CH_2CH_2^{Cer}$, H-3e*ax*, H-6^{Cer}, H-6'^{Cer}, 14Ac), 1.25 (m, 50H, 25-CH_2-), 0.88 (m, 6H, 2-CH_3^{Cer}); ^{13}C-NMR (150 MHz, $CDCl_3$) δ 172.7, 171.3, 171.2, 171.1, 170.5, 170.4, 170.3, 170.3, 169.9, 169.7, 169.3, 168.1, 166.1, 166.0, 165.9, 164.9, 164.3, 133.6, 133.3, 133.2, 130.4, 130.1, 130.0, 129.8, 129.7, 129.5, 129.4, 129.0, 128.8, 128.7, 128.6, 128.5, 128.4, 128.2, 124.7, 101.0, 100.7, 100.4, 99.2, 98.3, 97.5, 76.5, 75.2, 74.0, 73.6, 73.3, 72.4, 72.1, 72.0, 72.0, 71.7, 71.3, 70.5, 70.4, 70.0, 69.6, 68.7, 67.3, 67.0, 66.5, 63.5, 63.0, 62.8, 62.2, 61.4, 53.1, 51.8, 50.0, 49.3, 36.8, 36.1, 32.3, 31.9, 29.7, 29.7, 29.5, 29.5, 29.4, 29.3, 29.3, 28.8, 25.6, 23.4, 23.2, 23.2, 23.1, 22.9, 22.8, 22.7, 22.6, 22.2, 22.0, 21.8, 21.1, 20.9, 20.8, 20.7, 20.6, 20.5, 20.4, 14.1. HRMS (ESI) *m/z*: found $[1/2M+Na]^+$ 1436.6014, $C_{143}H_{190}N_4O_{54}$ calcd for $[1/2M+Na]^+$ 1436.6015.



GalNAc-GM1b: 2-acetamido-2-deoxy-β-D-galactopyranosyl-(1→4)-[5-acetamido-3,5-dideoxy-D-glycero-α-D-galacto-2-nonulopyranosylonic acid-(2→3)]-β-D-galactopyranosyl-(1→3)-2-acetamido-2-deoxy-β-D-galactopyranosyl-(1→4)-β-D-galactopyranosyl-(1→4)-β-D-glucopyranosyl-(1'→1)-(2S,3R,4E)-2-octadecanamido-4-octadecene-1,3-diol. To a solution of **33** (15.0 mg, 5.31 μmol) in MeOH/THF (1:1, 532 μL) was added NaOMe (28% solution in MeOH, 102 μg, 0.531 μmol) at 0 °C. After stirring for 6 d at rt as the reaction was monitored by TLC (5:4:1 $CHCl_3$–MeOH–10 mM aq $ZnCl_2$), water (10 μL) was added to the mixture. After stirring for 8 d at rt, the reaction was neutralized with Dowex (H^+) resin. The resin was filtered through cotton and the filtrate was then evaporated. The residue was purified by gel filtration column chromatography (LH-20) using $CHCl_3$–MeOH as eluent followed by silica gel column chromatography (5:4:0.5 $CHCl_3$–MeOH–H_2O) to give the target **GalNAc-GM1b** (8.2 mg, 88%). $[\alpha]_D$ +12.5° (c 0.2, 1:1 $CHCl_3$–MeOH); ^{1}H-NMR (600 MHz, 1:1 $CDCl_3$–CD_3OD) δ 5.70 (m, 1H, H-5^{Cer}), 5.45 (dd, 1H, $J_{3,4}$ = 7.6 Hz, $J_{4,5}$ = 15.1 Hz, H-4^{Cer}), 2.73 (br d, 1H, H-3e*eq*), 2.18 (m, 2H, C(=O)CH_2), 2.05–2.01 (m, 11H, 3Ac, H-6^{Cer}, H-6'^{Cer}), 1.85 (br t, 1H, H-3e*ax*), 1.59 (m, 2H, C(=O)$CH_2C\mathit{H}_2$), 1.37–1.19 (m, 50H, 25-CH_2-), 0.89 (m, 6H, 2-CH_3); ^{13}C-NMR (150 MHz, 1:1 $CDCl_3$–CD_3OD) δ 174.8, 174.6, 173.7, 173.4, 134.4, 129.7, 129.5, 128.0, 104.4, 103.8, 103.1, 102.0, 79.0, 76.2, 75.2, 75.0, 74.7, 74.5, 73.8, 73.6, 73.5, 72.1, 71.9, 71.3, 69.6, 69.5, 68.7, 68.6, 68.2, 64.6, 62.0, 61.5, 60.4, 60.2, 53.3, 53.1, 52.6, 51.8, 47.7, 36.4, 32.4, 32.0, 29.7, 29.6, 29.6, 29.6, 29.4, 29.3, 26.1, 22.7, 22.7, 22.0, 13.8. HRMS (ESI) *m/z*: found $[M-H]^-$ 1747.9487, $C_{81}H_{144}N_4O_{36}$ calcd for $[M-H]^-$ 1747.9488.

4. Conclusions

In this study, we investigated the development of a GlcCer cassette acceptor that was both readily accessible and highly reactive. We designed and prepared a novel cassette acceptor bearing electron-donating PMB groups at C2 and C3 of the glucose residue. Various types of linkers and their effect on the stereoselectivity of intramolecular glycosylation were examined. Although varying the linker did not significantly increase β-selectivity, the use of a nitrile solvent gave predominantly the desired β-product. Considering the accessibility of the acceptor, we opted for the succinyl linker. In the experiment on coupling the cassette acceptor and oligosaccharide donor, we found that the use of PMB groups as protecting groups at C2 and C3 positions of the glucose residue did not enhance the reactivity as a GlcCer cassette acceptor. This interesting finding should provide useful information for the future design of glycosyl acceptors. Furthermore, we extended the generality of the GlcCer cassette approach by applying it to the efficient total synthesis of the ganglioside GalNAc-GM1b. Our

laboratory is now conducting further studies to evaluate the scope and limitations of the GlcCer cassette approach.

Acknowledgments

The iCeMS is supported by World Premier International Research Center Initiative (WPI), MEXT, Japan. This work was financially supported in part by MEXT of Japan (a Grant-in-Aid for Scientific Research (B) No. 22380067 to M.K. and a Grant-in-Aid for Young Scientists (A) No. 23688014 and Grant-in-Aid on Innovative Areas No. 24110505, Deciphering sugar chain-based signals regulating integrative neuronal functions, to H.A.). We thank Kiyoko Ito (Gifu University) for providing technical assistance.

Conflicts of Interest

The authors declare no conflict of interest.

References and Notes

1. Sonnino, S.; Mauri, L.; Chigorno, V.; Prinetti, A. Gangliosides as components of lipid membrane domains. *Glycobiology* **2006**, *17*, 1R–13R.
2. Todeschini, A.R.; Hakomori, S.I. Functional role of glycosphingolipids and gangliosides in control of cell adhesion, Motility, and growth, through glycosynaptic microdomains. *Biochim. Biophys. Acta* **2008**, *1780*, 421–433.
3. Lopez, P.H.; Schnaar, R.L. Gangliosides in cell recognition and membrane protein regulation. *Curr. Opin. Struct. Biol.* **2009**, *19*, 549–557.
4. Posse de Chaves, E.; Sipione, S. Sphingolipids and gangliosides of the nervous system in membrane function and dysfunction. *FEBS Lett.* **2010**, *584*, 1748–1759.
5. Kusunoki, S.; Kaida, K. Antibodies against ganglioside complexes in Guillain–Barré syndrome and related disorders. *J. Neurochem.* **2011**, *116*, 828–832.
6. Koga, M.; Yuki, N.; Hirata, K. Subclass distribution and the secretory component of serum IgA anti-ganglioside antibodies in Guillain–Barré syndrome after *Campylobacter jejuni* enteritis. *J. Neuroimmunol.* **1999**, *96*, 245–250.
7. Ando, H.; Imamura, A. Proceedings in synthetic chemistry of sialo-glycoside. *Trends Glycosci. Glycotechnol.* **2004**, *16*, 293–303.
8. Hanashima, S. Recent strategies for stereoselective sialylation and their application to the synthesis of oligosaccharides. *Trends Glycosci. Glycotechnol.* **2011**, *23*, 111–121.
9. Synthetic Studies on Sialoglycoconjugates, Part 159. For Part 158; Tamai, H.; Ando, H.; Ishida, H.; Kiso, M. First synthesis of a pentasaccharide moiety of ganglioside GAA-7 containing unusually modified sialic acids through the use of *N*-Troc-sialic acid derivative as a key unit. *Org. Lett.* **2012**, *14*, 6342–6345.
10. Murase, T.; Kameyama, A.; Kartha, K.P.R.; Ishida, H.; Kiso, M.; Hasegawa, A. A facile, Regio and stereoselective synthesis of ganglioside GM4 and its position isomer. *J. Carbohydr. Chem.* **1989**, *8*, 265–283.

11. Murase, T.; Ishida, H.; Kiso, M.; Hasegawa, A. A facile, Regio- and stereo-selective synthesis of ganglioside GM3. *Carbohydr. Res.* **1989**, *188*, 71–80.
12. Imamura, A.; Ando, H.; Ishida, H.; Kiso, M. Ganglioside GQ1b: Efficient total synthesis and the expansion to synthetic derivatives to elucidate its biological roles. *J. Org. Chem.* **2009**, *74*, 3009–3023.
13. Fujikawa, K.; Nohara, T.; Imamura, A.; Ando, H.; Ishida, H.; Kiso, M. A cyclic glucosyl ceramide acceptor as a versatile building block for complex ganglioside synthesis. *Tetrahedron Lett.* **2010**, *51*, 1126–1130.
14. Fujikawa, K.; Nakashima, S.; Konishi, M.; Fuse, T.; Komura, N.; Ando, T.; Ando, H.; Yuki, N.; Ishida, H.; Kiso, M. The first total synthesis of ganglioside GalNAc-GD1a, a target molecule for autoantibodies in Guillain–Barré syndrome. *Chem. Eur. J.* **2011**, *17*, 5641–5651.
15. Nakashima, S.; Ando, H.; Saito, R.; Tamai, H.; Ishida, H.; Kiso, M. Efficiently synthesizing lacto-ganglio-series gangliosides by using a glucosyl ceramide cassette approach: The total synthesis of ganglioisde X2. *Chem. Asian J.* **2012**, *7*, 1041–1051.
16. Tamai, H.; Ando, H.; Tanaka, H.-N.; Hosoda-Yabe, R.; Yabe, T.; Ishida, H.; Kiso, M. The total synthesis of the neurogenic ganglioside LLG-3 isolated from the starfish *Linckia laevigata*. *Angew. Chem. Int. Ed.* **2011**, *50*, 2330–2333.
17. Fügedi, F.; Garegg, P.J. A novel promoter for the efficient construction of 1,2-*trans* linkages in glycoside synthesis, using thioglycosides as glycosyl donors. *Carbohydr. Res.* **1986**, *149*, C9–C12.
18. Kanie, O.; Kiso, M.; Hasegawa, A. Glycosylation using methylthioglycosides of *N*-acetylneuraminic acid and dimethyl(methylthio)sulfonium triflate. *J. Carbohydr. Chem.* **1988**, *7*, 501–506.
19. Itoh, T.; Li, Y.-T.; Li, S.-C.; Yu, R.K. Isolation and characterization of a novel monosialosylpentahexosyl ceramide from Tay-Sachs brain. *J. Biol. Chem.* **1981**, *256*, 165–169.
20. Müthing, J.; Schwinzer, B.; Peter-Katalinic, J.; Egge, H.; Mühlradt, P.F. Gangliosides of murine T lymphocyte subpopulations. *Biochemistry* **1989**, *28*, 2923–2929.
21. Yuki, N.; Taki, T.; Handa, S. Antibody to GalNAc-GD1a and GalNAc-GM1b in Guillain–Barré syndrome subsequent to Campylobacter jejuni enteritis. *J. Neuroimmunol.* **1996**, *71*, 155–161.
22. Ilyas, A.A.; Li, S.-C.; Chou, D.K.H.; Li, Y.-T.; Jungalwala, F.B.; Dalakas, M.C.; Quarles, R.H. Gangliosides GM2, IV^{4}GalNAcGM1b, and IV^{4}GalNAcGD1a as antigens for monoclonal immunoglobulin M in neuropathy associated with gammopathy. *J. Biol. Chem.* **1988**, *263*, 4369–4373.
23. Ortiz, N.; Rosa, R.; Gallardo, E.; Illa, I.; Tomas, J.; Aubry, J.; Santafé, M. IgM monoclonal antibody against terminal moiety of GM2, GalNAc-GD1a and GalNAc-GM1b from a pure motor chronic demyelinating polyneuropathy patient: Effects on neurotransmitter release. *J. Neuroimmunol.* **2001**, *119*, 114–123.
24. Sugimoto, M.; Fujikura, K.; Nunomura, S.; Ito, Y.; Ogawa, T. Total synthesis of an extended *ganglio*-ganglioside, IV^{4}GalNAcβGM1b. *Tetrahedron Lett.* **1990**, *31*, 1435–1438.
25. Murakata, C.; Ogawa, T. Stereoselective total synthesis of the glycosyl phosphatidylinositol (GPI) anchor of *Trypanosoma brucei*. *Carbohydr. Res.* **1992**, *235*, 95–114.
26. Fuse, T.; Ando, H.; Imamura, A.; Sawada, N.; Ishida, H.; Kiso, M.; Ando, T.; Li, S.-C.; Li, Y.-T. Synthesis and enzymatic susceptibility of a series of novel GM2 analogs. *Glycoconj. J.* **2006**, *23*, 329–343.

27. Konradsson, P.; Mootoo, D.R.; McDevitt, R.E.; Fraser-Reid, B. Iodonium ion generated *in situ* from *N*-iodosuccimide and trifluoromethanesulphonic acid promotes direct linkage of 'disarmed' pent-4-enyl glycosides. *J. Chem. Soc. Chem. Commun.* **1990**, 270–272.
28. Veeneman, G.H.; van Leeuwen, S.H.; van Boom, J.H. Iodonium ion promoted reactions at the anomeric centre. II An efficient thioglycoside mediated approach toward the formation of 1,2-*trans* linked glycosides and glycosidic esters. *Tetrahedron Lett.* **1990**, *31*, 1331–1334.
29. Matsuzaki, Y.; Ito, Y.; Nakahara, Y.; Ogawa, T. Synthesis of branched poly-*N*-acetyl-lactosamine type pentaantennary pentacosasaccharide: Glycan part of a glycosyl ceramide from rabbit erythrocyte membrane. *Tetrahedron Lett.* **1993**, *34*, 1061–1064.
30. Schmidt, R.R. New methods for the synthesis of glycosides and oligosaccharides–Are there alternatives to the Koenigs-Knorr method? [new synthetic methods (56)]. *Angew. Chem. Int. Ed.* **1986**, *25*, 212–235.
31. Yeung, B.K.S.; Hill, D.C.; Janicka, M.; Petillo, P.A. Synthesis of two hyaluronan trisaccharides. *Org. Lett.* **2000**, *2*, 1279–1282.

Sample Availability: Samples of the compounds are not available from the authors.

Int. J. Mol. Sci. **2014**, *15*, 3768-3783; doi:10.3390/ijms15033768

OPEN ACCESS
International Journal of
Molecular Sciences
ISSN 1422-0067
www.mdpi.com/journal/ijms

Review

Three-Dimensional Structural Aspects of Protein–Polysaccharide Interactions

Masamichi Nagae and Yoshiki Yamaguchi *

Structural Glycobiology Team, Systems Glycobiology Research Group,
RIKEN-Max Planck Joint Research Center, RIKEN Global Research Cluster,
2-1 Hirosawa, Wako, Saitama 351-0198, Japan; E-Mail: mnagae@riken.jp

* Author to whom correspondence should be addressed; E-Mail: yyoshiki@riken.jp;
Tel.: +81-48-467-9619; Fax: +81-48-467-9620.

Received: 26 December 2013; in revised form: 17 February 2014 / Accepted: 21 February 2014 / Published: 3 March 2014

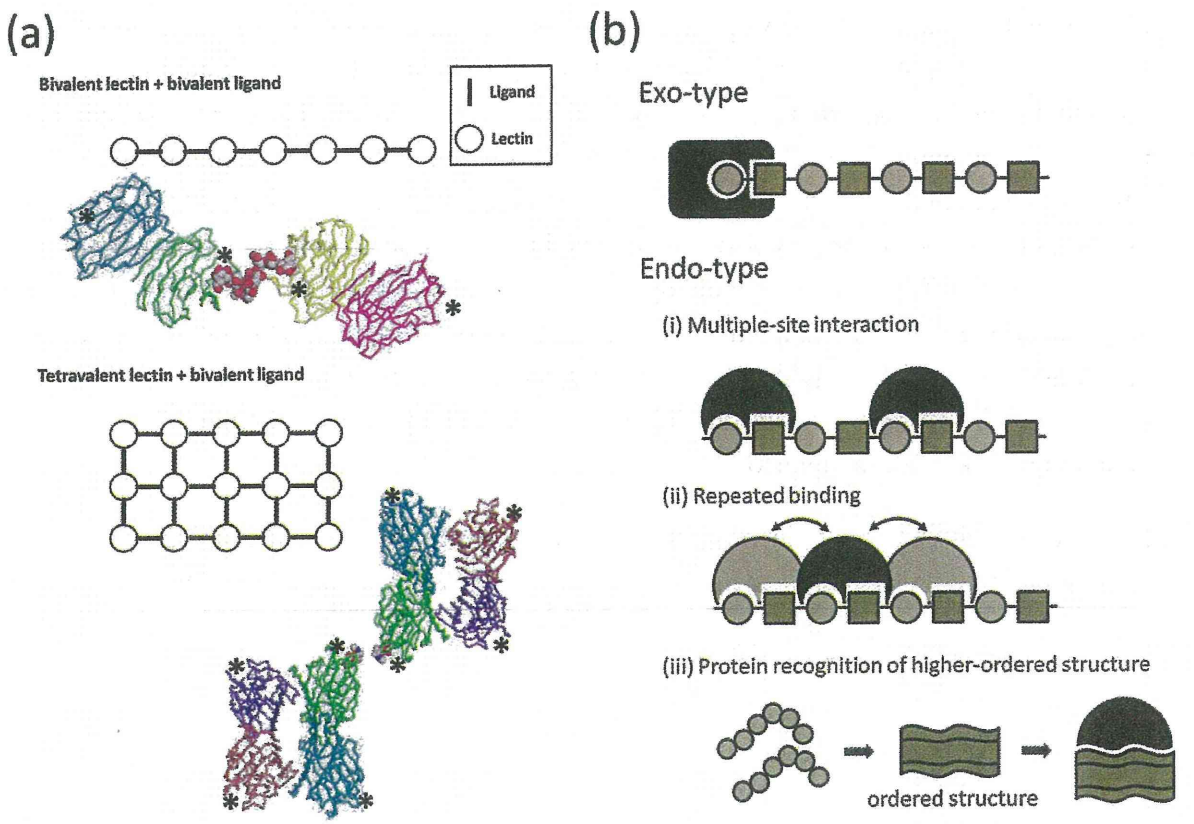
Abstract: Linear polysaccharides are typically composed of repeating mono- or disaccharide units and are ubiquitous among living organisms. Polysaccharide diversity arises from chain-length variation, branching, and additional modifications. Structural diversity is associated with various physiological functions, which are often regulated by cognate polysaccharide-binding proteins. Proteins that interact with linear polysaccharides have been identified or developed, such as galectins and polysaccharide-specific antibodies, respectively. Currently, data is accumulating on the three-dimensional structure of polysaccharide-binding proteins. These proteins are classified into two types: exo-type and endo-type. The former group specifically interacts with the terminal units of polysaccharides, whereas the latter with internal units. In this review, we describe the structural aspects of exo-type and endo-type protein-polysaccharide interactions. Further, we discuss the structural basis for affinity and specificity enhancement in the face of inherently weak binding interactions.

Keywords: polysaccharide; lectin; 3D structure; polylactosamine; galectin; carbohydrate binding module; β-glucan; polysialic acid; antibody; affinity; hyaluronan; CD44

1. Introduction

Compared with protein–protein interactions, the binding between proteins and individual, simple carbohydrate ligands is weak. Carbohydrate-protein interaction often requires affinity enhancement in order to attain biological significance. Binding affinity can be substantially increased by multivalent binding of carbohydrate ligand and/or carbohydrate-binding protein through the cluster effect [1]. Here, multivalent carbohydrate chains and clustered carbohydrate-binding proteins form lectin-carbohydrate lattices. Examples include bovine galectin-1 and soybean agglutinin (SBA) [2]. Dimeric bovine galectin-1 complexed with biantennary complex-type glycan displays multiple cross-links through its branched glycan termini [3]. SBA is a tetrameric legume lectin that uniquely complexes with multiantennary carbohydrates [4,5] (Figure 1a). The biological importance of the galectin/lectin lattice has been reviewed [6].

Figure 1. (**a**) Schematic representations of bivalent lectin-bivalent ligand (type I) and tetravalent lectin-bivalent ligand (type II) complexes. Crystal packings are shown for bivalent galectin-1 dimer in complex with biantennary complex-type glycan (type I, upper panel, PDB code 1SLB), and tetravalent soybean agglutinin tetramer in complex with binantennary complex-type glycan (type II, lower panel, PDB code 1SBE). Protein and carbohydrate molecules are shown in wire and sphere models, respectively, and arbohydrate binding sites are shown with asterisks; (**b**) Schematic representations of exo-type (upper) and endo-type (lower) interactions. The affinity enhancement strategies of endo-type lectins: (i) multiple-site interaction; (ii) repeated binding; and (iii) protein recognition of ordered/higher-ordered polysaccharide structures.



The mechanism of affinity enhancement among linear repeating polysaccharides, such as polylactosamine, hyaluronan, β-glucan, and polysialic acid, which are ubiquitous among living organisms, is quite distinct. Linear polysaccharides serve as a scaffold for crosslinking to polysaccharide-binding proteins, thereby enhancing the affinity of specific binding partners. Lectins that bind to linear polysaccharides are classified as either exo- or endo-types (Figure 1b). Exo-type lectins interact with terminal units of polysaccharides, and bind glycan chains in partially sealed clefts. Lectins of this type specifically bind to the terminal cap structure. In contrast, endo-type lectins interact with internal units of polysaccharides using open clefts. Endo-type lectins enhance binding affinity to polysaccharide ligands by three mechanisms: (1) Multiple-site interactions; (2) Repeated binding; (3) Recognition of ordered/higher-ordered polysaccharide structure.

Tighter binding of endo-type proteins to longer polysaccharides is often achieved by combining all three mechanisms. Multiple-site interaction occurs when a long polysaccharide binds two or more lectins simultaneously. Such interactions can be identified by isothermal titration calorimetry and the stoichiometry of binding analyzed using a Scatchard plot. Repeated binding implies that one lectin repeatedly dissociates and rebinds or slides on the polysaccharide. The apparent affinity is enhanced by slowing the dissociation of the protein. This mechanism can be identified by kinetic analysis, using techniques such as surface plasmon resonance. This mode of affinity enhancement requires that the protein has ligand-binding pockets open at both ends. The third mechanism is called the conformational epitope hypothesis, but remains speculative. In this interaction, a polysaccharide of a certain chain length forms a higher-ordered conformation, such as a helix, that assists in protein recognition and tighter binding. The chain length at which the polysaccharide forms a higher-ordered structure will depend on the glycosidic linkage and sugar type.

The shorter sugar chains, the oligosaccharides, are highly soluble in water as are many of the longer ones, the polysaccharides. However, some polysaccharides are insoluble—these include curdlan (linear β1–3 glucan), cellulose (β1–4 glucan), and chitin/chitosan (β1–4 linked GlcNAc/GlcN polymer). Insoluble polysaccharides such as cellulose and chitin play structural roles in plants, fungi, and insects. Interaction between proteins and cellulose has been studied extensively [7] and reviewed [8,9]. Type-A cellulose-binding modules (CBMs) use a flat surface populated with aromatic residues to bind to crystalline cellulose whereas type-B CBMs use a deep groove to bind individual twisted glucan chains found in disordered cellulose. Human chitin-binding proteins have been reported recently. Human YKL-39 and -40, also designated as cartilage glycoproteins, bind preferentially to chitooligosaccharides [10–12]. The specificity of YKL-40 for chitin is critically dependent on the length of the oligosaccharide. Although we now have 3D structural details of the chitooligosaccharide–protein interaction, the physiological role of these proteins remains poorly understood.

In this review, we highlight four types of soluble oligo/polysaccharides—polylactosamine, hyaluronan, short β1–3 glucans, and α2–8 polysialic acids—and summarize their three-dimensional (3D) structures and modes of protein interaction.

2. Atomic Details of the Polylactosamine and Galectin Interaction

Polylactosamine (polyLacNAc) contains repeating *N*-acetyllactosamine (LacNAc) units $(3Galβ1–4GlcNAcβ1–)_n$ as glycan extensions of cell-surface glycolipids and glycoproteins. Molecular

modeling of low-energy polyLacNAc conformers demonstrates that the β1–4 linkage adopts an extended conformation [13]. Galectins comprise a family of evolutionarily conserved β-galactoside-specific lectins [14] and are binding partners of polyLacNAc. Galectins form multivalent complexes with cell-surface glycoconjugates and in so doing transmit various signals intracellularly to regulate cellular activation, differentiation, and survival [6]. The binding affinity between galectins and polyLacNAc has been comprehensively analyzed by frontal affinity chromatography (FAC) [15]. Galectin-3, the *C*-terminal carbohydrate recognition domain (CRD) of galectin-8, and the *N*- and *C*-terminal CRDs (NCRD and CCRD) of galectin-9 bind longer oligolactosamines tighter, while galectin-1, -2, and -7 and the NCRD of galectin-8 bind with equal strength to any length [15]. Accordingly, the former group is classified as endo-type, while the latter as exo-type. A structural comparison between endo- and exo-type binding highlights the structural differences that define their different interaction modes (see below).

Crystal structures of human galectin-9 NCRD complexed with LacNAc dimers (LN2) and LacNAc trimers (LN3) were solved with two different types of crystal packing [16]. The overall structure of galectin-9 NCRD is a β-sandwich fold, and the protein interacts with glycan on its concave surface (Figure 2a). Galectin-9 NCRD interacts with a LacNAc unit in the LN2 and LN3 complexes in two ways (Figure 2b). In the LN2 complex, one molecule (molecule A) interacts with the reducing end of the LacNAc repeat, while the other molecule (molecule B) interacts with the non-reducing end (PDB codes 2ZHK and 2ZHL). In the LN3 complex, one molecule binds to the second LacNAc unit from the non-reducing end, while the other binds a third LacNAc unit (PDB codes 2ZHM and 2ZHN). Thus, human galectin-9 NCRD interacts with three consecutive sugar residues (GlcNAcβ1–3Galβ1–4GlcNAc) irrespective of their positions within the polyLacNAc chain. Accordingly, it is characterized as an endo-type lectin for polyLacNAc. The dissociation constants of human galectin-9 NCRD for LN2, LN3, and LacNAc pentamers (LN5) are 3.0, 0.81, and 0.12 μM, respectively [15]. The increased affinity for longer chains is explained by both multiple-site interaction and repeated binding.

A structural comparison between galectin-9 NCRD and other galectins highlights the difference between endo- and exo-type lectins. Human galectin-3 CRD (PDB code 1A3K) [17] can be structurally superimposed onto the human galectin-9 NCRD-LN2 or -LN3 complex without structural conflicts between galectin-3 and the respective oligosaccharides (Figure 3a). This may explain galectin-3's function as an endo-type lectin for polyLacNAc. In contrast, structural superposition of galectin-7 CRD (PDB code 1BKZ) [18] onto the LN3 complex shows significant steric clash between LN3 and the Gln42 side chain; thus, galectin-7 can only bind to terminal polyLacNAc units (Figure 3b). Molecular dynamics simulations suggest that galectin-8 CCRD interacts with internal LacNAc units of a lactosamine chain [19]. Actually structural superposition of galectin-8 CCRD (PDB code 3VKM) [20] and galectin-9 NCRD LN complexes shows no apparent steric clash.

Galectin-1 primarily interacts with the terminal LacNAc unit of a polyLacNAc chain [21,22]. Galectin-1 CRD monomer superimposes on galectin-9 NCRD-LN2 and -LN3 complexes without steric clash. However, the galectin-1 CRD dimer sterically interferes with the LN chain (Figure 3c). Possibly, formation of a galectin-1 dimer converts it into an exo-type lectin for polyLacNAc, a form that cannot accommodate a longer LN chain. The monomer-dimer equilibrium of galectin-1 in solution may regulate the binding to longer LN chain.

Figure 2. Polylactosamine recognition by galectins. (**a**) Structure of human galectin-9 *N*-terminal carbohydrate recognition domain (NCRD)–LacNAc trimers (LN3) complex (PDB code 2ZHN). The protein molecule is shown in semi-transparent surface and ribbon models. The amino-acid residues that interact with carbohydrate residues are shown in rod models. Hydrogen bonds are indicated by red dotted lines; (**b**) Schematic representation of the interaction modes of LN2 and LN3 complexes. Galactose residues that interact with conserved key residues in galectin-9 NCRD are shown as black triangles. Carbohydrate residues that interact with galectin-9 NCRD are indicated by double head arrows.

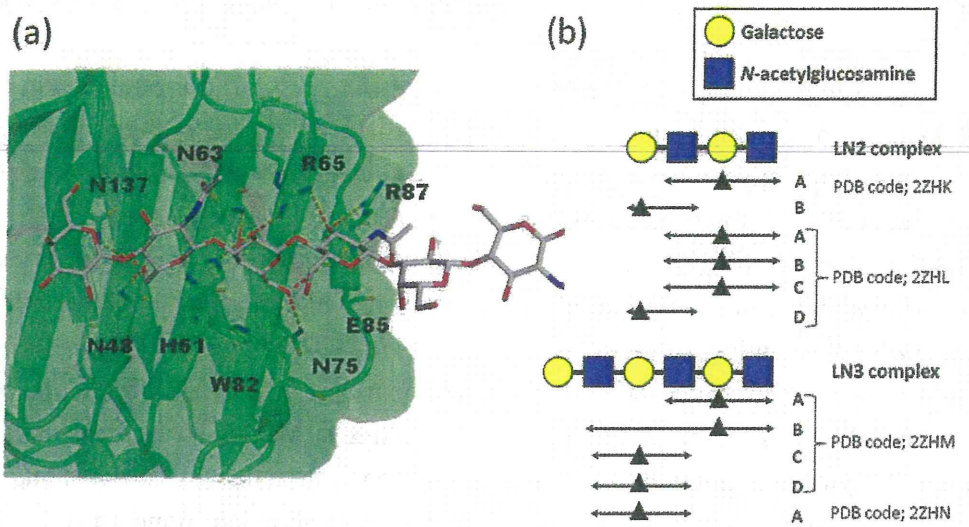
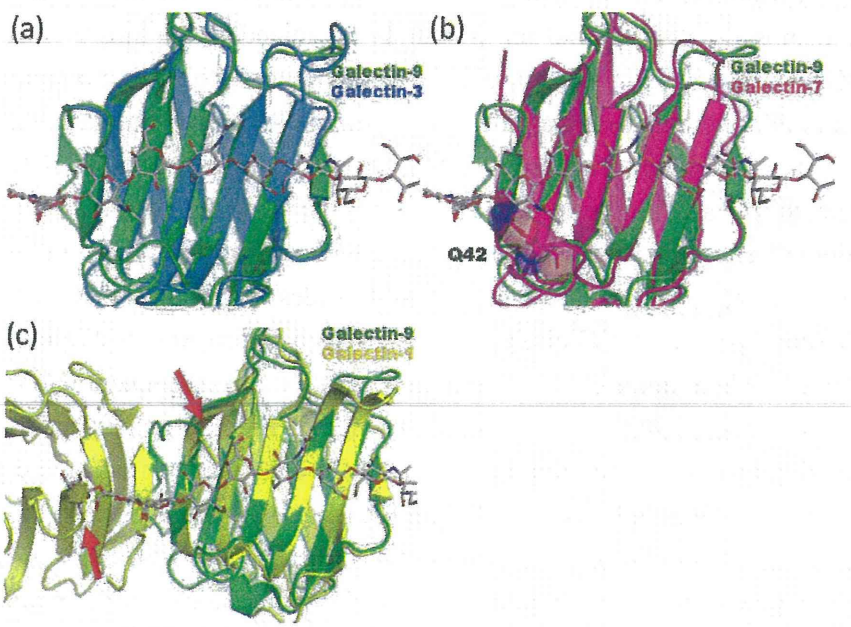


Figure 3. Structural comparison of galectin CRDs. (**a**) Structural superpositions between hypothetical galectin-9 NCRD-LN5 complex (green) and human galectin-3 CRD (cyan; PDB code 1A3K) [17]; (**b**) Structural comparison with human galectin-7 CRD (magenta; PDB code 1BKZ) [18]. The Gln42 side chain, which collides with LN5, is shown in rod model; (**c**) Structural comparison with human dimeric galectin-1 CRD (yellow; PDB code 3OYW) [23]. Possible steric clashes are indicated with red arrows.



3. Structural Studies of CD44 and Hyaluronan (HA) Complex

Hyaluronan is an unbranched polysaccharide comprising repeating disaccharides of GlcNAc and glucoronic acid (GlcUA) connected by β1–3 and β1–4 linkages, $(-3GlcNAc\beta1-4GlcUA\beta1-)_n$. Hyaluronan is localized in the extracellular compartment of most tissues and is a major component of cartilage and synovial fluid. Hyaluronan has diverse biological roles in vertebrates, such as being a vital structural component of connective tissue, the formation of loose hydrated matrices, immune cell adhesion and activation, and a role in intracellular signaling. Hyaluronan is a unique glycosaminoglycan since it is not sulfated. Hyaluronan is likely to be highly dynamic in solution, adopting a large number of low energy states [24].

Numerous hyaluronan-binding proteins have been identified [25]. The proteins contain a common structural domain, termed a Link module, which is involved in ligand binding. The Link module, also referred to as a proteoglycan tandem repeat, was first identified in the link protein isolated from cartilage and classified into three subgroups (types A, B, and C) on the basis of the size of their hyaluronan-binding domains [26]. The size of the binding domain appears to correlate broadly with the length of hyaluronan recognized. Structural analyses of HA binding proteins have been limited to tumor necrosis factor stimulating gene-6 (TSG-6) (type A) and CD44 (type B) (Figure 4a).

CD44 is a principal cell surface receptor for hyaluronan and has diverse functions including the attachment, organization, and turnover of extracellular matrix at the cell surface and mediation of the migration of lymphocytes during inflammation [27]. CD44 exists in numerous isoforms due to alternative splicing of 10 variant exons in different combinations. The most abundant standard isoform of CD44 is composed of an extracellular hyaluronan binding domain (HABD), a membrane proximal stalk region, a transmembrane region, and an intracellular *C*-terminal region. The *C*-terminal region is known to interact with cytoskeletal proteins and to link extracellular CD44-HA association with intracellular signaling cascades.

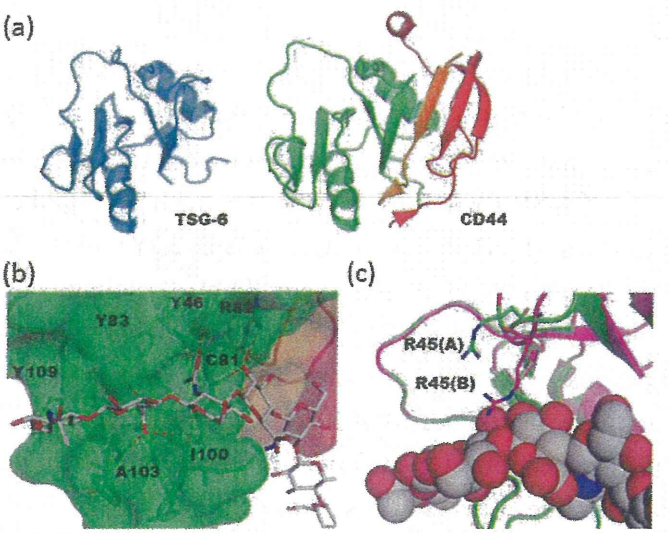
Structural analyses of CD44 HABD have been independently reported by two groups [28–31]. The three-dimensional structures of CD44 HABD in ligand-free state were revealed by X-ray (PDB code 1UUH (human), 2JCP (mouse)) and NMR studies (PDB code 1POZ (human)). CD44 HABD comprises two orthogonally disposed α-helices and ten β-strands organized into a single long, curving β-sheet with a pair of disulfide bridges (Figure 4a). Its fold shows resemblance to tumor necrosis factor stimulating gene-6 (*TSG-6*) Link module but possesses additional *N*- and *C*-terminal extended regions (Figure 4a).

In CD44 HABD-HA_8 complex (PDB code 2JCQ (type A complex) and 2JCR (type B complex)), seven carbohydrate residues of the octasaccharide could be modeled into the crystal structure—the GlcUA at the reducing end could not be accommodated [31]. The HA chain straddles the cavity of CD44. CD44 interacts strongly with two repeats (tetrasaccharides) at the non-reducing end. These residues are buried in a shallow groove of CD44 and stabilized by a hydrogen-bonded network (Figure 4b). All amino acid residues which make contacts with the carbohydrate are located within the Link module and there is no contribution from other β-strands to ligand binding. There are two crystal forms, termed type A and B complexes. The most significant difference between the two crystal forms is the location of the Arg45 side chain relative to HA_8 (Figure 4c). In the type A complex structure, there is no contact between Arg45 and HA_8, while in the type B there is a hydrogen bond between them. It was suggested that type A and type B represent low-affinity and high-affinity states, respectively.

Solution NMR analysis shows that the *C*-terminal segment of HABD has enhanced flexibility in the HA-bound state [30], although the corresponding region of a crystal structure of the HABD-HA_8 complex adopts an ordered conformation. Moreover, the solution conformation of Arg45 is not uniform even in the presence of HA oligomer. The conformations of the *C*-terminal segment in the unbound and HA-bound states were designated as "ordered" and "partially disordered" conformations, respectively. The transition between "ordered" to "partially disordered" conformations may account for switching from a low to a high affinity form under conditions of shear flow, as required for normal leukocyte trafficking [32]. Further studies will clarify the relationship between these conformation transitions and affinity regulation.

Chain-length-dependent interaction is also reported for CD44. Monovalent binding is observed for chain lengths between HA_6 and HA_{18}, whereas divalent binding occurs in lengths of more than HA_{20} [33]. Interestingly, CD44 shows increased affinity for hyaluronan when the latter is in complex with the TSG-6 link module [34]. It seems protein-protein interactions reinforce protein–carbohydrate interactions of CD44 although the structural basis is still unknown.

Figure 4. (**a**) Overall structures of tumor necrosis factor stimulating gene-6 (TSG-6) (PDB code 1O7B, [35], left panel) and CD44 (PDB code; 2JCQ, [31], right panel) Link modules. The extended *N*- and *C*-terminal regions of CD44 are colored in orange and red, respectively; (**b**) Close up view of CD44–HA_8 complex (PDB code 2JCQ). The protein molecule is shown in semi-transparent surface and ribbon models. The amino-acid residues that interact with carbohydrate residues are shown in rod models. Hydrogen bonds are indicated by red dotted lines; (**c**) Structural superposition of two crystal forms of CD44-HA_8 complex (crystal form A; green, 2JCQ, crystal form B; magenta, 2JCR). HA_8 molecule is shown in sphere model.



4. Atomic Characterization of the Interaction between Carbohydrate-Binding Modules and β1–3 Glucan

β1–3 Glucan is a major component of the fungal cell wall and is not synthesized in humans. The innate immune system in humans targets β-glucan as a non-self component. β-glucan assumes a

triple-helical structure according to data from X-ray fiber diffraction [36]. In a widely accepted hydrogen-bonding model, one β1–3 glucan chain forms intramolecular hydrogen bonds with the other two chains perpendicular to the axis of the triple helix [37]. Modeling studies of laminarin demonstrated an energetically favorable helical configuration with φ and ϕ angles of −72° and 108°, respectively [38]. Carbohydrate-binding module (CBM) families 4, 6, 39, 43, 52, and 56 interact with β1–3 glucans (CAZy database, http://afmb.cnrs-mrs.fr/CAZY/index.html). To date, 3D structural information on β-glucan binding is limited to CBM4, 6, and 39.

CBM4 binds to xylan, amorphous cellulose, β1–3/β1–4 mixed glucan, and β1–3 glucan but not to crystalline cellulose. The crystal structure of CBM4 from *Thermotoga maritima* Lam16A (*Tm*CBM4-2) in complex with laminariheptaose has been reported (PDB code 1GUI) [39]. The *Tm*CBM4-2 structure includes a β-jelly roll with ten β-strands and forms a long, shallow groove (Figure 5a). The carbohydrate-binding site is a U-shaped depression with high-sided walls formed by loops between the β-strands. *Tm*CBM4-2 interacts with glucose residues located in the middle of the glucan chain. Thus, it acts as an endo-type lectin for the β-glucan chain.

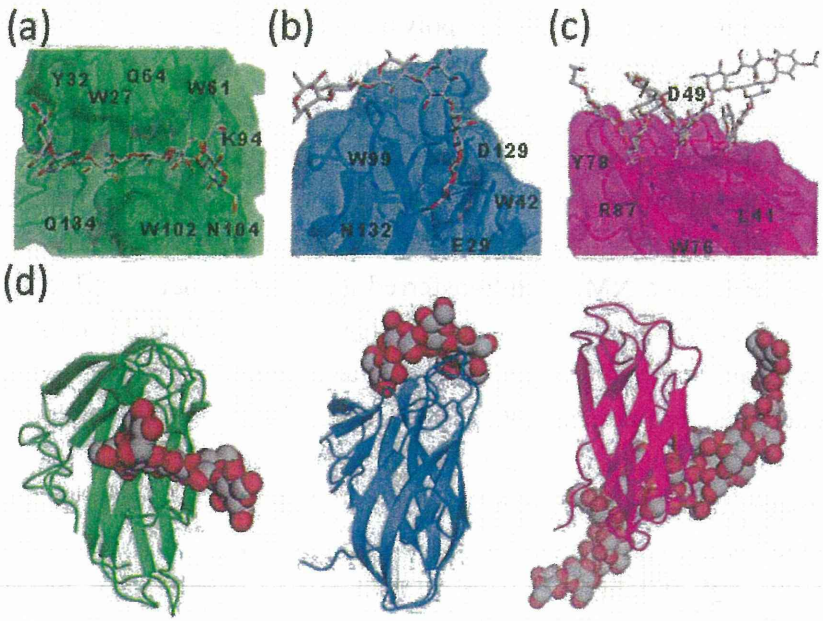
CBM6 is found in various enzymes with activities against xylan, mannan, agarose, arabinans, and β-glucans. BH0236 from *Bacillus halodurans* has laminariase activity and possesses CBM6 (*Bh*CBM6) in the *C*-terminal region. The crystal structure of *Bh*CBM6 in complex with laminarihexaose was previously reported (PDB code 1W9W) [40]. The overall structure of *Bh*CBM6 adopts a β-sandwich fold with five- and four-stranded β-sheets (Figure 5b). The binding cleft is occluded at one end, which prevents binding of linear polysaccharides and proper orientation of the sugar at this site, thus conferring specificity for the non-reducing ends of glycans. The terminal glucose residue at the non-reducing end is sandwiched between Trp42 and Trp99 and is recognized via a hydrogen-bond network. In addition, the OH3 group at the non-reducing end interacts with the Asn132 side chain; thus, *Bh*CBM6 interacts with the non-reducing end of the β1–3 glucan chain and acts as an exo-type lectin for the β-glucan chain.

The β1–3 glucan-recognition protein (β-GRP) is classified as a CBM39 and contains an *N*-terminal β-glucan-binding domain and a *C*-terminal β1–3 glucanase-like domain that lacks enzymatic activity [41]. Solution and crystallographic analyses of β-GRP have been performed [42–44]. The crystal structures of the *N*-terminal domain of βGRP (βGRP-N) in complex with laminarihexaose provided evidence for the recognition mechanism (PDB codes 3AQX and 3AQY) [43]. The overall structures of *Plodia interpunctella* and *Bombyx mori* βGRP-N are immunoglobulin (Ig)-like β-sandwich folds composed of three- and five-stranded β-sheets (Figure 5c). One βGRP-N simultaneously interacts with three structured laminarihexaoses via six glucose residues (two from each chain). Since the spatial arrangement of the laminarihexaose bound to βGRP-N is nearly identical to that of a β1–3 glucan triple-helical structure, βGRP-N exhibits high affinity for "triple-helical" β-glucans.

A structural comparison of three CBMs reveals distinct interaction modes despite a common overall fold and binding preference for β1–3 glucans. CBM4 accepts single-chain ligands, while CBM6 interacts with the termini of ligands (Figure 5d). In contrast, CBM39 binds to a conformational epitope formed by triple-helical β-glucans through the flat surface on the convex side. The various interaction modes may reflect the differing physiological roles of CBMs. The conformational-epitope mode of recognition has also been proposed for Dectin-1, which is a C-type lectin-like receptor for longer β1–3 glucans (degree of polymerization (DP) > 10) [45]. Nuclear magnetic resonance (NMR) of medium-sized

β-glucan oligomers (DP ≈ 25) suggested the presence of weak hydrogen bonds, which may contribute to the formation of the helix recognized by Dectin-1 [46]. Banana lectin (Banlec) is a dimeric plant lectin from the jacalin-related family and also known to interact with β-glucan. Interestingly, Banlec binds to laminaribiose (Glcβ1–3Glc) mainly through the reducing terminus [47]. A variety of recognition modes of β-glucan by lectins implies this carbohydrate has various physiological roles.

Figure 5. Interaction between cellulose-binding module (CBM) and β1–3 glucan. (**a**) Close-up view of the ligand-binding site of *Tm*CBM4-2 in complex with laminariheptaose (PDB code 1GUI) [39]. Protein molecule is shown in semi-transparent surface and ribbon models. Carbohydrate is shown in rod model; (**b**) Close-up view of the ligand-binding site of *Bh*CBM6 in complex with laminarihexaose (PDB code 1W9W) [40]; (**c**) Close-up view of the ligand-binding site of βGRP-N in complex with laminarihexaose (PDB code 3AQX) [43]; Hydrogen bonds are indicated by red dotted lines; (**d**) Structural comparison of three CBMs, *Tm*CBM4-2 (left), *Bh*CBM6 (middle) and βGRP-N (right). Protein and carbohydrate are shown in ribbon and sphere models, respectively.



5. Interaction between Polysialic Acid and Antibodies

Polysialic acid (polySia) is a linear homopolymer of α2–8- or α2–9-linked sialic acid with a DP of 3 to >400 residues [48,49]. PolySia was discovered as an abundant carbohydrate component in the developing mammalian brain [50] and plays a major role in the development, morphogenesis, and function of various neural systems. A number of antibodies that are specific for oligo/polysialic acid have been developed and are classified into three groups (groups I–III) based on their DP-dependent antigenic specificity [49]. Group I are antibodies that recognize chains of α2–8-linked polySia with a DP ≥ 8. Group II antibodies (termed anti-oligo plus polySia antibodies) recognize di/oligoSia with a DP of two to seven as well as polySia chains. Group III antibodies (termed anti-di/oligoSia antibodies) recognize di- and oligoSia with a DP of 2–4, but not polySia. There is limited structural information on the interactions between oligo/polySia and antibodies. Recently,

biochemical and structural analyses of the anti-oligo and polySia antibodies and their fragments 735 (group I), 12E3 (group II), and A2B5 (group III) provided atomic details of antigen recognition [51,52].

A murine monoclonal antibody mAb735 (group I) has a unique preference for longer polySia polymers (DP > 10). Based on a model of polySia docked in a mAb735 Fab, it was proposed that helical polySia fits into a crevice formed by the CDR loop regions [53]. The crystal structure of the single-chain variable fragment of mAb735 (scFv735) in complex with octasialic acid (PDB code 3WBD) [52] was recently reported. In the asymmetric unit, two scFv735 molecules associate with a single octasialic acid (Figure 6a). In both complexes of the unit, all complementarity-determining regions except for L3 interact with three of eight consecutive sialic acid residues. scFv735 forms hydrogen bonds with oligosialic acid residues by direct and water-mediated interactions (Figure 6b). scFv735 interacts with ligands by an endo-type mechanism, and several scFv735 molecules can bind simultaneously to the long polysialic acid chain. The dissociation constants for scFv735 with a series of oligo/polySia were 1.7×10^{-3}, 6.5×10^{-4}, 3.0×10^{-4}, and 3.5×10^{-6} for DP4, DP5, DP6, and polysialic acid (DP = 80–130), respectively. The increase in affinity from DP4 to 6 likely results from the decreased off-rate of the protein. On average, six molecules bind to a single polysialic acid chain; thus, multiple binding contributes to the affinity for polysialic acid. The contribution of helix formation of polysialic acid should be carefully considered, since the interaction between scFv735 and polysialic acid is enthalpically driven (ΔH: 80.1 kcal/mol) with an unfavorable entropic contribution ($T\Delta S$: −72.7 kcal/mol). The unfavorable entropy may result from the loss of conformational flexibility of the polysialic acid. In fact, the dihedral angles of the trisialic acid unit directly interacting with scFv735 are not uniform, indicating that mAb735 does not strictly favor the helical conformation, as once thought.

Saturation transfer difference NMR and transferred nuclear Overhauser effect (NOE) were used to analyze the binding modes of two anti-oligosialic acid IgMs, 12E3 (group II) and A2B5 (group III) [51]. A2B5 favors a trisialic acid unit, while 12E3 favors ligands with more than five successive sialic-acid residues. A2B5 predominantly interacts with sialic-acid residues at the non-reducing end, while 12E3 binds to internal sialic-acid residues through C4–C8 moieties and *N*-acetyl groups. Thus, it is likely that A2B5 has exo-type and 12E3 endo-type interactions. scFv735 is similar to 12E3 in that it likely binds internal residues and directly interacts with OH4, OH7, OH8, a carboxyl group, and *N*-acetyl oxygen atoms. Notably, scFv735 interacts with all OH4 groups of the triplet, whereas 12E3 primarily interacts with *N*-acetyl groups of three sequential sialic-acid residues.

The functions of the polysialic acids are likely closely related to their three-dimensional structure; however, the conformations of polysialic acids remain a topic of debate [54–57]. Further analyses of the structure of polysialic acid-binding proteins in complex with their ligands and their dynamics will help to clarify the conformation of polysialic acid chains.

6. Additional Modifications

The interactions between linear polysaccharides and proteins can be regulated by additional modifications such as sulfation or branch formation of polysaccharide chains. Although the atomic details have been reported for certain heparin-binding proteins including antithrombin [58], FGF2 [59], and annexin [60], 3D structural information on the recognition mode of sulfated polysaccharides remains limited. The interaction with sulfated glycan is mostly electrostatic, attained by basic residues,