Electronic Supplementary Material (ESI) for Chemical Communications

Supporting Information

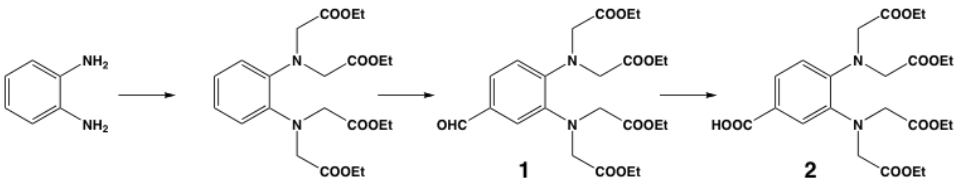
Lanthanide-assisted NMR evaluation of a dynamic ensemble of oligosaccharide conformations

Sayoko Yamamoto, Ying Zhang, Takumi Yamaguchi, Tomoshi Kameda, and Koichi Kato

Preparation of tagged GM3 trisaccharide

General: Reagents and solvents were commercially available and used without any further purification unless otherwise noted. Column chromatography was performed on Silica Gel 60N purchased from Kanto Chemical Co., Inc., Wakosil 40C18 from Wako Pure Chemical Industries, Ltd. or Waters Sep-Pak C18. Elemental analysis (EA) and high resolution MS measurement were performed on Yanaco MT-6 and JEOL JMS-777V, respectively, at Instrument Center, IMS. NMR spectra were recorded on JEOL JNM ECA-600 spectrometer equipped with a 5-mm FG/HCN probe. TMS (in $CDCl_3$) served as internal standard for ^{1}H and ^{13}C NMR measurements.

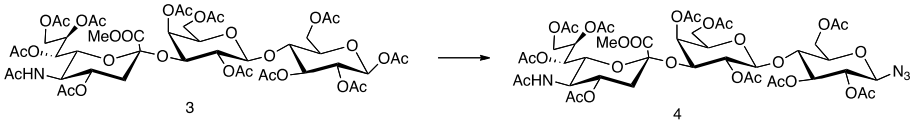
Preparation of metal chelating unit **2**:



The precursor **1** was synthesized from *o*-phenylenediamine in 2 steps[1].NaH_2PO_4 (2.6 g, 21.8 mmol) and $NaClO_2$ (2.4 g, 32.8 mmol) were dissolved in H_2O (20 mL) and then transferred to the solution of **1** (1.3 g, 2.7 mmol) and 2-methyl-2-butene (2.9 mL, 27.2 mmol) in *tert*-butyl alcohol (20 mL). The mixture was stirred at RT for 7 h, and then concentrated and neutralized with HCl (1 M) aqueous. This mixture was extracted with EtOAc, dried with Na_2SO_4 and concentrated. The residue was purified on a silica gel column with EtOAc/Hexane (1:1) to give

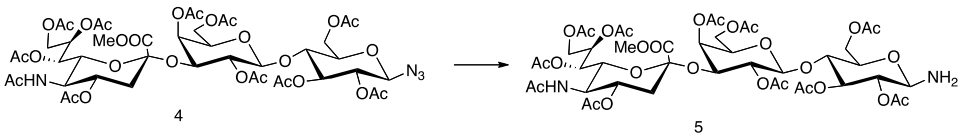
2 (1 g, 74%). ^{1}H-NMR ($CDCl_3$, 300 K): δ = 7.80 (d, *J* = 1.38Hz, 1H, Ar*H*), 7.70 (dd, *J* = 8.3, 1.4 Hz, 1H, Ar*H*), 7.02 (d, *J* = 8.9 Hz, 1H, Ar*H*), 4.39 (s, 4H, N(CH_2)$_2$), 4.26 (s, 4H, N(CH_2)$_2$), 4.11(m, 8H, COOCH_2), 1.20 (dd, *J* = 14.1, 7.13Hz, 12H, CH_2CH_3); ^{13}C-NMR ($CDCl_3$, 300 K): δ = 171.9, 171.0, 147.5, 141.0, 126.0, 124.1, 126.0, 120.5, 61.0, 52.4, 14.1; HRMS (FAB): Calcd for $C_{23}H_{33}N_2O_{10}$ [M+H$^+$]: 497.2135; Found: 497.2130.

Preparation of azide **4**:



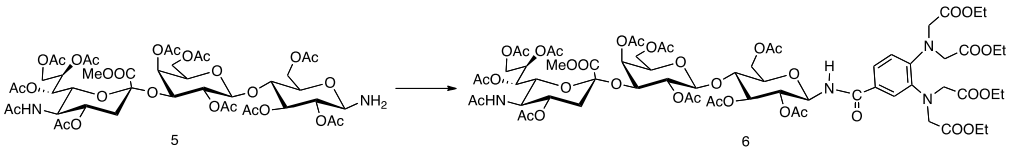
Trimethylsiylazide (179 μL, 1.36 mmol) and $SnCl_4$ (83.75 μL, 0.715 mmol) was added at 0 °C to a solution of compound **3** (371 mg, 0.335 mmol) in CH_2Cl_2 (5 mL). The mixture was stirred at RT for 12 h. Then the reaction mixture was diluted with $CHCl_3$, washed with saturated aqueous $NaHCO_3$, H_2O, dried with Na_2SO_4 and concentrated. The residue was purified by column chromatography on silica gel ($CHCl_3$/MeOH 20:1) to give **4** (320 mg, 88%).^{1}H-NMR ($CDCl_3$, 300 K): δ = 5.52 (m, 1H, Neu8), 5.38 (dd, *J* = 9.65, 2.86 Hz, 1H, Neu7), 5.18 (t, *J* = 9.37 Hz, 1H, Glc3), 5.03 (d, *J* = 10.4 Hz, 1H, N*H*), 4.94-4.81 (m, 4H, Gal2, 4, Neu4, Glc2), 4.67 (d, *J* = 7.60 Hz, 1H, Gal1), 4.62 (d, *J* = 8.26 Hz, 1H, Glc1), 4.51 (dd, *J* = 10.3, 3.44 Hz, 1H, Gal3), 4.47 (dd, *J* = 11.7, 2.04 Hz, 1H, Gal6a), 4.41 (dd, *J* = 12.4, 2.70 Hz, 1H, Gal6b), 4.18 (dd, *J* = 11.7, 5.56 Hz, 1H, Glc6a), 4.00 (m, 4H, Glc6b, Neu5, 9), 3.89 (t, *J* = 9.65 Hz, 1H, Glc4), 3.83 (4H, COOMe, Gal5), 3.69 (m, 1H, Glc5), 3.61 (dd, *J* = 11.0, 2.74 Hz, Neu6), 2.56 (dd, *J* = 12.7, 4.48 Hz, 1H, Neu3a), 2.23-1.99 (m, 33H, OAc), 1.84 (s. 3H, NHAc), 1.66 (t, *J* = 12.4 Hz, 1H, Neu3b); ^{13}C-NMR ($CDCl_3$, 300 K): δ = 171.7 171.3, 171.0, 170.2, 169.4, 168.5, 101.2, 97.1, 87.8, 76.1, 75.0, 74.9, 73.3, 72.2, 71.4, 71.2, 70.6, 70.0, 69.9, 69.4, 67.8, 67.3, 67.0, 62.3, 62.1, 61.5, 53.3, 49.3, 37.5, 23.3, 21.6, 21.0, 20.7; HRMS (FAB): Calcd for $C_{44}H_{61}N_4O_{28}$ [M+H$^+$]:1093.3472; Found:1093.3483.

Preparation of amine **5**:



Compound **4** (270 mg, 0.247 mmol) was dissolved in MeOH (5 mL), to this solution was added Pd/C (4.8 mg). The mixture was stirred in hydrogen atmosphere at RT for 12 h. The reaction mixture was filtered through celite and concentrated, then purified by column chromatography on silica gel ($CHCl_3$/MeOH 20:1) to give **5** (200 mg, 76%). ^{1}H-NMR ($CDCl_3$, 300 K): δ = 5.51 (m, 1H, Neu8), 5.38 (dd, *J* = 9.64, 2.70 Hz, 1H, Neu7), 5.19 (t, *J* = 9.31 Hz, 1H, Glc3), 5.03 (d, *J* = 10.4 Hz, 1H, N*H*), 4.90 (m, 3H, Gal4, Neu4, Gal2), 4.71 (t, *J* = 9.57 Hz, 1H, Glc2), 4.62 (d, *J* = 8.23 Hz, 1H, Gal1), 4.50 (dd, *J* = 9.67, 3.35 Hz, 1H, Gal3), 4.40 (m, 2H, Gal6), 4.14 (m, 2H, Glc1, 6a), 3.99 (m, 4H, Glc6b, Neu5, 9), 3.83 (br, 4H, Gal5, COOMe), 3.78 (t, *J* = 9.57 Hz, 1H, Glc4), 3.60 (m, 2H, Neu6, Glc5), 2.54 (dd, *J* = 12.7, 4.45 Hz, 1H, Neu3a), 1.99-2.23 (m, 33H, OAc), 1.85 (s. 3H, NHAc), 1.66 (t, *J* = 12.3 Hz, 1H, Neu3b); ^{13}C-NMR ($CDCl_3$, 300 K): δ = 171.4, 171.2, 171.1, 170.9, 170.2, 168.4 101.0, 97.5, 84.6, 76.6, 73.7, 73.5, 72.6, 72.1, 71.6, 70.5, 70.0, 69.4, 67.8, 67.4, 67.0, 62.6, 62.3, 61.7, 53.3, 49.1, 31.3, 23.3, 21.6, 21.0, 20.7. EA: Calcd for $C_{44}H_{62}N_2O_{28}$•$2H_2O$: C, 47.91; H, 6.03; N, 2.54. Found: C, 48.02; H, 5.89; N, 2.25.

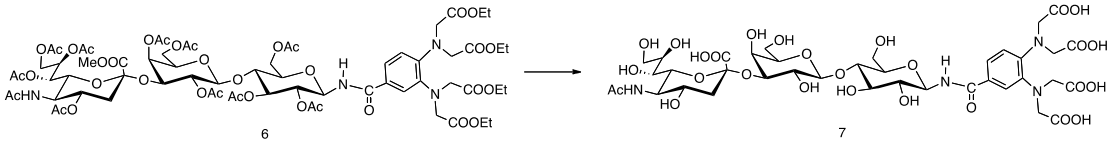
Preparation of compound **6**:



Compound **2** (100 mg, 0.20 mmol), DIPEA (0.35 mL, 2 mmol) and HATU (77.1 mg, 0.20 mmol) was dissolved in DMF (5 mL) and the mixture was stirred at RT for 10 min. Subsequently, the mixture was transferred to a solution of compound **5** (180 mg, 0.169 mmol) and DIPEA (17.3 μL, 0.1 mmol) in DMF (1 mL) and this solution was stirred at RT for 12 h. Reaction mixture extracted with EtOAc, washed with H_2O, dried with Na_2SO_4 and concentrated.

The residue was purified on a silica gel column with $CHCl_3$/MeOH (25:1) to give **6** (36 mg, 25%). ^{1}H-NMR ($CDCl_3$, 300 K): δ = 7.51 (dd, *J* = 1.3 Hz, 1H, Ar*H*), 7.25 (1H, Ar*H*), 7.00 (dd, *J* = 8.2 Hz, 1H, Ar*H*), 6.82 (d, *J* = 8.85 Hz, 1H, N*H*CO), 5.47 (m, 1H, Neu8), 5.40 (dd, *J* = 9.44, 2.55 Hz, 1H, Neu7), 5.33 (m, 2H, Glc1, 3), 5.07 (d, *J* = 9.78 Hz, 1H, N*H*), 4.91 (m, 4H, Glc2, Gal2, 4,Neu4), 4.61 (d, *J* = 7.71 Hz, Gal1), 4.50 (dd, *J* = 10.3, 3.47 Hz, Gal3), 4.45-4.15 (m, 11H, N(CH_2)$_2$, Gal6, Glc6a), 4.1 (m, 8H, N(CH_2)$_2$), 4.0 (m, Neu5, 9, Glc6b), 3.85 (m, 6H, COOMe, Gal5, Glc4, 5), 3.62 (dd, *J* = 11.0, 2.75 Hz, 1H, Neu6), 2.56 (dd, *J* = 13.1, 4.90 Hz, 1H, Neu3a), 2.24-1.97 (m, 33H, OAc), 1.8 (s, 3H, NHAc), 1.66 (t, *J* = 12.3 Hz, 1H, Neu3b), 1.2 (m, 12H, CH_2CH_3); ^{13}C-NMR ($CDCl_3$, 300 K): δ = 172.6, 171.4, 171.2, 171.0, 168.5, 167.2, 163.0, 145.7, 141.7, 127.6, 122.0, 121.6, 120.7, 101.0, 78.9, 76.1, 74.6, 72.6, 72.1, 71.5, 70.5, 70.0, 69.4, 68.0, 67.8, 67.4, 66.8, 62.1, 61.8, 60.8, 52.1, 53.3, 53.1, 49.2, 53.3, 37.5, 23.2, 21.0, 21.6, 20.8, 14.2, HRMS (FAB): Calcd for $C_{67}H_{93}N_4O_{37}$ [M+H$^+$]:1545.5519; Found: 1545.5511.

Preparation of modified trisaccharide **7**:



Compound **6** (60 mg) was dissolved in MeOH and small aliquots of NaOH (1 M) aqueous solution were added until the reaction was complete (TLC). The reaction mixture was purified by ODS column to give **7** (32 mg, 83%). ^{1}H-NMR (600 MHz, D_2O, 300 K): δ =7.32 (s, 1H, Ar*H*), 7.29 (d, *J* = 8.27 Hz, 1H, Ar*H*), 6.81 (d, *J* = 8.17 Hz, 1H, Ar*H*), 5.11 (d, *J* = 9.59 Hz, Glc1), 4.49 (d, *J* = 7.89 Hz, Gal1), 4.13 (m, 4H, N(CH_2)$_2$), 4.05 (dd, *J* = 9.7, 6.8 Hz, 1H, Gal3), 3.93-3.86 (m, 6H, N(CH_2)$_2$, Gal4, Glc6a), 3.82-3.76 (m, 4H, Neu5, 8, 9a, Glc6b), 3.70-3.62 (m, 7H, Gal5, 6, Glc3, 4, 5, Neu4), 3.56-3.52 (m, 5H, Gal2, Glc2, Neu6, 7, 9b), 2.68 (dd, *J* = 12.4, 7.88 Hz, Neu3a), 1.95 (s, 3H, NHAc), 1.72 (t, *J* = 12.1 Hz, 1H, Neu3b); ^{13}C-NMR (150 MHz, D_2O, 300 K): δ = 178.8, 175.3, 174.2, 171.7, 146.2, 140.2, 125.4, 123.1, 121.1, 102.6, 100.1, 80.1, 77.8, 76.6, 75.6, 75.3, 75.2, 73.0, 71.9, 71.6, 69.4, 68.5, 68.2, 67.6, 62.6, 61.0, 59.9, 54.6, 54.1, 51.6, 39.5, 21.8.

PCS observation of the tagged sugar

Compound **7** (2 mg) was dissolved in D_2O (0.6 mL) and pH was increased to 8.0 by adding NaOD solution. This solution was titrated with D_2O solution of MCl_3 (250 mM; M = La^{3+}, Tm^{3+} or Tb^{3+}) for NMR measurements. For PCS observation, ^{1}H-^{13}C HSQC spectra were recorded at 300 K with 512 (t_1) and 1024 (t_2) complex points. NMR spectra were processed and analyzed with the programs NMRPipe[2] and Sparky[3].

Generation of conformational ensemble of the GM3 trisaccharide

MD simulations of the sugar moiety of GM3:

All-atom molecular dynamics simulations of the GM3 trisaccharide were employed using the Sander module of the Amber11 package[4] with the GLYCAM_06 force field.[5] To create the initial structure and topology file of the GM3 trisaccharide, the tLeap module of the AmberTools1.5 program[4] was used. TIP3P waters were added to the solvent layer to ensure a depth of at least 8 Å from any atom. Ten Na^+ ion and nine Cl^- ion was added to neutralize the system. Before MD runs were performed, the entire system was energy minimized by 500 steps of steepest descent followed by 500 steps of conjugate gradient. The system was heated to 300 K with a 2-fs time step in the NPT ensemble[6] at 1 atm over 50 ps using isotropic position scaling. Production MD simulations were performed for 12 ns at 300 K with a 2-fs time step in the NPT ensemble. The initial velocities are randomized. Scaling of nonbonded 1–4 van der Waals and electrostatic interactions was not performed (i.e., SCEE=SCNB=1.0). All bonds involving hydrogen atoms were constrained with the SHAKE algorithm,[7] and long-range electrostatics were treated by the particle mesh Ewald method.[8] Snapshots were collected every 1 ps. Ten MD trajectories excluding the first 2 ns were combined into one. Analyses of the trajectories were performed using the PTRAJ module of the AmberTools1.5 program, and molecular graphics images were produced using VMD.[9]

Definition of the paramagnetic center:

We performed MD simulations of the lanthanum-chelating tag attached to a terminal glucose residue to consider the motion of the tag moiety. The initial structures were built by modifying a previously reported crystal structure of $[Fe(1,2\text{–}(N(CH_2COO)_2)_2C_6H_4\cdot H_2O]^-$, and their torsion angles for the rotatable C–C bond between benzene and the amide group were set to 0° or 180°.

The Antechamber program[10], in combination with the general Amber force field (GAFF),[11] was used to generate parameters for the glucose-attached tag without a lanthanum ion. The charge of the glucose and tag component were assigned by the AM1-BCC method.[12] For the La^{3+} ion, previously reported parameters were used.[13] The topology file for the molecule with the La^{3+} ion was created with the tLeap module. TIP3P waters were added to the solvent layer to ensure a depth of at least 8 Å from any atom. Ten Na^{+} ions and nine Cl^{-} ions were added to neutralize the system. The distance between the La^{3+} ion and the carboxly oxygen or diamine nitrogen atoms of the tag was restrained to be 2.3 or 2.5 Å, respectively. Before MD runs were performed, the entire system was energy minimized by 500 steps of steepest descent followed by 500 steps of conjugate gradient and then heated to 300 K with a 2-fs time step in the NPT ensemble at 1 atm over 50 ps. Production MD simulations were performed for 30 ns at 300 K with a 2-fs time step in the NPT ensemble. The simulations were carried out twice for each initial structure. All bonds involving hydrogen atoms were constrained by the SHAKE algorithm and long-range electrostatics were treated by the particle mesh Ewald method. Snapshots were collected every 1 ps. By averaging the coordinate of the La^{3+} ion over the trajectories except the first 5 ns, we define the position of the paramagnetic center relative to the six-membered ring of glucose.

Tensor determination

Two thousand conformers were extracted from the combined trajectory of the GM3 trisaccharide every 50 ps, and the averaged paramagnetic center was added by aligning each glucose ring. The $\Delta\chi$ tensor for the ensembles incorporating either the Tm^{3+} ion or Tb^{3+} ion was determined by a modified version of Mspin.[14] Every conformer was estimated to contribute equally to the PCSs.

References:

1. J. Wang and X. Qian, *Org. Lett.*, 2006, **8**, 3721-3724.
2. F. Delaglio, S. Grzesiek, G. W. Vuister, G. Zhu, J. Pfeifer and A. Bax, *J. Biomol. NMR*, 1995, **6**, 277-293.
3. T. D. Goddard and D. G. Kneller, *SPARKY 3*, University of California, San Francisco.
4. D. A. Case, T. A. Darden, T. E. Cheatham, C. L. Simmerling, J. Wang, R. E. Duke, R. Luo, R. C. Walker, W. Zhang, K. M. Merz, B. P. Roberts, B. Wang, S. Hayik, A. Roitberg, G. Seabra, I. Kolossvai, K. F. Wong, F. Paesani, J. Vanicek, J. Liu, X. Wu, S. R. Brozell, T. Steinbrecher, H. Gohlke, Q. Cai, X. Ye, J. Wang, M.-J. Hsieh, G. Cui, D. R. Roe, D. H. Mathews, M. G. Seetin, C. Sagui, V. Babin, T. Luchko, S. Gusarov, A. Kovalenko and P. A. Kollman, *Amber11*, (2010) University of California, San Francisco.
5. K. N. Kirschner, A. B. Yongye, S. M. Tschampel, J. González-Outeiriño, C. R. Daniels, B. L. Foley and R. J. Woods, *J. Comput. Chem.*, 2008, **29**, 622-655.
6. H. J. C. Berendsen, J. P. M. Postma, W. F. v. Gunsteren, A. DiNola and J. R. Haak, *J. Chem. Phys.*, 1984, **81**, 3684-3691.
7. J.-P. Ryckaert, G. Ciccotti and H. J. C. Berendsen, *J. Comput. Phys.*, 1977, **23**, 327-341.
8. T. Darden, D. York and L. Pedersen, *J. Chem. Phys.*, 1933, **98**, 10089-10093.
9. W. Humphrey, A. Dalke and K. Schulten, *J. Molec. Graphics*, 1996, **14**, 33-38.
10. J. Wang, W. Wang, P. A. Kollman and D. A. Case, *J. Mol. Graph. Model.*, 2006, **25**, 247-260.
11. J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, *J. Comput. Chem.*, 2004, **25**, 1157-1174.
12. A. Jakalian, B. L. Bush, D. B. Jack and C. I. Bayly, *J. Comput. Chem.*, 2000, **21**, 132-146.
13. M. Baaden, F. Berny, G. Wipff and C. M. J., *J. Phys. Chem. A*, 2000, **104**, 7659.
14. V. M. Sánchez-Pedregal, R. Santamaría-Fernández and A. Navarro-Vázquez, *Org. Lett.*, 2009, **11**, 1471-1474.

Table S1. ^{1}H and ^{13}C chemical shifts of GM3 with the tag complexed with La^{3+}, Tm^{3+}, and Tb^{3+}.

	La^{3+}		Tm^{3+}		Tb^{3+}	
	$\delta^{1}H$/ppm	$\delta^{13}C$/ppm	$\delta^{1}H$/ppm	$\delta^{13}C$/ppm	$\delta^{1}H$/ppm	$\delta^{13}C$/ppm
Glc1	5.14	79.94	2.63	77.11	0.85	75.15
2	3.56	71.53	1.83	69.66	0.08	68.01
3	3.69	75.14	2.58	73.93	1.46	72.80
4	3.69	77.75	2.71	76.63	1.74	75.69
5	3.68	76.65	2.32	75.20	1.20	74.14
6	3.90	59.98	2.86	58.86	2.07	57.98
	3.80	59.97	2.82	58.86	1.96	57.98
Gal 1	4.49	102.70	3.93	102.08	3.31	101.59
2	3.53	69.48	3.11	69.03	2.62	68.57
3	4.04	75.58	3.74	75.24	3.38	74.91
4	3.90	67.60	3.67	67.29	3.36	66.94
5	3.65	75.27	3.31	74.91	2.87	74.50
6	3.68	61.17	3.45	60.90	3.09	60.46
	3.68	61.16	3.45	60.90	2.98	60.47
Neu5Ac 3	2.70	39.78	2.53	39.63	2.27	39.34
3	1.75	39.78	1.55	39.62	1.27	39.34
4	3.62	68.47	3.48	68.29	3.34	68.14
5	3.78	51.80	3.61	51.66	3.41	51.46
6	3.57	72.97	3.42	72.78	3.23	72.60
7	3.52	68.22	3.38	68.02	3.21	67.83
8	3.82	71.88	3.57	71.64	3.29	71.45
9	3.80	62.68	3.59	62.49	3.38	62.28
	3.57	62.69	3.40	62.48	3.20	62.28

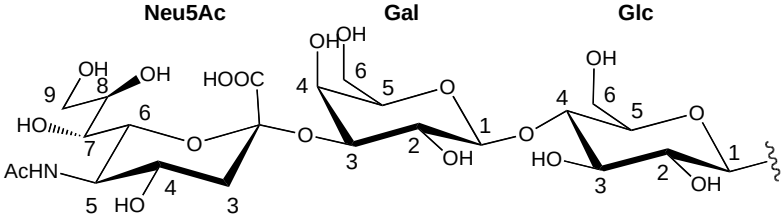


Table S2. *Q* and Δχ values of **7** complexed with Tm^{3+} and Tb^{3+}.

	Tm^{3+}	**Tb^{3+}**
[a]Q	0.05	0.06
[b]q_{ave}	0.08	0.08
[b]q_{high}	0.44	0.58
[b]q_{low}	0.04	0.04
$\Delta\chi_{ax}$ ($\times 10^{-23}$ m^{-3})	8.9	16.1
$\Delta\chi_{rh}$($\times 10^{-23}$ m^{-3})	3.6	3.9
[c]α	28.1	26.6
[c]β	20.1	16.3
[c]γ	-2.4	-6.0
[d]$Q_{selected}$	0.17	0.11

[a] $Q = \mathrm{rms}(\Delta\delta_{calc} - \Delta\delta_{obs})/\mathrm{rms}(\Delta\delta_{obs})$.$\Delta\delta_{calc}$ is given by following equation; $\Delta\delta_{calc} = \sum_{i=1}^{N}\left(p_i \cdot 1/12\pi r_i^3 \cdot [\Delta\chi_{ax}(3\cos^2\vartheta_i - 1) + 3/2 \cdot \Delta\chi_{rh}(\sin^2\vartheta_i \cos 2\varphi_i)]\right)$, where p_i is populations of each structure (set to 0.0005), N is number of each conformers, and (r_i, θ_i, φ_i) defines the position vector for conformer i of the nuclear in polar coordinates with respect to the metal center and principal axis of Δχ tensor.

[b]average, highest and lowest value of q_i. $q_i = \mathrm{rms}(\Delta\delta_{i,calc} - \Delta\delta_{obs})/\mathrm{rms}(\Delta\delta_{obs})$, where $\Delta\delta_{i,calc}$ is back-calculated PCSs for individual conformer as follows: $\Delta\delta_{i,calc} = 1/12\pi r_i^3 \cdot [\Delta\chi_{ax}(3\cos^2\vartheta_i - 1) + 3/2 \cdot \Delta\chi_{rh}(\sin^2\vartheta_i \cos 2\varphi_i)]$.

[c] The Euler angles for principal axis of Δχ tensor.

[d] Q values are for a combination of selected conformers. The torsion angles for these structures were set to averaged values of torsion angles in each conformational cluster populated by more than 5%.Exact torsion values for Φ(Glc–Gal), Ψ(Glc–Gal), Φ(Gal–Neu5Ac), and Ψ(Gal–Neu5Ac) were 43, 1, −175, and −87; 43, 1, −89, and −89; 43, 1, −70, and −87 with a relative incidence estimated at 1:2:2, respectively.

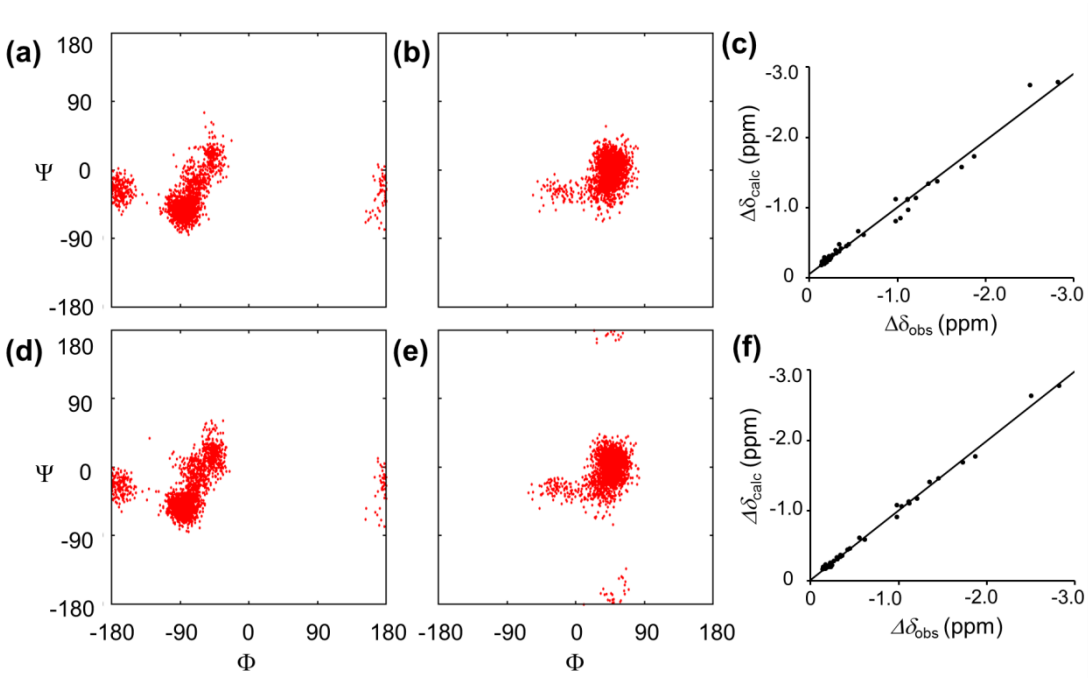


Figure S1. Influence of ensemble bias on the correlation between experimental and calculated PCS data. Scattered plots of torsion angles of (a and d) the Neu5Ac–Gal linkage and (b and e) the Gal–Glc linkage of the ensembles and (c and f) the correlations between the experimentally observed PCS values with Tm^{3+} and back-calculated PCS values. (a, b and c) 2,000 conformers from one trajectories (12 ns) gave *Q* value of 0.10. (d, e and f) 2,000 conformers from all 10 trajectory gave *Q* value of 0.05.