

A rigid lanthanide binding tag to aid NMR assignments of a 70 KDa homodimeric coat protein of human norovirus

Alvaro Mallagaray,^a Gema Domínguez,^b Thomas Peters^{*a} and Javier Pérez-Castells^{*b}

Materials and Methods	pp
1. General remarks	1
2. Protein expression and purification of GII.4 Saga P dimers	2
2.1 Unlabeled samples	2
2.2 U-[²H, ¹⁵N] labelling	2
3. Paramagnetic NMR experiments	3
3.1. General	3
3.2. NMR chelating studies	3
3.2.1 Titration of 4 with DyCl₃	3
3.2.2 PCS observed for the L-fucose moiety of 4	3
3.3 Ligand experiments	5
3.3.1 Aggregation study of 4 in the absence of metal	5
3.3.2 Aggregation study of 4 in the presence of LaCl₃	5
3.3.3 STD NMR experiments	6
3.3.3.1 Calculation of K_D by STD NMR	6
3.3.3.2 Epitope Mapping with STD NMR	7
3.4 Protein NMR experiments & modeling	8
3.4.1 Chemical shift titrations for K_D calculation	8
3.4.2 Binding of methyl α-L-fucopyranoside vs. 4/La³⁺	11
3.4.3 Model of 4 based on available crystal structures	12
3.4.4 Pseudo contact shifts (PCS)	12
3.4.5 Paramagnetic Relaxation Enhancement (PRE)	14
4. Notes and references	16
5. Organic Synthesis	17
5.1 Synthesis of new products	17
Tetraethyl 2,2',2'',2'''-((4-ethynyl-1,2-phenylene)bis(azanetriyl))tetraacetate, 3.	
General Procedure for the preparation of tetraethyl 2,2',2'',2'''-((4-(1-(glycosyl)-1H-1,2,3-triazol-4-yl)-1,2-phenylene)bis(azanetriyl))tetraacetates	
Tetraethyl 2,2',2'',2'''-((4-(1-(2,3,4-tri-O-acetyl-α-L-fucosyl)-1H-1,2,3-triazol-4-yl)-1,2-phenylene) bis(azanetriyl))tetraacetate	
Tetraethyl 2,2',2'',2'''-((4-(1-(2,3,4-tri-O-acetyl-β-L-fucosyl)-1H-1,2,3-triazol-4-yl)-1,2-phenylene)bis (azanetriyl))tetraacetate	
2,2',2'',2'''-((4-(1-((2R,3S,4R,5S,6S)-3,4,5-trihydroxy-6-methyltetrahydro-2H-pyran-2-yl)-1H-1,2,3-triazol-4-yl)-1,2-phenylene)bis(azanetriyl))tetraacetic acid, 4	
5.2 Spectra of new products	19

1. General remarks

Preparative high performance liquid chromatography (HPLC) was performed using a Merck Hitachi D-7000 HPLC system equipped with a Merck-Hitachi L-7400 UV detector. The following mobile phases were used for separation: 0.1 % trifluoroacetic acid (TFA) (v/v) in MP-H₂O (solvent A) and 0.1 % TFA in acetonitrile (solvent B). HPLC was performed using a reversed-phase (RP) column Machery-Nagel VP 250/21 Nucleosil 100-7 C18 (250 x 21 mm).

¹H NMR and ¹³C NMR routine spectra were acquired on a Bruker Avance 400 MHz, spectrometer. Chemical shifts (TM) are in parts per million relative to tetramethylsilane at 0.00 ppm. IR spectra were determined by a FT-IR Perkin-Elmer 2000 spectrometer. TLC analyses

were performed on commercial aluminium sheets bearing 0.25 mm layer of Merck Silica gel 60F254. Silica gel Acros Organics 0.035-0.070 mm, 60 Å was used for column chromatography. Elemental analyses were carried out at the Elemental Analysis Center of the Complutense University of Madrid, using a Perkin Elmer 2400 CHN. Electrospray ionisation mass spectrometry (ESI-MS) analyse was obtained on an Esquire 3000 (Bruker) spectrometer. Toluene was refluxed over calcium hydride and *i*-Pr₂NH over KOH. All reagents were bought from Aldrich or Acros Organics.

2. Protein expression and purification of GII.4 Saga P dimers

Saga GII.4 P dimers were expressed and purified as reported elsewhere.¹ Briefly:

2.1 Unlabeled samples

GII.4 Saga P domains (residues 224 to 538) [Genbank accession number AB447457] were overexpressed in *Escherichia coli*. The codon-optimized gene was cloned into a modified expression vector (pMalc2X) and transformed into *E. coli* BL21 (DE3) cells (Novogen, Darmstadt, Germany) for protein expression. Transformed cells were grown for 3 h at 37 °C in modified Terrific Broth (TB) medium (12 g tryptone, 24 g yeast extract and 40 mL glycerol per liter culture) supplemented with M9 minimal medium components (0.5 g of NaCl, 3.3 g of KH₂PO₄, 16.6 g of Na₂HPO₄·12H₂O, 1 g of NH₄Cl, 1 mL of 1 M MgSO₄, 1 mL of 0.1 M CaCl₂ and 0.2 % glucose per liter culture), 0.4 % casamino acids and 100 µg/mL ampicillin. Overexpression was induced with 1 mM isopropylthiolactoside at OD₆₀₀ between 1.2–1.5, and the incubation was continued at 17 °C for 44 to 48 h. To maintain the antibiotic pressure constant, 100 µg/mL ampicillin were added after the first 24 h. Cells were harvested by centrifugation at 9000 g for 15 min and the bacterial pellet was resuspended in 25 mM PBS buffer (pH 7.4). To the suspension 4 µl of 1 mg/mL Aprotinin and Leupeptin solutions (Roth), 25 µl of a 10 mg/mL chicken egg white lysozyme solution (Novagen) and 0.1 µl of a 25 U/µl Benzonase solution (Novagen) per gram of wet pellet were added, the suspension was incubated for 30 min at 4 °C and passed twice through a French Pressure cell at 14,000 psi followed by an ultracentrifugation at 125,000 g for 1 h. The protein was purified following a published protocol.² Briefly, the His-tagged P domain protein was purified from the supernatant by a Ni column (Qiagen) and digested with HRV-3C protease (Novagen) overnight at 4 °C. Cleaved P domains were separated on the Ni column and dialyzed in gel filtration buffer (25 mM Tris-HCl, pH 7.4 and 300 mM NaCl) overnight at 4 °C. The P domains were purified by size-exclusion chromatography, concentrated to 2 to 5 mg/mL and stored in gel filtration buffer at 4 °C. Average yield: 130 mg GII.4 Saga P dimers per litre culture.

2.2 U-[²H, ¹⁵N] labeling

The U-[²H, ¹⁵N] GII.4 Saga P domains were expressed following a modified version of a previously reported protocol.³ Briefly, *E. coli* BL21(DE3) containing the codon-optimized P domain gene were grown in 40 mL of supplemented TB medium at 37 °C for 16 h as described above. A volume containing cells enough to give an OD₆₀₀ of 0.1 in a 30 mL culture was spun down (1,200 g at room temperature), the supernatant was discarded and the pellet was resuspended in 30 mL of sterile M9/D₂O minimal medium.¹ The culture was incubated at 37 °C until the cells reach an OD₆₀₀ of about 0.4-0.6. Like before, a culture volume for a final OD₆₀₀ of 0.1 in 100 mL was spun down, the supernatant was discarded, the pellet resuspended in 100 mL of sterile M9/D₂O minimal medium and incubated at 37 °C until OD₆₀₀ of 0.4-0.6. The culture was diluted to a volume of 200 mL with sterile M9/D₂O minimal medium and grown at 37 °C until OD₆₀₀ was 0.4-0.5. The culture was diluted one last time up to 1 L with sterile M9/D₂O minimal medium and incubated at 37 °C until the OD₆₀₀ was 0.15-0.2. The culture was cooled down to 17 °C for 15 min and the overexpression was induced with 1 mM isopropylthiolactoside in D₂O. The incubation was continued at 17 °C until an OD₆₀₀ of 1.9-2.1 was reached. To maintain the antibiotic pressure constant, 100 µg/mL ampicillin were added

every 24 h. Cells were harvested and the P protein was purified as described above. Average yield: 17.5 mg U-[^2H , ^{15}N] GII.4 Saga P dimers per litre culture.

3. Paramagnetic NMR experiments

3.1. General

NMR experiments were recorded on a Bruker AV III 500 MHz NMR spectrometer equipped with a TCI cryogenic probe or on a Bruker AV I 700 MHz NMR spectrometer equipped with a TXI cryogenic probe. Spectra were recorded at 298 K. The NMR data were processed and analyzed with Topspin 3.1 if not otherwise stated.

3.2. NMR chelating studies

3.2.1 Titration of **4** with DyCl_3

A sample containing 1 mM of **4** in Tris-*d11*-HCl 25 mM pH 7.4, 0.3 M NaCl and 100 μM DSS-*d6* as internal reference in D_2O was titrated with increasing amounts of DyCl_3 . The concentration of **4** was kept constant during the whole titration. Only two sets of ^1H NMR signals were observed at 75% DyCl_3 , clearly indicating slow exchange in the NMR time-scale, and therefore thight binding affinity of the PhDTA moiety towards the metals (Figure S1).

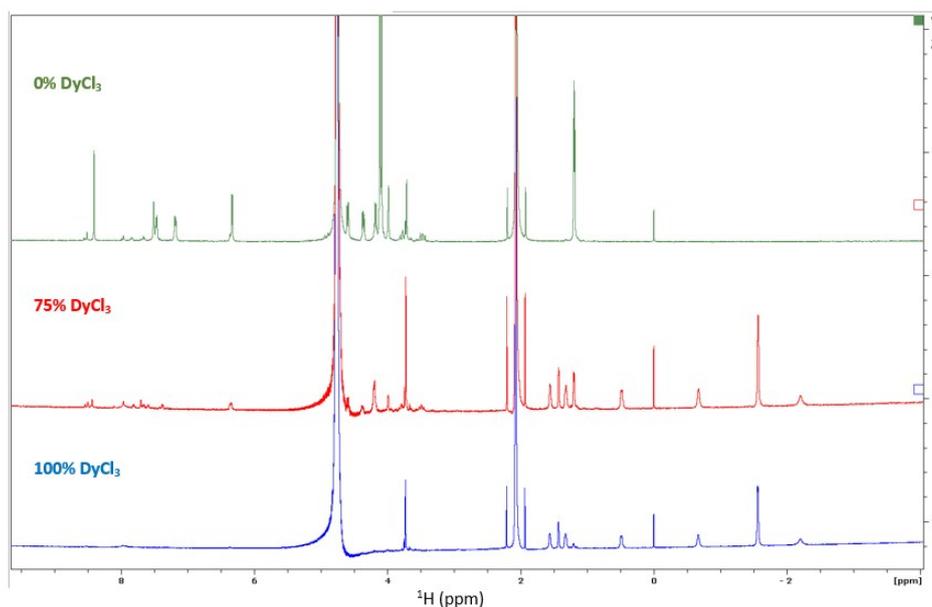


Figure S1: Series of ^1H NMR spectra of **4** in the presence of 0%, 75% and 100% ratios of DyCl_3 (relative to **4**). Only two set of peaks were observed. 500 MHz.

3.2.2 PCS observed for the L-fucose moiety of **4**

A standard phase sensitive ^1H , ^{13}C -HSQC pulse sequence from Bruker pulse program library was used with 16 dummy scans, 10 scans (3h) and a FID resolution of 7.32 Hz and 68.73 Hz for the direct and the indirect dimensions, respectively. The frequency offset and spectral width were 2.0 ppm and 15.0 ppm for the ^1H dimension and 60.0 ppm and 140.0 ppm for the ^{13}C dimension. Samples were prepared at 1 mM **4** with 1 mM of either LaCl_3 or DyCl_3 in Tris-*d11*-HCl 25 mM pH 7.4, 0.3 M NaCl and 100 μM DSS-*d6* as internal reference in D_2O . Large PCS were observed for every ^1H and ^{13}C nucleus of the L-fucose moiety (Figure S2 and Table S1).

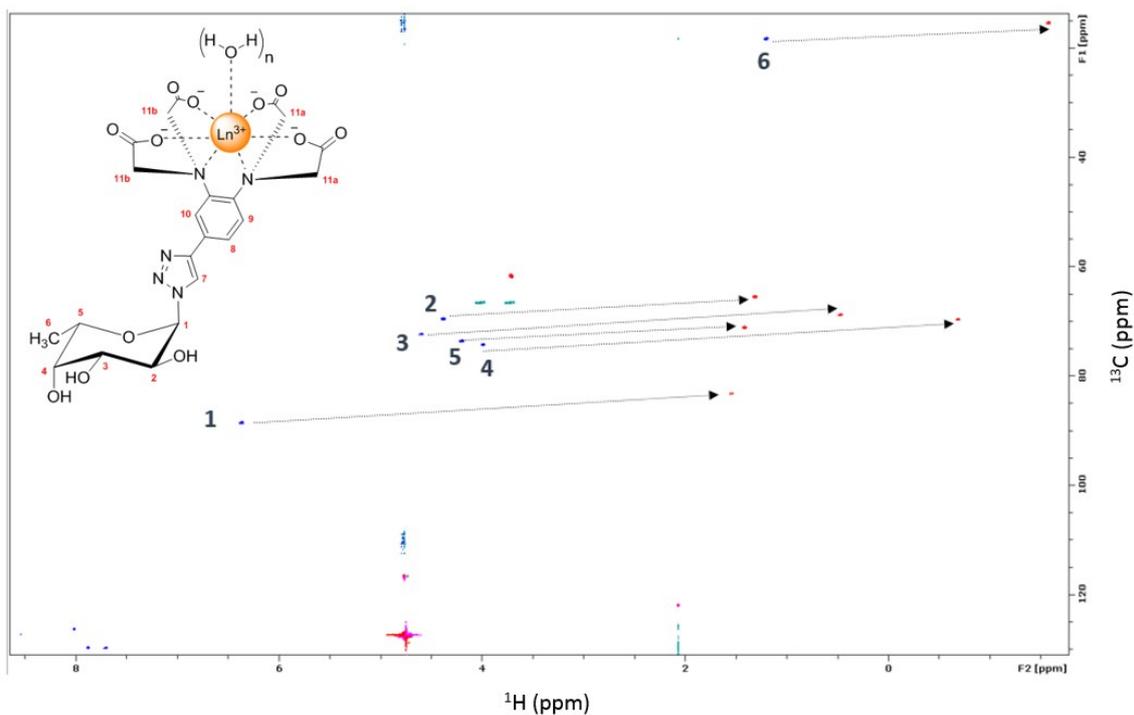


Figure S2: Superimposition of ^1H , ^{13}C HSQC spectra of **4** acquired under isotropic (blue, LaCl_3) and anisotropic (red, DyCl_3) conditions. 500 MHz.

Table 1S: PCS observed over the L-fucose moiety of **4** in the presence of DyCl_3 .

Nucleus	PCS (ppm)
H-1	4.82
H-2	3.07
H-3	4.12
H-4	4.68
H-5	2.78
H-6	2.78
C-1	5.33
C-2	4.07
C-3	3.47
C-4	4.53
C-5	2.38
C-6	3.02

3.3 Ligand experiments

3.3.1 Aggregation study of 4 in the absence of metal

A sample of 30 μM saga P dimers in Tris-*d*11-HCl pH 7.4 containing 0.3 M NaCl and 100 μM DSS-*d*6 in D_2O was titrated with 4. The protein concentration was kept constant during the whole experiment. At each titration point one ^1H NMR spectrum was acquired. The concentration of 4 ranged from 0.3 to 4 mM. Only the signals corresponding to the PhDTA and the triazole moiety experienced shifts, indicating aggregation through the chelating unit (figure S3).

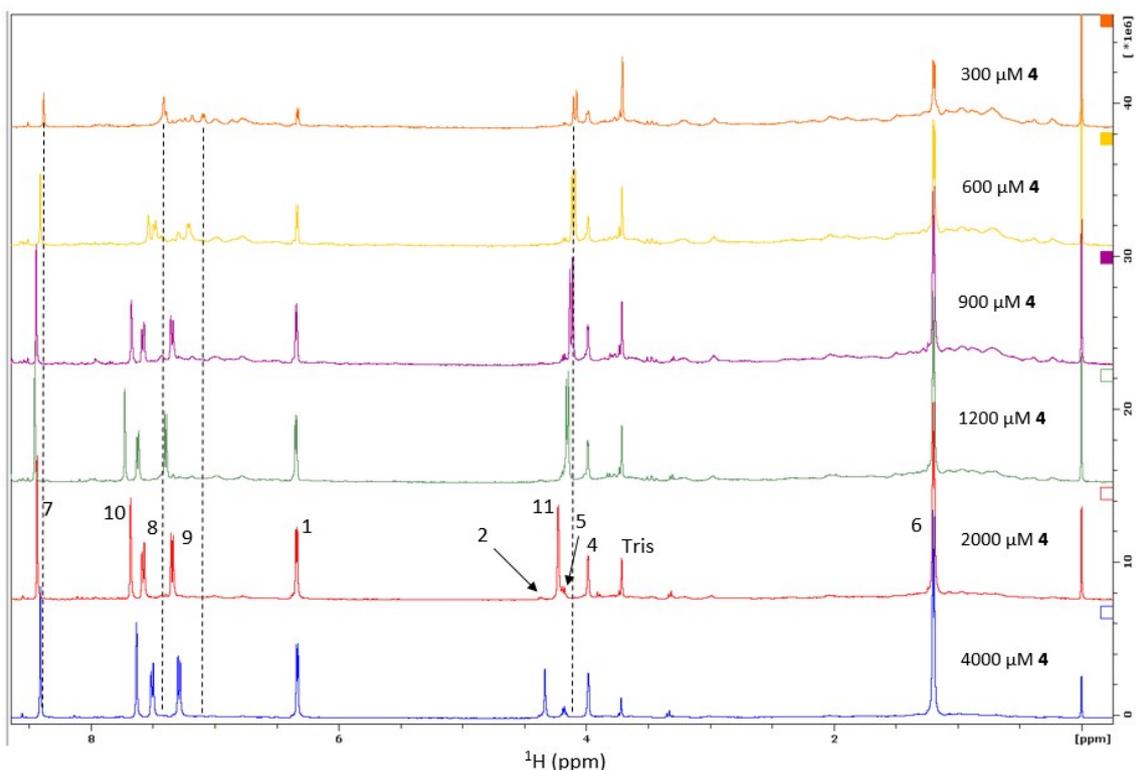


Figure S3: Series of ^1H spectra acquired in the presence of 30 μM SAGA P dimers and increasing concentrations of 4. No peak shifting was observed for the protons on the L-fucose moiety, indicating aggregation through the chelating tag (500 MHz).

3.3.2 Aggregation study of 4 in the presence of LaCl_3

A sample of 30 μM saga P dimers in Tris-*d*11-HCl pH 7.4 containing 0.3 M NaCl and 100 μM DSS-*d*6 in D_2O was titrated with 4/ La^{3+} . The protein concentration was kept constant during the whole experiment. At each titration point one ^1H NMR spectrum and one STD NMR experiment (see below) were acquired. The concentration of 4/ La^{3+} ranged from 0.3 to 4 mM. No signals shifted up to 2 mM 4/ La^{3+} . However, at 3 mM 4/ La^{3+} aggregation occurs with loss of metal ion, as can be observed in the new set of signals observed in figure S4.

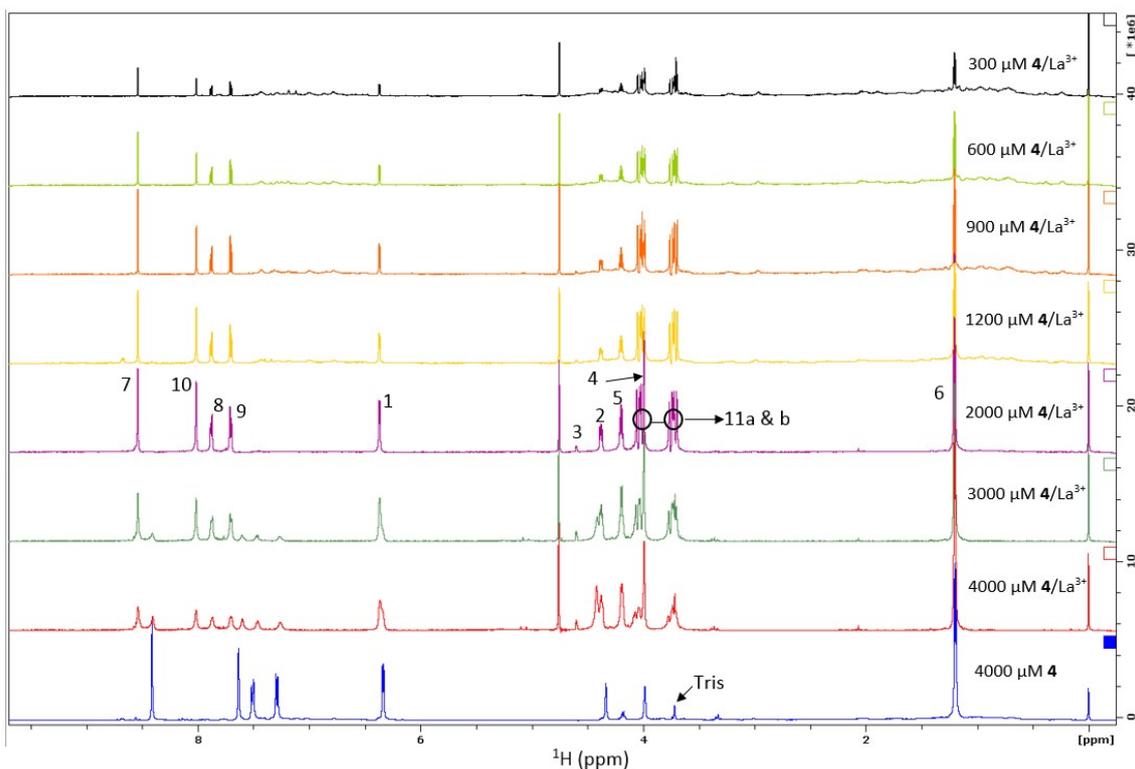


Figure S4: Series of ^1H spectra acquired in the presence of $30\ \mu\text{M}$ SAGA P dimers and increasing concentrations of $4/\text{La}^{3+}$. A new set of signal corresponding to **4** without metal can be observed over $2\ \text{mM}$ $4/\text{La}^{3+}$ (500 MHz).

3.3.3 STD NMR experiments

A train of 50 ms Gaussian-shaped radio frequency pulses separated by 1 ms for a total duration of 2 s and a power level of 40 db was used for protein irradiation. Unwanted protein signals were suppressed *via* a 30 ms spinlock filter, and the water signal was suppressed using an excitation sculpting sequence with gradients. The acquisition time was set to 2.34 s with an additional relaxation delay of 6 s. On resonance was selected at $-4\ \text{ppm}$ for 500 MHz and $-5\ \text{ppm}$ for 700 MHz. The off resonance was kept constant at 200 ppm. The number of scans was ranging from 2.6k to 2k.

3.3.3.1 Calculation of K_D by STD NMR

The STD enhancements were expressed as the STD amplification factor (AF), described by equation (S1):

$$STD\ AF = \frac{I_0 - I_{sat}}{I_0} \times \text{ligand excess} \quad (\text{S1})$$

With I_0 and I_{sat} being the signal intensities in the off- and on-resonance spectra, respectively. Only points up to $900\ \mu\text{M}$ $4/\text{La}^{3+}$ were plotted due to direct irradiation at higher concentrations. The curve was fitted using ORIGIN 2015 to equation (S2):

$$STD\ AF([L]) = \frac{[L]}{K_D + [L]} \times STD\ AF_{max} \quad (\text{S2})$$

Being $[L]$ the total ligand concentration, K_D the dissociation constant and $STD\ AF_{max}$ the maximum STD AF. The dissociation constant was calculated from figure S5 as an estimate, being $K_D = 0.68 \pm 0.12\ \text{mM}$ with a $R^2/\chi^2 = 0.9976/0.00037$.

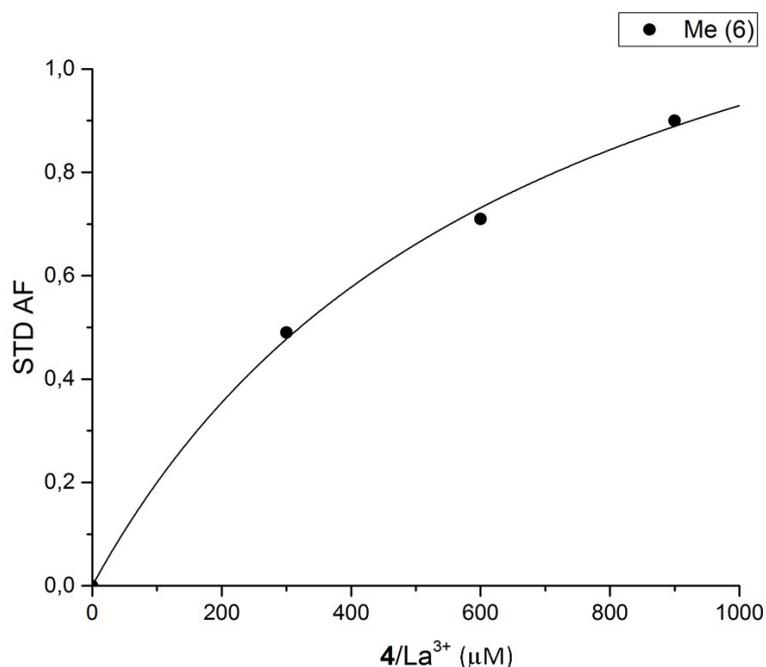


Figure S5: STD NMR titration curve for 4/La³⁺. The concentration of P dimers was 30 μM. Shown is the STD amplification factor as a function of ligand concentration.

3.3.3.2 Epitope Mapping with STD NMR

STD NMR epitope mapping was carried out with a constant saturation time of 2s. Epitopes were calculated by setting the largest STD value at 100%. The result of the epitope mapping at 900 μM 4/La³⁺ concentration is shown in figure S6.

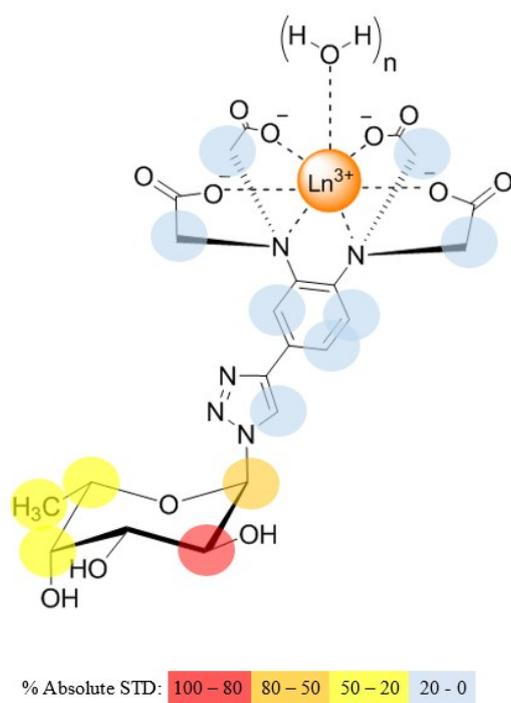


Figure S6: Binding epitopes for $4/\text{La}^{3+}$. It is clear that the fucose moiety is mediating the key contacts with the binding pockets. Proton H-3 was not accessible to the determination of STD values due to signal overlap.

3.4 Protein NMR experiments

A standard phase sensitive $^1\text{H}, ^{15}\text{N}$ -TROSY-HSQC pulse sequence from the Bruker pulse program library was used with 16 dummy scans, 40 scans (3h and 25 min) and a FID resolution of 7.82 Hz and 23.74 Hz for the direct and the indirect dimensions, respectively. The frequency offset and spectral width were 4.68 ppm and 16.0 ppm for the ^1H dimension and 130.0 ppm and 60.0 ppm for the ^{15}N dimension. Spectra were processed using Topspin 3.1 and analyzed with the CCPNMR 2.4.0 software suite.

3.4.1 Chemical shift titrations for K_D calculation

A sample containing 300 μM U- $^{2}\text{H}, ^{15}\text{N}$] GII.4 Saga P dimers, 25 mM Tris-HCl (pH 7.3), 0.3 M NaCl, 100 μM DSS-*d6* and 10% D_2O was titrated with $4/\text{La}^{3+}$ at the following concentrations: 0, 0.25, 0.5, 1, 1.5, 1.75, 2, 2.25 and 2.5 mM. At every step one $^1\text{H}, ^{15}\text{N}$ -HSQC-TROSY spectrum was acquired. 242 cross peaks out of 310 were unambiguously monitored during the whole titration. Average Euclidian distances⁴ (d) were calculated for each cross peak at each titration point. ^{15}N scaling factor (α) was measured from the spectra as 0.14. In order to discriminate meaningful data from noise, only cross peaks with $d \geq$ standard deviation at 2.25 mM were selected and plotted d against ligand concentration (figure S7). These conditions were fulfilled by 72 cross peaks.

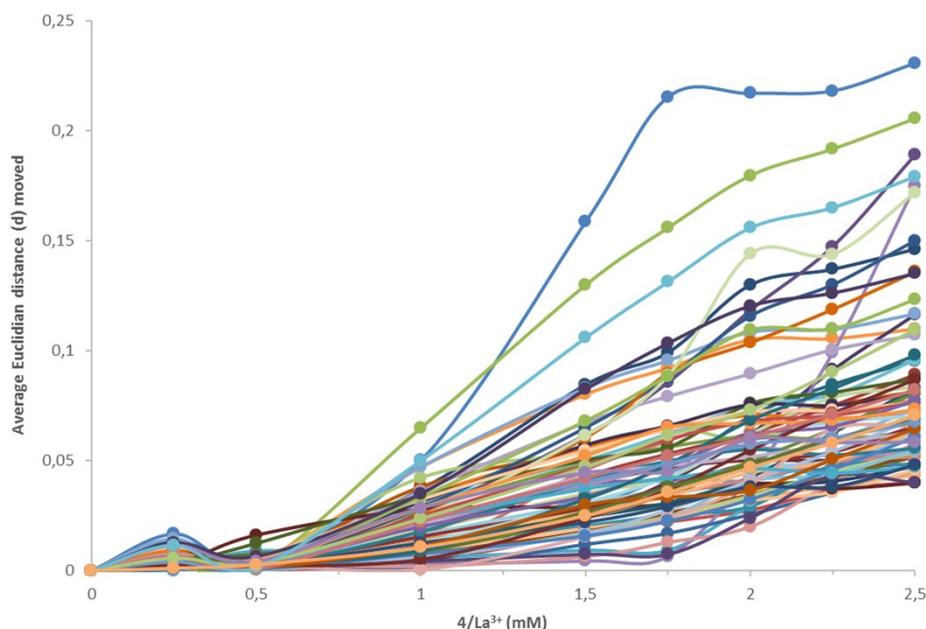


Figure S7: Average Euclidian distances calculated from the chemical shifts perturbations of individual peaks of the $^1\text{H}, ^{15}\text{N}$ -TROSY-HSQC spectra of GII.4 Saga P dimers as a function of increasing amounts of $4/\text{La}^{3+}$. Only the 72 cross-peaks with a $d \geq$ standard deviation at 2.25 mM $4/\text{La}^{3+}$ are shown. The curves shown only interpolate between data points and represent no fitted curves.

It has been described that GII.4 Saga P-dimers exhibit at least 4 binding sites for HBGAs with strong cooperativity, leading to a step-wise binding.¹ The binding curves will therefore reflect these steps. From a visual inspection it is easy to identify a first binding event which is almost saturated over 2 mM $4/\text{La}^{3+}$, and a second binding event which starts over 1.75 mM $4/\text{La}^{3+}$. Based on that, the curves were divided in three groups: 32 cross peaks clearly exhibited a first binding event saturated over 2 mM $4/\text{La}^{3+}$ (figure S8, a). Four cross peaks were only affected by the second binding event, which starts over 1.75 mM $4/\text{La}^{3+}$ (figure S8, b). Finally, 36 cross peaks reflect an ambiguous combination of both binding events (figure S8, c).

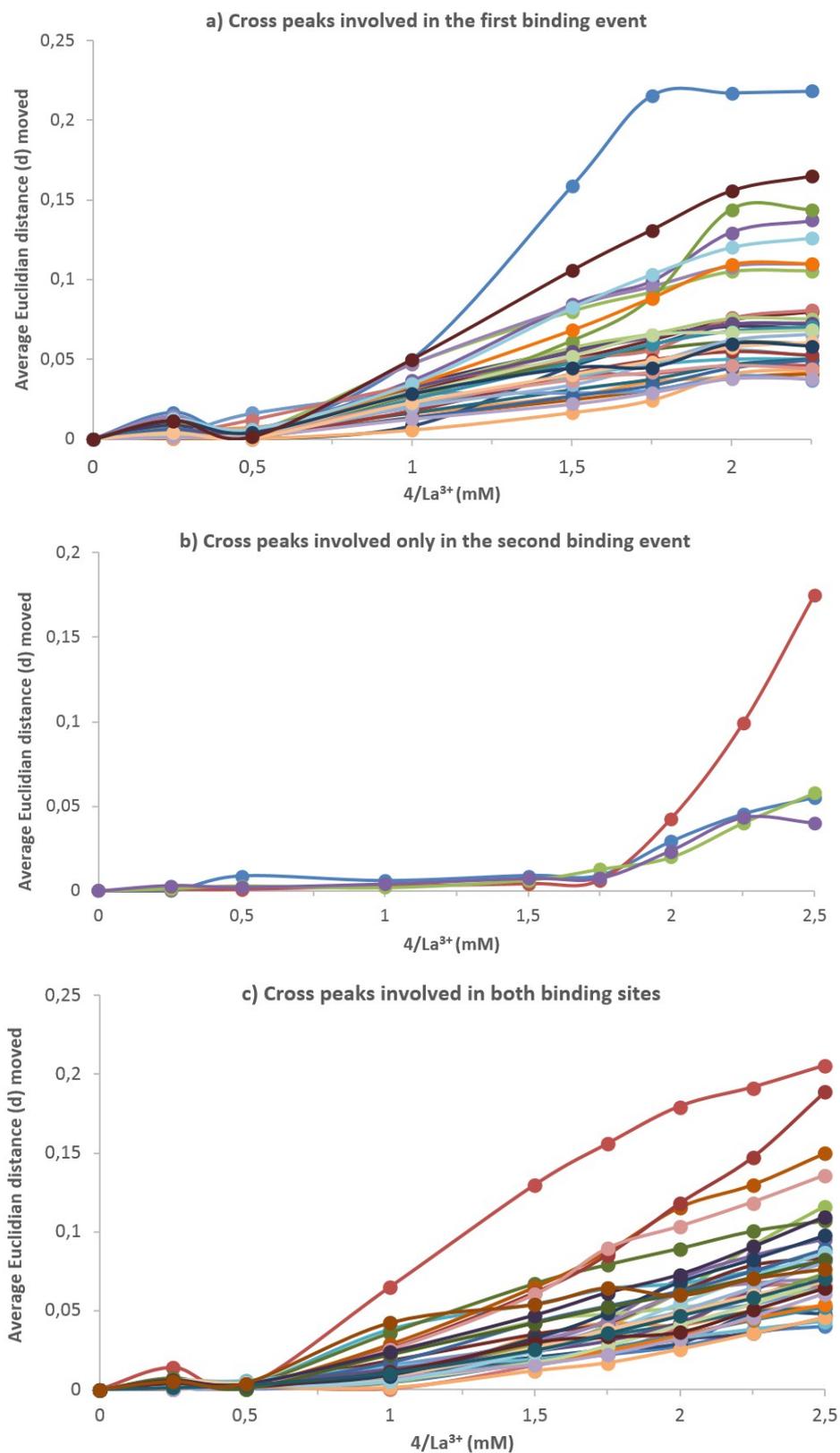


Figure S8: Average Euclidian distances calculated from the chemical shifts perturbations of individual peaks of the $^1H, ^{15}N$ -TROSY-HSQC spectra of GII.4 Saga P dimers as a function of increasing amounts of $4/La^{3+}$. Cross peaks were grouped as: a) involved only in the first binding event, b) involved only in the second binding event and c) involved in both binding events. The curves shown only interpolate between data points and represent no fitted curves.

The shifts corresponding to the 32 cross peaks involved only in the first binding event were subjected to global data analysis using equation (S3) and Origin 2015.

$$d_{obs} = \frac{[L]^h}{(K_D)^h + [L]^h} \times d_{max} \quad (S3)$$

Being d_{obs} the average Euclidian distance observed, $[L]$ the free ligand concentration, K_D the dissociation constant, h the Hill coefficient and d_{max} the maximum average Euclidian distance. The dissociation constant was calculated as $K_D = 0.94 \pm 0.04$ mM with a $R^2/\chi^2 = 0.9878/0.00003$. An overview of the fitted curves is shown in figure S9.

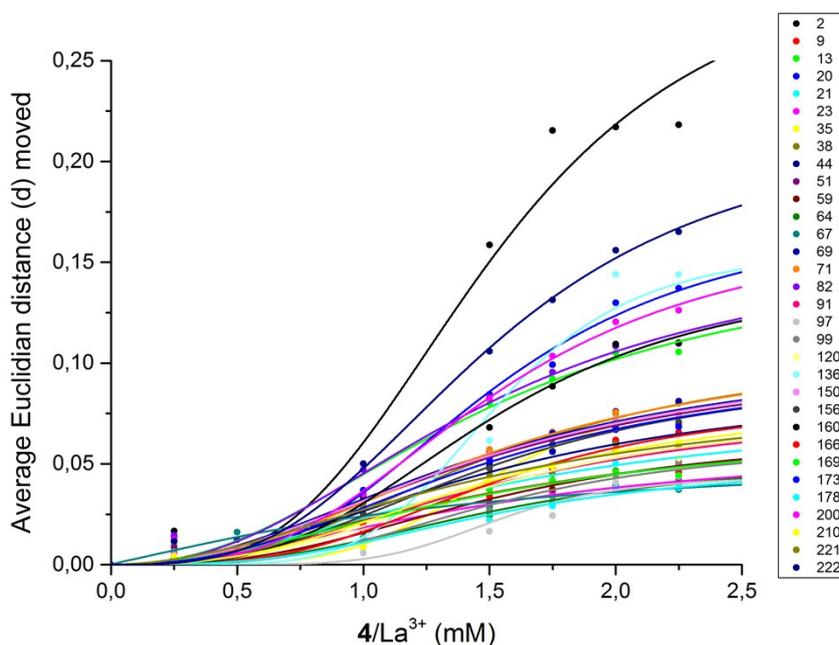


Figure S9: Average Euclidian distances of individual peaks from the $^1\text{H}, ^{15}\text{N}$ -TROSY-HSQC spectra of GII.4 Saga P dimers as a function of increasing amounts of $4/\text{La}^{3+}$. Curves were obtained from fitting equation (S3) to the average Euclidian distances of the 32 cross peaks exhibiting shifts only for the first binding event. Cross peaks are numbered since no assignment is available yet.

3.4.2 Binding of methyl α -L-fucopyranoside vs. $4/\text{La}^{3+}$

A sample containing 300 μM U- ^{2}H , ^{15}N] GII.4 Saga P dimers, 25 mM Tris-HCl (pH 7.3), 0.3 M NaCl, 100 μM DSS-*d6* and 10% D_2O was titrated with methyl α -L-fucopyranoside at the following concentrations: 0, 0.5, 1 and 2 mM. At every step one ^1H , ^{15}N -HSQC-TROSY spectrum was acquired. A superposition with the ^1H , ^{15}N -HSQC-TROSY spectra acquired at 0, 0.5, 1 and 1.25 mM $4/\text{La}^{3+}$ concentrations is shown in figure S10.

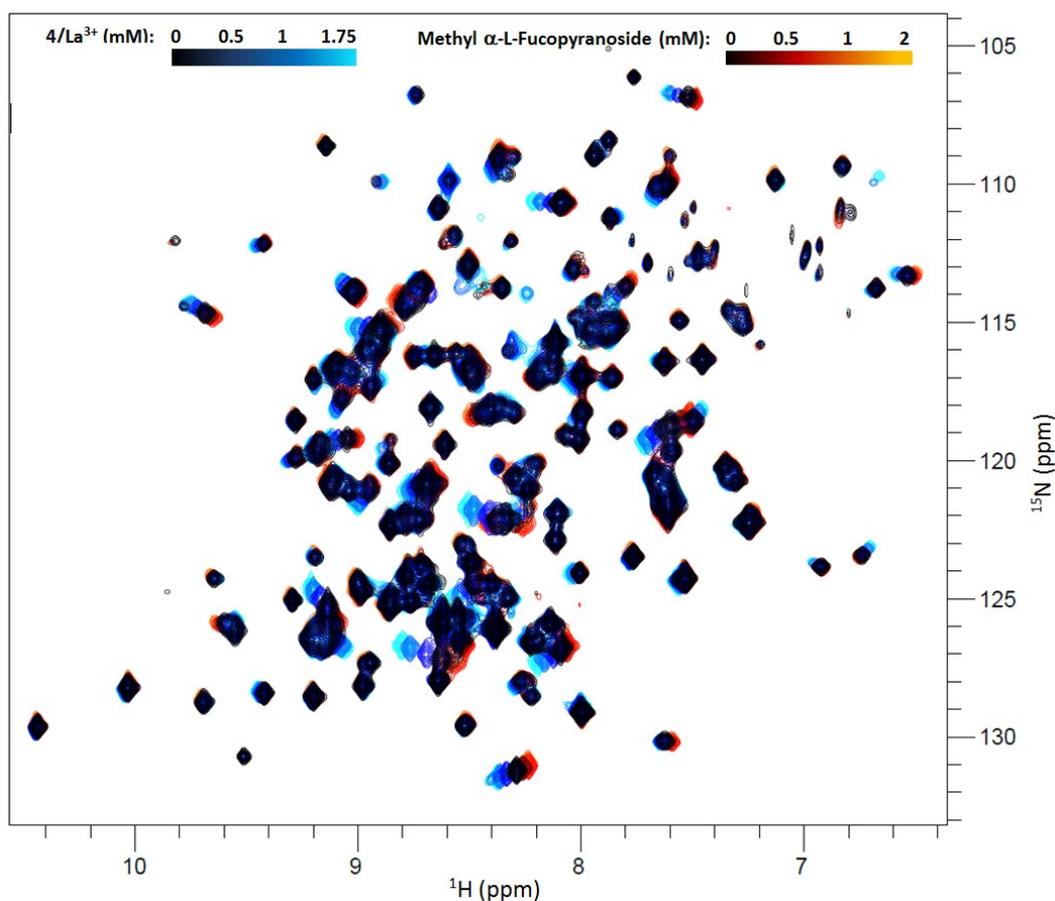


Figure S10: Superposition of ^1H , ^{15}N -TROSY-HSQC spectra of U- ^{2}H , ^{15}N] GII.4 Saga P dimers showing the shifts observed during the titration of $4/\text{La}^{3+}$ (blue palette) and methyl α -L-fucopyranoside (orange palette). Titration points: 0, 0.5, 1 and 1.75 mM for $4/\text{La}^{3+}$ and 0, 0.5, 1 and 2 mM for methyl α -L-fucopyranoside. The same peaks are affected with both ligands, which strongly suggests binding to the same site. The difference in distance and direction of the chemical shift perturbations observed for $4/\text{La}^{3+}$ and methyl α -L-fucopyranoside is due to the presence of the aromatic rings attached to the fucose moiety. 500 MHz.

3.4.3 Model of 4 based on available crystal structures

To estimate the position of the paramagnetic metal during the binding event, a basic model based on crystal structures was constructed. A crystal structure of GII.4 Saga P dimers in complex with A disaccharide (PDB code: 4X07) was used for modeling the α -L-fucose moiety in the binding pocket.⁵ The PhDTA chelating unit was modeled based on x-ray coordinates of PhDTA in complex with Fe^{3+} .⁶ Rotation around the bond between the triazole ring and the benzene moiety is possible. Thus, in order to calculate distances the metal ion coordinates were defined as the average between all possible metal positions.

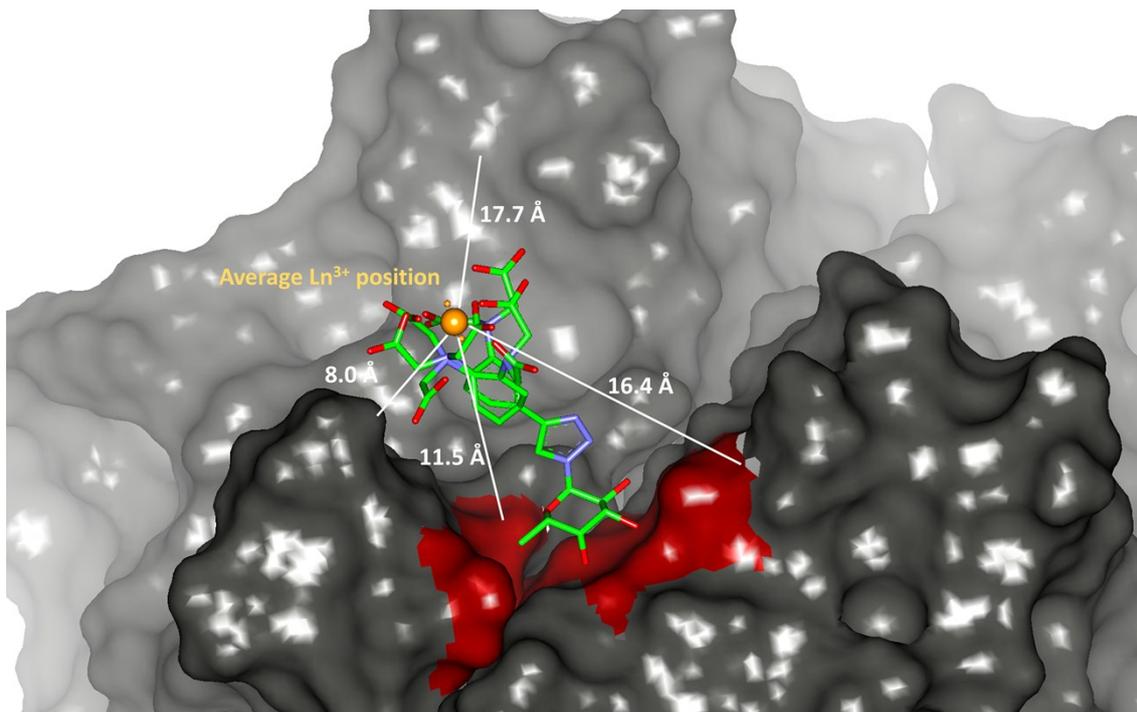


Figure S11: Model of 4 in complex with a paramagnetic metal in the binding site of GII.4 Saga P dimers showing the backbone H^{N} for amino acids in the binding pocket (red) at the closest and farthest distance (Gly 443, 11.5 Å and Asp 374, 16.4 Å, respectively) relative to the metal ion. The distances to the two closest loop structures are also depicted.

3.4.4 Pseudo contact shifts (PCS)

A sample containing 300 μM U- ^{2}H , ^{15}N] GII.4 Saga P dimers, 25 mM Tris-HCl (pH 7.3), 0.3 M NaCl, 100 μM DSS-*d6* and 10% D_2O was titrated at 0.25, 0.5, 1, 1.5 and 1.75 mM 4/ Dy^{3+} concentrations. At every step one ^1H , ^{15}N -HSQC-TROSY spectrum was acquired. For the PCS analysis, isotropic (4/ La^{3+}) and anisotropic (4/ Dy^{3+}) ^1H , ^{15}N -HSQC-TROSY spectra at the same ligand concentration were superimposed and compared. At the highest ligand concentration (1.75 mM 4/metal) 43 cross peaks were broadened beyond detection due to PRE, whilst for the remaining 199 cross peaks both positive and negative PCS were observed. A schematic visualization can be seen in figure S11.

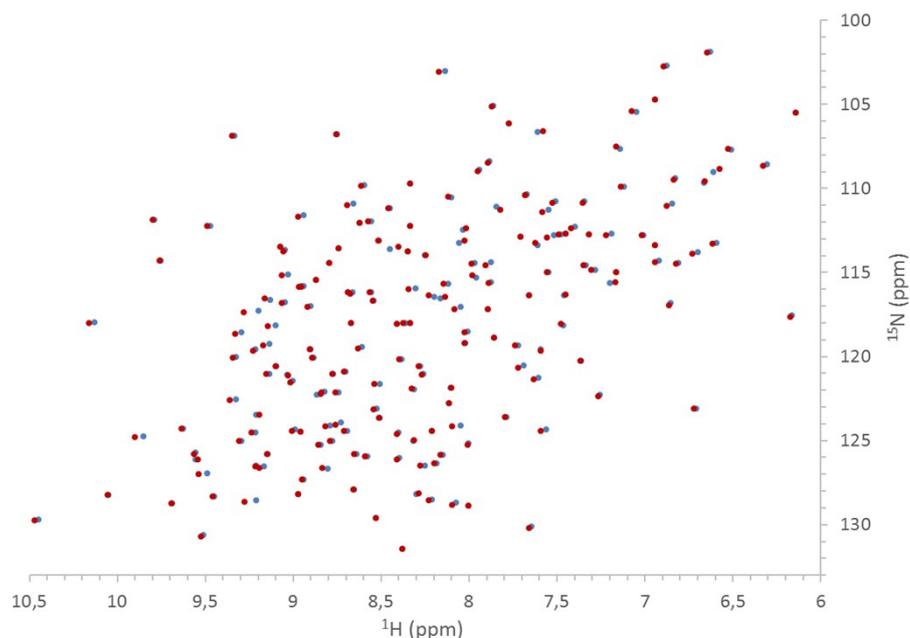


Figure S12: Schematic representation showing the position of the cross peaks extracted from $^1\text{H}, ^{15}\text{N}$ -HSQC-TROSY spectra of U- $[^2\text{H}, ^{15}\text{N}]$ GII.4 Saga P dimers in the presence of a diamagnetic (blue, 1.75 mM **4**/ La^{3+}) and a paramagnetic (red, 1.75 mM **4**/ Dy^{3+}) sample. The inaccuracy observed in the ^{15}N dimension is due to low FID resolution.

Interestingly, the cross peaks broadened beyond detection can be observed at lower **4**/metal concentrations, and can therefore be used for tensor calculation. Figure S12 shows a schematic representation of 24 cross peaks at 1 mM **4**/ Dy^{3+} concentration not visible at higher **4**/ Dy^{3+} concentrations.

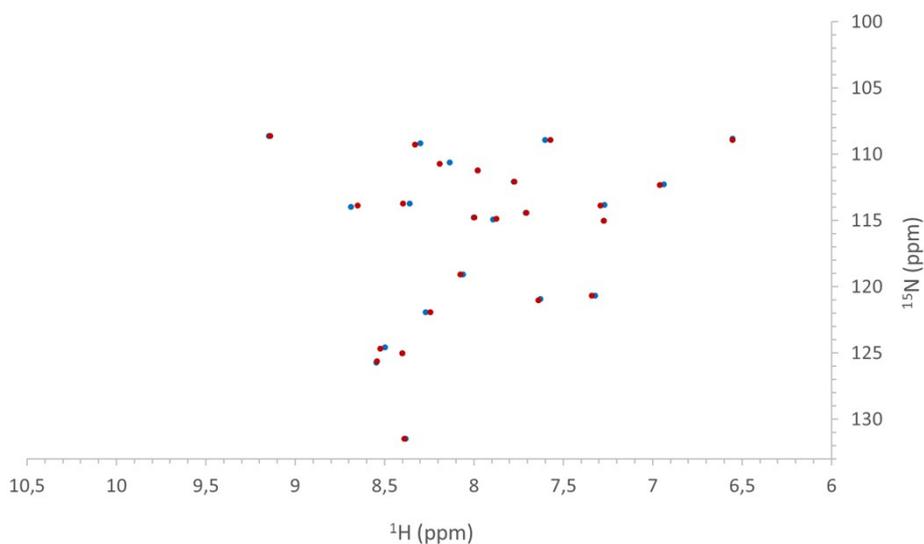


Figure S13: Schematic representation showing the position of the cross peaks extracted from $^1\text{H}, ^{15}\text{N}$ -HSQC-TROSY spectra of U- $[^2\text{H}, ^{15}\text{N}]$ GII.4 Saga P dimers in the presence of a diamagnetic (blue, 1 mM **4**/ La^{3+}) and a paramagnetic (red, 1 mM **4**/ Dy^{3+}) sample. Only the cross peaks broadened beyond detection at **4**/ Dy^{3+} 1.75 mM concentration are shown. The inaccuracy observed in the ^{15}N dimension is due to a low FID resolution.

3.4.5 Paramagnetic Relaxation Enhancement (PRE)

In order to define groups of cross-peaks corresponding to residues located in the binding pocket, a $^1\text{H},^{15}\text{N}$ -HSQC-TROSY spectrum of a sample containing 300 μM U- $^{2}\text{H},^{15}\text{N}$] GII.4 Saga P dimers in 25 mM Tris-HCl (pH 7.3), 0.3 M NaCl, 100 μM DSS-*d*6 and 10% D_2O and 2 mM of methyl α -L-fucopyranoside was acquired (fig S14, a). Next, another sample containing 300 μM U- $^{2}\text{H},^{15}\text{N}$] GII.4 Saga P dimers (same conditions as above) was titrated at 0.1, 0.25 and 0.5 mM $4/\text{Gd}^{3+}$ concentrations. At every step one $^1\text{H},^{15}\text{N}$ -HSQC-TROSY spectrum was acquired. For the PRE analysis, $4/\text{La}^{3+}$ and $4/\text{Gd}^{3+}$ $^1\text{H},^{15}\text{N}$ -HSQC-TROSY spectra at the same ligand concentration were superimposed and compared (figure S14, b-d).

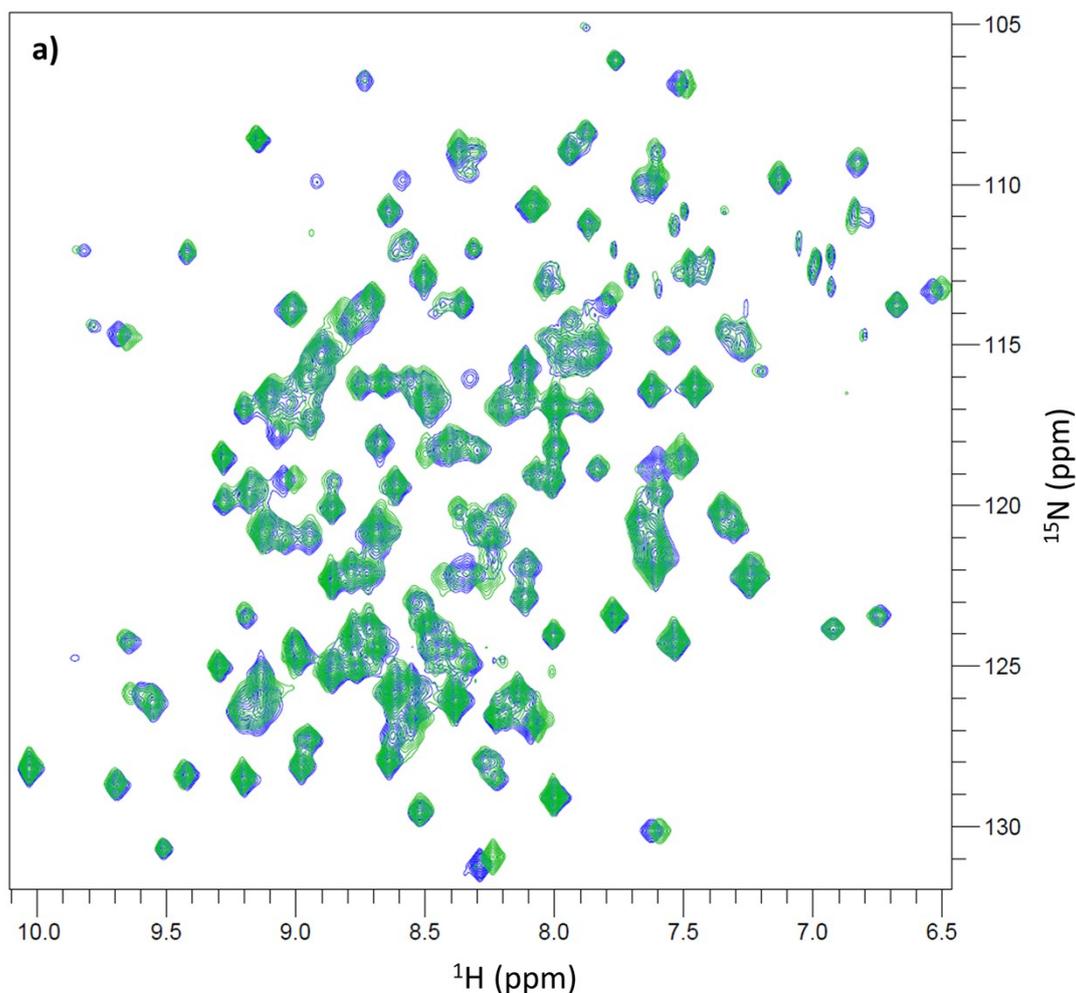
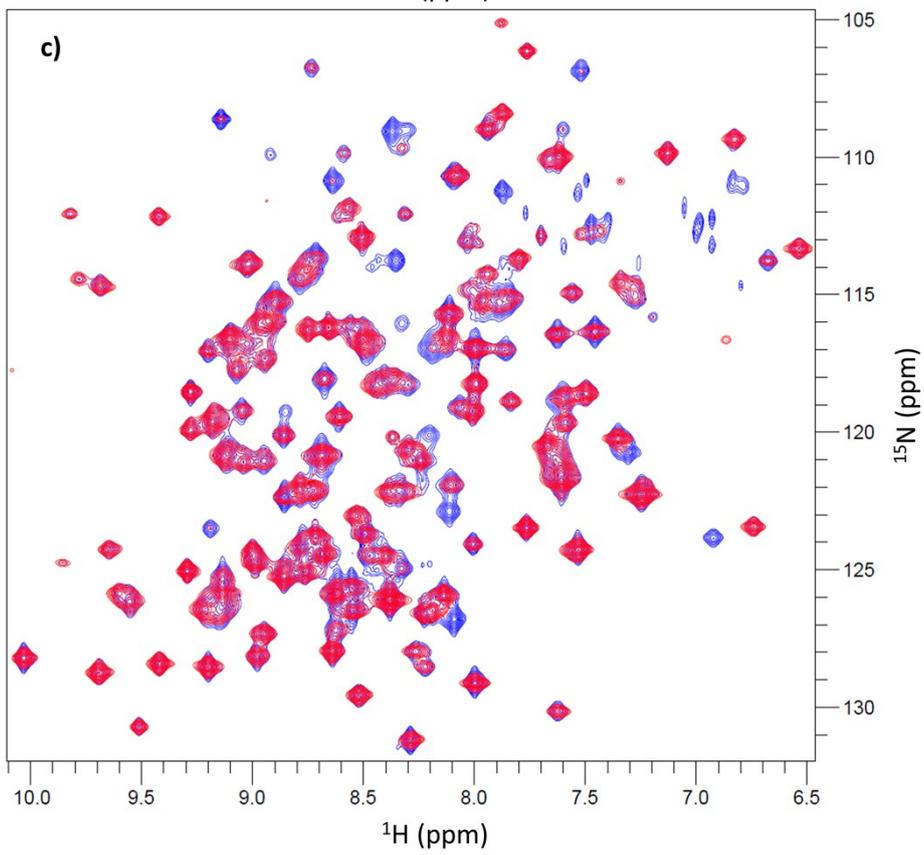
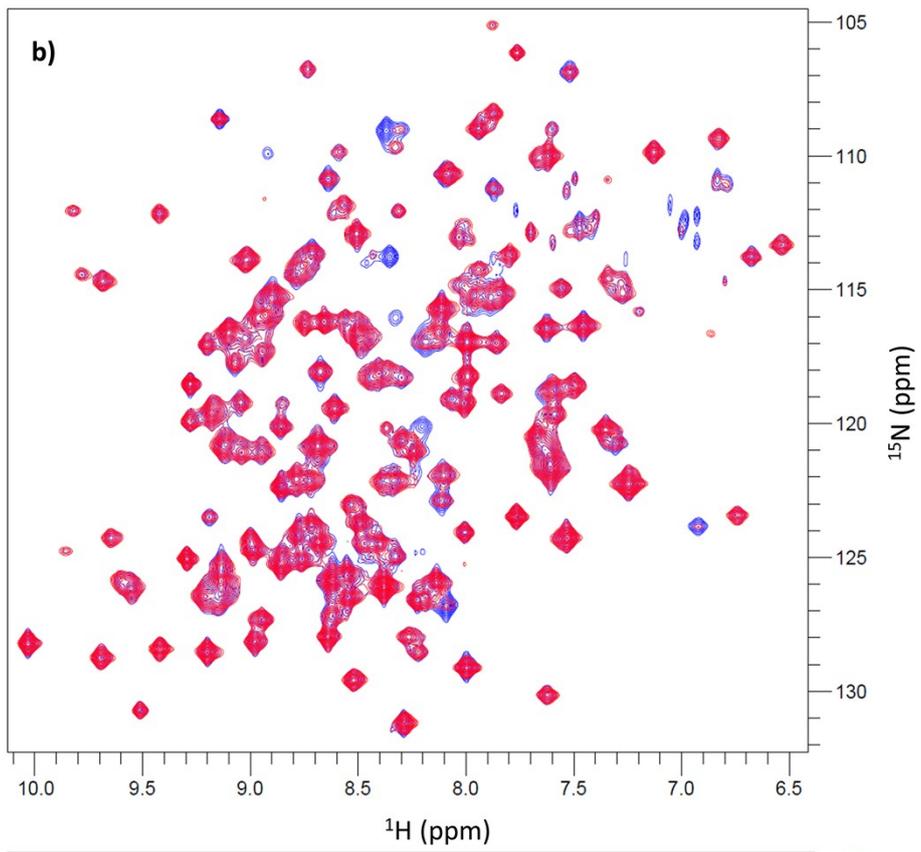


Fig. S14: a) Superposition of $^1\text{H},^{15}\text{N}$ -TROSY-HSQC spectra of U- $^{2}\text{H},^{15}\text{N}$] GII.4 Saga P dimers in the absence (blue) and presence (green) of 2 mM methyl α -L-fucopyranoside ligand concentration. Cross-peaks disturbed by the presence of methyl α -L-fucopyranoside are easy to identify from a visual inspection due to chemical shift perturbations. 500 MHz.



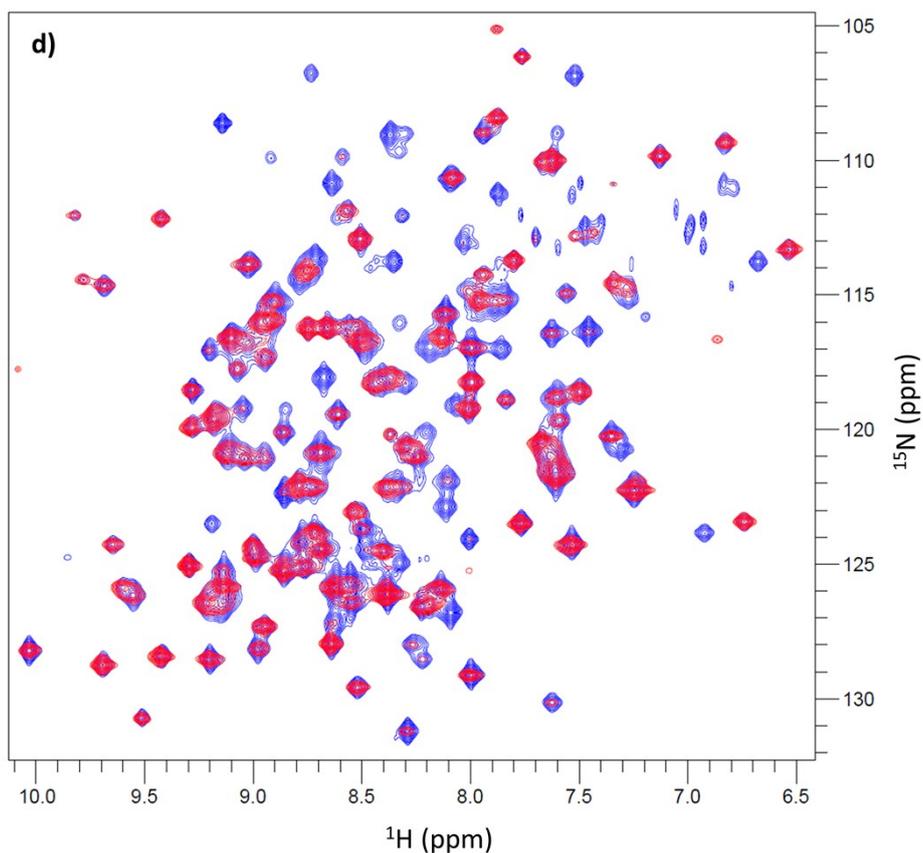


Figure S14 continued: Superposition of $^1\text{H}, ^{15}\text{N}$ -TROSY-HSQC spectra of U- $[^2\text{H}, ^{15}\text{N}]$ GII.4 Saga P dimers in the presence of **4**/ La^{3+} (blue) and **4**/ Gd^{3+} (red) at b) 0.1 mM, c) 0.25 mM and d) 0.5 mM ligand concentration. 500 MHz.

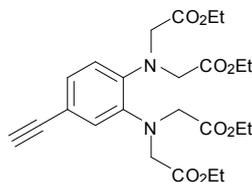
4. Notes and references

1. A. Mallagaray, J. Lockhauserbäumer, G. S. Hansman, C. Uetrecht and T. Peters, *Angew. Chem. Int. Ed.* 2015, **54**, 12014-12019.
2. G. S. Hansman, C. Biertümpfel, I. Georgiev, J. S. McLellan, L. Chen, T. Zhou, K. Katayama, P. D. Kwong, *J. Virol.*, 2011, **85**, 6687-6701.
3. V. Tugarinov, V. Kanelis, L. E. Kay, *Nature Protocols*, 2006, **1**, 749-754.
4. M. P. Williamson, *Prog. in NMR Spectr.*, 2013, **73**, 1-16.
5. B. K. Singh, M. M. Leuthold and G. S. Hansman, *J. Virol.*, 2015, **89**, 2024-2040.
6. M. Mizuno, S. Funahashi, N. Nakasuka, M. Tanaka, *Inorg. Chem.* 1991, **30**, 1550-1553.

5. Organic Synthesis

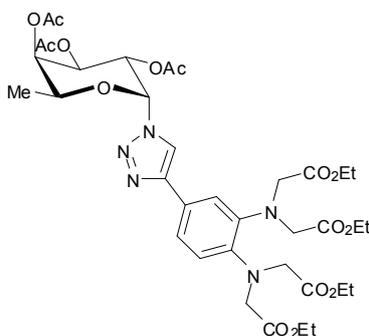
5.1 Synthesis of new products

Tetraethyl 2,2',2'',2'''-((4-ethynyl-1,2-phenylene)bis(azanetriyl))tetraacetate, 3.



To a solution of tetraethyl 2,2',2'',2'''-((4-bromo-1,2-phenylene)bis(azanetriyl))tetraacetate (0.740 g, 1.39 mmol) in dioxane (10 mL) under argon, Pd(PPhCN)₂Cl₂ (53 mg, 0.14 mmol, 10 mol%), CuI (19 mg, 0.10 mmol, 7 mol %), P(*t*-Bu)₃ (0.28 mL of a 1 M solution in toluene, 20 mol %), *i*-Pr₂NH (0.23 mL, 1.67 mmol) and (trimethylsilyl)acetylene (0.79 mL mg, 6.95 mmol) were added. This mixture was stirred for 6 d at 60 °C. After the aryl bromide has been consumed, the reaction mixture was diluted with toluene (25 mL), filtered through a small pad of celite™ and concentrated under vacuum. The crude product was dissolved in 10 mL of THF and TBAF (2.0 mL of 1M in THF) was added at 0 °C. The mixture was stirred for 90 minutes at RT. Finally, the solvent was removed *in vacuo* and the residue was dissolved in EtOAc (25 mL) and washed with water (25 mL). The organic layer was dried over MgSO₄ and the solvent was evaporated at reduced pressure. The product was purified by column chromatography (Hex: EtOAc, 9:1). Yield: 0.470 g (71%), white waxy solid; IR (neat) 2985, 2985, 2921, 1747, 1601 cm⁻¹; ¹H NMR (400 MHz, CDCl₃) δ 1.18-1.23 (m, 12H), 4.08-4.15 (m, 8H), 4.26 (s, 4H), 4.31 (s, 4H), 6.94 (d, *J* = 8.3 Hz, 1H), 7.09 (dd, *J*₁ = 8.3 Hz, *J*₂ = 1.8 Hz, 1H), 7.16 (d, *J* = 1.8 Hz, 1H); ¹³C NMR (100 MHz, CDCl₃) δ 14.1, 29.6, 52.2, 52.3, 60.58, 60.60, 76.1, 83.8, 116.3, 121.1, 125.5, 127.2, 141.0, 142.4, 170.56, 170.61; Anal. calcd. For C₂₄H₃₂N₂O₈: C, 60.49; H, 6.77; N, 5.88. Found: C, 60.29; H, 6.90; N, 5.57.

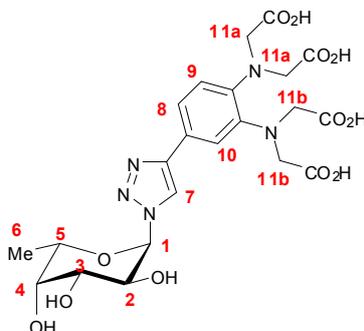
Tetraethyl 2,2',2'',2'''-((4-(1-(2,3,4-tri-O-acetyl- α -L-fucosyl)-1H-1,2,3-triazol-4-yl)-1,2-phenylene)bis(azanetriyl))tetraacetate



To a solution of tetraethyl 2,2',2'',2'''-((4-ethynyl-1,2-phenylene)bis(azanetriyl))tetraacetate (0.42 mmol, 0.200 g) and the corresponding 1-azido-2,3,4-tri-O-acetyl- α -L-fucose (0.38 mmol, 0.120 g) in a 1:1:1 mixture of THF (3 mL), H₂O (3 mL) and *t*-BuOH (3 mL) was added copper(II)sulfate (0.19 mmol, 0.95 mL of freshly prepared 0.2 M solution in H₂O), followed by sodium ascorbate (0.76 mmol, 3.8 mL of freshly prepared 0.2 M solution in H₂O). The reaction mixture was vigorously stirred overnight at room temperature in darkness. After completion (monitored by TLC) the reaction mixture was concentrated *in vacuo* and H₂O (15 mL) was added and then it was extracted with EtOAc (3x 10 mL). The organic layers were combined and washed with brine (15 mL), dried (MgSO₄), filtered, and evaporated. The crude was purified by flash chromatography (60 mesh silica gel, hexane/EtOAc 2:1) to afford the final compound. Yield: 0.256 g (85%); white solid, m.p. = 72-74 °C; [α]_D²⁵ = -60.68 (*c* = 0.049, CHCl₃); IR (KBr) 2983, 2932, 1747, 1610, 1574 cm⁻¹; ¹H NMR (400 MHz, CDCl₃) δ 1.13 (d, *J* = 6.4 Hz, 3H), 1.21 (t, *J* = 7.1 Hz, 12 Hz), 1.88 (s, 3H), 2.02 (s, 3H), 2.22 (s, 3H), 4.11 (q, *J* = 7.1 Hz, 8H), 4.34 (s, 4H), 4.35 (s, 4H), 4.53 (q, *J* = 6.4 Hz, 8.0 Hz, 1H), 5.49-5.54 (m, 2H), 6.25 (dd, *J*₁ = 10.8 Hz, *J*₂ = 3.5 Hz, 1H), 6.38 (d, *J* = 6.1

Hz, 1H), 7.10 (d, $J = 8.3$ Hz, 1H), 7.40 (dd, $J_1 = 8.3$ Hz, $J_2 = 1.6$ Hz, 1H), 7.57 (d, $J = 1.6$ Hz, 1H), 7.73 (s, 1H); ^{13}C NMR (100 MHz, CDCl_3) δ 14.1, 16.1, 20.5, 20.62, 20.64, 29.7, 52.4, 52.5, 60.60, 60.62, 67.2, 68.3, 68.8, 70.7, 82.1, 119.2, 120.6, 121.2, 121.7, 124.7, 141.7, 141.8, 147.0, 169.5, 170.4, 170.7, 170.8; MS (ESI): $m/z = 792$ $[\text{M} + \text{H}]^+$, 793 $[\text{M} + 2\text{H}]^{2+}$; Anal. calcd. For $\text{C}_{36}\text{H}_{49}\text{N}_5\text{O}_{15}$: C, 54.61; H, 6.24; N, 8.84. Found: C, 54.76; H, 6.01; N, 8.99.

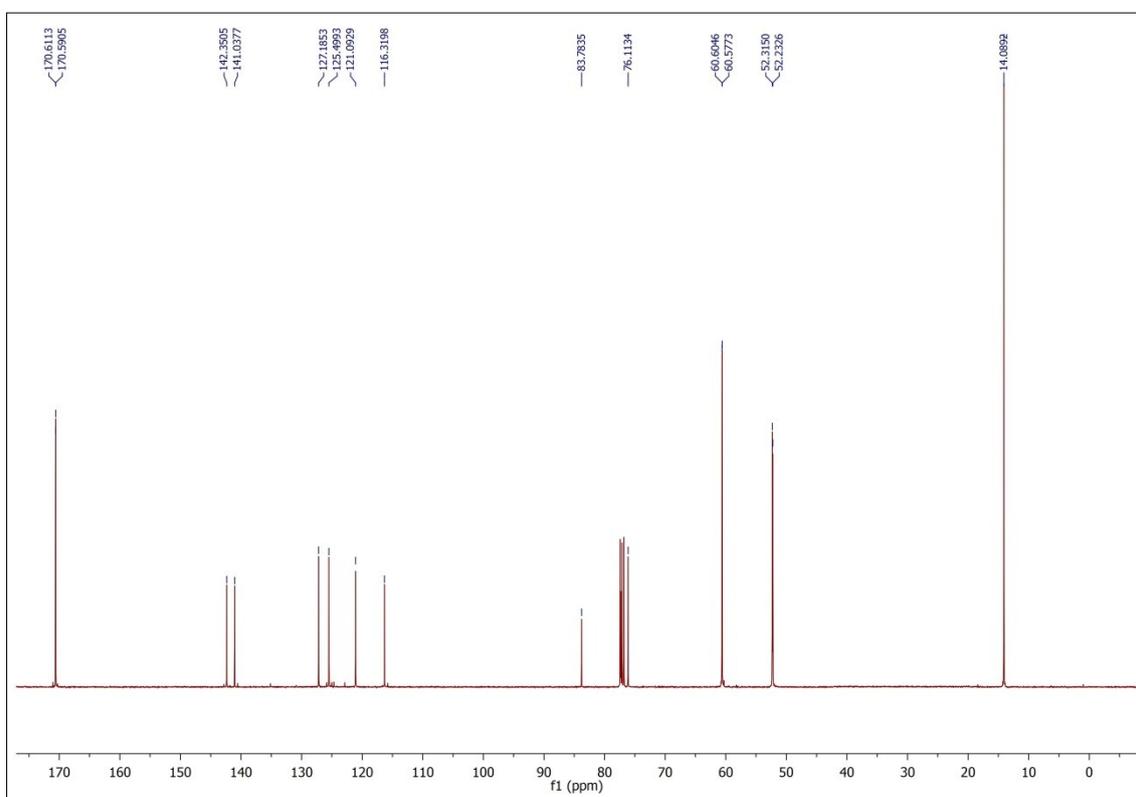
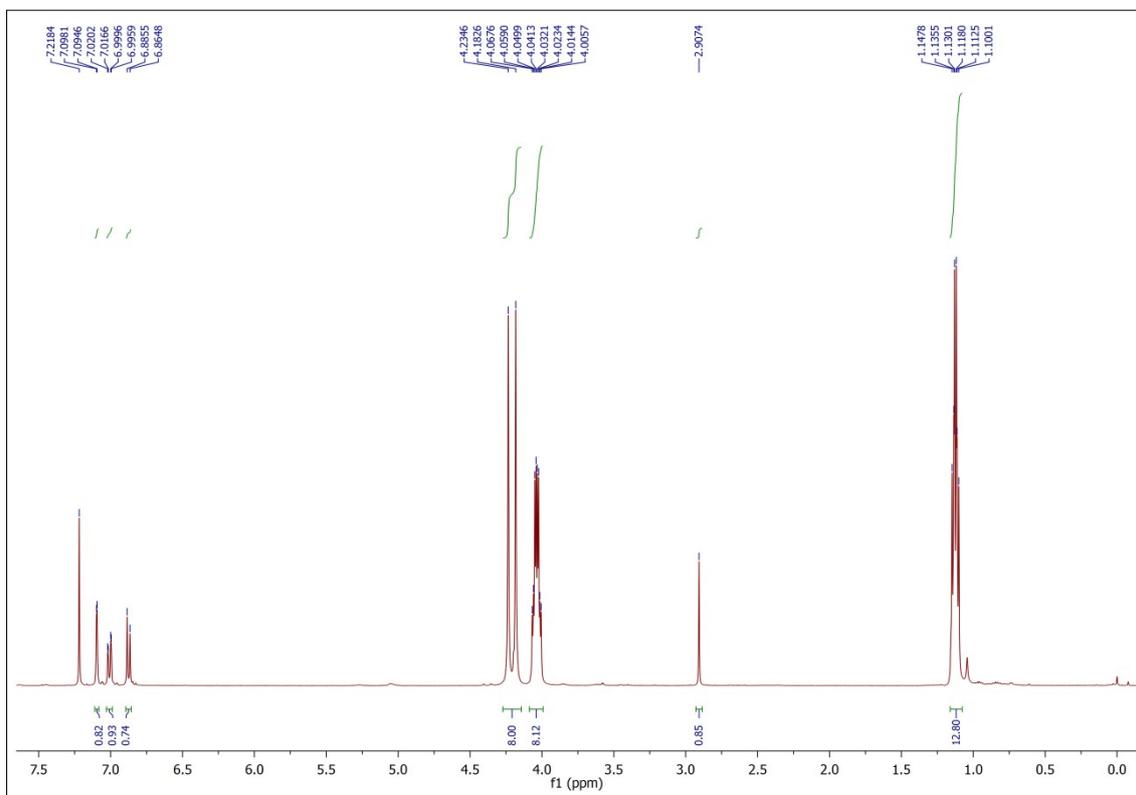
2,2',2'',2'''-((4-(1-((2R,3S,4R,5S,6S)-3,4,5-trihydroxy-6-methyltetrahydro-2H-pyran-2-yl)-1H-1,2,3-triazol-4-yl)-1,2-phenylene)bis(azanetriyl))tetraacetic acid, 4



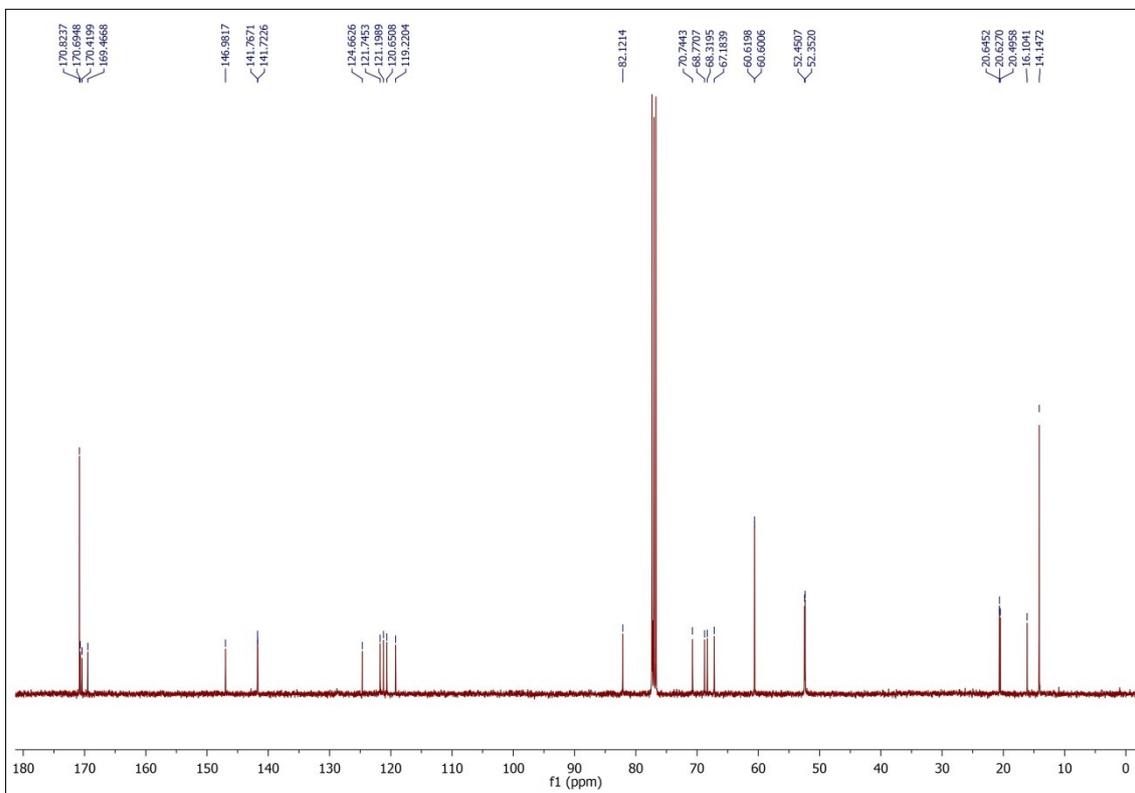
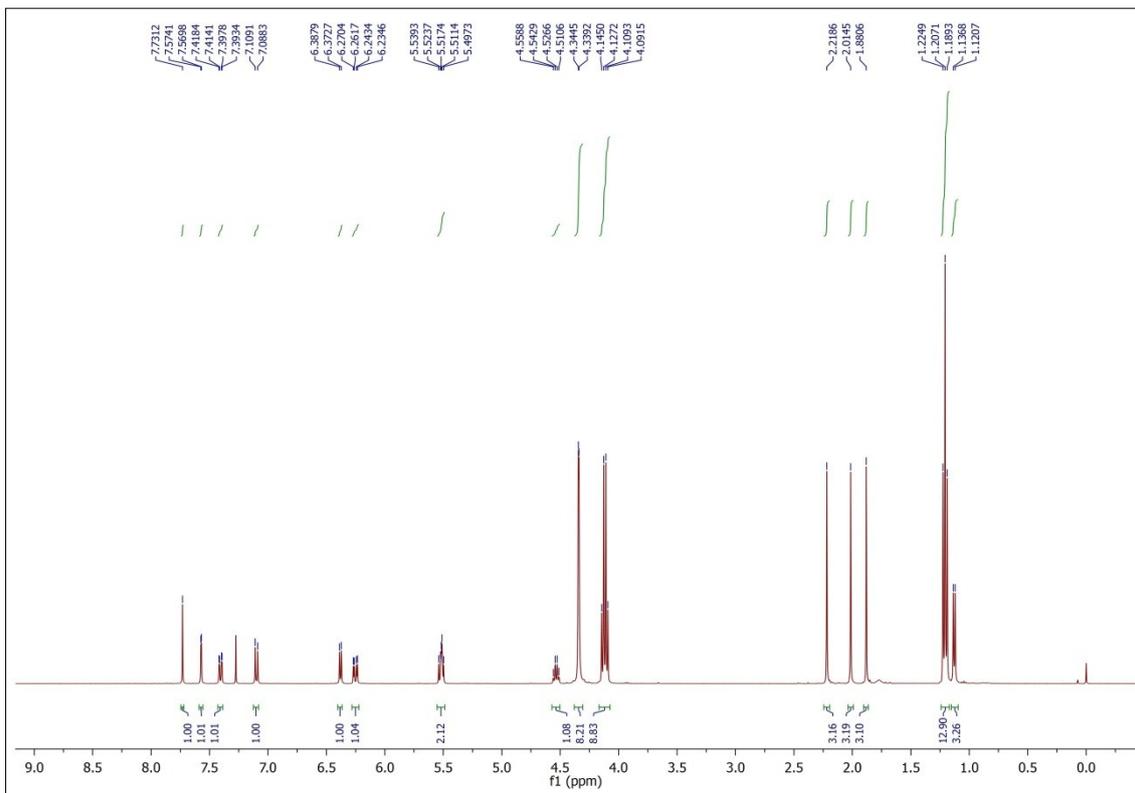
Over a solution of Tetraethyl 2,2',2'',2'''-((4-(1-(2,3,4-tri-O-acetyl- β -L-fucosyl)-1H-1,2,3-triazol-4-yl)-1,2-phenylene)bis(azanetriyl))tetraacetate (0.076 mmol, 60.0 mg) in methanol (1.5 mL), 400 μl of a 4 M NaOH solution in methanol were added, and the reaction was stirred for 3 h. After completion (monitored by TLC, MeOH/ H_2O 4:1) 1 mL MP- H_2O was added followed by Amberlite IR-120 in small portions until pH 1-2 was reached. The yellow solution was filtered, and the product was purified by RP-HPLC (gradient solvent B 0% for 5 min, then gradient to a final B 50% in 80 min, RT ~ 33-34 min) to afford the pure compound. The combined fractions showed pure material were lyophilized, affording **4** (30.2 mg, 54.6 μmol , 72%) as a yellow solid. M.p. = 145–148 $^\circ\text{C}$ (with decomposition); $[\alpha]_{\text{D}}^{25} = -54.56$ ($c = 94.5 \cdot 10^{-3}$, H_2O); IR (KBr) 3138, 2924, 1713, 1497 cm^{-1} ; ^1H NMR (500 MHz, Tris-*d*11 25 mM pH 7.4, NaCl 0.3 M, D_2O) δ 1.19 (d, $J = 6.5$ Hz, 3xH-6), 3.98 (d, $J = 3.5$ Hz, H-4), 4.17 (q, $J = 6.5$ Hz, H-5), 4.37 (dd, $J_1 = 10.5$ Hz, $J_2 = 6.1$ Hz, H-2), 4.41 (s, 4xH-11a), 4.41 (s, 4xH-11b), 4.59 (dd, $J_1 = 10.5$ Hz, $J_2 = 3.5$ Hz, H-3), 6.33 (d, $J = 6.1$ Hz, H-1), 7.24 (d, $J = 8.4$ Hz, H-9), 7.44 (dd, $J_1 = 8.4$ Hz, $J_2 = 1.7$ Hz, H-8), 7.59 (d, $J = 1.8$ Hz, H-10), 8.38 (s, H-7); ^{13}C NMR (125 MHz, CDCl_3) δ 18.4, 55.1, 55.1, 69.5, 72.2, 73.4, 74.1, 88.3, 121.1, 123.6, 124.2, 126.2, 127.0, 143.8, 144.0, 149.2, 177.8, 177.8; Anal. calcd. For $\text{C}_{22}\text{H}_{27}\text{N}_5\text{O}_{12}$: C, 47.74; H, 4.92; N, 12.65. Found: C, 47.81; H, 4.83; N, 12.45.

5.2 Spectra of new products

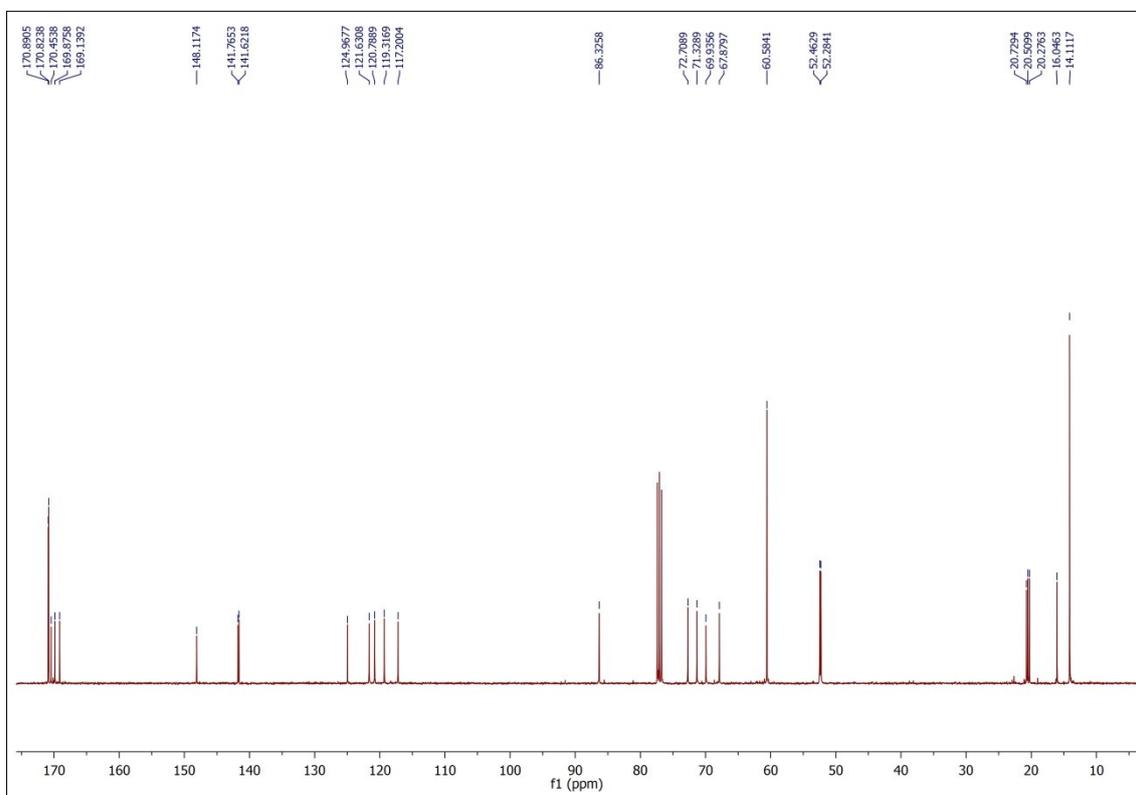
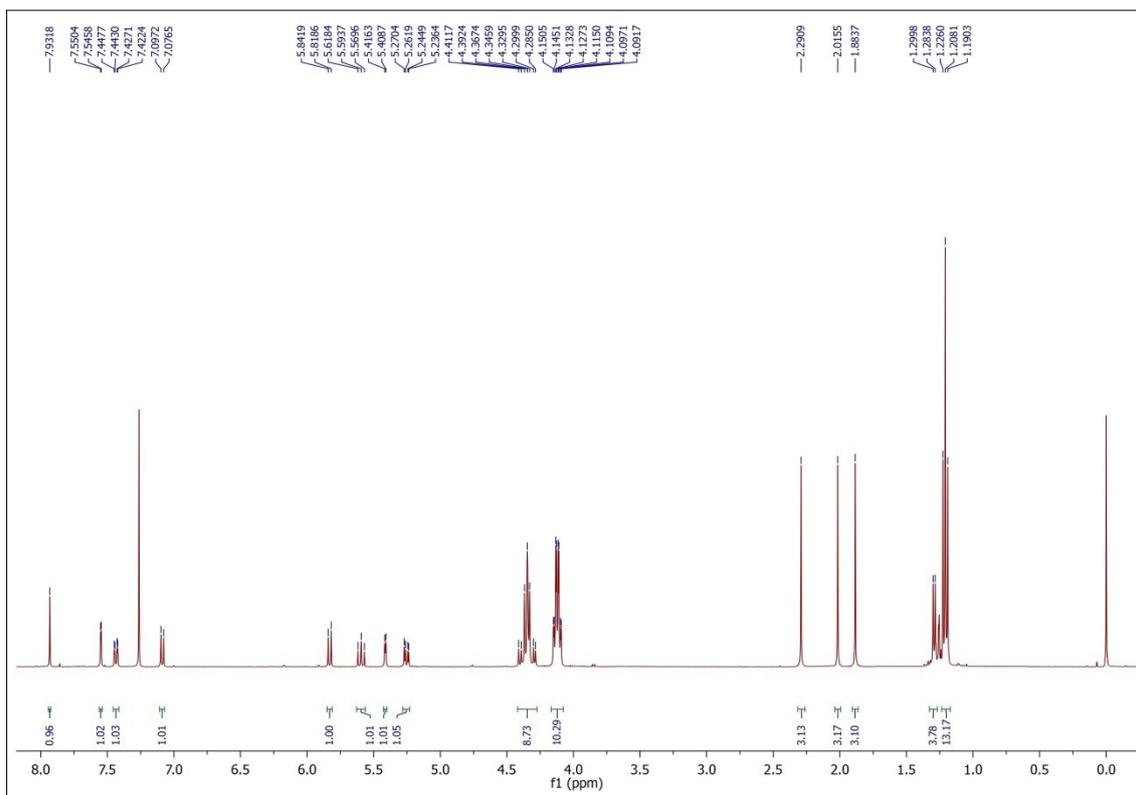
Tetraethyl 2,2',2'',2'''-((4-ethynyl-1,2-phenylene)bis(azanetriyl))tetraacetate, **3**.



Tetraethyl 2,2',2'',2'''-((4-(1-(2,3,4-tri-O-acetyl- α -L-fucosyl)-1H-1,2,3-triazol-4-yl)-1,2-phenylene)bis(azanetriyl))tetraacetate



Tetraethyl 2,2',2'',2'''-(((4-(1-(2,3,4-tri-O-acetyl-β-L-fucosyl)-1H-1,2,3-triazol-4-yl)-1,2-phenylene)bis(azanetriyl))tetraacetate



2,2',2'',2'''-((4-(1-((2R,3S,4R,5S,6S)-3,4,5-trihydroxy-6-methyltetrahydro-2H-pyran-2-yl)-1H-1,2,3-triazol-4-yl)-1,2-phenylene)bis(azanetriyl))tetraacetic acid, **4**.

