Electronic Supplementary Material (ESI) for New Journal of Chemistry. This journal is © The Royal Society of Chemistry and the Centre National de la Recherche Scientifique 2023

SUPPORTING INFORMATION – Part A

Synthesis, Conformational Analysis and GalNAc-Lectin Interactions of constrained C-glycoside analogue of T_N antigen

Juliette Dourdan, Florian Rouzier, Thanh Thao Huynh, Sullivan Bricaud, Arnaud Nourry,* Stéphane Guillarme*

Institut des Molécules et des Matériaux du Mans, UMR 6283 CNRS and Le Mans Université, Avenue O. Messiaen, 72085 Le Mans, France.

Table of Contents

General Information	2
Synthesis of T_N antigen derivative 1	2
Experimental data on T_N antigen derivative 1	3
Experimental data on constrained <i>C</i> -glycoside 2	5
¹ H NMR spectra of analogue 2 at different temperatures	6
¹ H and ¹³ C NMR spectra of compounds 1 and 2	7
Experimental part on STD experiments	21

1- General Information

All experiments were carried out under argon with anhydrous solvents in dried glassware. THF was dried over activated alumina on a dry station purchased from Innovative Technologies. Commercially available materials were used without further purification. Flash chromatographies were performed on silica gel (40-63 µm from Macherey-Nagel) using Reveleris X² Grace apparatus. Analytical TLCs were carried out on pre-coated silica gel 60 F254 from Macherey-Nagel. Optical rotations were measured using a Jasco P2000 at the sodium D line (Λ = 589 nm) with a 1-dm path length cell at 25°C. Melting points were measured using Buchi B-545 apparatus. NMR spectra were recorded with a Bruker Avance 400 or a Bruker Avance Neo 500 spectrometer. Chemical shifts are reported in ppm from TMS as the internal standard for the ¹H NMR spectrum and from the residual peaks of the solvent (CDCl₃) for ¹³C NMR spectrum. Structural assignments of the isolated compounds were based on ¹H, ¹³C, COSY and HSQC NMR experiments. High resolution mass spectroscopy was performed with a Bruker microTOF QIII mass spectrometer using ESI techniques.

For NMR assignment, following numbering was chosen:



2- Synthesis of T_N antigen derivative 1

Tn antigen derivative **1** was isolated from 3,4,6-tri-O-Benzyl-2-nitro-galactal following methodology developed by Schmidt et al.¹ As first attempt, Michael addition of the N-Ac serine derivative **8a** led to an inseparable mixture of the two anomers of desired compounds in 74 % yield. Reduction of nitro group and acylation of this mixture did not allow the separation of both anomers. Then, the addition of the N-Boc serine derivative **8b** onto 2-nitrogalactal led to a separable mixture of the two anomers (Scheme 1).



Scheme 1

Nitro group of α -anomer **9b** was reduced and the amine was treated with acetic anhydride leading to **10** in 68% yield. Boc deprotection, acylation and hydrogenolysis of benzyl groups afforded **1** (Scheme 2).

¹ G. A. Winterfeld , Y. Ito , T. Ogawa , R. R. Schmidt , *Eur. J. Org. Chem.* **1999**, 1167 –1171



Scheme 2

3- Experimental data on T_N antigen derivative 1

a) Compound 9b

Under argon, to a solution of 3,4,6-tri-*O*-benzyl-2-nitro-D-galactal (5 g, 10.8 mmol) in dry THF (25mL) was added a solution of (N-Boc, NHMe)-serine **7** (2.1 g 9.8 mmol) in dry THF (60 mL). The reaction mixture was cooled to 0 °C, and a solution of potassium *tert*-butylate in dry THF (0,188 M, 13 mL, 2.45 mmol) was added dropwise. The solution was stirred at room temperature for 23 h, and after concentration, the crude product was purified by column chromatography (Cyclohexane/EtOAc, 1:0 to 8:2) to afford α -anomer **9b** (2.71 g, 4,69 mmol, 43%), β -anomer **9b** (0.4 g, 0.69 mmol, 6%) and a mixture of the two anomers (1.85 g, 3.2 mmol, 30%).

α-anomer **9b**. White solid. $R_f = 0.47$ (cyclohexane/EtOAc, 5:5). $[α]_D^{20} = +91.4$ (*c* 1.04, CH₂Cl₂). **HRMS** (ESI) Calculated for C₃₆H₄₅N₃NaO₁₀ ([M+Na]⁺), m/z = 702.2997; found: 702.2985. ¹H NMR (400 MHz, CDCl₃): δ 7.38 – 7.20 (m, 15H, H_{Ar}), 6.12 – 6.09 (m, 1H, NHCH₃), 5.40 – 5.35 (m, 2H, NHBoc, H-1), 5.00 (dd, *J* = 10.7; 4.2 Hz, 1H, H-2), 4.83 (d, *J* = 11.3 Hz, 1H, CH₂Ph), 4.70 (s, 2H, CH₂Ph), 4.52 (d, *J* = 11.6 Hz, 1H, CH₂Ph), 4.47 (d, *J* = 11.3 Hz, 1H, CH₂Ph), 4.42 – 4.36 (m, 2H, CH₂Ph, H-3), 4.18 (br s, 1H, H-2'), 4.02 – 3.95 (m, 3H, H-1', H-4, H-5), 3.70 – 3.66 (m, 1H, H-1'), 3.59 – 3.50 (m, 2H, H-6), 2.75 (d, *J* = 4.8 Hz, 3H, NHCH₃), 1.46 (s, 9H, Boc). ¹³C NMR (100.6 MHz, CDCl₃): δ 170.0 (C3'), 155.4 (C=O (Boc)), 137.9 (Cq_{Ar}), 137.6 (Cq_{Ar}), 128.7 (CH_{Ar}), 128.5 (CH_{Ar}), 128.3 (CH_{Ar}), 128.2 (CH_{Ar}), 128.1 (CH_{Ar}), 97.2 (C1), 84.5 (C2), 80.6 (Cq (Boc)), 75.2 (CH₂-Ph), 75.1 (C3), 73.9 (CH₂-Ph), 73.1 (CH₂-Ph), 73.0 (C4), 70.4 (C5), 69.2 (C1'), 68.8 (C6), 54.2 (C2'), 28.5 (CH₃ (Boc)), 26.4 (NH-CH₃).

b) Compound 10

 α -anomer **9b** (1.5 g, 2.22 mmol, 1 eq) was dissolved in a 2:1 THF/Water mixture (225 mL) and cooled down between 0 °C and -5 °C. Concentrated hydrochloric acid (6.8 mL, 100 eq), acetic acid (40 mL, 320 eq) and freshly activated zinc (3.5 g, 53.3 mmol, 24 eq) were added and the solution was stirred, at 0 °C, until the TLC showed complete reaction. The mixture was filtered on Celite® and the filtrate was diluted in DCM. After separation of the two layers, the organic layer was washed with water, a saturated aqueous NaHCO₃ solution and then, water. The organic layer was then dried over MgSO₄,

concentrated under reduced pressure and then co-evaporated with toluene. The crude product was dissolved in dry DCM (70 mL) and DMAP (13.5 mg, 0.111 mmol, 0.05 eq), anhydride acetic (1.68 mL, 17.8 mmol, 8 eq) and Et₃N (2.09 mL, 15.54 mmol, 7 eq) were added. The solution was stirred at room temperature during 4 h and the reaction was diluted with water. The aqueous phase was extracted with DCM, then the organic phase was dried over MgSO₄, concentrated under reduced pressure and co-evaporation with toluene. Purification by column chromatography on silica gel (EtOAc/MeOH, 100:0 to 90:10) gave compound **10** (1.34 g, 1.94 mmol, 87%).

White powder. $R_f = 0.31$ (CH₂Cl₂/MeOH, 95:5). $[\alpha]_D^{20} = + 71.3$ (*c* 1, CH₂Cl₂). **HRMS** (ESI) Calculated for C₃₈H₄₉N₃NaO₉ ([M+Na]⁺), m/z = 714.3361; found: 714.3352. ¹H NMR (400 MHz, CDCl₃) δ 7.40 – 7.27 (m, 15H, ArH), 6.42 (d, *J* = 4.5 Hz, 1H, NH-CH₃), 6.02 (s, 1H, NH-Boc), 5.50 (d, *J* = 8.3 Hz, 1H, NH-Ac), 4.94 and 4.55 (AB syst., *J* = 11.7 Hz, 2H, CH₂-Ph), 4.90 (d, *J* = 3.7 Hz, 1H, H-1), 4.71 and 4.47 (AB syst., *J* = 12.2 Hz, 2H, CH₂-Ph), 4.68 – 4.62 (m, 1H, H-2), 4.59 and 4.36 (AB syst., *J* = 11.5 Hz, 2H, CH₂-Ph), 4.32 – 4.26 (br s, 1H, H-2'), 4.23 (d, *J* = 11.3 Hz, 1H, H-1'), 3.93 – 3.83 (m, 2H, H-4 and H-5), 3.70 – 3.61 (m, 2H, H-1' and H-6), 3.57 (dd, *J* = 11.1, 1.6 Hz, 1H, H-3), 3.45 (dd, *J* = 9.6, 4.6 Hz, 1H, H-6), 2.59 (d, *J* = 4.5 Hz, 3H, NH-CH₃), 1.93 (s, 3H, CH₃ (Ac)), 1.46 (s, 9H, CH₃ (Boc)). ¹³C NMR (100.6 MHz, CDCl₃) δ 170.6 (C=O (Ac)), 170.2 (C3'), 155.7 (C=O (Boc)), 138.3 (Cq_{Ar}), 138.1 (Cq_{Ar}), 137.5 (Cq_{Ar}), 128.7 (CH_{Ar}), 128.5 (CH_{Ar}), 128.4 (CH_{Ar}), 128.3 (CH_{Ar}), 128.0 (CH_{Ar}), 127.9 (CH_{Ar}), 127.7 (CH_{Ar}), 100.6 (C1), 80.4 (Cq (Boc)), 77.4 (C3), 74.5 (CH₂-Ph), 74.1 (CH₂-Ph), 72.8 (C4), 72.4 (C1'), 71.9 (CH₂-Ph), 70.9 (C5), 70.5 (C6), 55.1 (C2'), 49.1 (C2), 28.5 (CH₃ (Boc)), 26.2 (CH₃ (CH₃-N)), 23.5 (CH₃ (Ac)).

c) Compound 11

To a solution of compound **10** (236 mg, 0.341 mmol, 1 eq) in dry DCM (23 mL) was added TFA (7.6 mL). The solution was stirred at room temperature during 20 min and then concentrated under reduced pressure and co-evaporated with toluene. The crude amine was dissolved in pyridine (0.791 mL, 30 eq) and anhydrid acetic (1.85 ml, 60 eq) was added. After stirring during 2 h, the solution was concentrated under reduced pressure. The crude product was purified by column chromatography on silica gel (DCM/MeOH, 100:0 to 90:10) to afford compound **11** (170 mg, 0.265 mmol, 78%).

White powder. $R_f = 0.66 (CH_2Cl_2/MeOH, 9:1). [\alpha]_D^{20} = + 86.1 (c 0.4, MeOH). HRMS (ESI) Calculated for C_{35}H_{43}N_3NaO_8 ([M+Na]⁺), m/z = 656.2942; found: 656.2948. ¹H NMR (400 MHz, CDCl_3) <math>\delta$ 7.40 – 7.27 (m, 15H, HAr), 7.02 (d, *J* = 8.7 Hz, 1H, AcN*H*-C'2), 6.34 (d, *J* = 4.8 Hz, 1H, N*H*Me), 5.51 (d, *J* = 8.7 Hz, 1H, AcN*H*-C2), 4.95 (d, *J* = 11.7 Hz, 1H, CH₂-Ph), 4.90 (d, *J* = 3.8 Hz, 1H, H-1), 4.72 (d, *J* = 12.1 Hz, 1H, CH₂-Ph), 4.68 – 4.61 (m, 1H, H-2), 4.58 – 4.53 (m, 3H, H-2' and CH₂-Ph), 4.50 (d, *J* = 10.2 Hz, 1H, CH₂-Ph), 4.40 (d, *J* = 11.4 Hz, 1H, CH₂-Ph), 4.18 (dd, *J* = 11.3, 4.2 Hz, 1H, H-6), 3.96 – 3.88 (m, 2H, H-5 and H-4), 3.69 (dd, *J* = 9.7, 7.5 Hz, 1H, H-1'), 3.62 (ddd, *J* = 11.3, 8.2, 3.8 Hz, 2H, H-3 and H-6), 3.42 (dd, *J* = 9.7, 4.4 Hz, 1H, H-1'), 2.66 (d, *J* = 4.8 Hz, 3H, N-CH₃), 1.97 (s, 3H, CH₃ (Ac)), 1.93 (s, 3H, CH₃(Ac)). ¹³C NMR (100.6 MHz, CDCl₃) δ 170.3 (2 * C=O (Ac)), 170.1 (C3'), 138.3 (Cq_{Ar}), 138.0 (Cq_{Ar}), 137.3 (Cq_{Ar}), 128.8 (CH_{Ar}), 128.7 (CH_{Ar}), 128.4 (CH_{Ar}), 128.1 (CHAr), 128.0 (CH_{Ar}), 127.8 (CH_{Ar}), 100.8 (C1), 77.4 (C3), 74.5 (CH₂-Ph), 74.2 (CH₂-Ph), 72.7 (C4), 72.3 (C1'), 72.0 (CH₂-Ph), 71.0 (C5), 70.7 (C6), 53.6 (C2'), 49.5 (C2), 26.3 (CH₃-N), 23.6 (CH₃ (Ac)), 23.4 (CH₃ (Ac)).

d) T_N antigen derivative 1

Compound **11** (325 mg, 0.53 mmol, 1 eq) was dissolved in EtOH (4.2 mL). Cyclohexene (2.1 mL, 4 eq) and Pd(OH)₂/C (20%w/w) were then added. The solution was stirred at reflux during 6 h, then filtered on Silice/Celite® and the filtrate was concentrated under reduced pressure. Purification by column chromatography on silica gel (DCM/MeOH, 100:0 to 70:30) gave compound **1** (138 mg, 0.38 mmol, 71%).

White powder. M.p. 260-261°C. $R_f = 0.11$ (DCM/MeOH, 8:2). $[\alpha]_D^{20} = +146$ (c 0.49, H₂O). **HRMS** (ESI) Calculated for $C_{14}H_{25}N_3NaO_8$ ([M+Na]⁺), m/z = 386.1534, found: 386.1532. ¹**H NMR** (400 MHz, CDCl₃) δ 4.92 (d, J = 3.8 Hz, 1H, H-1), 4.57 (t, J = 5.4 Hz, 1H, H-5), 4.18 (dd, J = 11.1, 3.8 Hz, 1H, H-2), 4.01 (d, J = 3.1 Hz, 1H, H-4), 3.97 – 3.88 (m, 3H, H-6, H-3 and H-2'), 3.85 (dd, J = 10.9, 5.4 Hz, 1H, H-6), 3.81 – 3.75 (m, 2H, H-1'), 2.78 (s, J = 5.6 Hz, 3H, NH-CH₃), 2.11 (s, 3H, CH₃ (Ac)), 2.07 (s, 3H, CH₃ (Ac)). ¹³**C NMR** (100.6 MHz, CDCl₃) δ 174.4 (C=O (Ac)), 174.4 (C=O (Ac)), 171.9 (C3'), 97.8 (C1), 71.3 (C2'), 68.4 (C4), 67.6 (C3), 67.1 (C6), 61.2 (C1'), 53.9 (C5), 49.8 (C2), 26.0 (CH₃-N), 22.0 (CH₃ (Ac)), 21.8 (CH₃ (Ac)).

4- Experimental data on constrained C-glycoside 2

a) Compound 6

A solution of deprotected isoxazolidine² (175 mg, 0.29 mmol) in pyridine (3.5 mL) and acetic anhydride (0.35 mL, 3.7 mmol) was stirred overnight. The reaction was diluted with EtOAc (15 mL) and water (20 mL). After separation of layers, the aqueous one was extracted with EtOAc (2×10 mL). Organic phase was then washed with water (2×15 mL), dried over MgSO₄ and concentrated under reduced pressure. Crude product was purified by column chromatography on silica gel (DCM/MeOH, 98:2 to 96:4) to afford compound **6** (134 mg, 0.208 mmol, 72%).

Colorless oil. $R_f = 0.44$ (DCM/MeOH, 95:5). $[\alpha]_D^{20} = + 3.6$ (*c* 1.0, CH₂Cl₂). **HRMS** (ESI) Calculated for C₃₆H₄₃N₃NaO₈ ([M+Na]⁺), m/z = 668.2942, found: 668.2948. ¹**H NMR** (400 MHz, CDCl₃, 323 K) δ 7.36–7.20 (m, 15H, ArH), 6.50 (br s, 1H, NH), 4.72 (d, *J* = 11.8 Hz, 1H, CH₂Ph), 4.70-4.64 (m, 1H, H-3'), 4.63 (d, *J* = 11.8 Hz, 1H, CH₂Ph), 4.53 (d, *J* = 12.0 Hz, 1H, CH₂Ph), 4.53-4.48 (m, 2H, CH₂Ph), 4.43 (d, *J* = 11.8 Hz, 1H, CH₂Ph), 4.28 (ddd, *J* = 7.3, 5.1, 3.3 Hz, 1H, H-1'), 4.25-4.19 (m, 1H, H-5), 4.18-4.13 (m, 1H, H-2), 4.09 (dd, *J* = 11.6, 9.3 Hz, 1H, H-6), 4.02-3.86 (m, 2H, H-1, H-3), 3.77 (dd, *J* = 6.1, 2.9 Hz, 1H, H-4), 3.67 (br d, *J* = 11.6 Hz, 1H, H-6), 3.00-2.60 (m, 2H, H-2'), 2.80 (d, *J* = 4.7 Hz, 3H, CH₃-N), 2.02 (br s, 6H, 2*CH₃-CO). ¹³C NMR (100.6 MHz, CDCl₃) δ 170.8 (C=O), 170.3 (C=O), 138.2 (Cq_{Ar}), 137.8 (2 Cq_{Ar}), 128.5 (2 CHAr), 128.4 (CHAr), 127.9 (CHAr), 127.7 (CHAr), 127.4 (CHAr), 80.8 (C1'), 75.3 (C5), 74.2 (C3), 73.3 (2 CH₂), 73.2 (C4), 71.4 (CH2), 68.3 (C1), 65.0 (C6), 61.4 (C3'), 53.4 (C2), 38.1 (C2'), 26.3 (CH₃-N), 23.1 (CH₃-CO), 21.2 (CH₃-CO).

b) Constrained C-glycoside 2

Under argon, to a solution compound **6** (175 mg, 0.271 mmol) in ethanol (2.5 mL) were added cyclohexene (1.2 mL) and Pearlman's catalyst (35 mg). The solution was heated under reflux for 6.5 h. After filtration onto Celite[®], volatiles were removed under reduced pressure and crude product was purified by column chromatography on silica gel (DCM/MeOH, 90:10 to 80:20) to give *C*-glycoside analogue **2** (64.5 mg, 0.172 mmol, 64%).

White solid. M.p.: 153.5-154°C. $R_f = 0.25$ (DCM/MeOH, 80:20). $[\alpha]_D^{20} = +88.3$ (*c* 0.5, H₂O). **HRMS** (ESI) Calculated for $C_{15}H_{25}N_3O_8Na$ ($[M+Na]^+$), m/z = 398.1534; found: 398.1534. ¹H NMR (500 MHz, D₂O, 383 K) δ 4.80 (m, 1H, H-3'), 4.62-4.56 (m, 1H, H-1'), 4.25 (dd, *J* = 9.0, 5.7 Hz, 1H, H-2), 4.12 (dd, *J* = 5.7, 5.7Hz, 1H, H-1), 4.03-4.00 (m, 1H, H-4), 3.97 (dd, *J* = 9.0, 3.3 Hz, 1H, H-3), 3.96-3.92 (m, 1H, H-5), 3.79 (dd, *J* = 12.0, 7.6 Hz, 1H, H-6), 3.69 (dd, *J* = 12.0, 7.6 Hz, 1H, H-6), 2.74 (s, 3H, CH₃-N), 2.73-2.62 (m, 1H, H-2'), 2.47 (ddd, *J* = 13.0, 6.0, 2.6 Hz, 1H, H-2'), 2.11 (s, 3H, CH₃ (Ac)), 2.01 (s, 3H, CH₃ (Ac). ¹³C NMR (125.7 MHz, D₂O, 383 K) δ 175.1 (C=O), 171.8 (C=O), 79.4, (C1'), 76.2 (C5), 70.7 (C1), 68.7 (C3), 68.0 (C4), 61.1 (C6), 59.1 (C3'), 49.9 (C2'), 26.4 (CH₃-N), 22.3 (CH₃-CO), 20.3 (CH₃-CO).

² F. Rouzier, R. Sillé, O. Montiège, A. Tessier, M. Pipelier, G. Dujardin, A. Martel, A. Nourry, S. Guillarme, *Eur. J. Org. Chem.* **2020**, 43, 6749–6757



5- ¹H NMR spectra of analogue 2 at different temperatures.

6- ¹H and ¹³C NMR spectra

6.1- Compound 9b α .





6.1.b- ^{13}C NMR spectrum of compound $9b\alpha$ (CDCl_3, 298 K)



6.2- Compound 10.

6.2.a-¹H NMR spectrum of compound **10** (CDCl₃, 298 K)



6.2.b-¹³C NMR spectrum of compound **10** (CDCl₃, 298 K)



6.3- Compound 11

6.3.a-¹H NMR spectrum of compound **11** (CDCl₃, 298 K)



6.3.b-¹³C NMR spectrum of compound **11** (CDCl₃, 298 K)



6.4- Compound 1

6.4.a- 1 H NMR spectrum of compound **1** (D₂O, 298 K)







6.4.c- 13 C NMR spectrum of compound **1** (D₂O, 298 K)



6.5- Compound 6

6.5.a- ¹H NMR spectra of compound **6** (CDCl₃, 323 K)



6.5.b-¹³C NMR spectrum of compound **6** (CDCl₃, 298 K)



6.6- Compound 2

6.6.a-¹H NMR spectra of compound **2** (D₂O, 353 K)



6.6.b- COSY spectra of compound $\boldsymbol{2}$ (D₂O, 353 K)



6.6.c- 13 C NMR spectra of compound **2** (D₂O, 353 K)



7- Experimental part on STD experiments

At 298 K, unless otherwise stated, STD spectra were acquired with 1024 scans, using a 50 ms Gaussian-shaped pulse for the selective saturation of the protein protons in the aliphatic region ($\delta = 1$ ppm). The off-resonance frequency was set at $\delta = 40$ ppm, the relaxation time D1 = 2 s and DS = 4.

Glycine max lectin (SBA) and dolichos biflorus lectin (DBA) were purchased at Glycodiag.

For SBA and DBA samples, a PBS (phosphate-buffered saline) solution [NaCl (137 mM), KCl (2.7 mM), Na₂HPO₄ (10 mM), KH₂PO₄ (1.7 mM), pH 7.3] in D₂O was used. A T1 rho spin-lock filter was used to suppress the protein signals.

- Glycine max lectin (SBA) :

STD spectra for the systems SBA/natural derivative and SBA/C-glycoside were acquired using samples of SBA (40 μ M) and the corresponding ligand (4 mM). Competition STD experiments between natural derivative and *C*-glycoside for SBA were ran by stepwise 4 addition of *C*-glycoside to a solution of SBA (40 μ M) and natural derivative (4 mM) until a final concentration of 12 mM of the *C*-glycosides. Five ratio T_N antigen derivative **1**/*C*-glycoside analogue **2** were acquired: 1:0, 1:0.5, 1:1, 1:1.5 and 1:2.

- Dolichos biflorus lectin (DBA) :

STD spectra for the system DBA/natural derivative and DBA/*C*-glycoside were acquired using samples of DBA (40 μ M) and the corresponding ligand (4 mM). DBA samples contained 40% of non deuterated water, therefore water-suppression filter was applied using excitation sculpting (stddiffgpes.3).³ Competition STD experiments between natural derivative and *C*-glycoside for SBA were ran by stepwise four additions of *C*-glycoside to a solution of SBA (40 μ M) and natural derivative (4 mM) until a final concentration of 12 mM of the *C*-glycosides. Five ratio T_N antigen derivative **1**/*C*-glycoside analogue **2** were acquired: 1:0, 1:0.5, 1:1, 1:1.5 and 1:2.

Spectra were analysed using the STD value A ($A=I_{STD}/I_{Off}$).

³ T. L. Hwang, A. J. Shaka, *J. Magn. Reson. A*, **1995**, *112*, 275-279