Supplementary Material

Big is not better: Comparing two alpha-Gal-bearing glycotopes in neoglycoproteins as biomarkers for *Leishmania (Viannia) braziliensis* infection

Alba L. Montoya[a,1], Eileni R. Gil[a,1], Irodiel Vinales[a,1], Igor L. Estevao[b], Paola Taboada[b], Mary Cruz Torrico[c], Faustino Torrico[c], Jorge Diego Marco[d], Igor C. Almeida[b,*], and Katja Michael[a,*]

[a] *Department of Chemistry and Biochemistry, and Border Biomedical Research Center, University of Texas at El Paso, 500 West University Avenue, El Paso, Texas, 79968, U.S.A.*

[b] *Department of Biological Sciences, and Border Biomedical Research Center, University of Texas at El Paso, 500 West University Avenue, El Paso, Texas, 79968, U.S.A.*

[c] *Universidad Mayor de San Simón, Faculty of Medicine, and Fundación CEADES, Cochabamba, Bolivia*

[d] *Universidad Nacional de Salta (UNSa)-Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Instituto de Patología Experimental, Facultad de Ciencias de la Salud, Universidad Nacional de Salta, Salta, Argentina*

* Corresponding authors.
E-mail addresses: kmichael@utep.edu (Katja Michael); icalmeida@utep.edu (Igor C. Almeida)

[1] These authors contributed equally to this work.

TABLE OF CONTENTS

GENERAL INFORMATION

All chemicals were purchased as reagent grade from Thermo Fisher Scientific, Sigma-Aldrich, or Acros Organic, and used without further purification. The ACS grade solvents used for reactions were obtained from Thermo Fisher Scientific and they were distilled from the appropriate drying agents. Molecular sieves (3 Å and 4 Å) were purchased from Alfa Aesar and Thermo Fisher Scientific, respectively, and activated under high vacuum and heat prior to use. Reactions were performed under an argon atmosphere, strictly anhydrous conditions and monitored by TLC on silica gel 60 F254 plates from EMD Millipore or Dynamic Adsorbents, Inc. Spots were detected under UV light (254 nm) and/or by charring with 4% sulfuric acid in ethanol. The purification of the compounds was performed by flash column chromatography on silica gel (40-60 μm) from Thermo Fisher Scientific, and the ratio between silica and crude product ranged from 50:1 to 120:1 (dry w/w). FPLC purifications were performed with an AKTA Purifier 100 FPLC system from Cytiva (former GE Healthcare) using a Resource RPC column with a stationary phase of 15 □m polystyrene/divinylbenzene beads, solvent A: 2% CH_3CN/H_2O; solvent B: 85% CH_3CN/H_2O. ^1H and ^{13}C NMR spectra were recorded on a Bruker Avance III HD 400 MHz NMR spectrometer at 400 and 101 MHz. Chemical shifts (in ppm) were determined relative to tetramethylsilane (δ 0.00 ppm) as an internal standard in $CDCl_3$, or relative to the $CDCl_3$ signal (δ 77.0 ppm) in ^{13}C NMR spectra. In case of spectra measured in D_2O, a solution of tetramethylsilane in $CDCl_3$ in a sealed capillary was used as an external standard for calibration. Coupling constant(s) [Hz] were measured from one-dimensional ^1H-NMR spectra. Full or partial assignments were made by 1D spectra as well as standard COSY, HSQC, and TOCSY experiments. In disaccharides and tetrasaccharides, protons of galactopyranose are labeled with an italized "*p*", protons of galactofuranose with an italized "*f*" and protons of mannose with "*m*" . Protons in the allyl group are labeled as "a" for the sp3-hybridized CH_2, "b" for the sp2-hybridized CH, and "c" for the terminal sp2-hybridized CH_2. MS analyses of the carbohydrate derivatives were performed on a high-resolution JEOL AccuTOF mass spectrometer using an electrospray ionization (ESI) source. The thiol-ene reactions were performed in a Rayonet RPR200 photochemical reactor (Southern New England Ultraviolet Company, Branford, CT) equipped with 16 UV lamps (350 nm). Bovine serum albumin (BSA) and BSA derivatives (neoglycoproteins (NGPs) and 2-mercaptoethanol-BSA) were measured by matrix-assisted laser/desorption/ionization mass spectrometer (MALDI)-TOF-MS (MALDI-8020, Shimadzu) using 10 mg/mL sinapinic acid, 0.1% trifluoroacetic acid, in 50% acetonitrile as a matrix. Polystyrene Nunc MaxiSorp 96-well ELISA plates, and chemiluminescent ELISA reagents were purchased from Thermo Fisher Scientific or Jackson ImmunoResearch, and chemiluminescence was recorded on a Luminoskan Ascent, Thermo Fisher Scientific. Optical rotations were measured on an ATAGO AP-300 Automatic Polarimeter.

ABBREVIATIONS USED

Å	angstrom
Abs	antibodies
Ac	acetyl
Ac2O	acetic anhydride
AcSH	thiolacetic acid
AgOTf	silver trifluoromethanesulfonate
AIBN	azobisisobutyronitrile
All	allyl
AllOH	allyl alcohol
Ar	argon
$BF_3 \cdot Et_2O$	boron trifluoride etherate
BMK	biomarker
BSA	bovine serum albumin
Bz	benzoyl
BzCl	benzoyl chloride
CCl_3CN	trichloroacetonitrile
CD	Chagas disease
cELISA	chemiluminescent enzyme-linked immunosorbent assay
CL	cutaneous leishmaniasis
DBU	1,8-diazabicyclo[5.4.0]undec-7-ene
DCM	methylene dichloride
DL	disseminated leishmaniasis
DMAP	4-Dimethylaminopyridine
DPAP	2,2-dimethoxy-2-phenylacetophenone
ELISA	enzyme-linked immunosorbent assay
equiv.	equivalent
ESI-TOF HRMS	electrospray ionization time-of-flight high-resolution mass spectrometry
Et_3N	triethylamine
EtOAc	ethyl acetate
EtSH	ethanethiol
FPLC	Fast protein liquid chromatography
GIPL	glycoinositolphospholipid
GPI	glycosylphosphatidylinositol
GU	glycan unit
MALDI-TOF	matrix-assisted laser desorption ionization time-of-flight
MeOH	methanol
ML	mucosal leishmaniasis
MS	molecular sieves
m/z	mass-to-charge ratio
NGP	neoglycoprotein
NHS	normal human serum
NIS	N-iodosuccinimide
NMR	nuclear magnetic resonance
PCR	polymerase chain reaction

PTLC	preparative thin-layer chromatography
PTSA	para-toluenesulfonic acid
quant.	quantitative
RLU	relative luminescence unit(s)
rt	room temperature
SL	subclinical asymptomatic leishmaniasis
tBu	tert-butyl
TCEP	tris(2-carboxyethyl)phosphine
TFA	trifluoroacetic acid
tGPI-MUC	*Trypanosoma cruzi* trypomastigote-derived glycosylphosphatidylinositol-anchored mucin
TL	tegumentary leishmaniasis
TLC	thin-layer chromatography
TolSH	para-thiocresol
THF	tetrahydrofuran

MALDI-TOF-MS of NGP28b

To determine the average molecular masses of BSA and **NGP28b**, 1 µL of BSA solution (0.4 mg/mL in H_2O) was mixed with 1 µL NGP31b solution (1 mg/mL in H_2O) and 2 µL matrix solution (10 mg/mL sinapinic acid in 50% acetonitrile with 0.1% TFA) in a 1.5-mL microcentrifuge tube. Mass spectra were obtained by matrix-assisted laser/desorption ionization (MALDI-TOF-MS) using a Shimadzu MALDI-8020 MS configured in linear mode. The instrument employed dithering across a scan range of 10,000 to 100,000 *m/z*. The parameters for data acquisition were set as follows: the laser power was adjusted to 110 units; the laser repetition rate was fixed at 50 Hz; each spectrum was the result of 5 accumulated shots, with an additional 2 blast shots to clean the plate; 200 spectral profiles were averaged; the pulse extraction was precisely set to a value of 66431; and a blanking mass was applied to exclude ions below *m/z* 15,000. For data processing, the Threshold Apex software by Shimadzu was utilized, applying a constant threshold value for peak detection, Gaussian smoothing with a filter width of 200 for noise reduction, and setting the peak width at 2 to finely resolve the spectral peaks. Calibration of the mass spectrometer was meticulously performed using the Pierce Bovine Serum Albumin Standard Ampules with a concentration of 2 mg/mL (provided by Thermo Fisher Scientific, catalog number 23209). An internal calibration was established by setting the mass of the $[BSA + H]^+$ ion at *m/z* 66,402 with an exceptionally tight mass tolerance of 5 ppm to ensure high precision in mass measurements (Fig. S1).

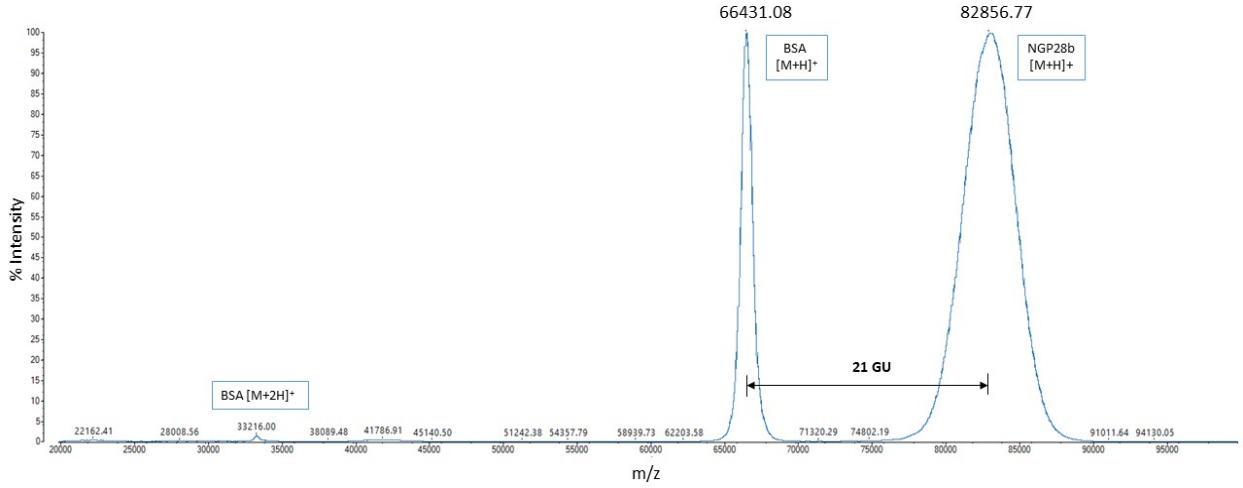

Figure S1. MALDI-TOF mass spectrum of **NGP28b** overlaid with the mass spectrum of pure BSA. *m/z*, mass-to-charge ratio. BSA $[M+H]^+$ = 66,431 *m/z*; for **NGP28b** $[M+H]^+$ = 82,856 *m/z*. The average glycan load of **G28** units (including the linkers) per BSA molecule was 21.

Cross-titration ELISA for NGP28b, NGP31b, and 2ME-BSA varying antigen quantity [ng/well] and dilution of pooled sera

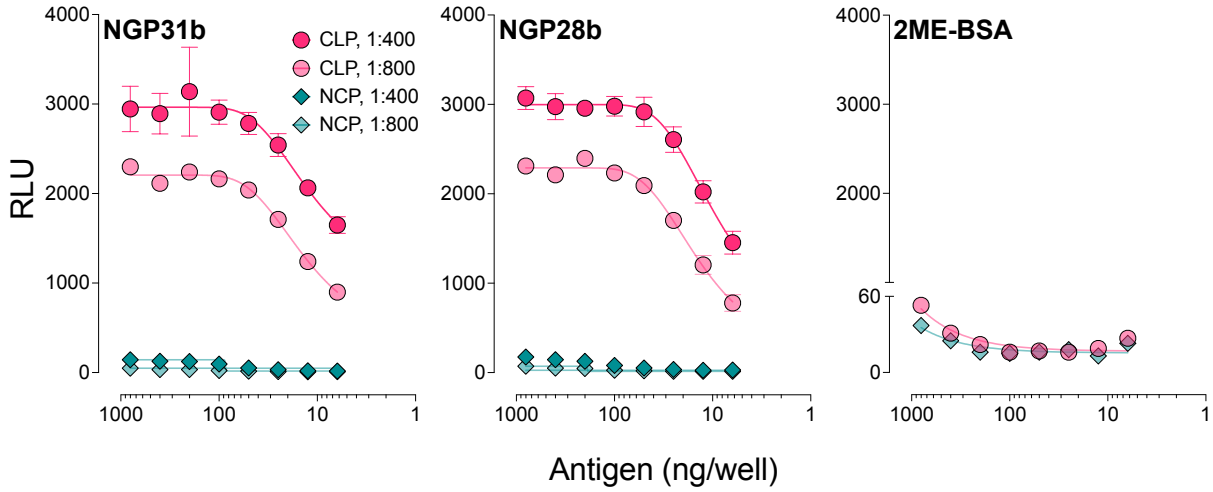

Figure S2. Initial antigen-serum cross-titration of **NGP31b**, **NGP28b**, and **2ME-BSA** measured by cELISA. Both NGPs and **2ME-BSA** (negative antigen control) were immobilized at 800, 400, 200, 100, 50, 25, 12.5, and 6.25 ng/well, and assayed in duplicate with pools of sera (n=10 each) from individuals with cutaneous leishmaniasis (CLP) or healthy negative controls (NCP) from endemic areas for CL and CD, at 1:400 and 1:800 dilutions. RLU, relative luminescence units.

^{1}H NMR spectrum of compound 4

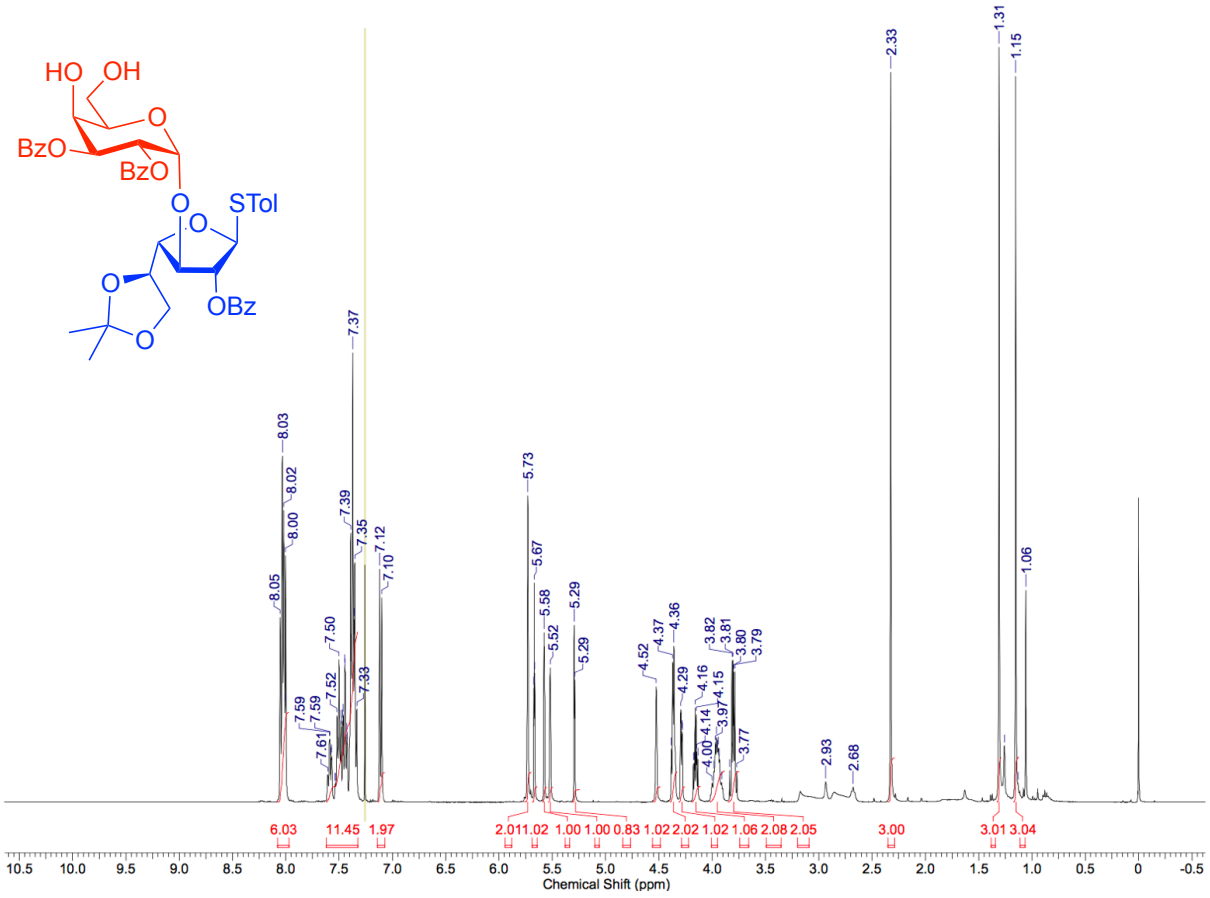

Figure S3. ^{1}H NMR spectrum of compound **4** in $CDCl_3$, 400 MHz.

^{13}C NMR spectrum of compound 4

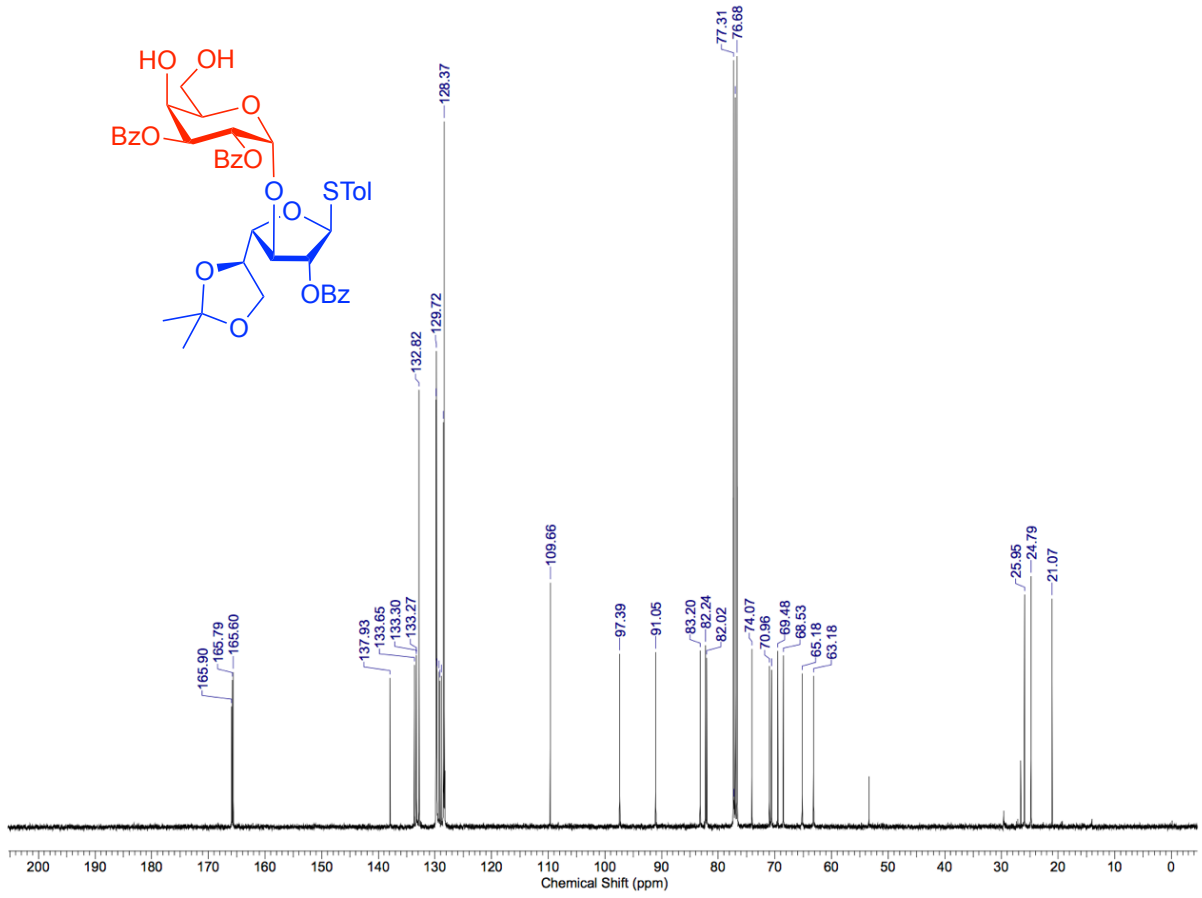

Figure S4. ^{13}C NMR spectrum of compound **4** in $CDCl_3$, 100 MHz.

Mass spectrum of compound 4

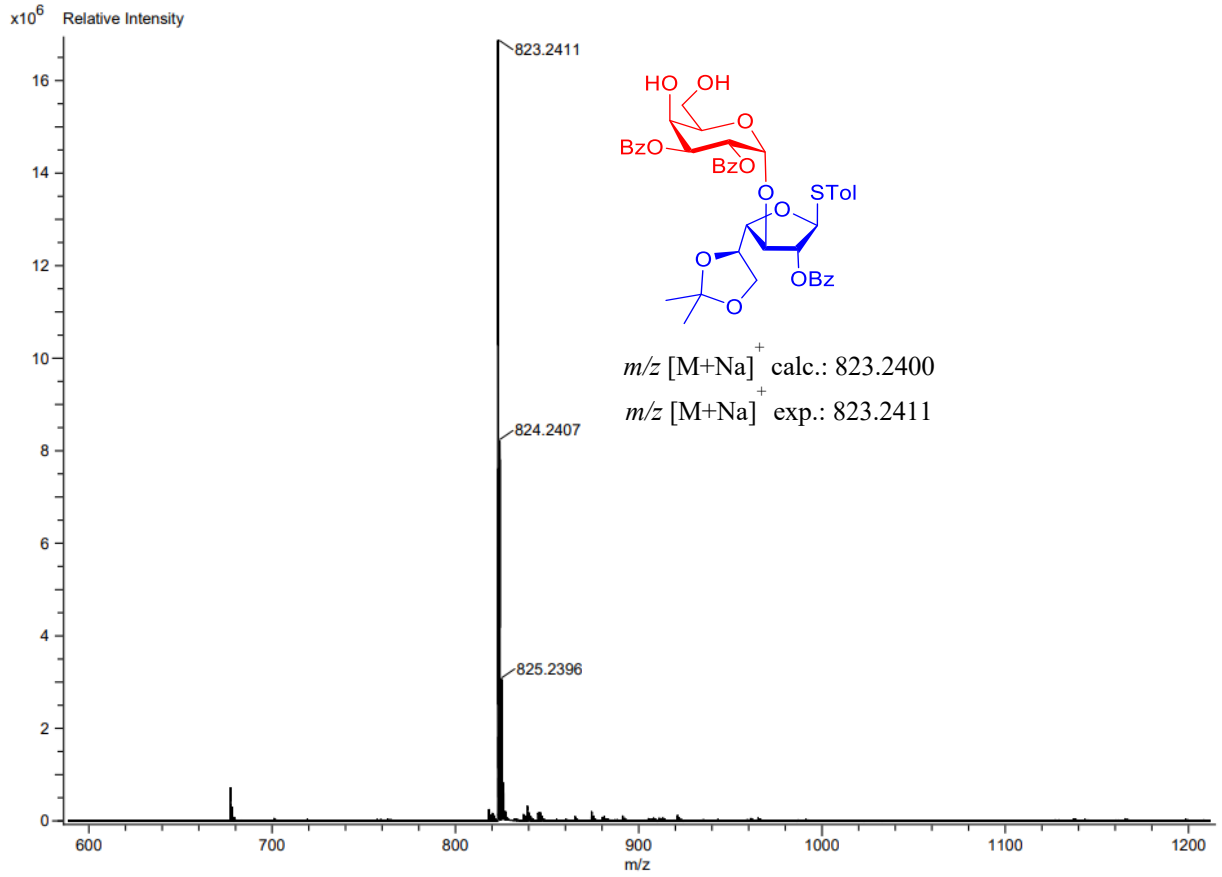

Figure S5. ESI-TOF HR MS of compound **4**

^{1}H NMR spectrum of compound 5

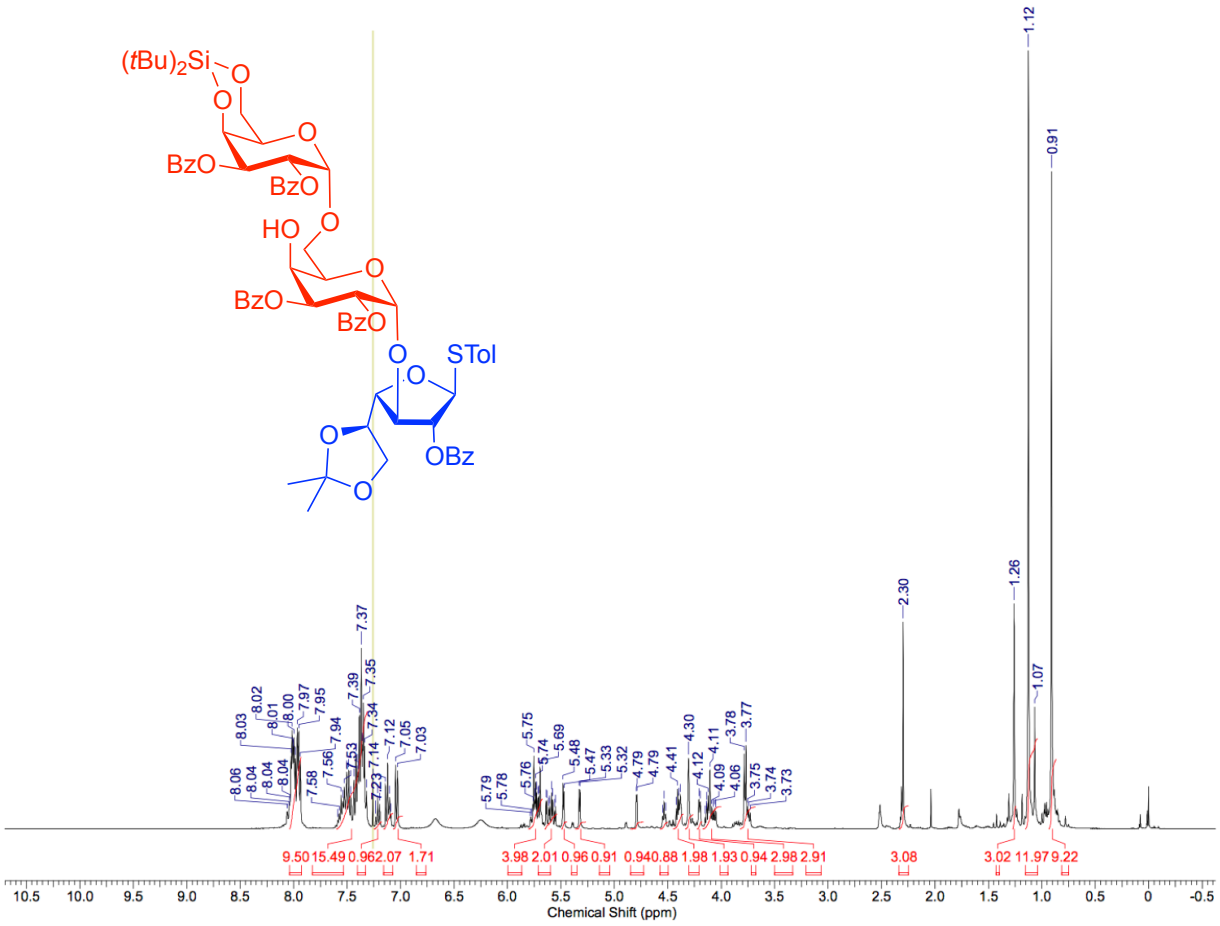

Figure S6. ^{1}H NMR spectrum of compound 5 in $CDCl_3$, 400 MHz.

^{13}C NMR spectrum of compound 5

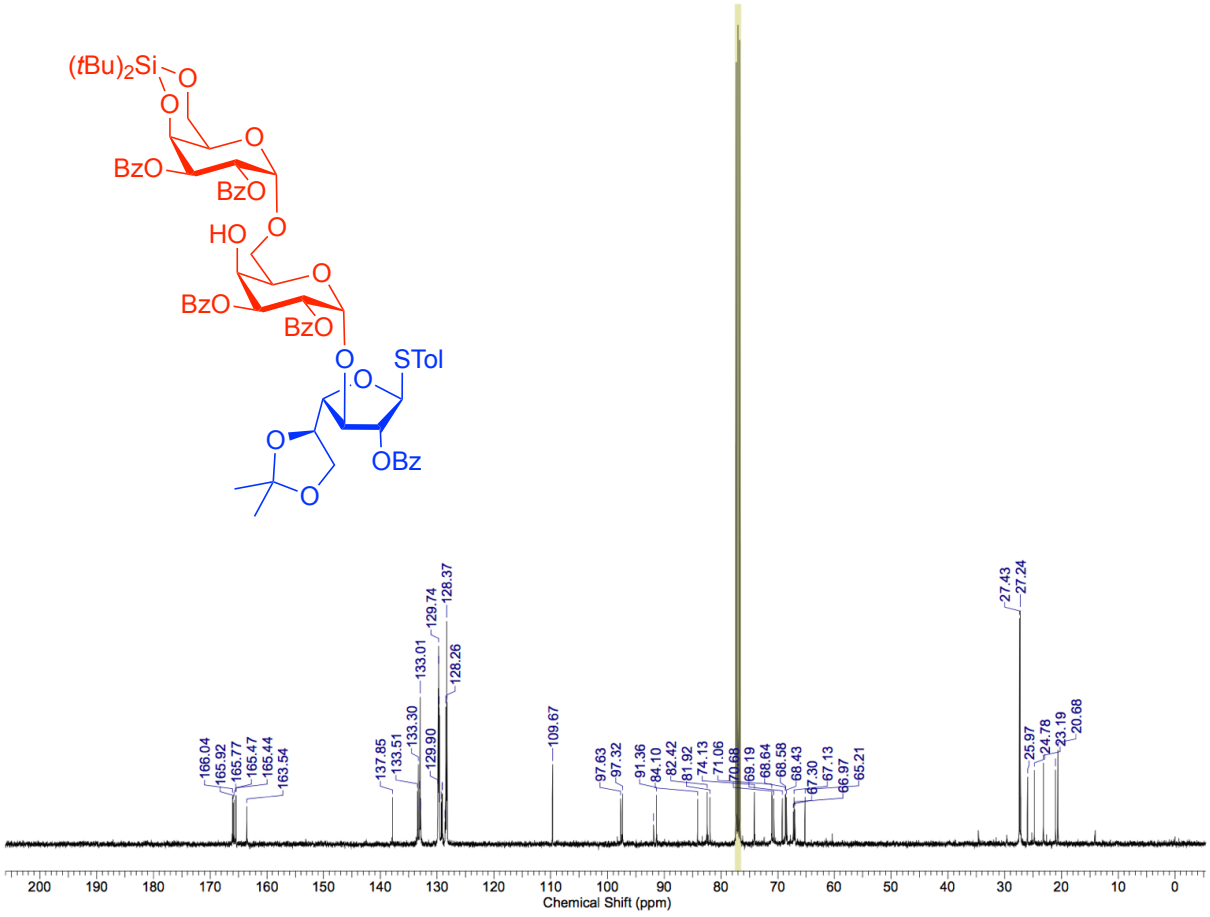

Figure S7. ^{13}C NMR spectrum of compound **5** in $CDCl_3$, 100 MHz.

Mass spectrum of compound 5

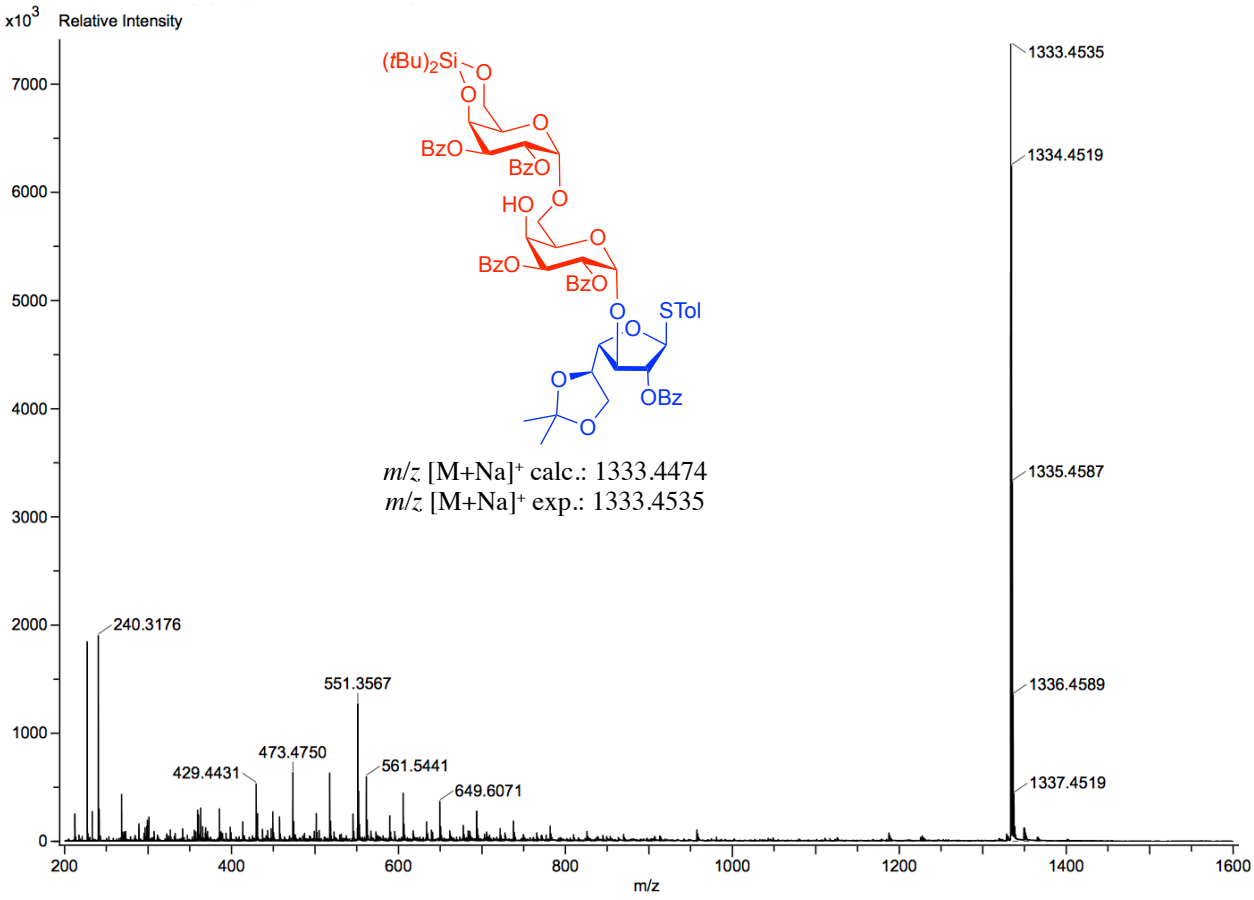

Figure S8. ESI-TOF HR MS of compound **5**.

^{1}H NMR spectrum of compound 7

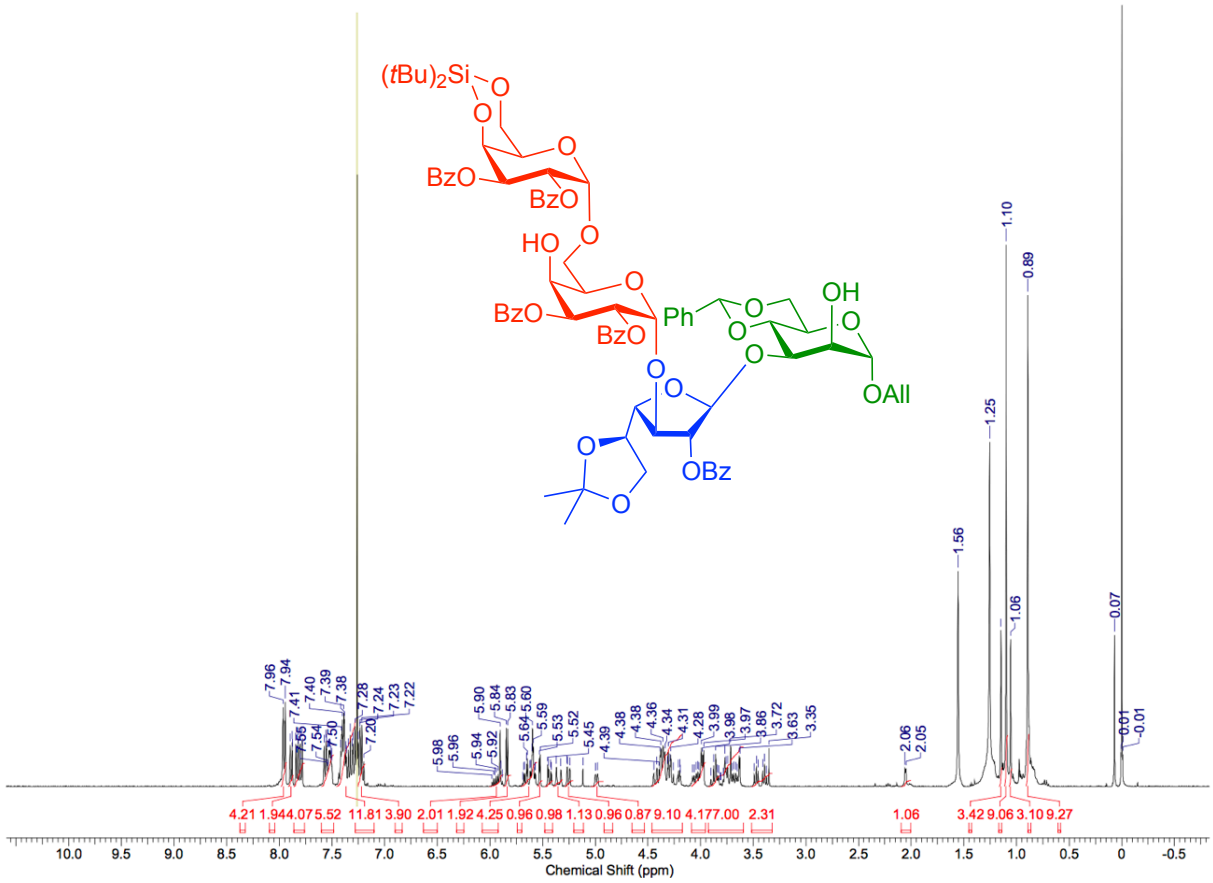

Figure S9. ^{1}H NMR spectrum of compound **7** in $CDCl_3$, 400 MHz.

^{13}C NMR spectrum of compound 7

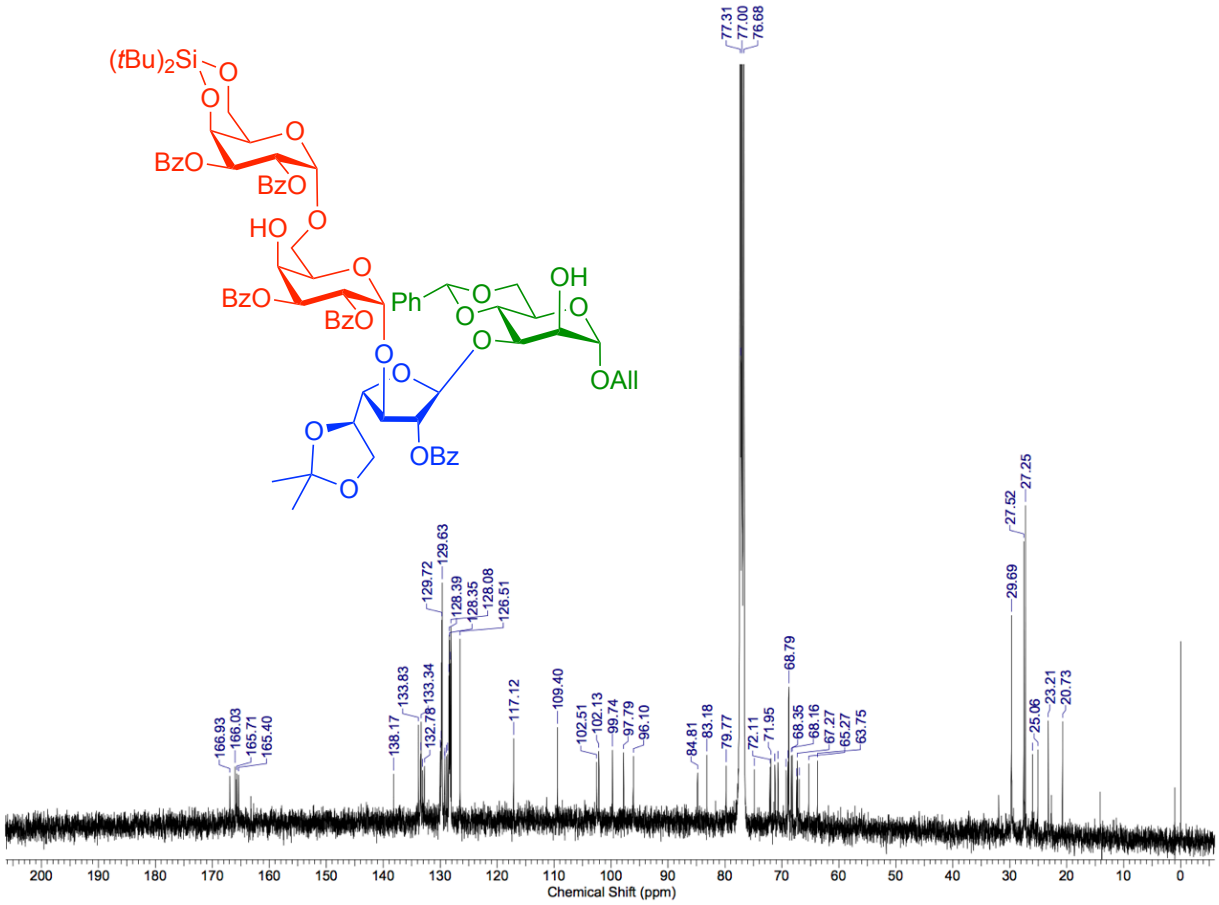

Figure S10. ^{13}C NMR spectrum of compound **7** in $CDCl_3$, 100 MHz.

Mass spectrum of compund **7**

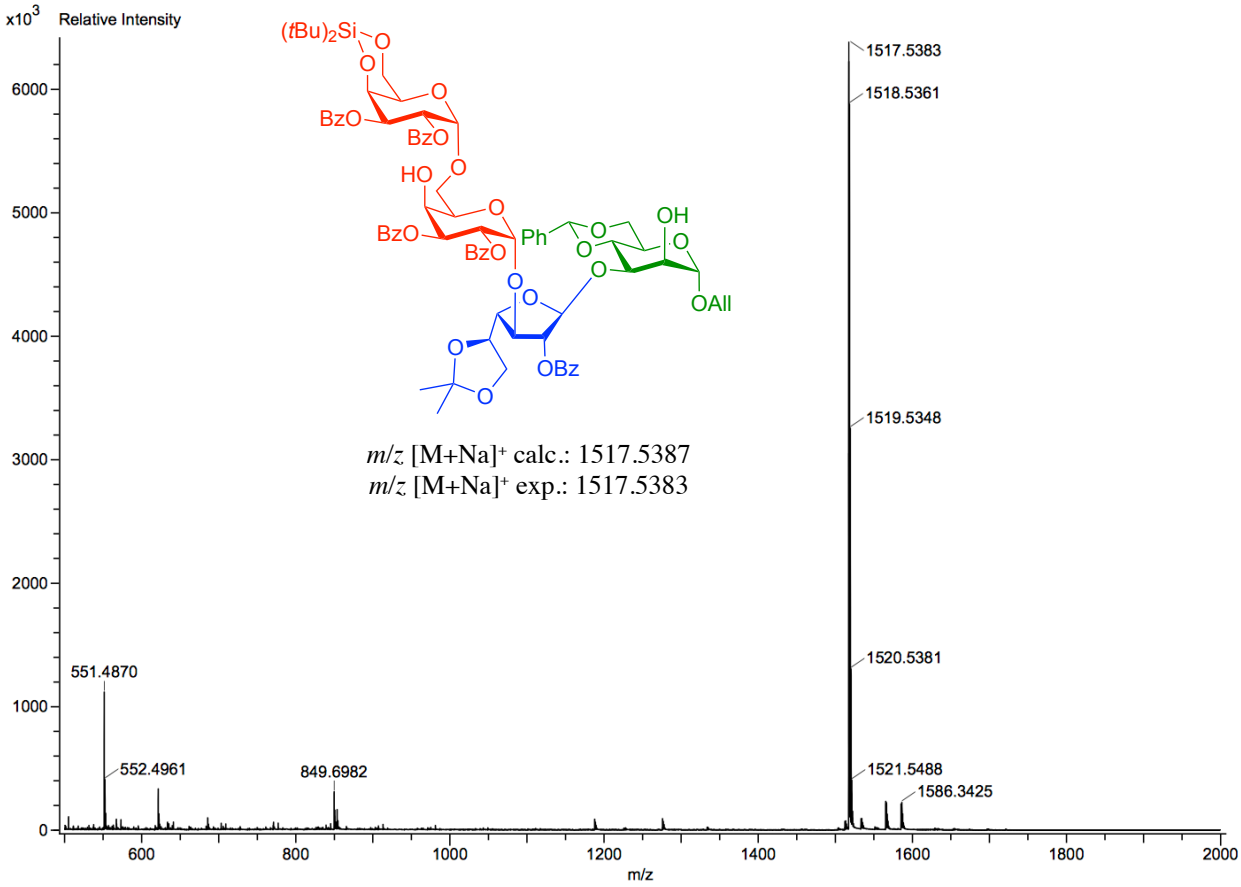

Figure S11. ESI-TOF HR MS of compound **7**.

^{1}H NMR spectrum of compound 8

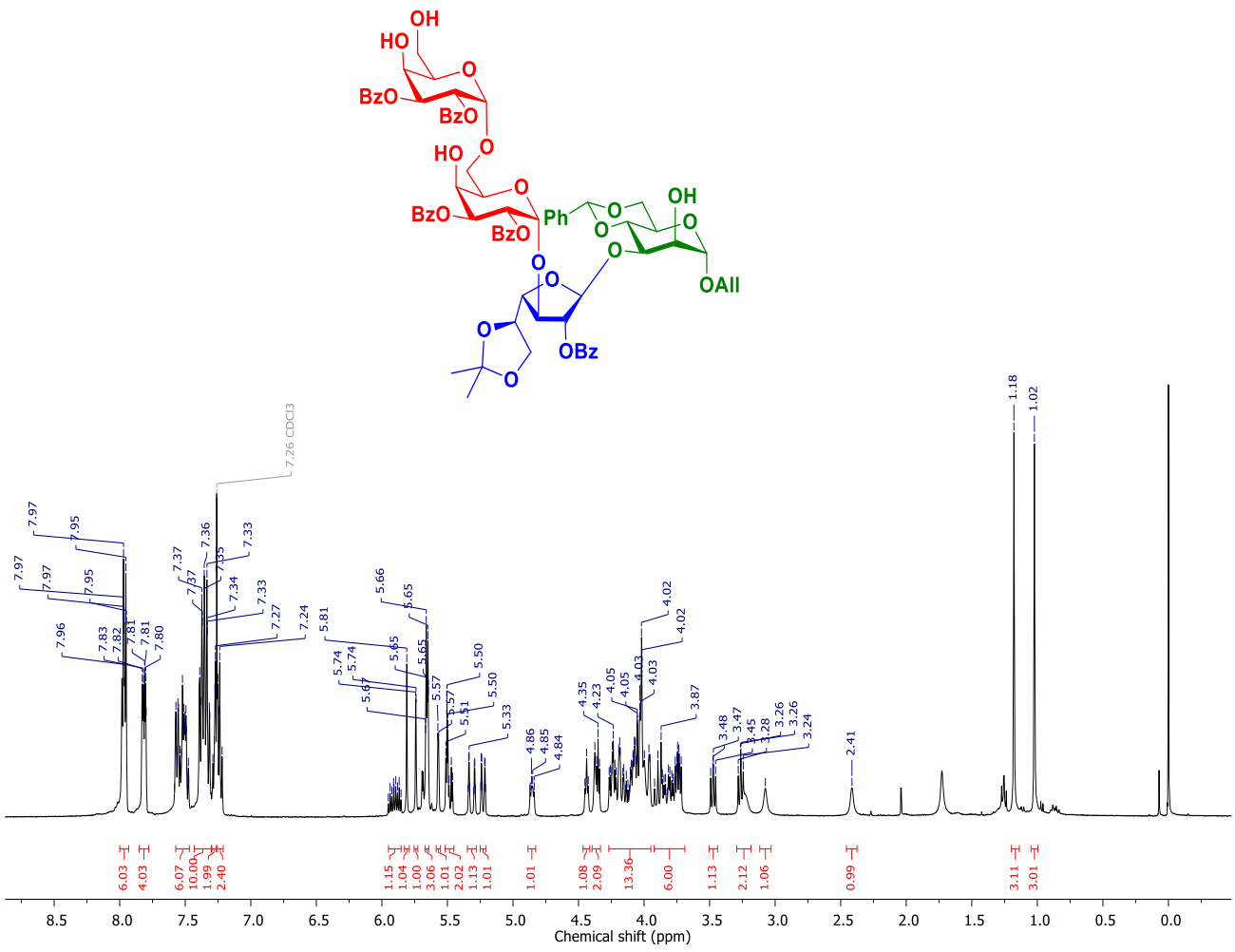

Figure S12. ^{1}H NMR spectrum of compound **8** in $CDCl_3$, 400 MHz.

^{13}C NMR spectrum of comound **8**

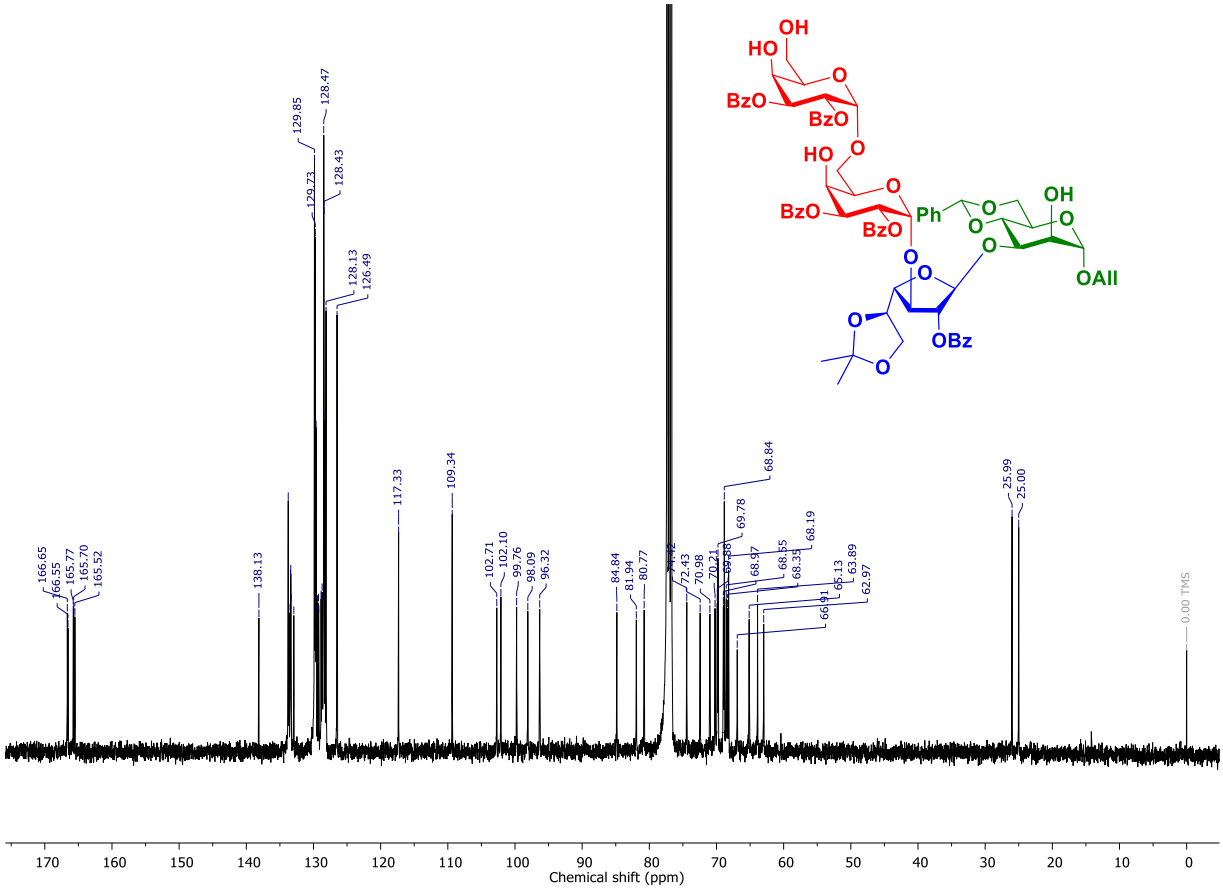

Figure S13. ^{13}C NMR spectrum of compound **8** in $CDCl_3$, 100 MHz.

COSY NMR spectrum of compound 8

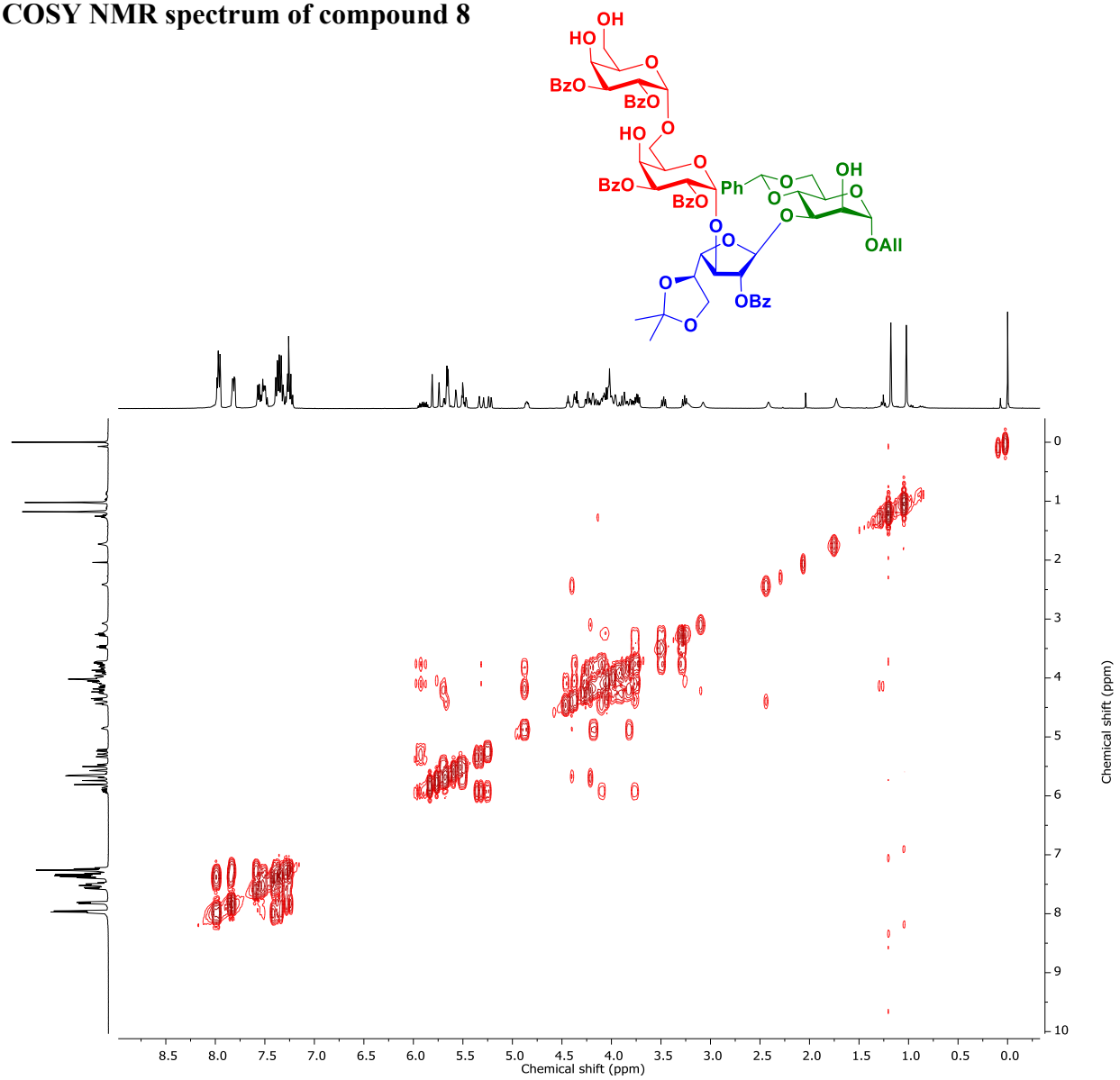

Figure S14. COSY NMR spectrum of compound **8** in $CDCl_3$, 400 MHz.

HSQC NMR spectrum of compound 8

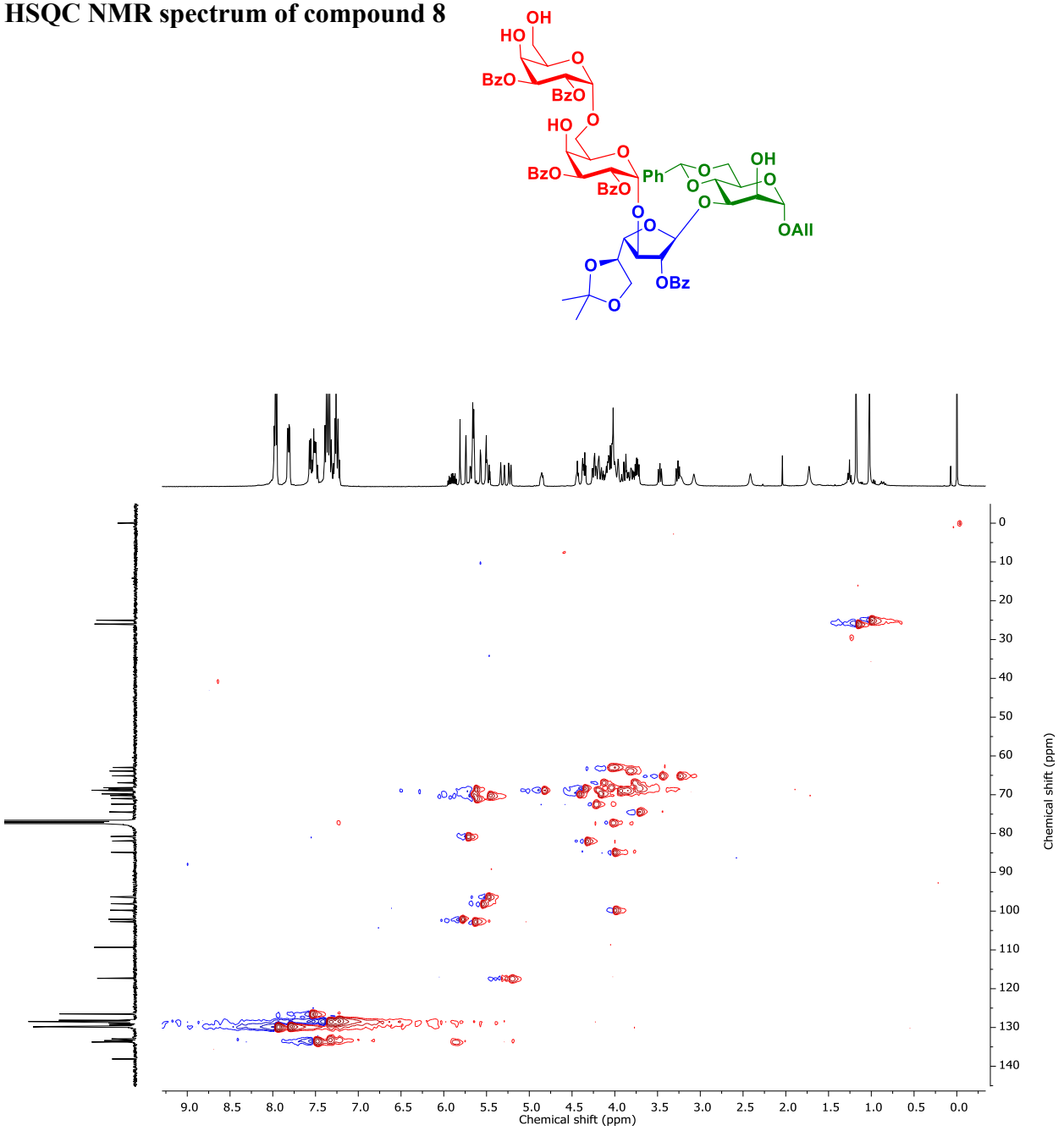

Figure S15. HSQC NMR spectrum of compound **8** in $CDCl_3$.

Mass spectrum of compound 8

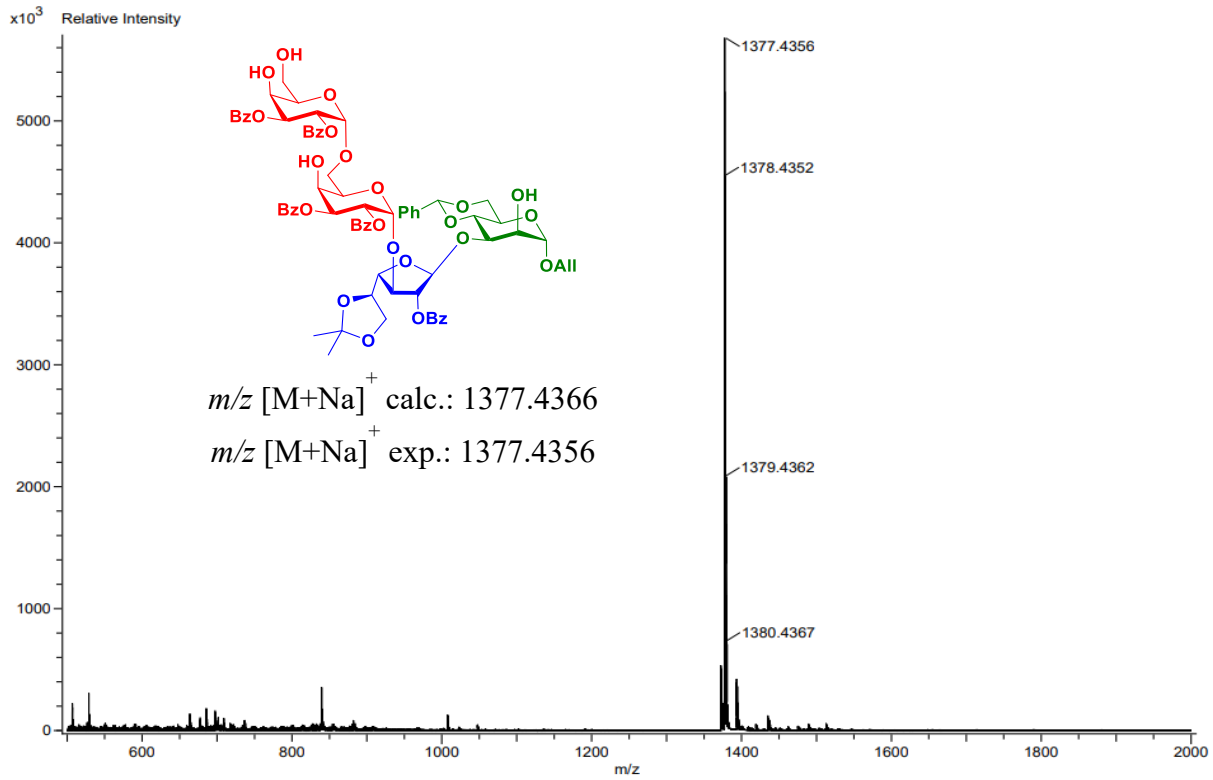

Figure S16. ESI-TOF HR MS of compound **8**.

^{1}H NMR spectrum of compound 9

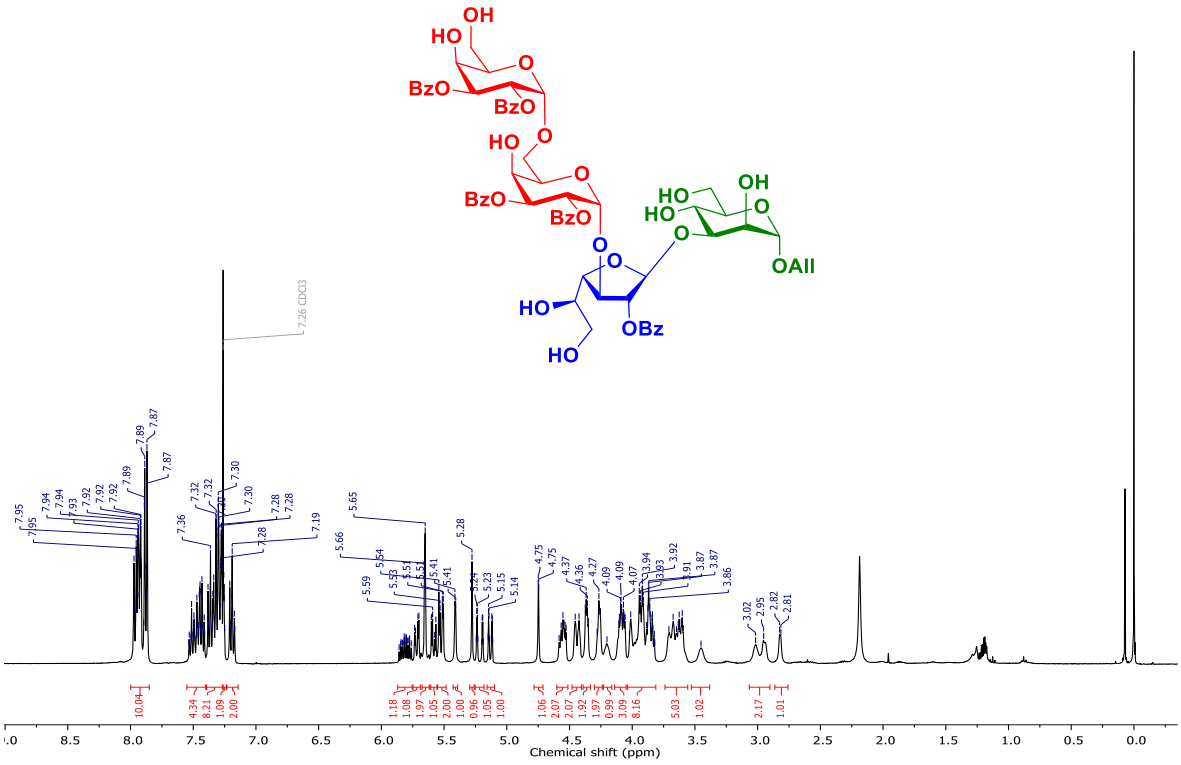

Figure S17. ^{1}H NMR spectrum of compound **9** in $CDCl_3$, 400 MHz.

^{13}C NMR spectrum of compound 9

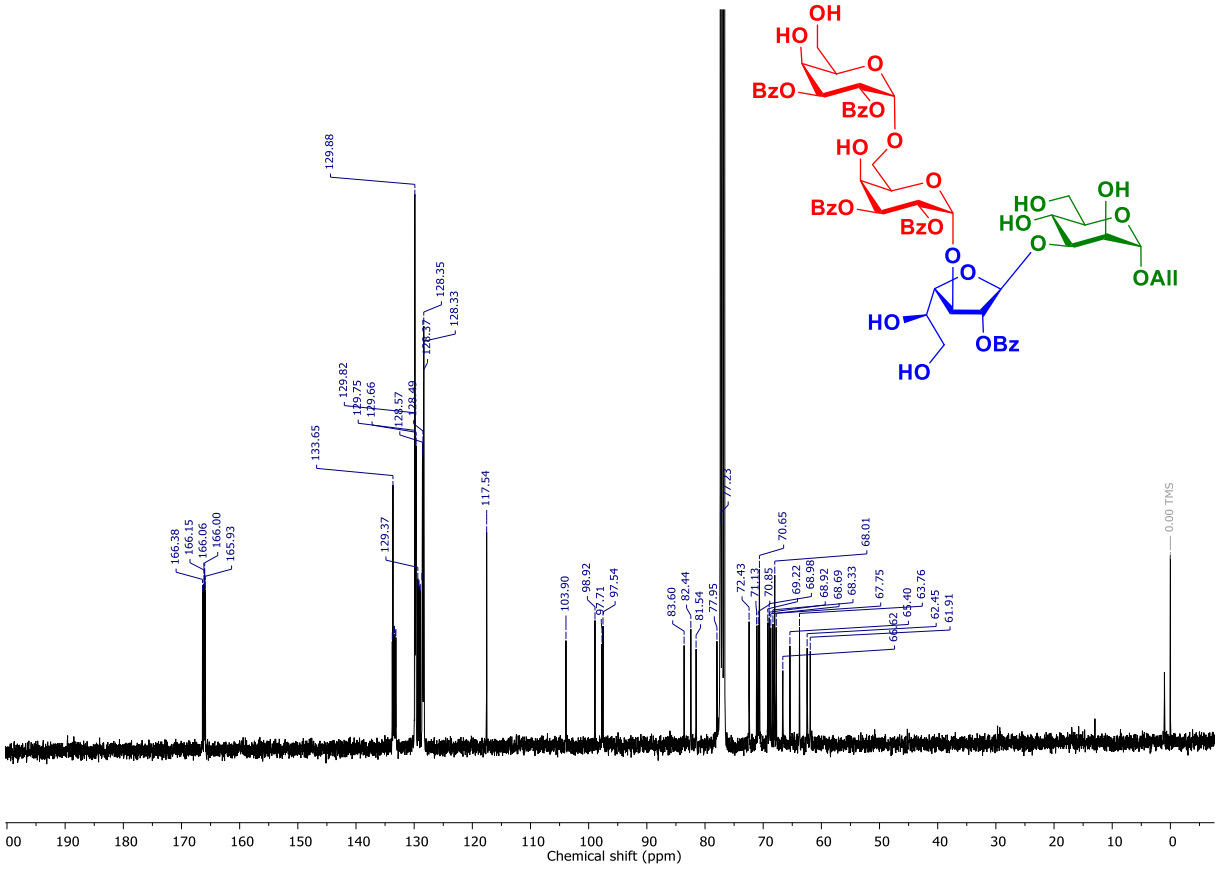

Figure S18. ^{13}C NMR spectrum of compound **9** in $CDCl_3$, 100 MHz.

HSQC NMR spectrum of compound 9

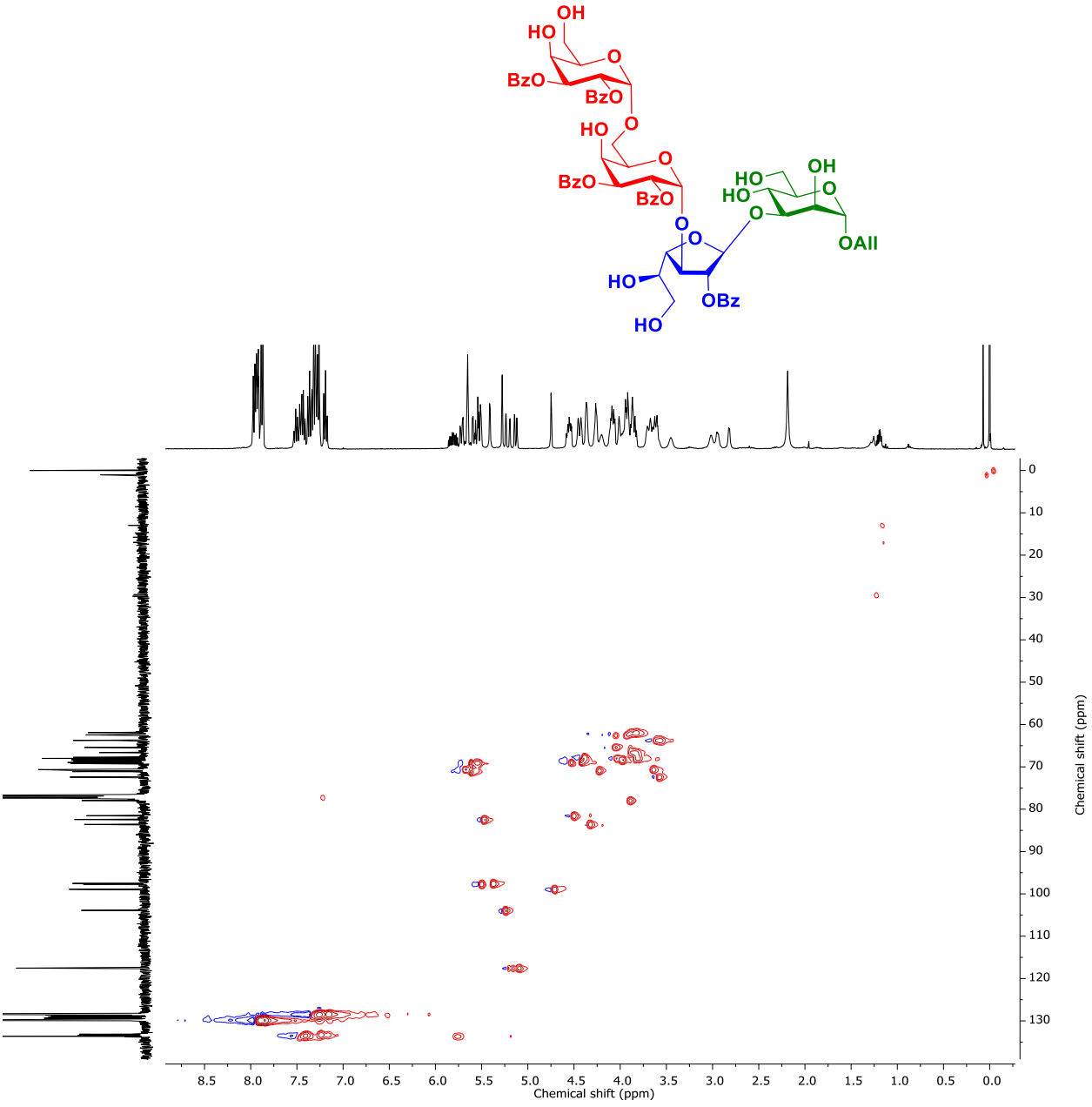

Figure S19. HSQC NMR spectrum of compound **9** in $CDCl_3$.

Mass spectrum of compound 9

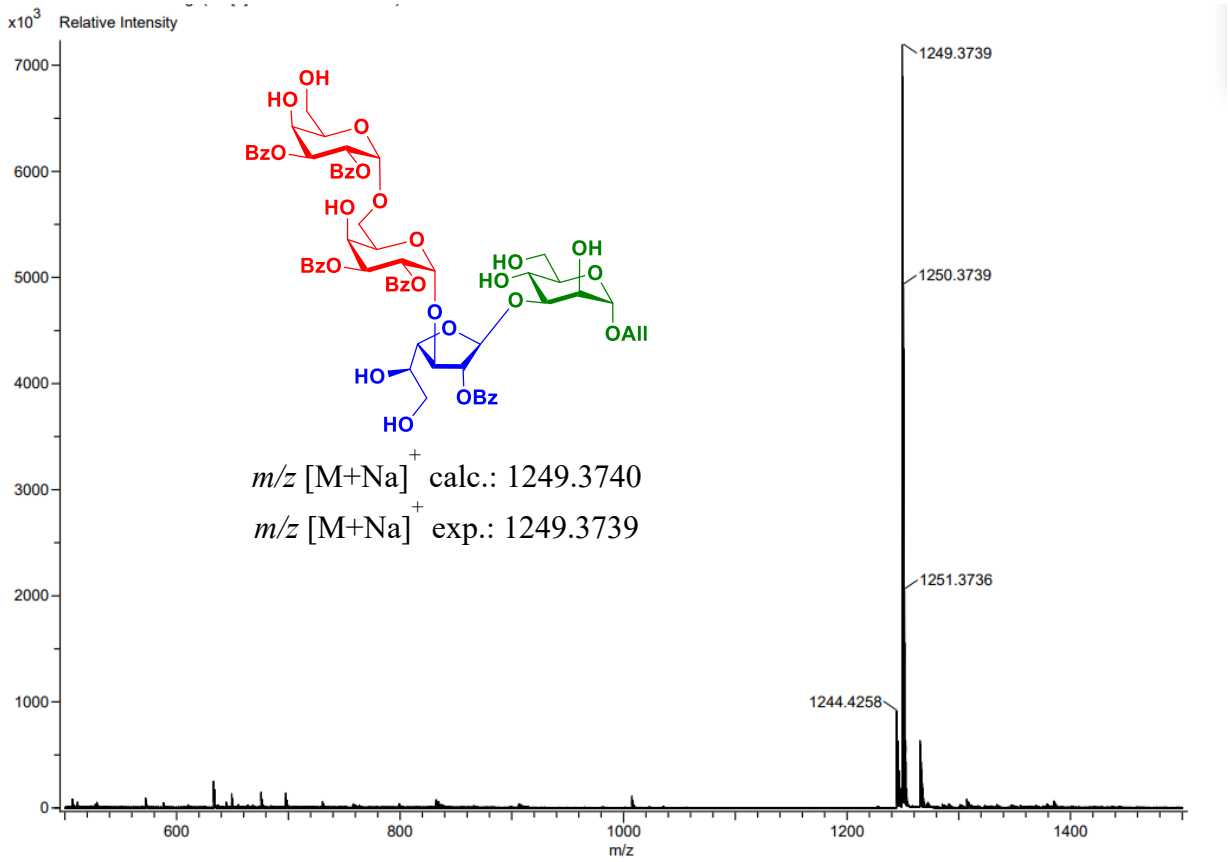

Figure S20. ESI-TOF HR MS of compound **9**

^{1}H NMR spectrum of compound 10

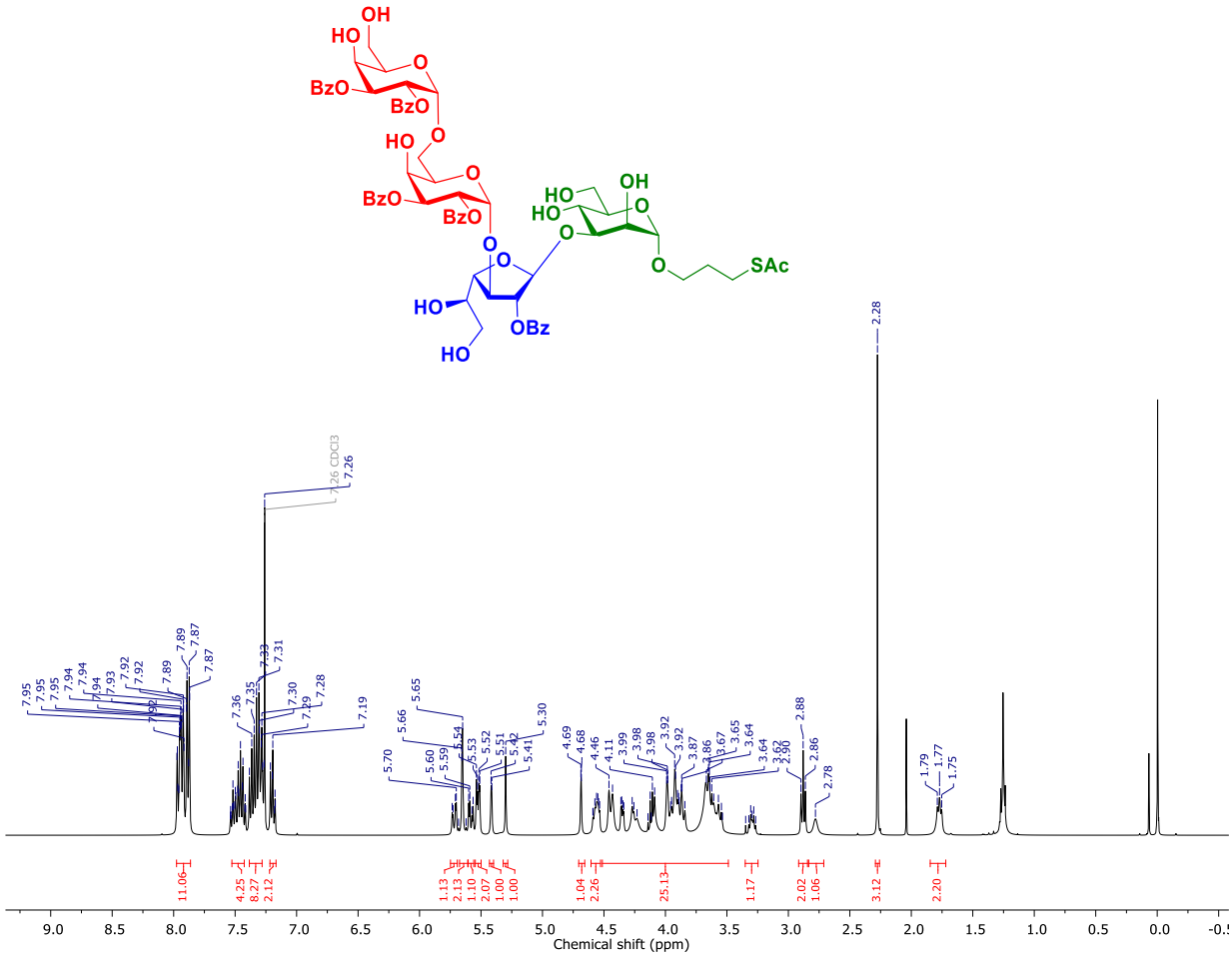

Figure S21. ^{1}H NMR spectrum of compound **10** in $CDCl_3$, 400 MHz.

^{13}C NMR spectrum of compound 10

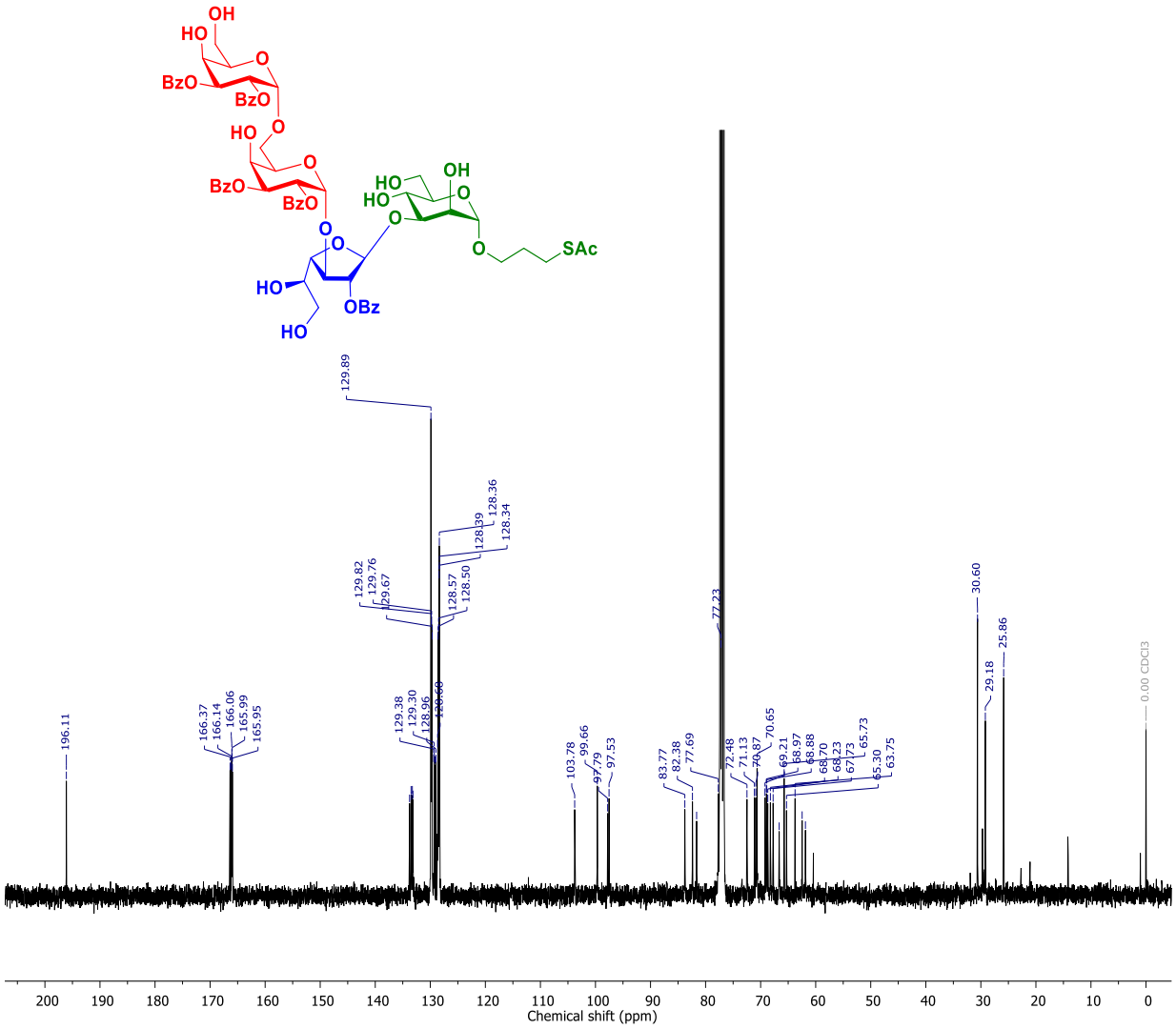

Figure S22. ^{13}C NMR spectrum of compound **10** in $CDCl_3$, 100 MHz.

HSQC NMR spectrum of compound 10

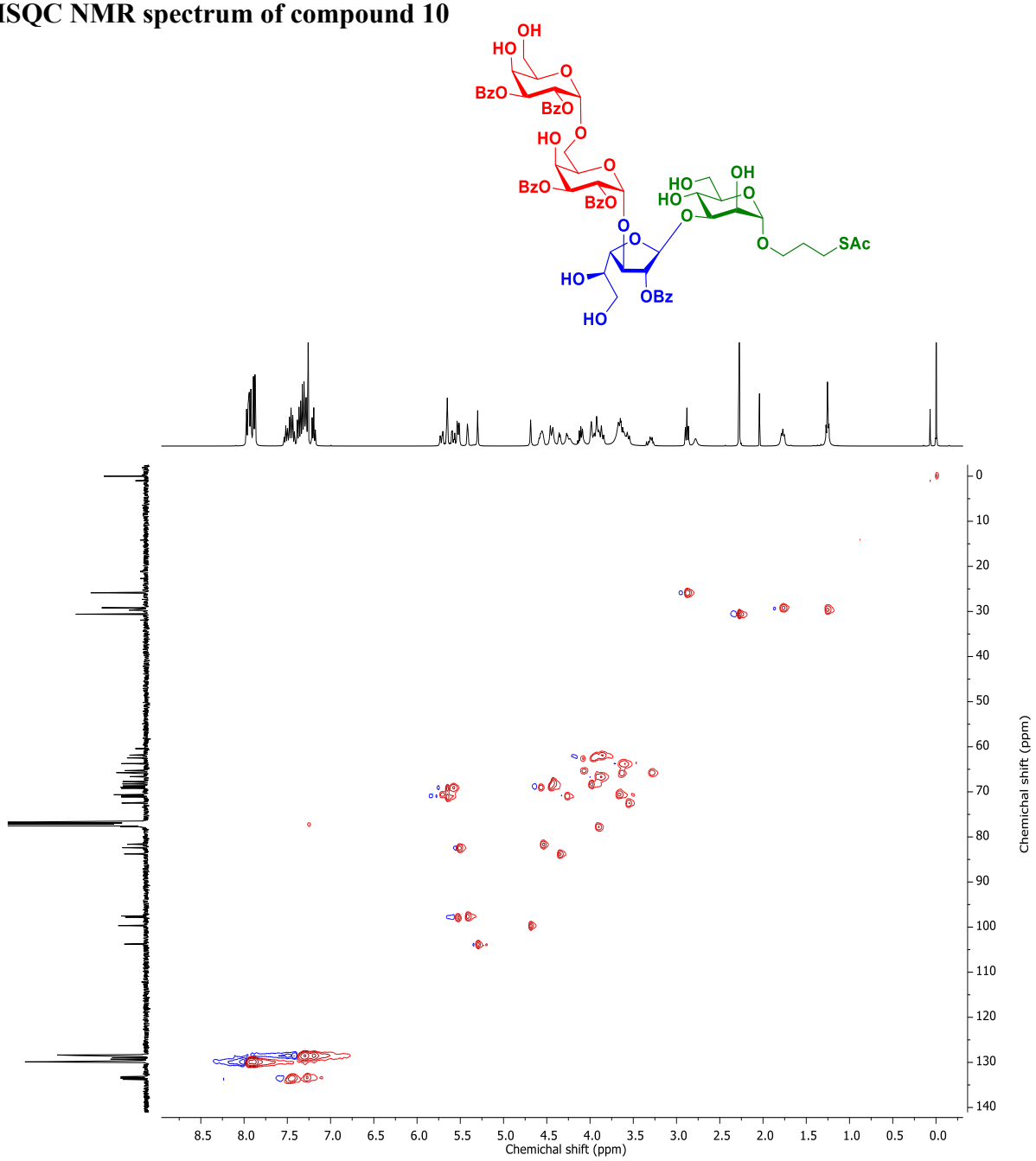

Figure S23. HSQC NMR spectrum of compound **10** in $CDCl_3$.

Mass spectrum of compound 10

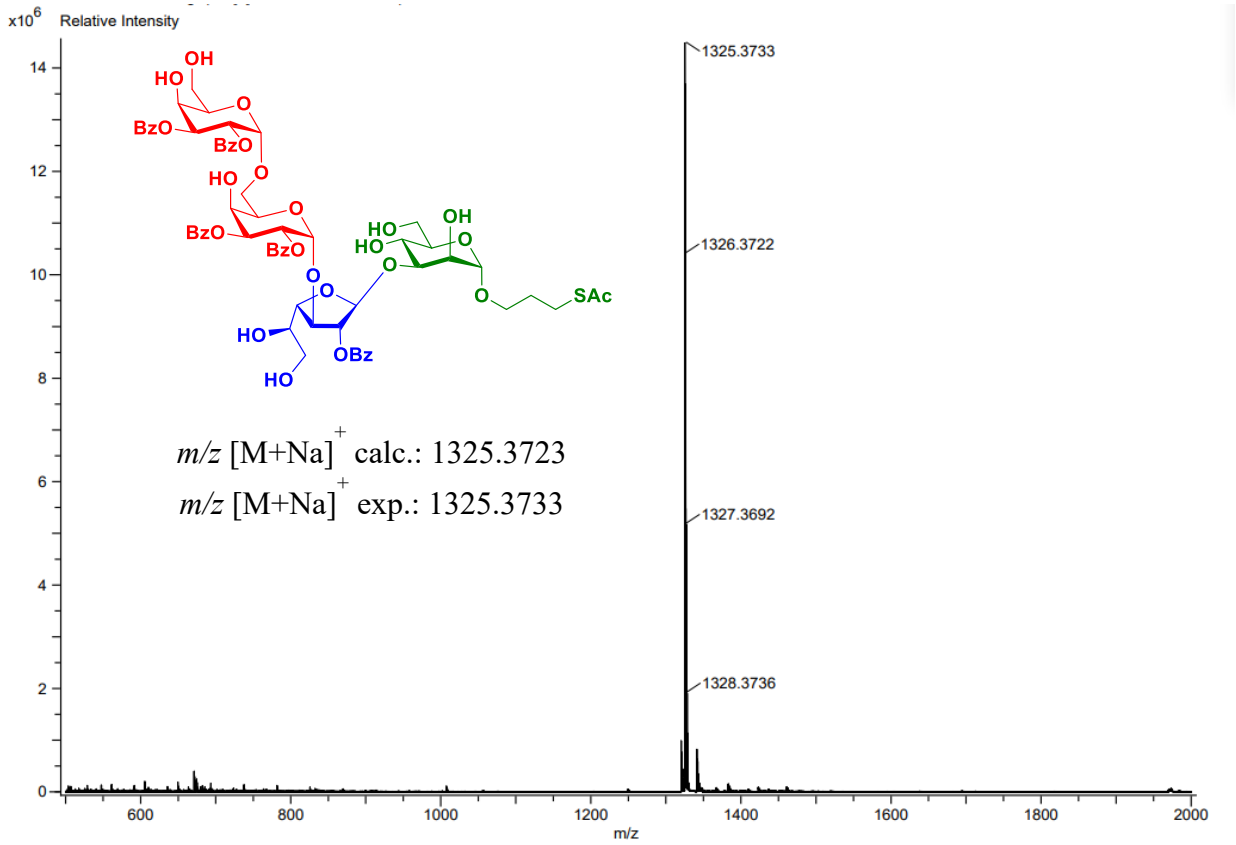

Figure S24. ESI-TOF HR MS of compound **10**

^{1}H NMR spectrum of compound G31$_{SH}$

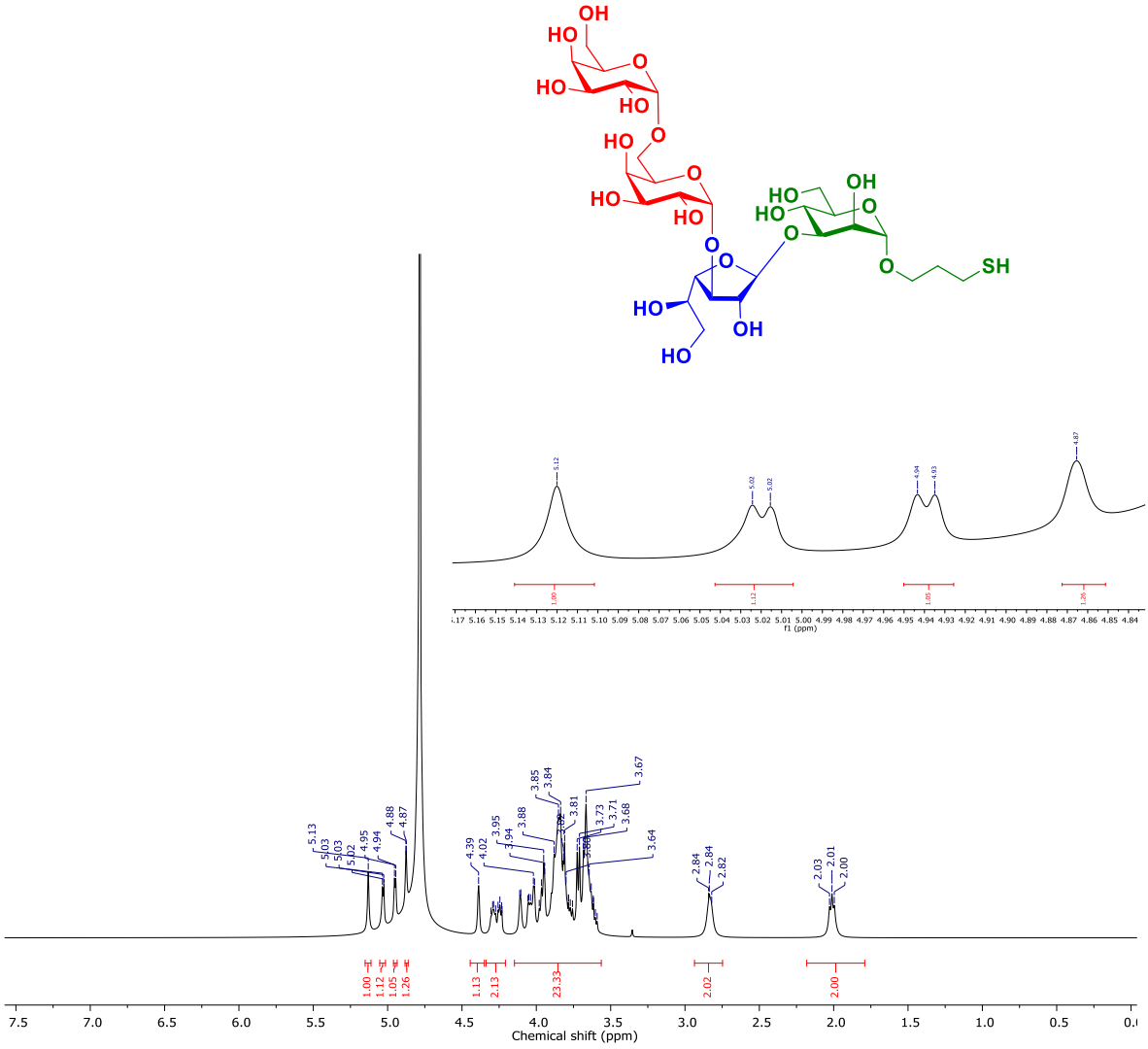

Figure S25. ^{1}H NMR spectrum of compound **G31$_{SH}$** in D_2O, 400 MHz. The insert shows the region of the anomeric protons. The disulfide form **(G31$_S$)$_2$** is likely to be present as well.

^{13}C NMR spectrum of compound G31$_{SH}$

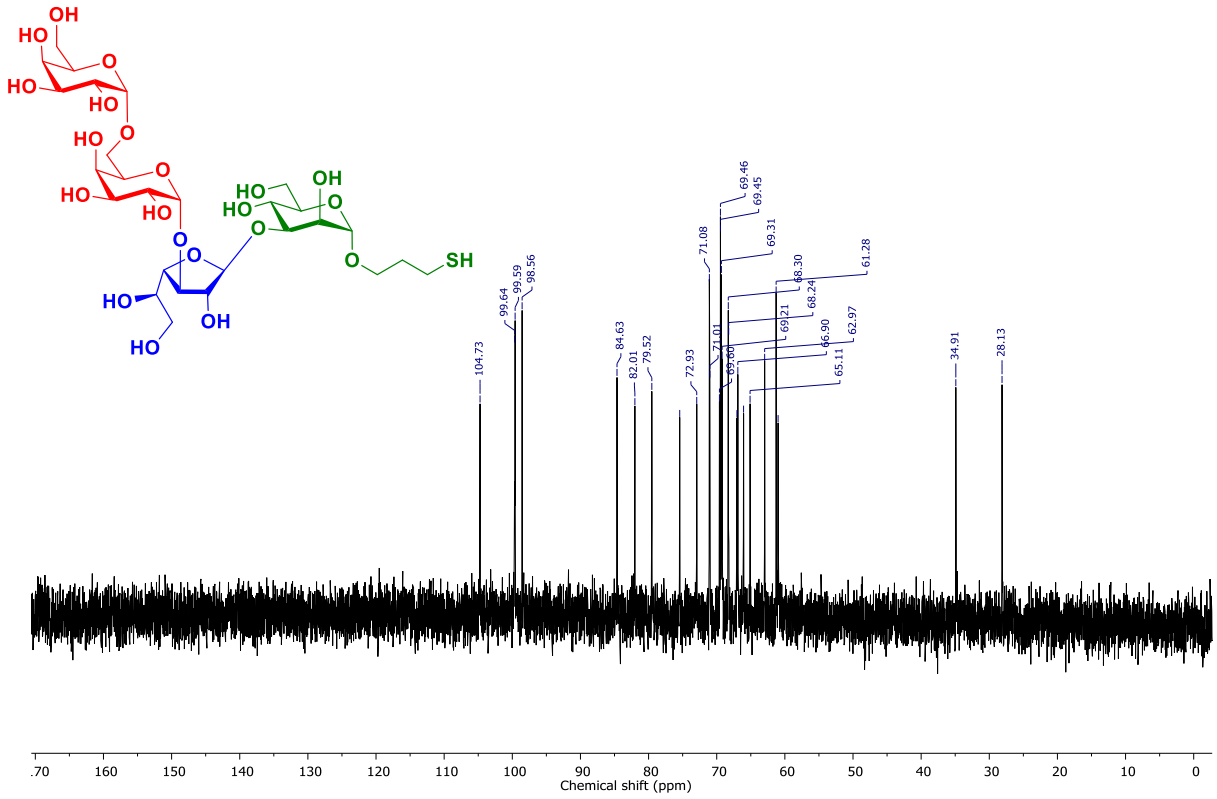

Figure S26. ^{13}C NMR spectrum of compound **G31$_{SH}$** in D_2O, 100 MHz. The disulfide form **(G31$_S$)$_2$** is likely to be present as well.

HSQC NMR sectrum of $G31_{SH}$

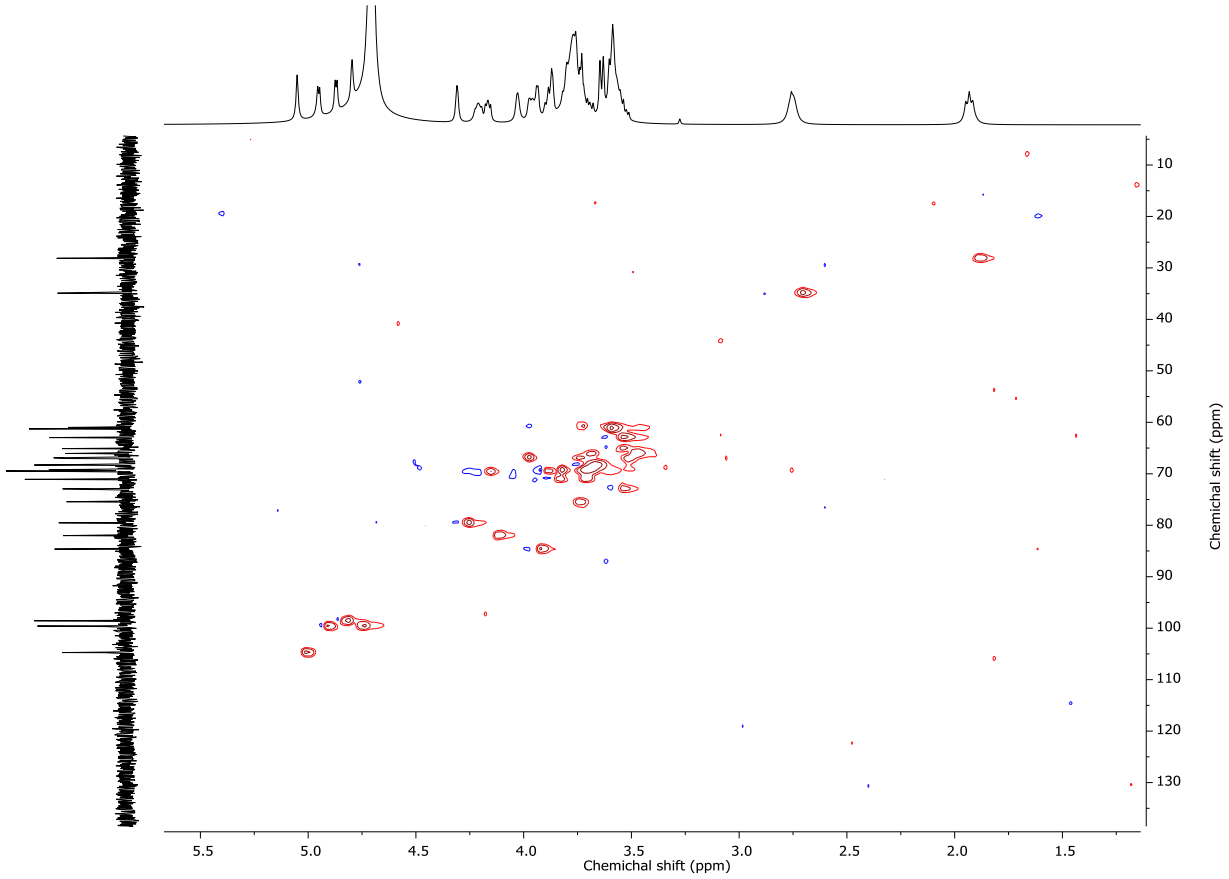

Figure S27. HSQC NMR spectrum of **$G31_{SH}$** in D_2O. The disulfide form **$(G31_S)_2$** is likely to be present as well.

Mass spectrum of compound G31$_{SH}$

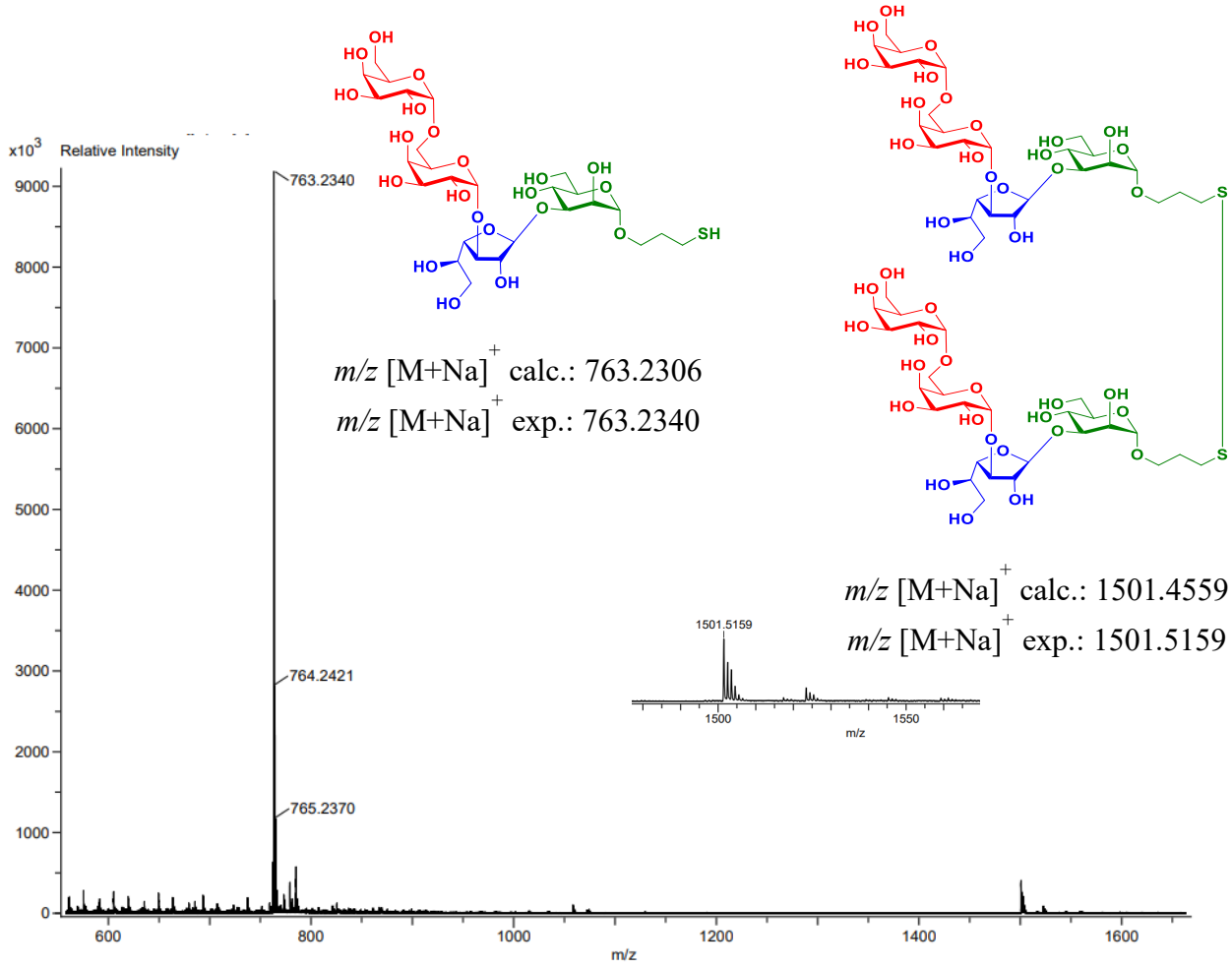

Figure S28. ESI-TOF HR MS of compound **G31$_{SH}$**. The disulfide **(G31$_S$)$_2$** is also present.

Supplementary Table 1. Tegumentary Leishmaniasis Cohort			
Sample #	**Gender**	**Age**	**Country**
X-Le002-A001-0066	F	30	Bolivia
X-Le002-A001-0112	F	30	Bolivia
6117	F	46	Argentina
6631	F	41	Argentina
X-Le002-A001-0008	M	38	Bolivia
X-Le002-A001-0011	M	21	Bolivia
X-Le002-A001-0012	M	37	Bolivia
X-Le002-A001-0013	M	36	Bolivia
X-Le002-A001-0024	M	24	Bolivia
X-Le002-A001-0038	M	49	Bolivia
X-Le002-A001-0108	M	37	Bolivia
5915	M	77	Argentina
5924	M	46	Argentina
6179	M	48	Argentina
6185	M	39	Argentina
6203	M	57	Argentina
6520	M	22	Argentina
6575	M	23	Argentina
6583	M	14	Argentina
6591	M	42	Argentina
6597	M	42	Argentina
6607	M	33	Argentina
6609	M	70	Argentina
6611	M	53	Argentina
6614	M	52	Argentina
6618	M	66	Argentina
6629	M	28	Argentina
6632	M	14	Argentina

Supplementary Table 2. Chagas Disease Cohort

Patient #	Gender	Age	Country
T2100505	F	37	Bolivia
N2100467	F	38	Bolivia
E2100478	F	25	Bolivia
V2100488	F	57	Bolivia
V2100491	F	48	Bolivia
B2100504	F	33	Bolivia
B2100505	F	35	Bolivia
R2100513	F	48	Bolivia
A2100523	F	40	Bolivia
F2100524	F	53	Bolivia
E2100438	F	44	Bolivia
R2100470	M	40	Bolivia
S2100485	M	52	Bolivia
M2100489	M	20	Bolivia
R2100493	M	48	Bolivia
R2100497	M	58	Bolivia
R2100496	M	46	Bolivia
S1900019	M	53	Bolivia
S2100427	M	62	Bolivia
A2100429	M	62	Bolivia
E2100432	M	37	Bolivia
B2100434	M	32	Bolivia
H2100435	M	47	Bolivia
A2100440	M	35	Bolivia
G2100493	M	35	Bolivia
F2100495	M	42	Bolivia
R2100497	M	75	Bolivia
S2100501	M	38	Bolivia

Supplementary Table 3. Negative Controls Cohort (non-TL, non-CD)

Patient #	Gender	Age	Country
V2100499	F	28	Bolivia
E2100508	F	30	Bolivia
C2100514	F	53	Bolivia
I1900024	F	18	Bolivia
F1900037	F	22	Bolivia
C1900094	F	18	Bolivia
H1900128	F	29	Bolivia
C1200439	F	31	Bolivia
C2100441	F	29	Bolivia
S2100526	F	43	Bolivia
Q2100408	F	27	Bolivia
R2100383	F	36	Bolivia
H2100384	F	16	Bolivia
P2100488	F	66	Bolivia
C2100474	M	28	Bolivia
M2100490	M	34	Bolivia
C2100502	M	29	Bolivia
T2100525	M	48	Bolivia
M1900165	M	24	Bolivia
J2100425	M	28	Bolivia
H2100433	M	30	Bolivia
M2100442	M	47	Bolivia
C2100498	M	32	Bolivia
M2100500	M	35	Bolivia
P2100520	M	28	Bolivia
R2100524	M	51	Bolivia
C2100540	M	34	Bolivia
M2100542	M	31	Bolivia