Major human schistosome species express different glycans with immunological and diagnostic implications.

Laudine M. C. Petralia[1,2], Angela van Diepen[1], Tao Zhang[3], Anna O. Kildemoes[1], Tom Veldhuizen[1], D. Linh Nguyen[1], Cornelis H. Smit[1], Peter G. Kremsner[4,5], Manfred Wuhrer[3], Jeremy M. Foster[2], Cornelis H. Hokke[1]*

Supplementary figures

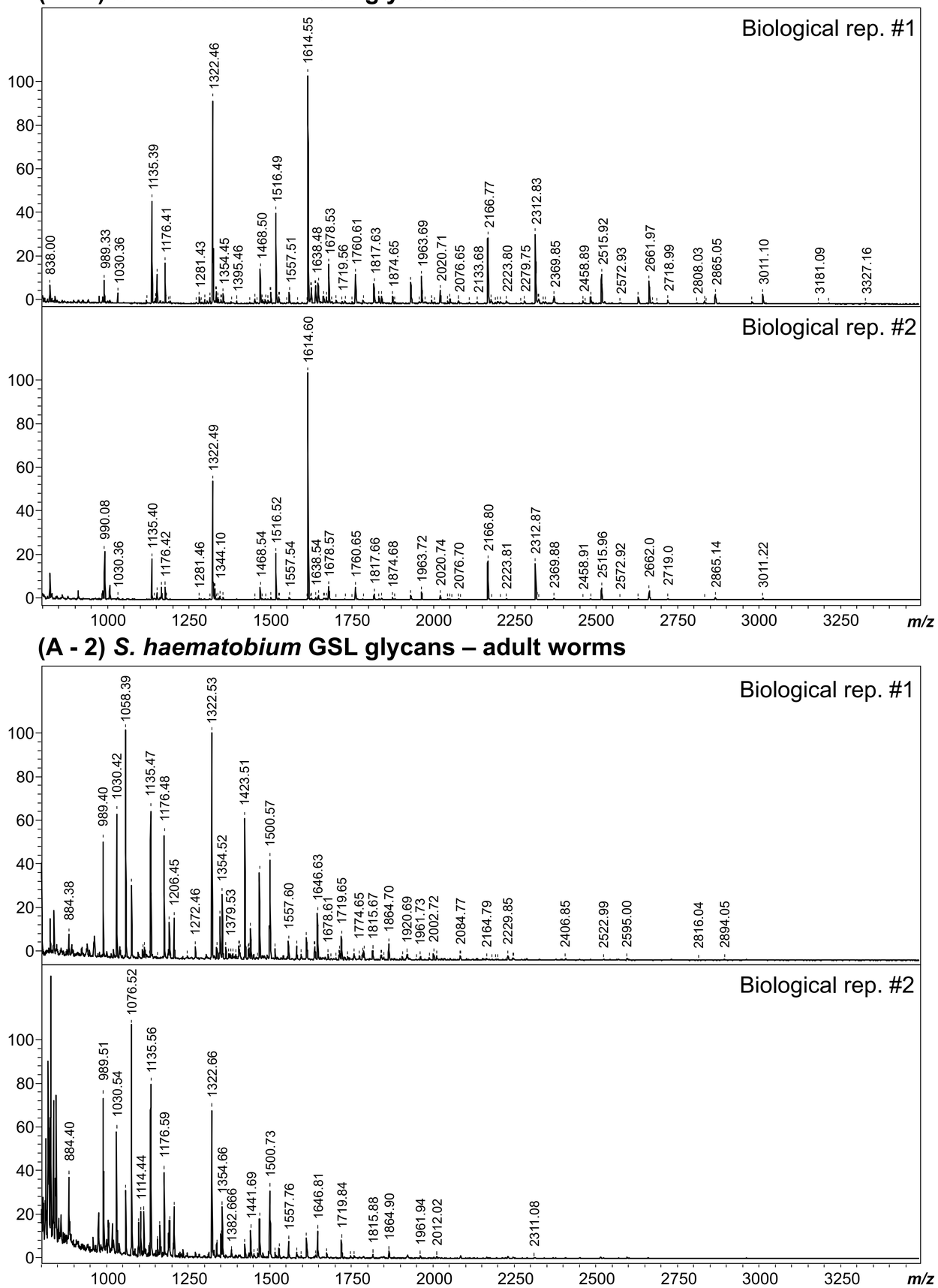
Figure S1 – MALDI-TOF-MS spectra of *S. haematobium* GSL derived, *N*-linked and *O*-linked glycans

S. haematobium (glyco-)lipids were treated with rEGCase II to release GSL glycans (A) while *N*-glycans were released from *S. haematobium* glycoproteins using PNGase F (B). Enzymatically released glycans were labeled with AA and analyzed using MALDI-TOF-MS. GSL glycan and *N*-glycan spectra were acquired in negative-ion reflectron mode. All signals are labeled with monoisotopic masses (*m/z,* $[M-H]^-$). Signal intensities in % are indicated on the Y-axis. Biological duplicates were generated for GSL and *N*-linked glycans of *S. haematobium* cercariae, adult worms and eggs (A-B).

O-glycans were released from *S. haematobium* glycoproteins by β-elimination and permethylated prior to MALDI-TOF-MS analysis (C). Spectra were recorded in positive-ion reflectron mode and monoisotopic masses of measured signals are indicated ($[M+Na]^+$). Signal intensities in % are indicated on the Y-axis. Technical duplicates were generated.

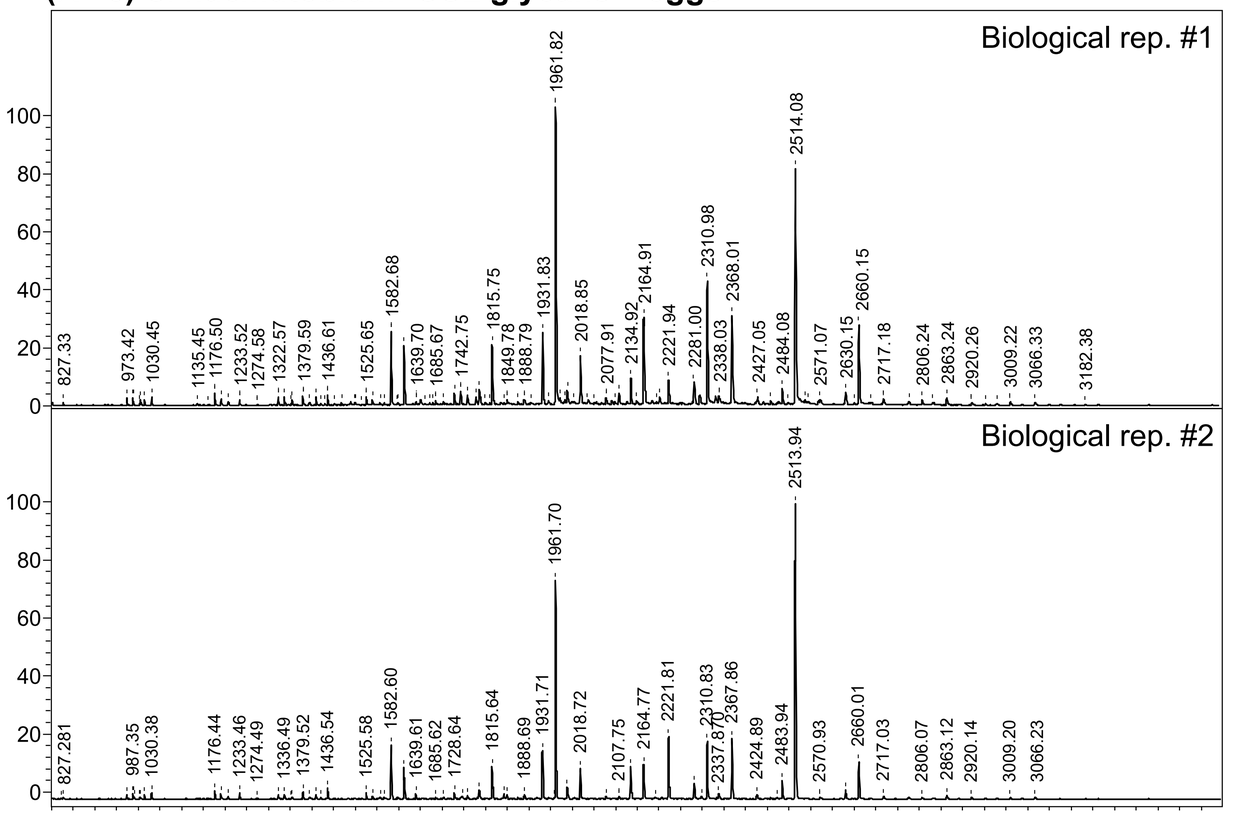
Glycan class and parasite life-stage from which glycans were extracted are indicated at the top of each panel: cercariae, adult worms or eggs. Known non-glycan signals are labeled with the # symbol. Raw MALDI-TOF-MS data for all spectra can be found in Supplementary Data 1.

(A - 1) *S. haematobium* GSL glycans – cercariae

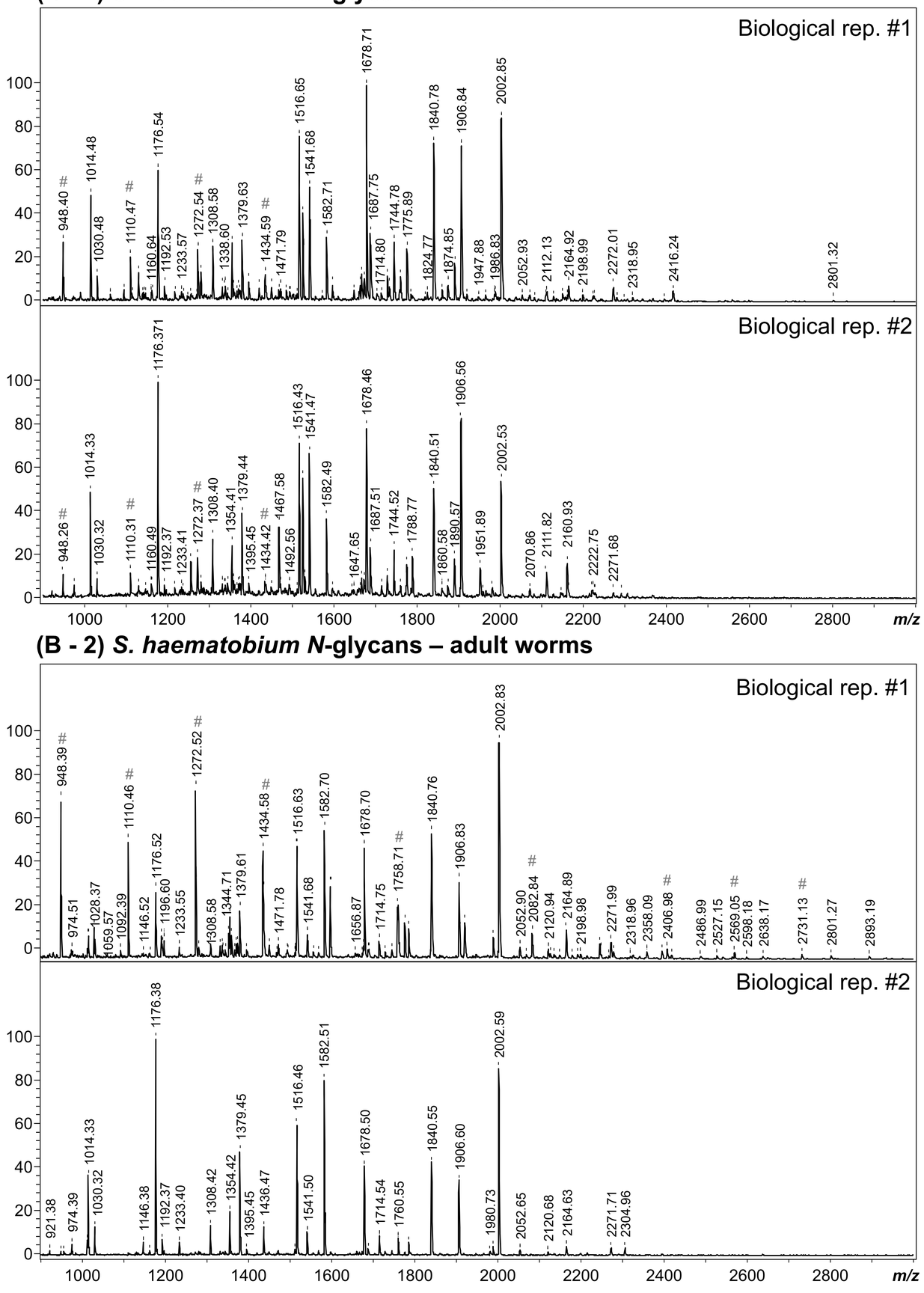


(A - 2) *S. haematobium* GSL glycans – adult worms

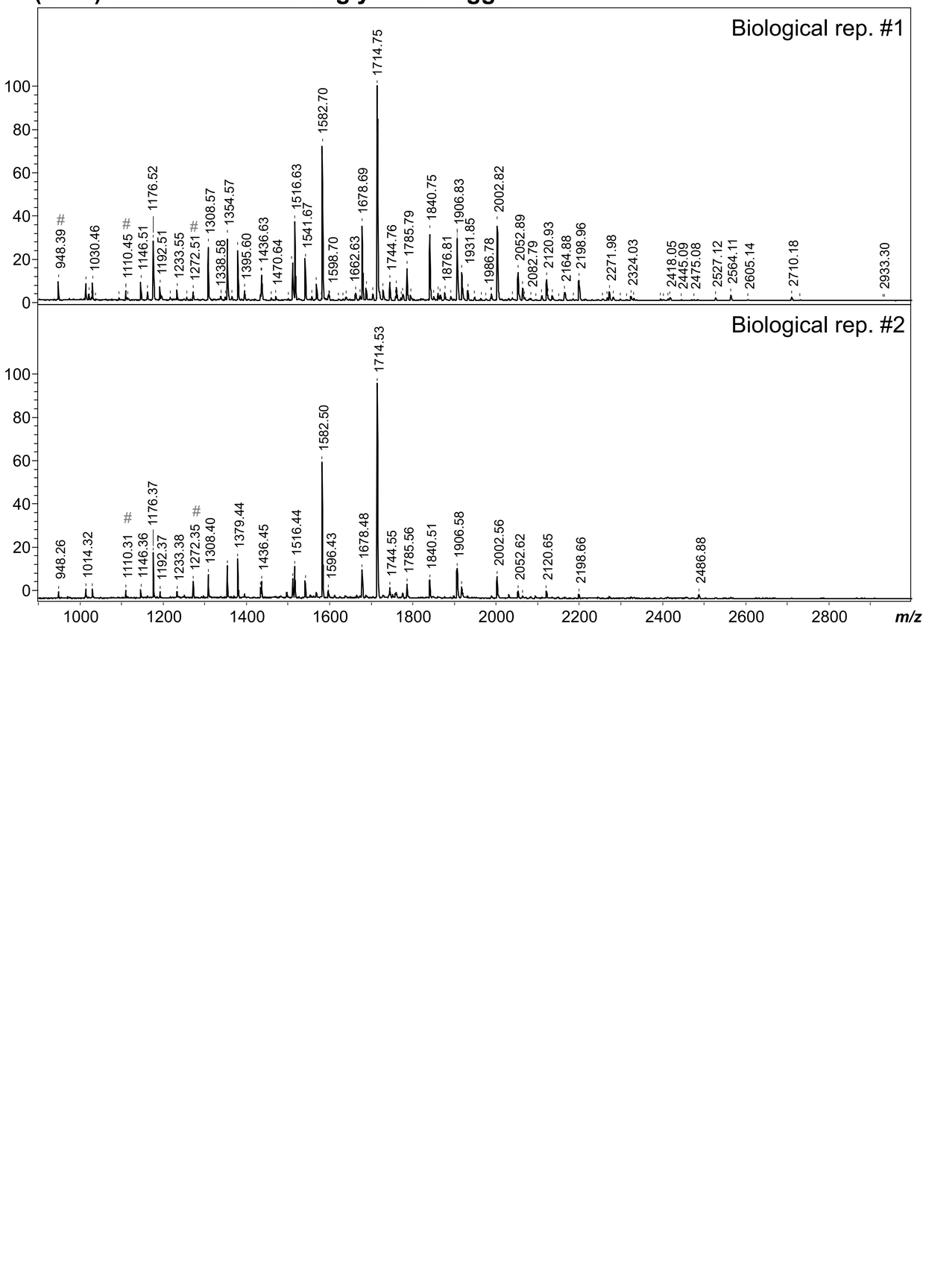
(A - 3) *S. haematobium* GSL glycans – eggs



(B - 1) *S. haematobium* *N*-glycans – cercariae

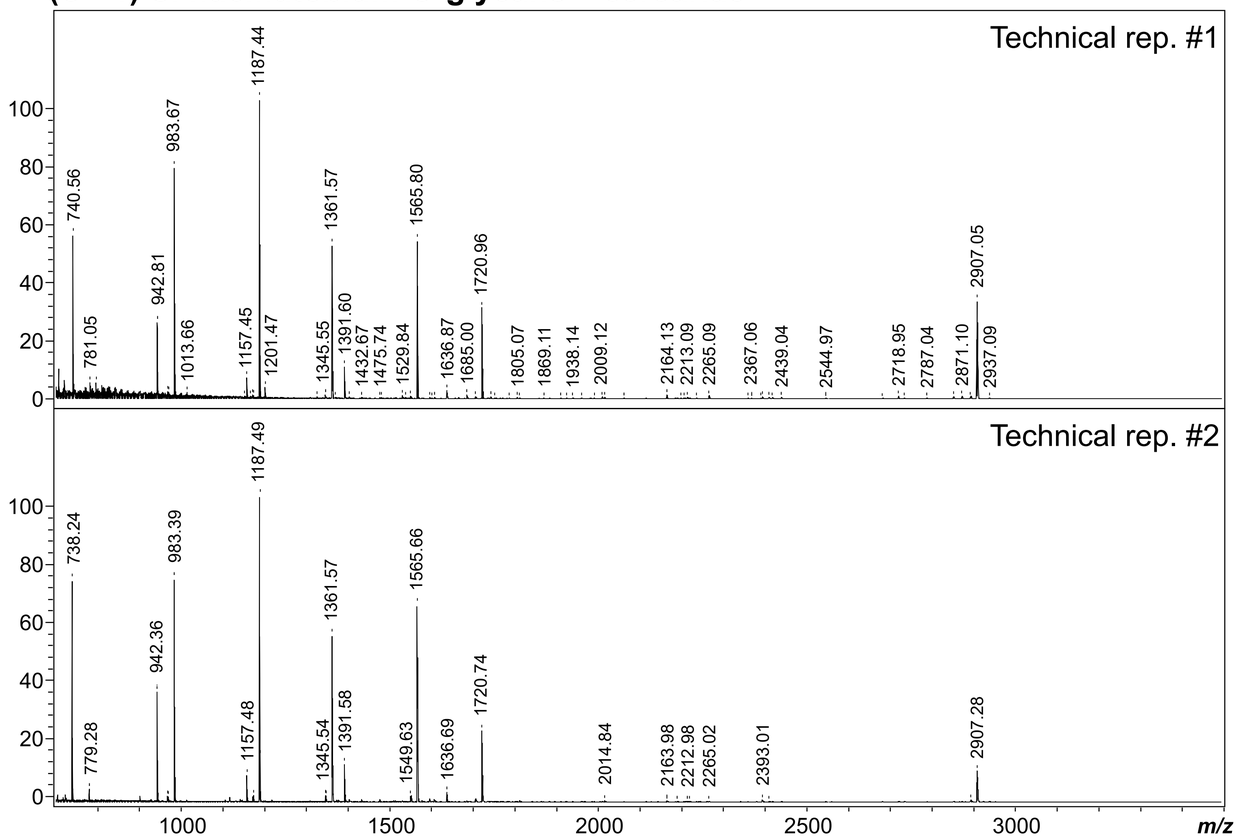


(B - 3) *S. haematobium* *N*-glycans – eggs



(B - 2) *S. haematobium* *N*-glycans – adult worms

(C - 1) *S. haematobium O*-glycans – cercariae



(C - 2) *S. haematobium O*-glycans – eggs

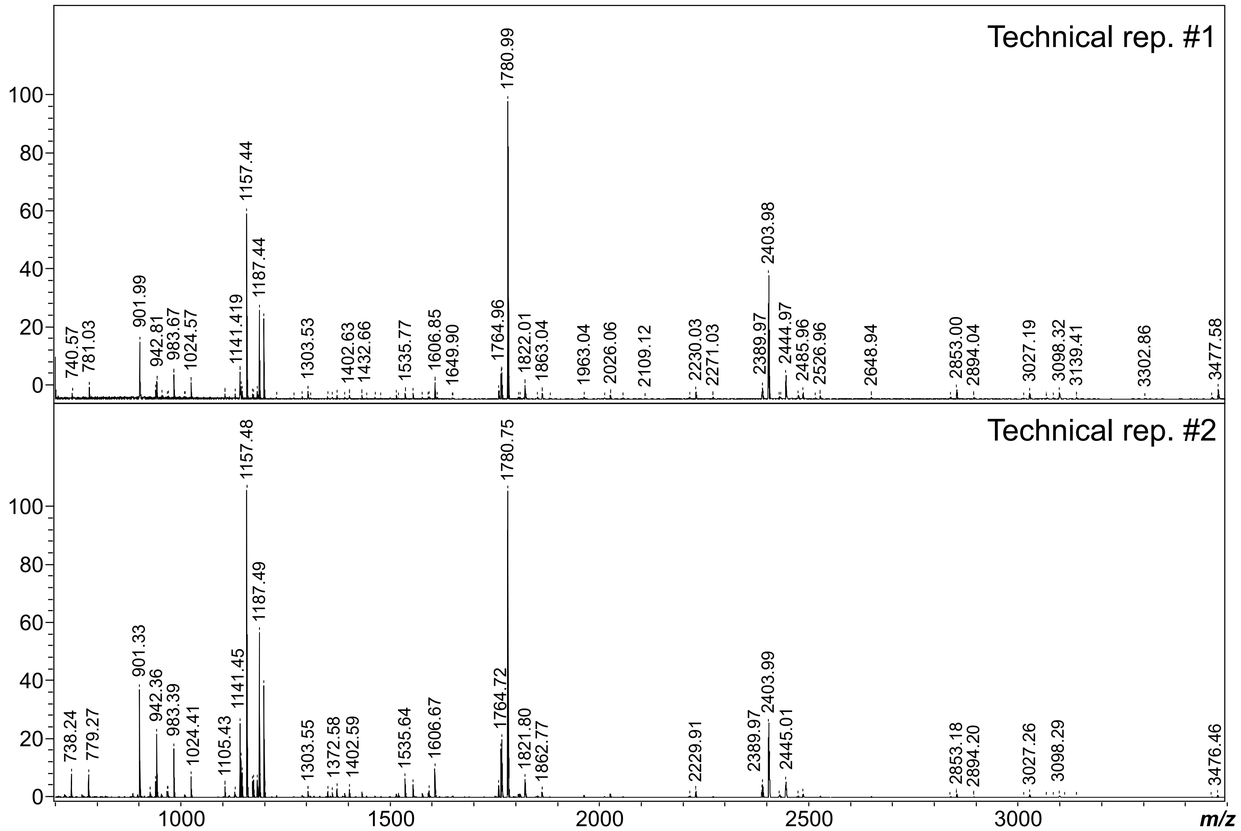


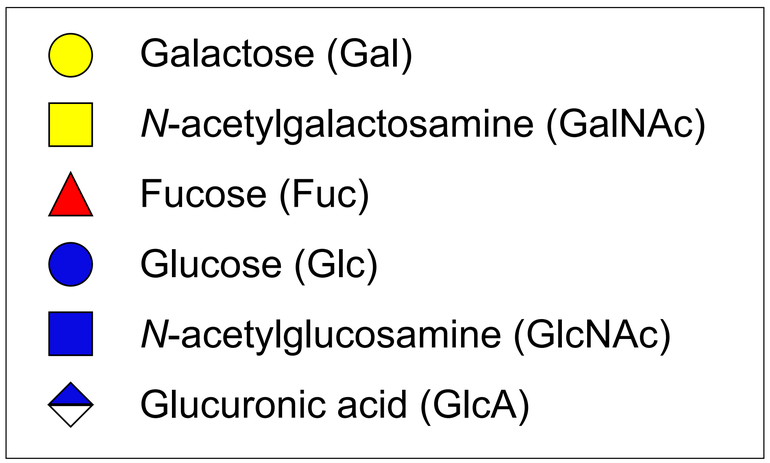
Figure S2 – Structural characterization of *S. haematobium* and *S. mansoni* GSL glycans

GSL glycans were released from their lipid carriers using rEGCase II prior to AA-labeling.

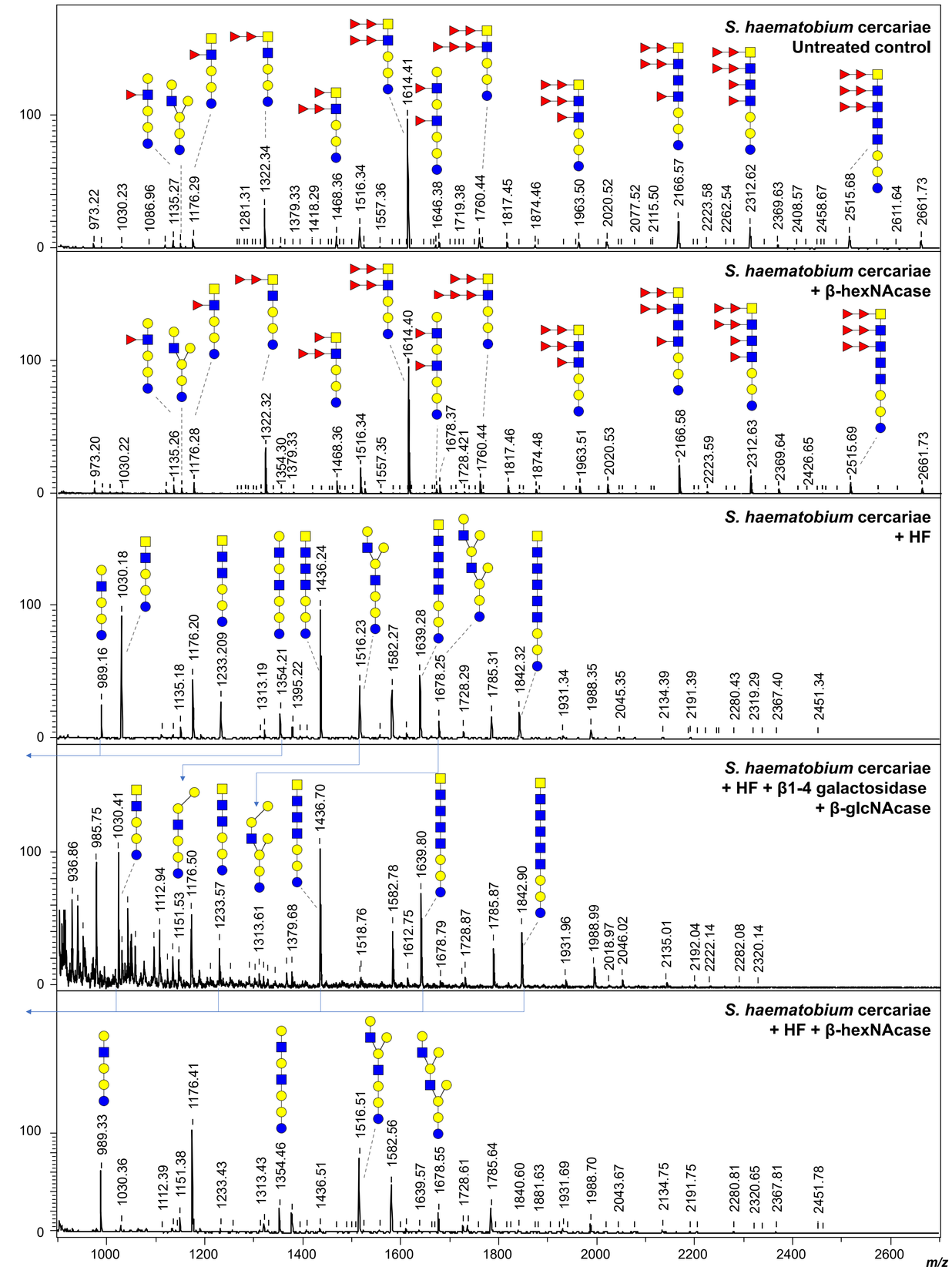
Glycan sequencing AA-labeled GSL glycans of *S. haematobium* cercariae (A), adult worms (B), eggs (C) and of *S. mansoni* eggs (acidic fraction only, E) were subjected to hydrofluoric acid (HF) treatment and/or to digestion with exoglycosidase(s). Enzymes were either applied simultaneously to the glycan sample, as a cocktail, or separately ("Sequential") as indicated. Reactions were performed in the conditions detailed in Supplementary Data 1 and in the Methods section. Treated samples and controls were then analyzed using MALDI-TOF-MS. All spectra were acquired in negative-ion reflectron mode and signals are labeled with monoisotopic masses (*m/z,* $[M-H]^-$). Signal intensities in percentages are indicated on the Y-axis. Sample type (untreated control or treated sample), parasite species (*S. haematobium* or *S. mansoni*) and life-stage (cercariae, adult worms or eggs) are indicated at the top of each panel. Blue arrows highlight products resulting from the aforementioned treatments. Contaminants introduced by the preparation and known non-glycan signals are labeled with the # symbol. Ions for which composition and structure were not determined are labeled **ND**.

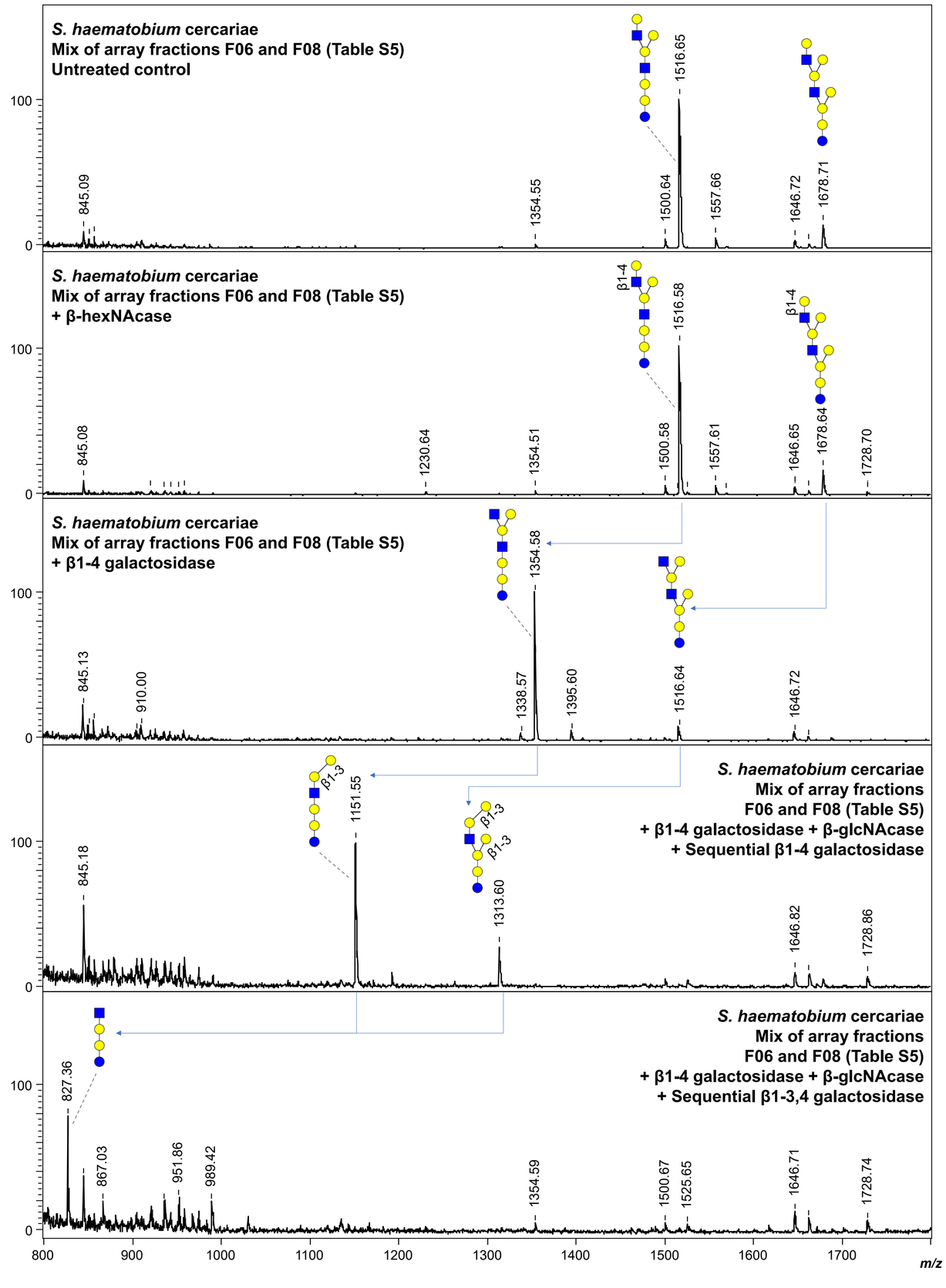
MALDI-TOF-MS of GSL glycans derived from *S. haematobium* immature and mature eggs (D), separated by Percoll gradient centrifugation as described in the Methods section.

All (putative) glycans are represented using the CFG nomenclature below.

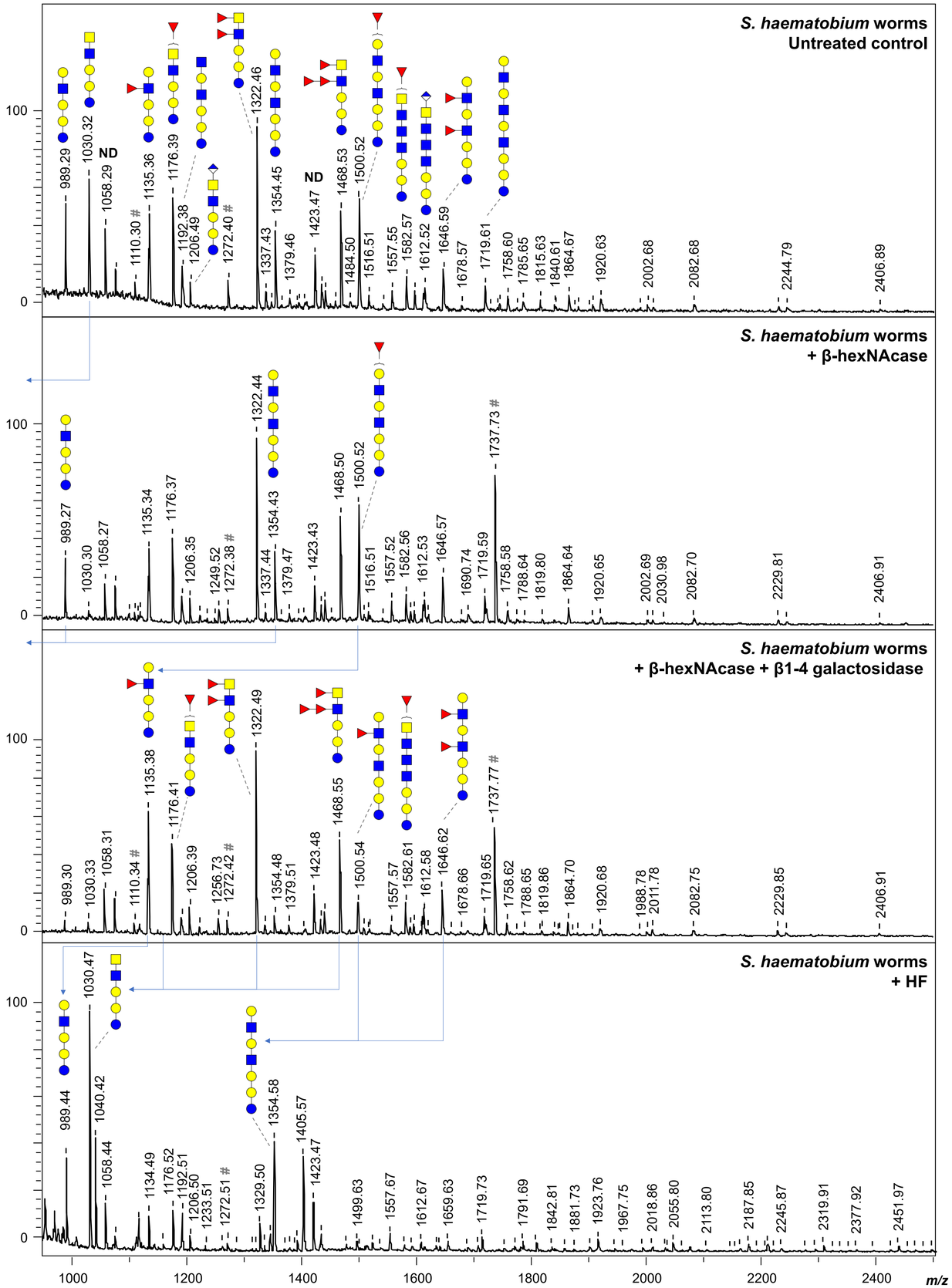


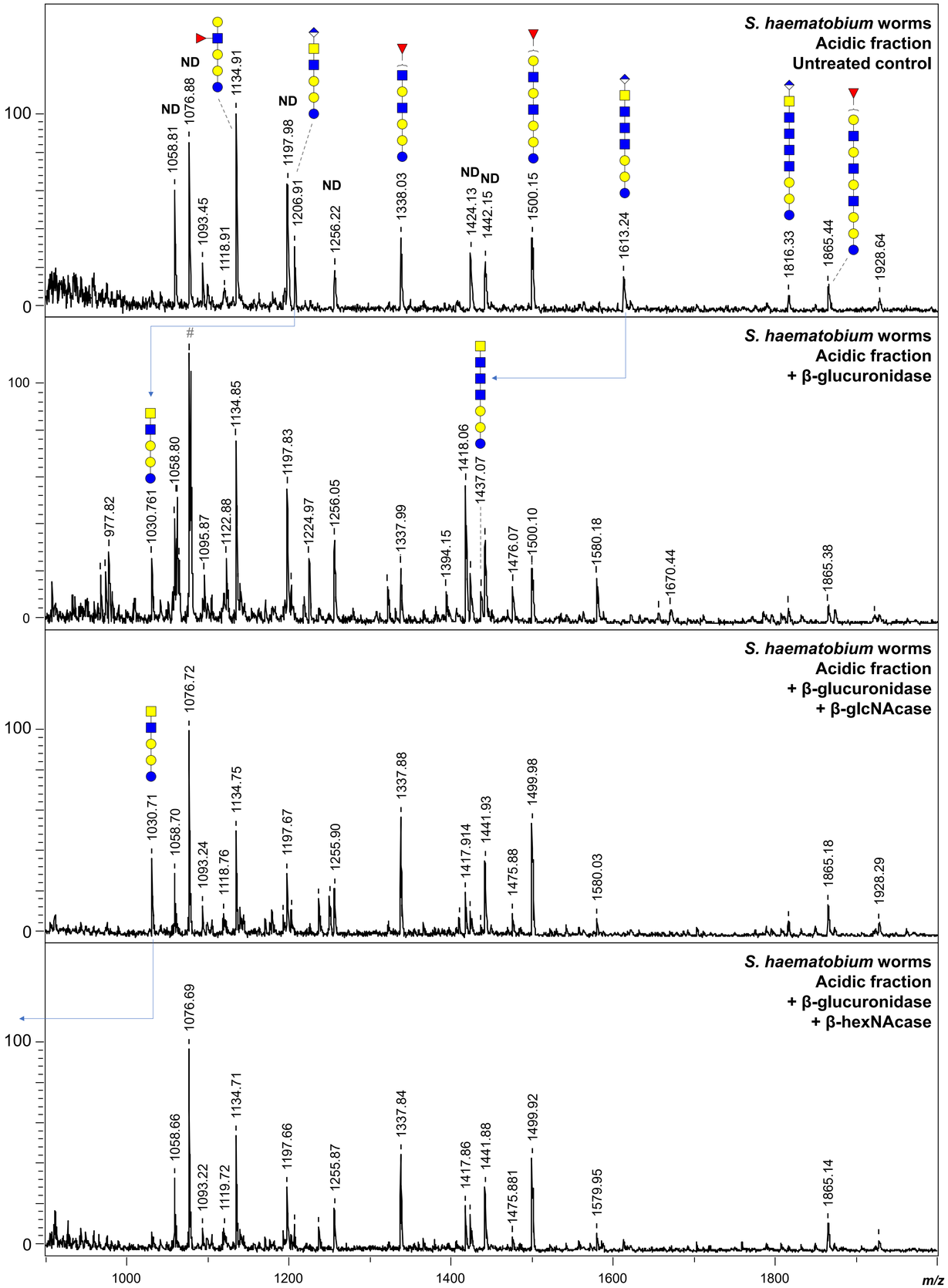
(A) *S. haematobium* cercariae – glycan sequencing



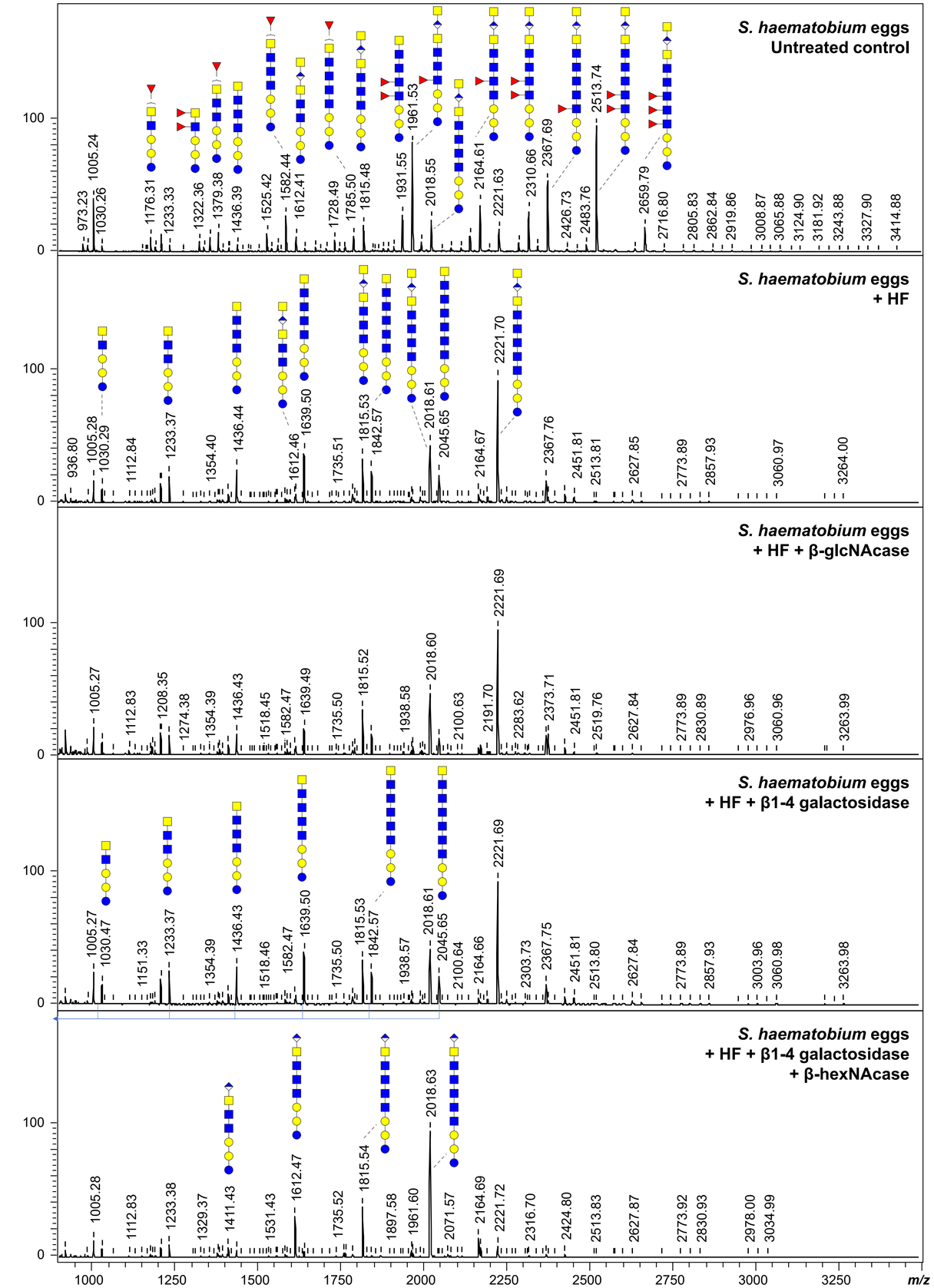


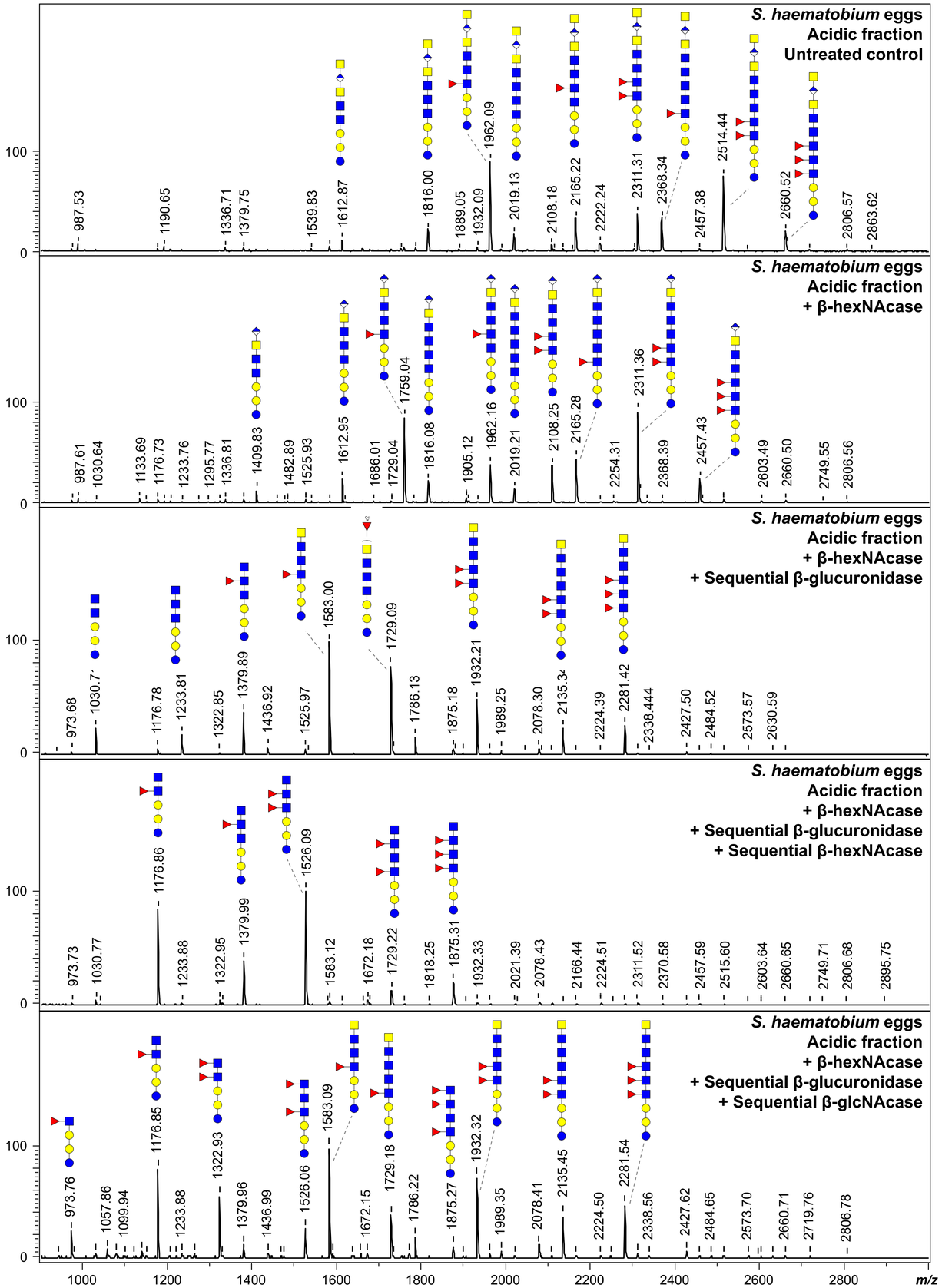
(B) *S. haematobium* adult worms – glycan sequencing



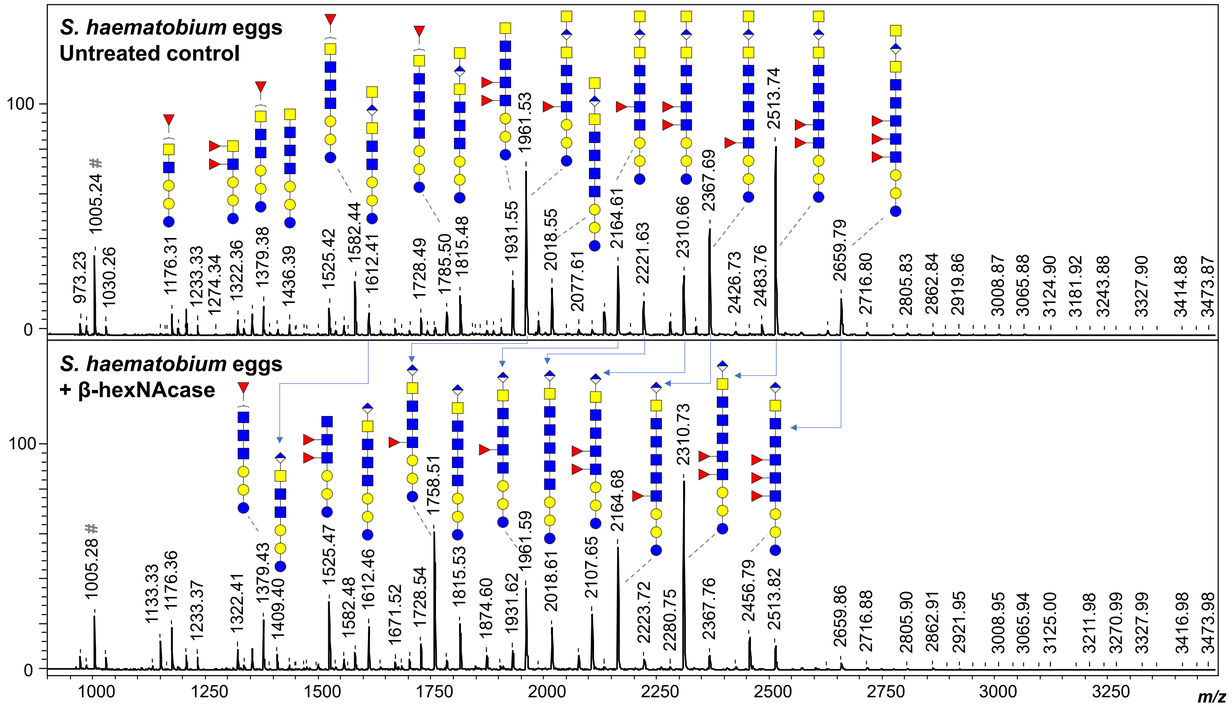


(C) *S. haematobium* eggs – glycan sequencing

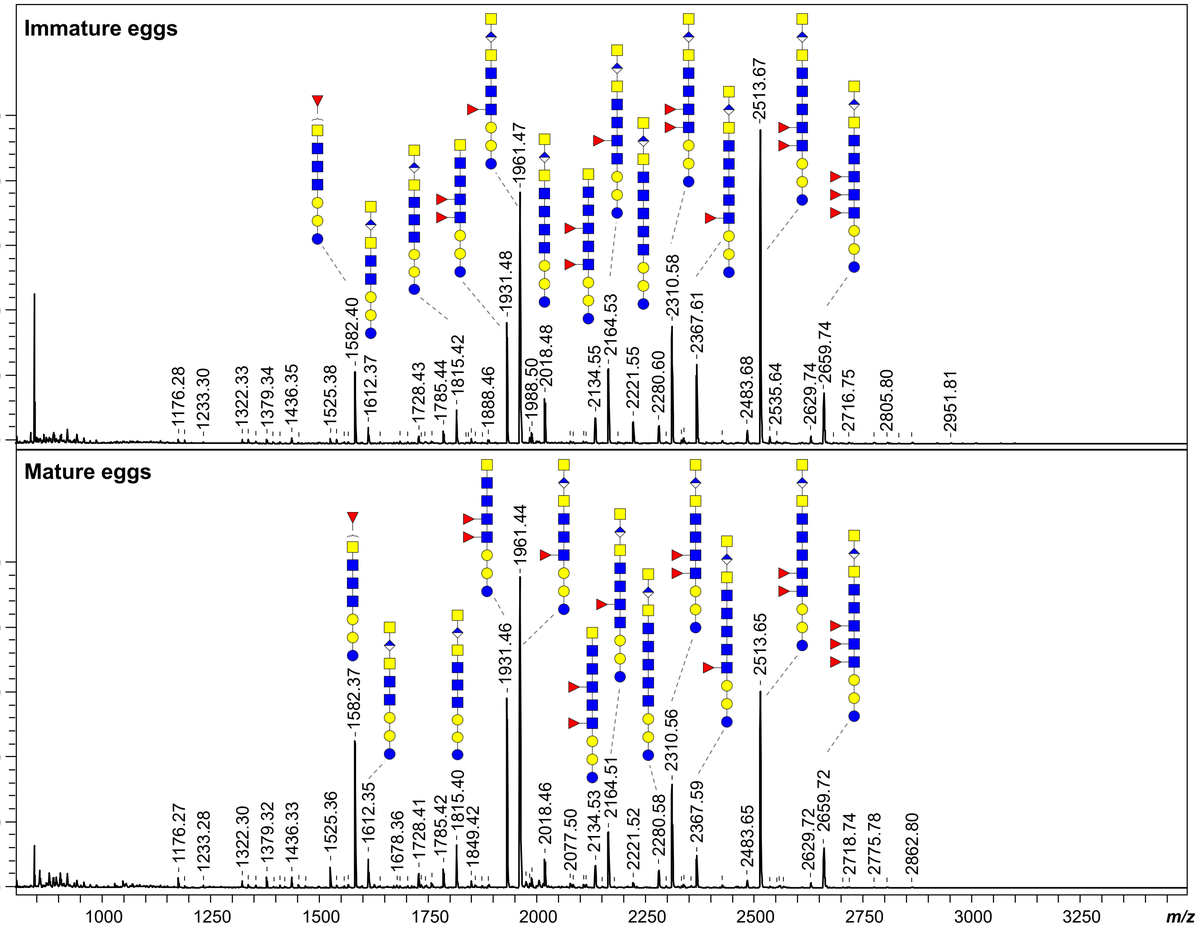




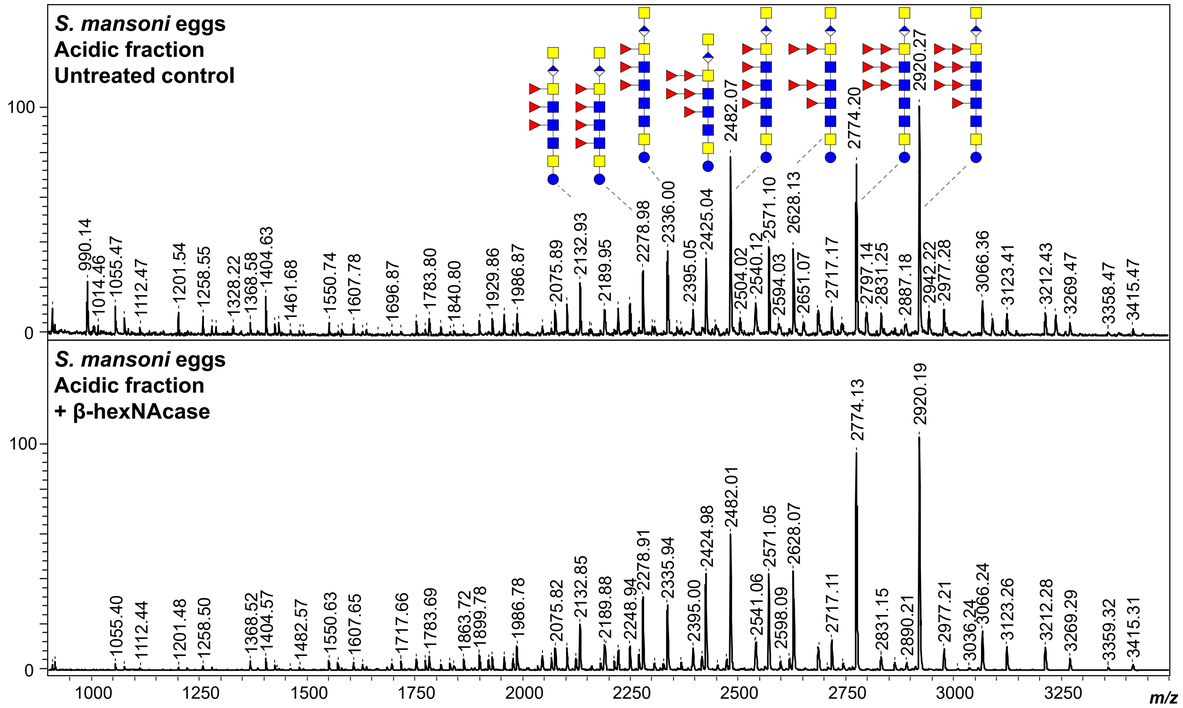
(C) *S. haematobium* eggs – glycan sequencing *(continued)*



(D) GSL glycans of *S. haematobium* immature and mature eggs



(E) Acidic GSL glycans of *S. mansoni* eggs – glycan sequencing



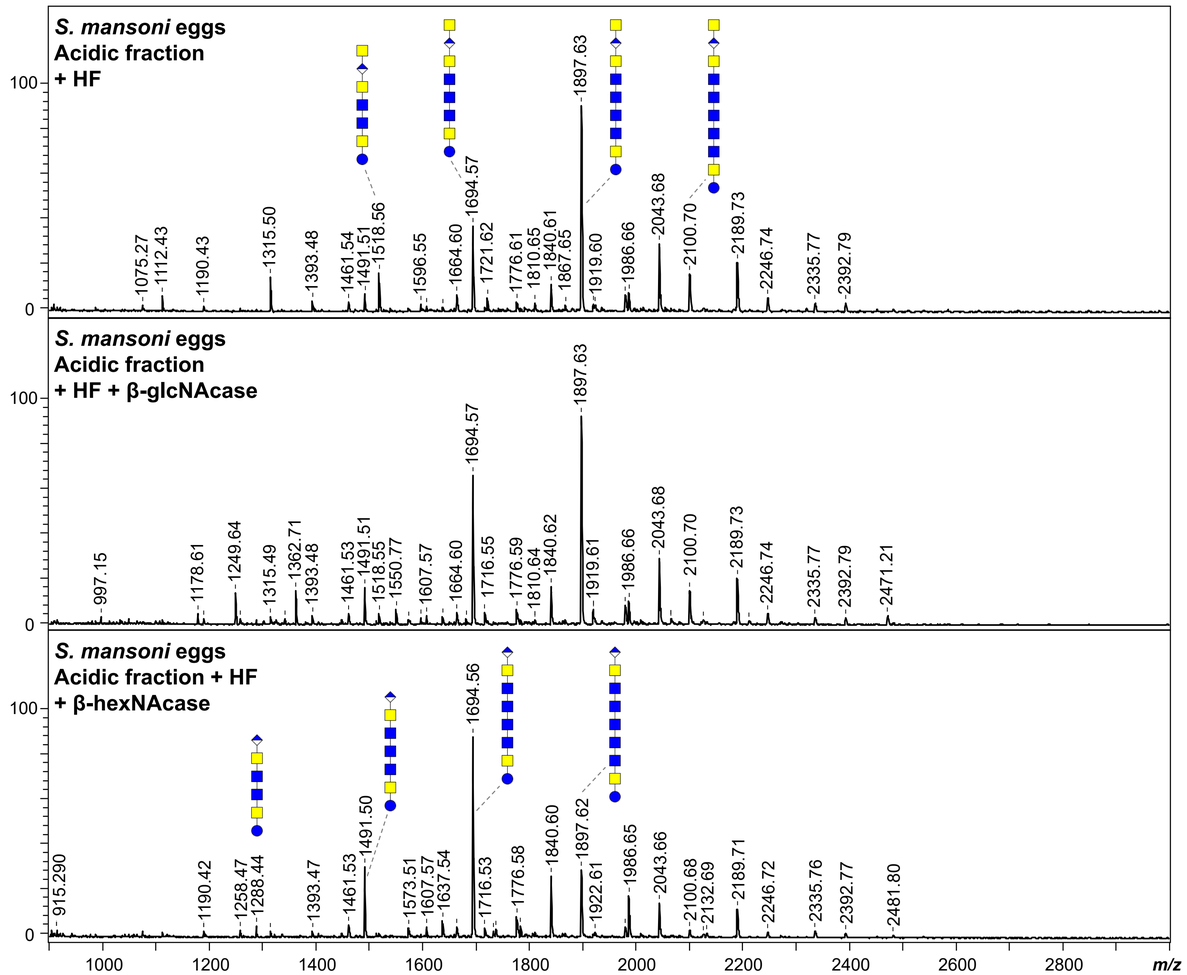
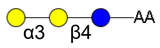
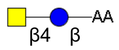
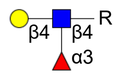
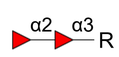
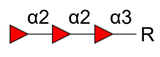
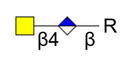
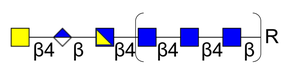
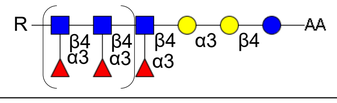
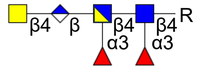
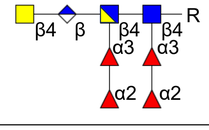
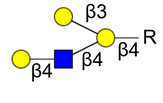


Figure S3 – MALDI-TOF-MS/MS of *S. haematobium* and *S. mansoni* GSL glycans

GSL glycans were released from their lipid carriers using rEGCase II prior to AA-labeling. MALDI-TOF-MS/MS was performed on selected ion species of the GSL glycans of *S. haematobium* cercariae (A), adult worms (B), eggs (C) and of *S. mansoni* eggs (acidic fraction only, D). Theoretical masses (Th. *m/z*, $[M-H]^-$) of fragmented parent ions are indicated on the upper left corner of each panel. Y-type fragment-ions, as defined by Domon and Costello (https://doi.org/10.1007/BF01049915) are represented, unless indicated otherwise (B = B type, C = C type, Z = Z type). Diagnostic ions and/or fragmentation patterns indicative of glycan traits of interest are listed in the Table below together with the corresponding panels in which they are observed. All (putative) glycans are represented using the CFG nomenclature (see inset below). Note that based on glycan sequencing using exoglycosidases (Figure S2), we identified instances where the HexNAc preceding the GlcA residue could be either GlcNAc or GalNAc (Supplementary Data 2). The GalNAc-containing isomer is represented in panels C-D.

Identified glycan trait		Diagnostic ion *m/z* (theoretical *m/z*) or Δ*m*	Present in	Panel
Trihexosyl GSL core		Y_2-H_2-AA (*m/z* 462.16) Y_3-H_3-AA (*m/z* 624.21)	*S. haematobium* cercariae *S. haematobium* adult worms *S. haematobium* eggs	A B C
Disaccharidic GSL core		Y_2-H_1N_1-AA (*m/z* 503.19)	*S. mansoni* eggs	D
LeX		Δ*m* = 511.19 ($H_1N_1F_1$)	*S. haematobium* cercariae *S. haematobium* adult worms	A B
DF		Δ*m* = 292.12 (F_2) Δ*m* = 495.20 (N_1F_2)	*S. haematobium* cercariae *S. haematobium* adult worms *S. mansoni* eggs	A B D
TF		Δ*m* = 438.174(F_3) Δ*m* = 641.253 (N_1F_3)	*S. haematobium* cercariae	A
Terminal HexNAc and subterminal GlcA		Δ*m* = 203.08 (N_1) followed by Δ*m* = 176.03 (A_1)	*S. haematobium* eggs *S. mansoni* eggs	C D
Subterminal GlcA in a chain of 1 to 4 unsubstituted HexNAc		B_{3-6}-$N_{2-5}A_1$ (*m/z* 581.18; 784.26; 987.34;1190.42)	*S. haematobium* eggs	C
Fuc in proximity with trihexosyl core		Y_4-$H_3N_1F_1$-AA (*m/z* 973.35) Y_5-$H_3N_2F_2$-AA (*m/z* 1322.49) Y_6-$H_3N_3F_3$-AA (*m/z* 1671.63)	*S. haematobium* eggs	C
F-LDN-F in proximity with subterminal GlcA		BY-$N_2A_1F_2$ (*m/z* 873.30) C-$N_2A_1F_4$ (*m/z* 1165.42)	*S. mansoni* eggs	D
DF and DF-LDN-DF in proximity with subterminal GlcA		B-$N_2A_1F_2$ (*m/z* 873.30) BY-$N_2A_1F_4$ (*m/z* 1165.42) C-$N_3A_1F_4$ (*m/z* 1386.51)	*S. mansoni* eggs	D
Branched Galactose		CY_3-N_2H_2 (*m/z* 747.27) Y_3Y_4-N_1H_3-AA (*m/z* 827.29)	*S. haematobium* cercariae	A

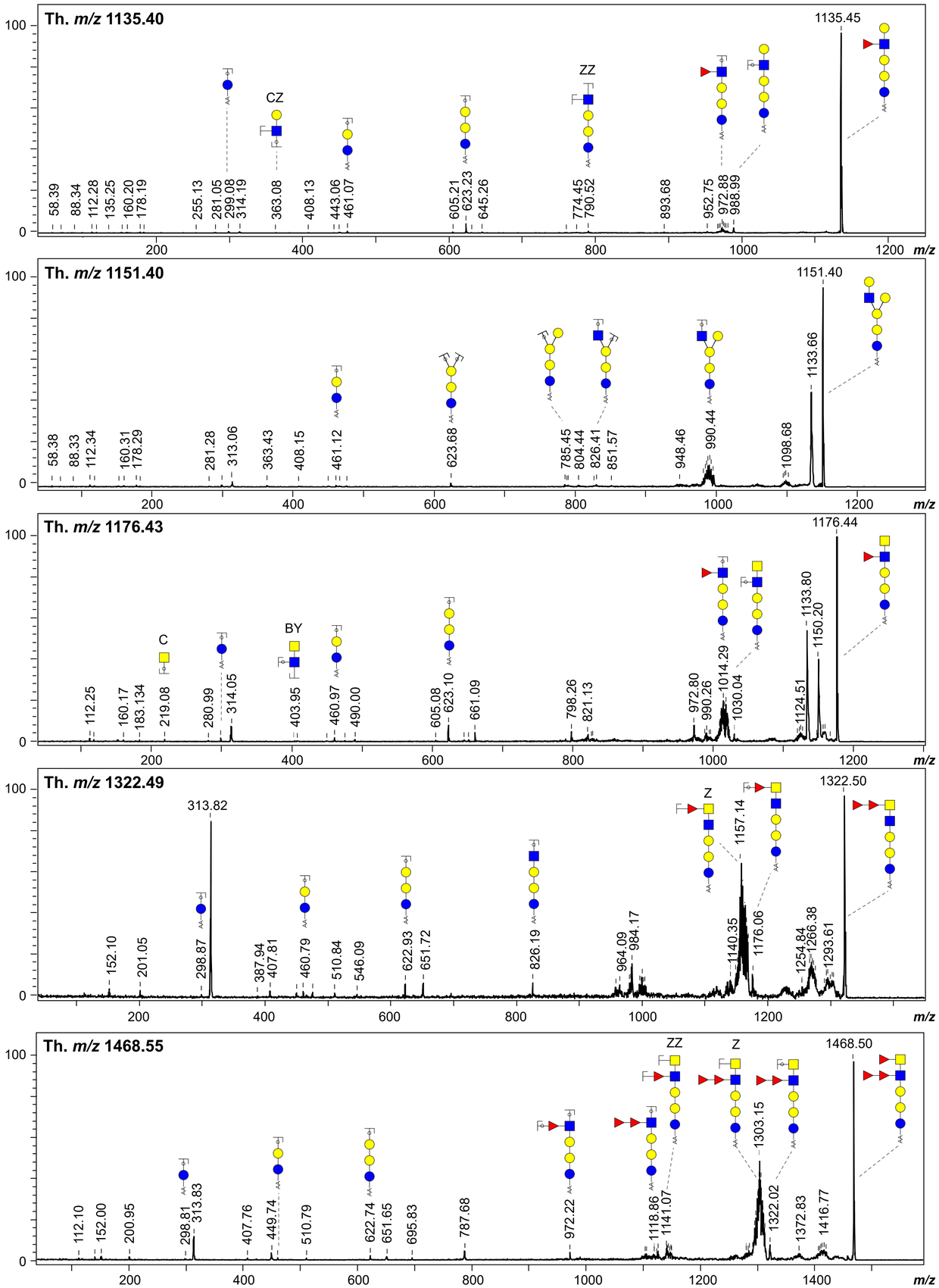
Monosaccharide nomenclature

- Galactose (Gal)
- *N*-acetylgalactosamine (GalNAc)
- Fucose (Fuc)
- Glucose (Glc)
- *N*-acetylglucosamine (GlcNAc)
- Glucuronic acid (GlcA)
- GlcNAc or GalNAc

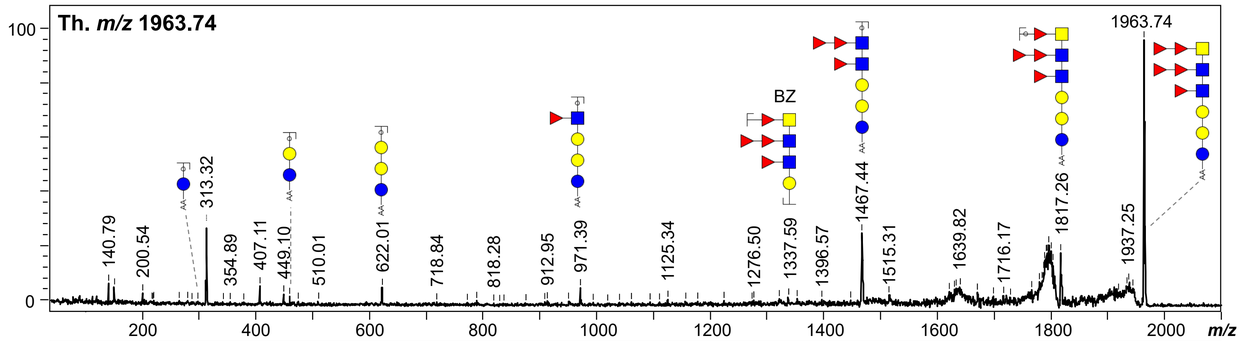
Fragment ion nomenclature

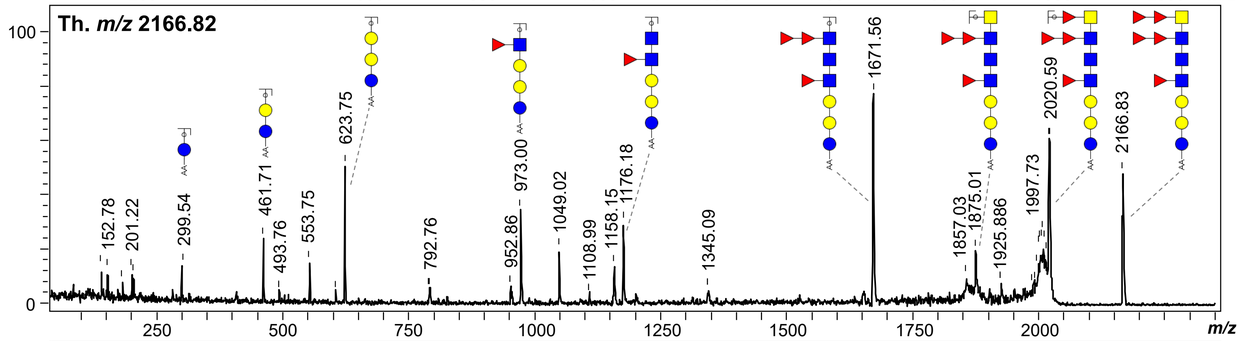
- B-type
- C-type
- Y-type
- Z-type

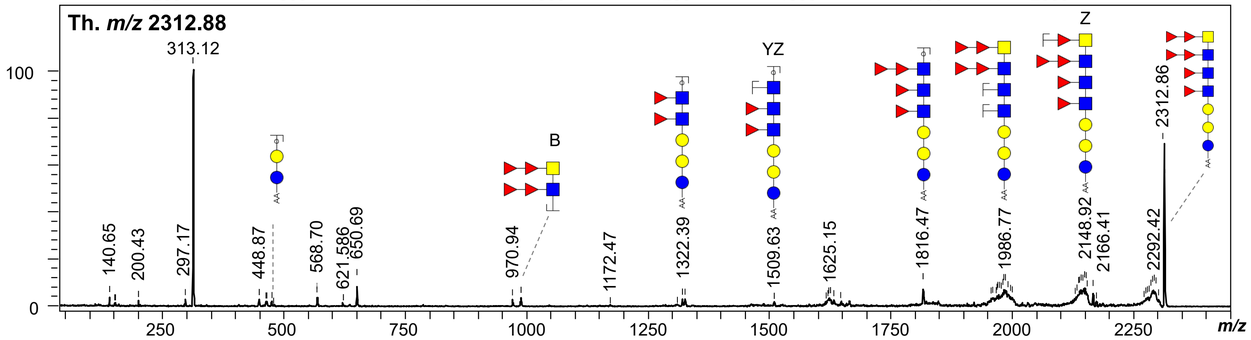
(A) *S. haematobium* cercariae

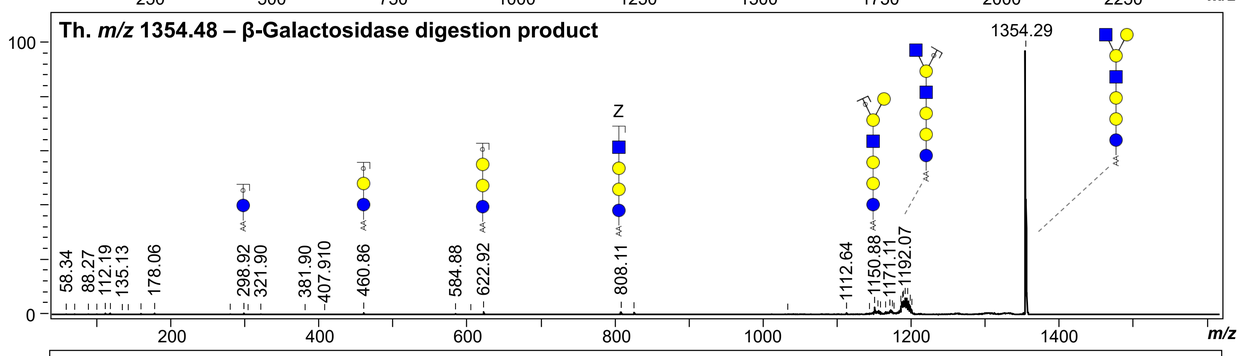


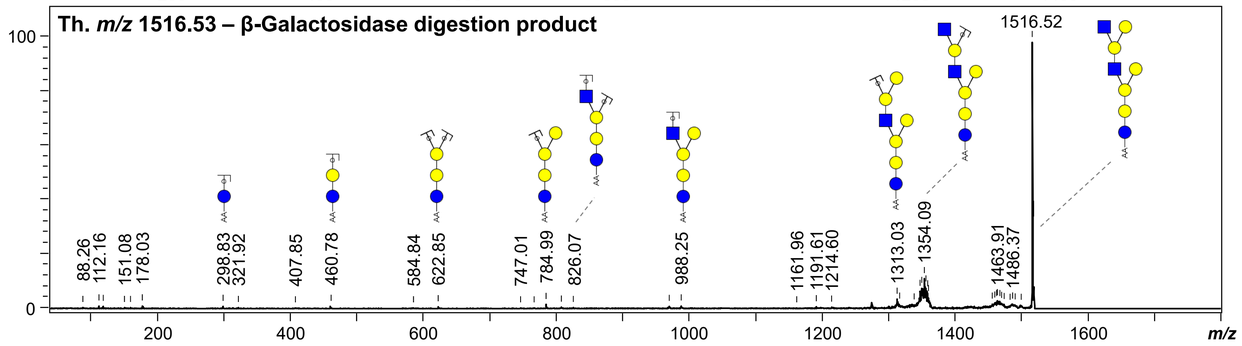
(A) *S. haematobium* cercariae *(continued)*



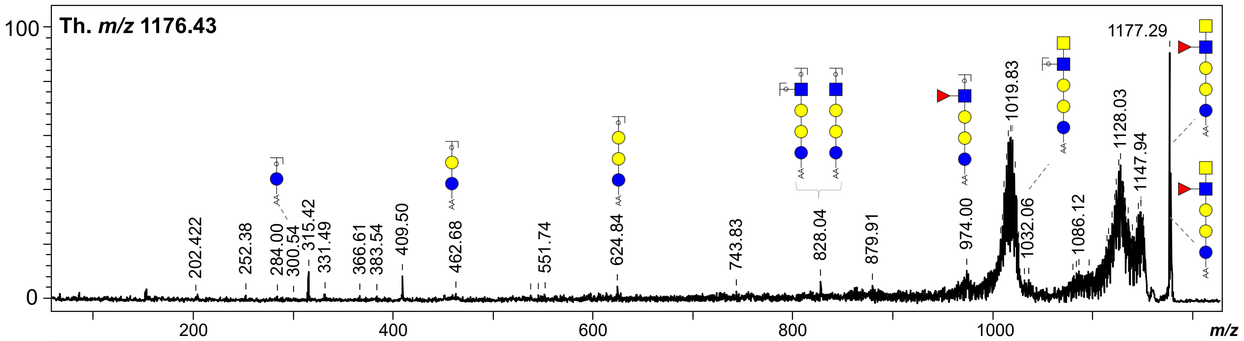


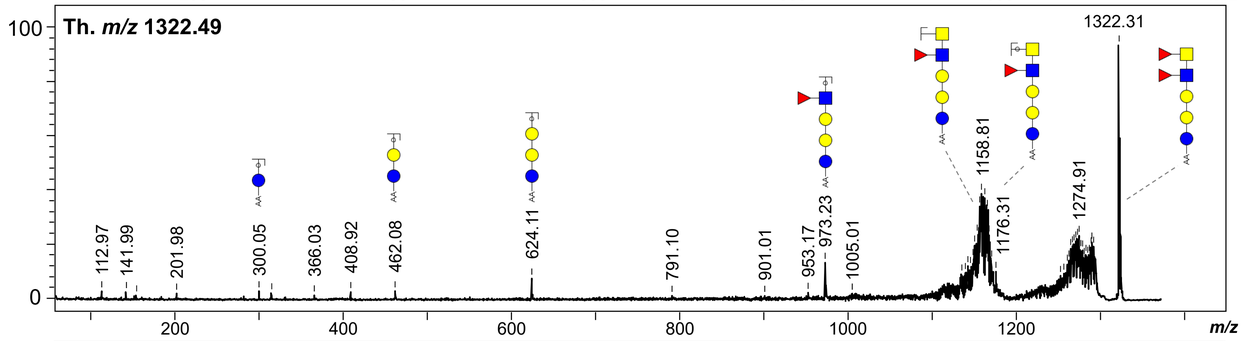


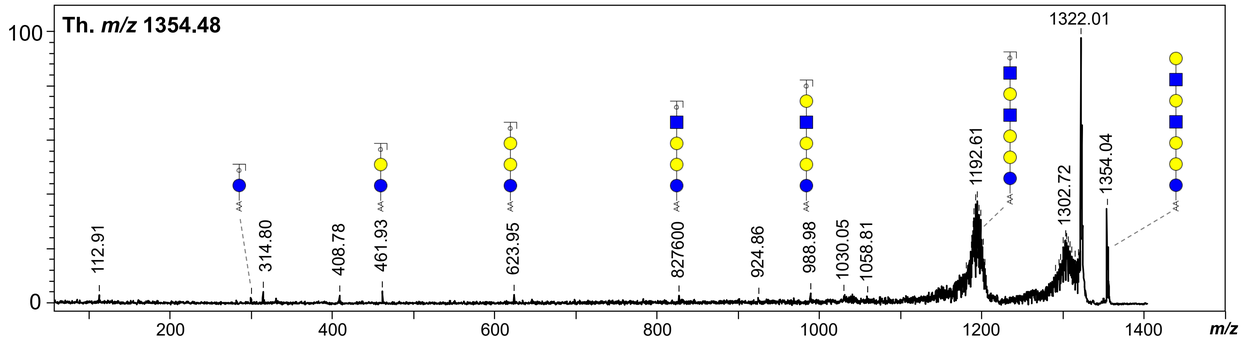


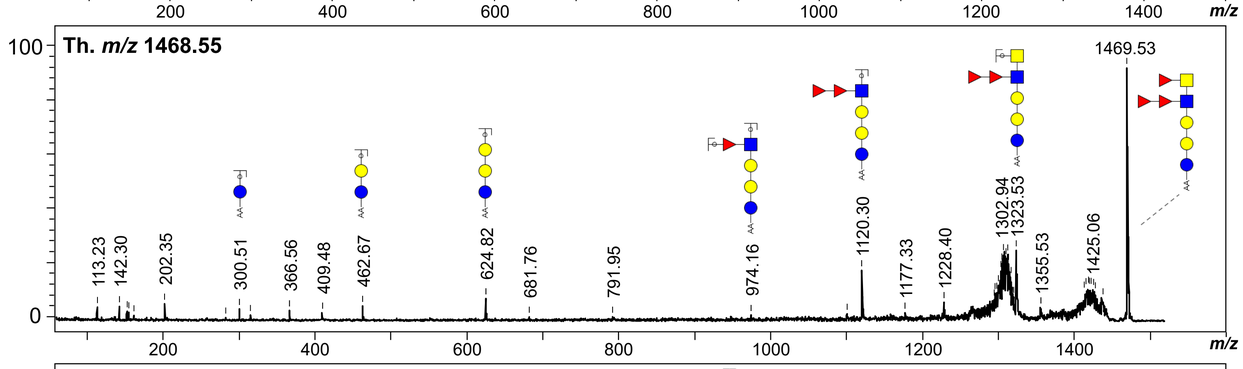


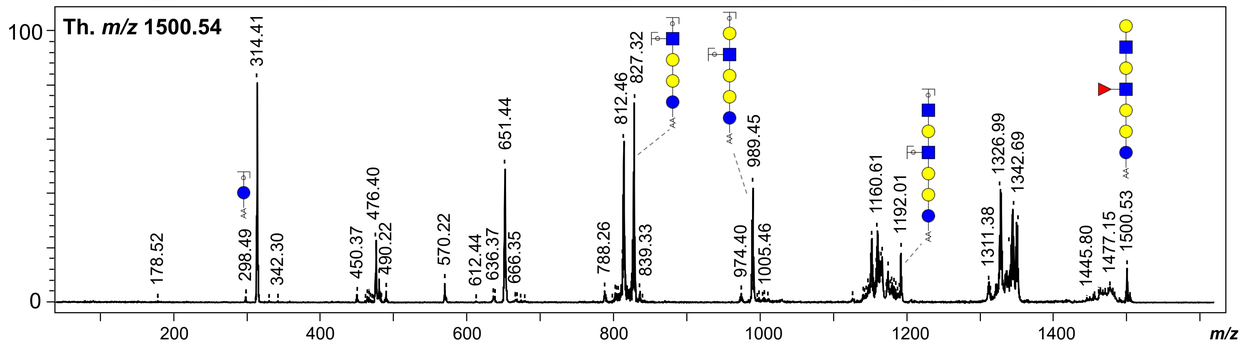
(B) *S. haematobium* adult worms



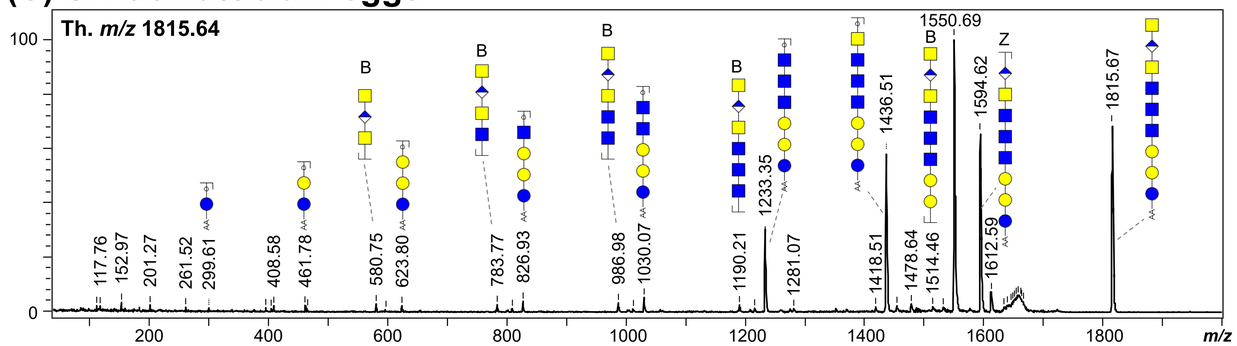


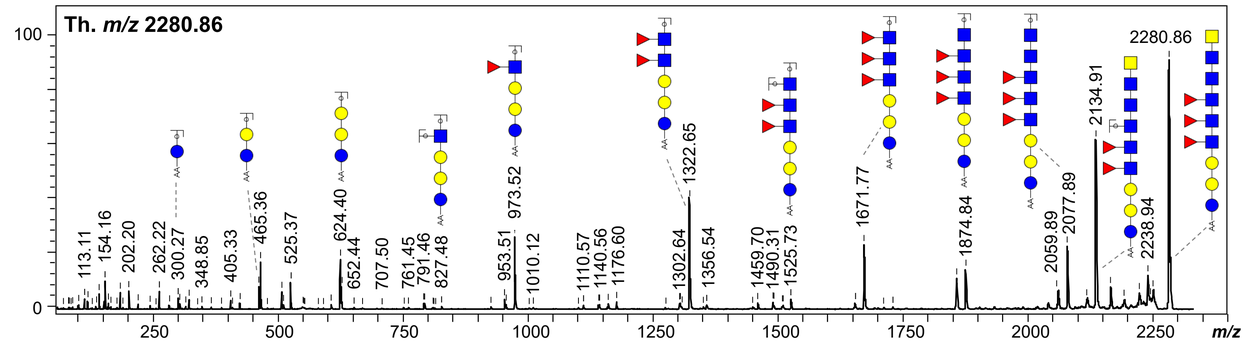


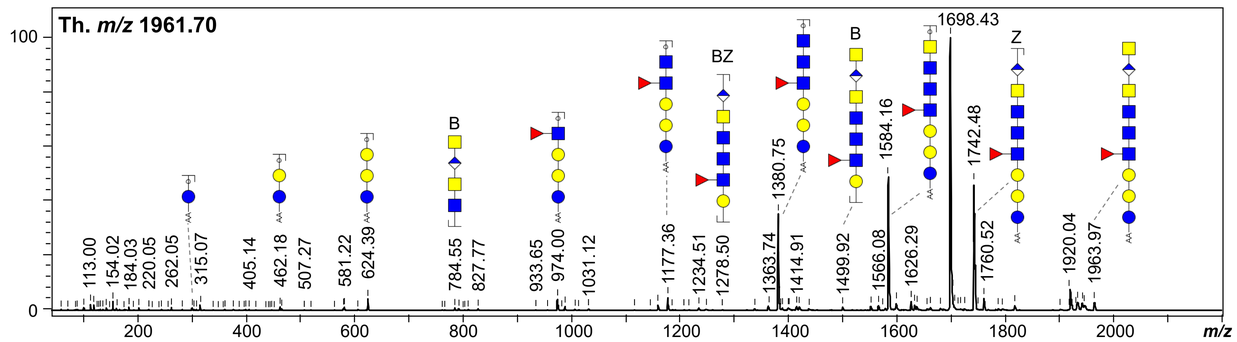


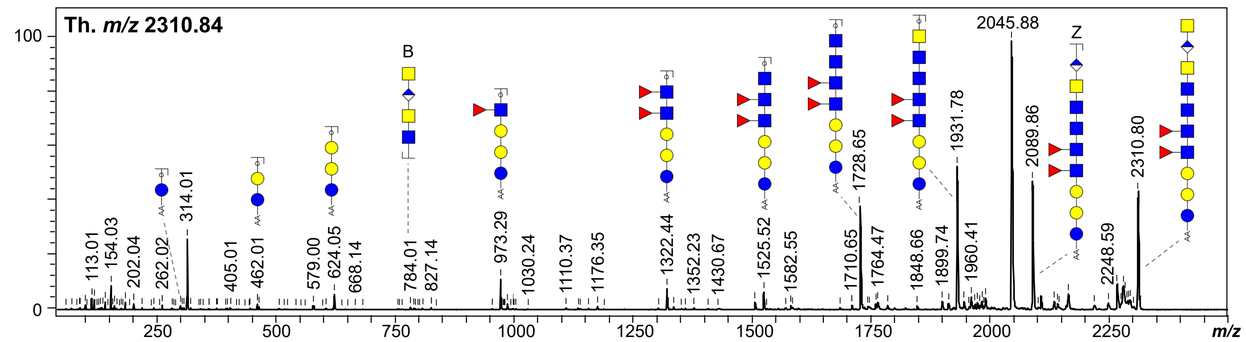


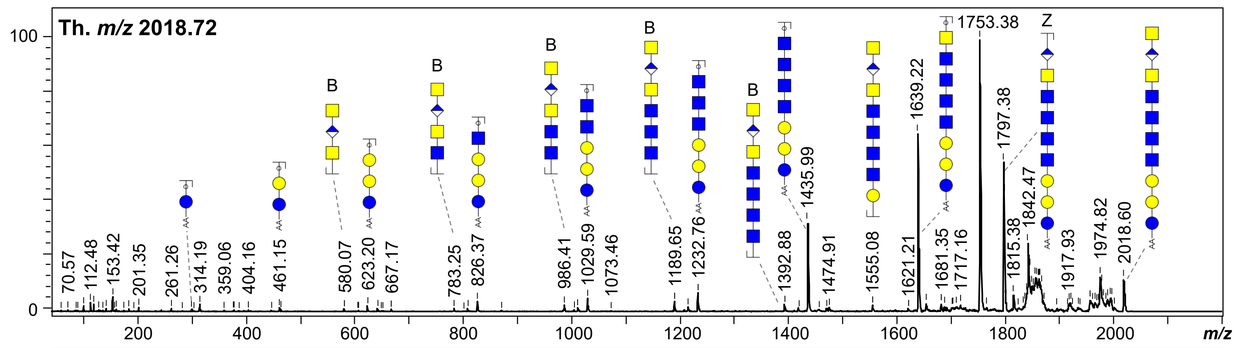
(C) *S. haematobium* eggs

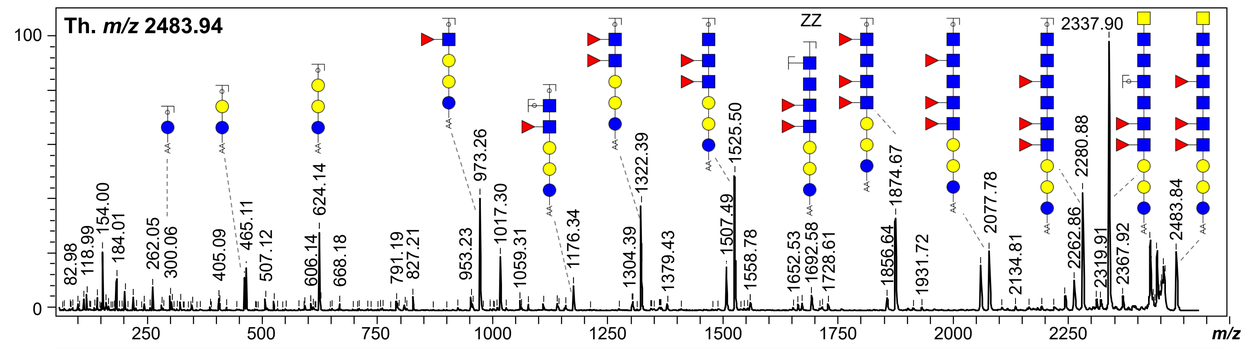


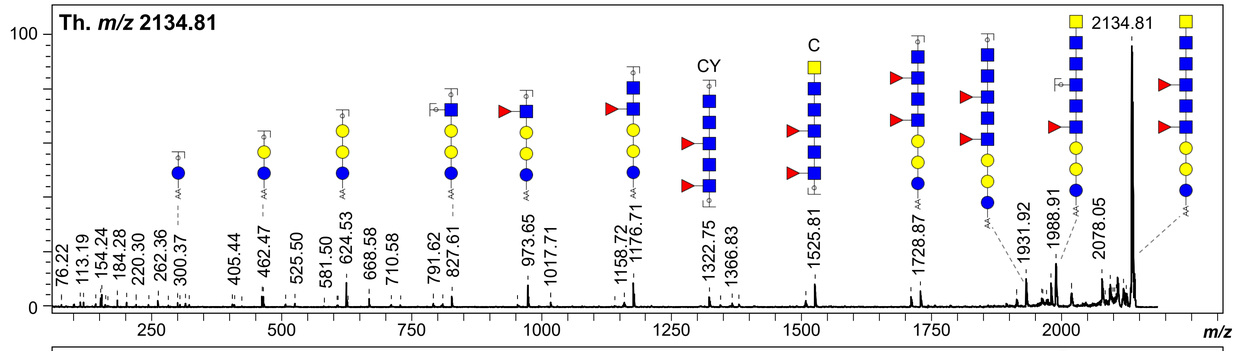


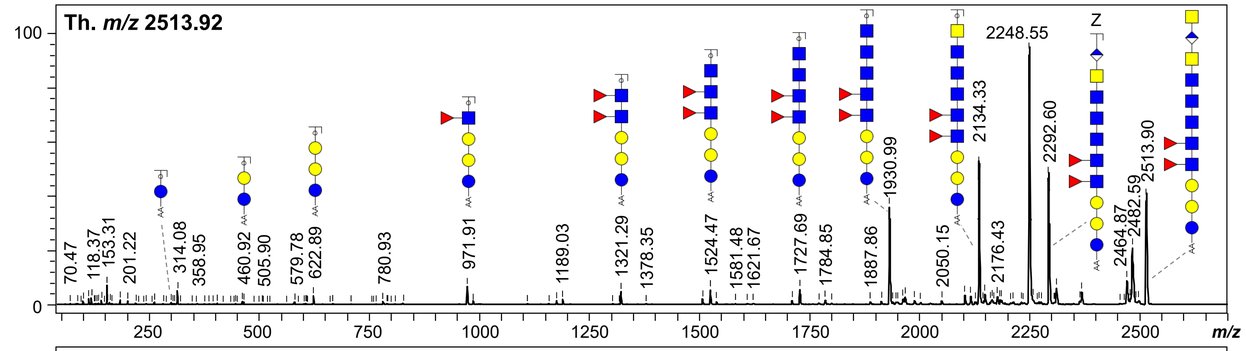


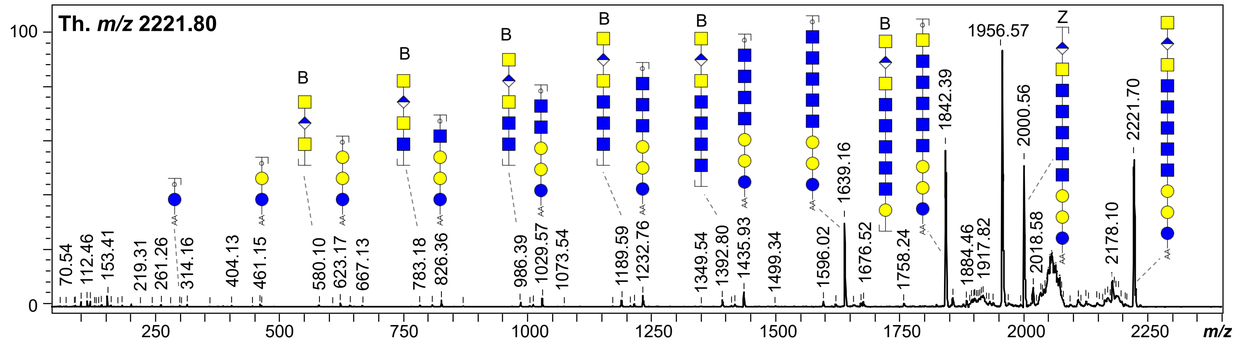


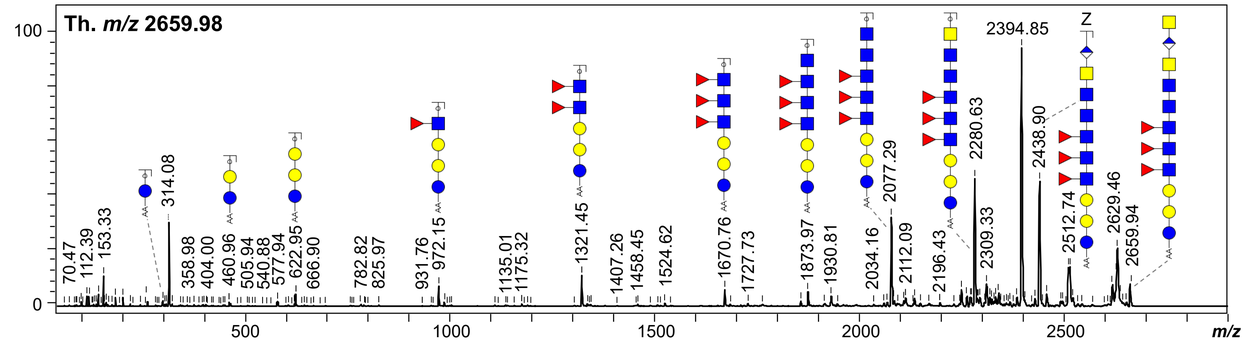




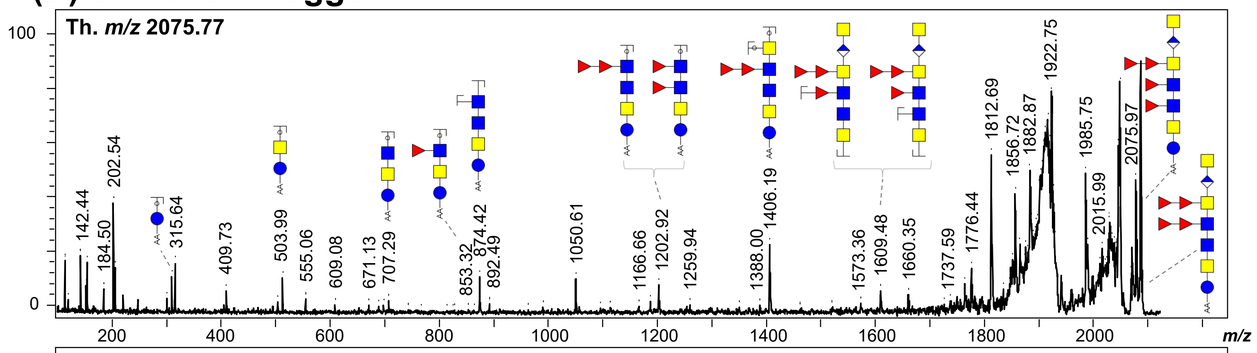


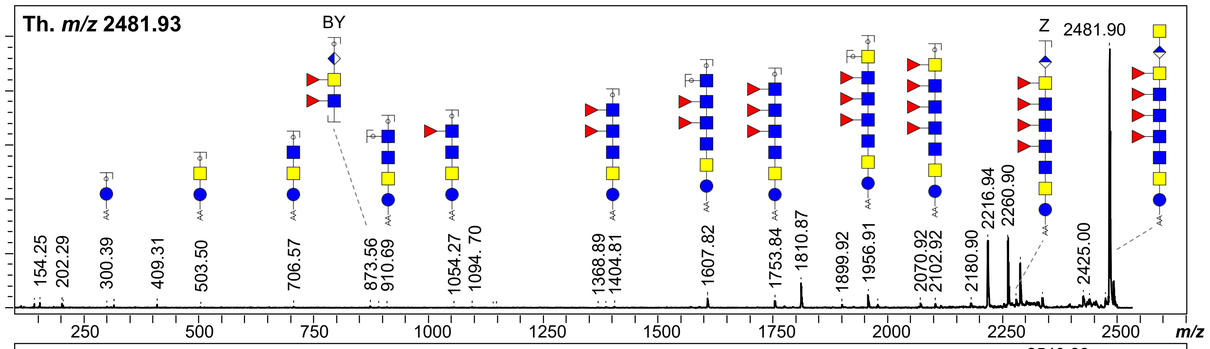


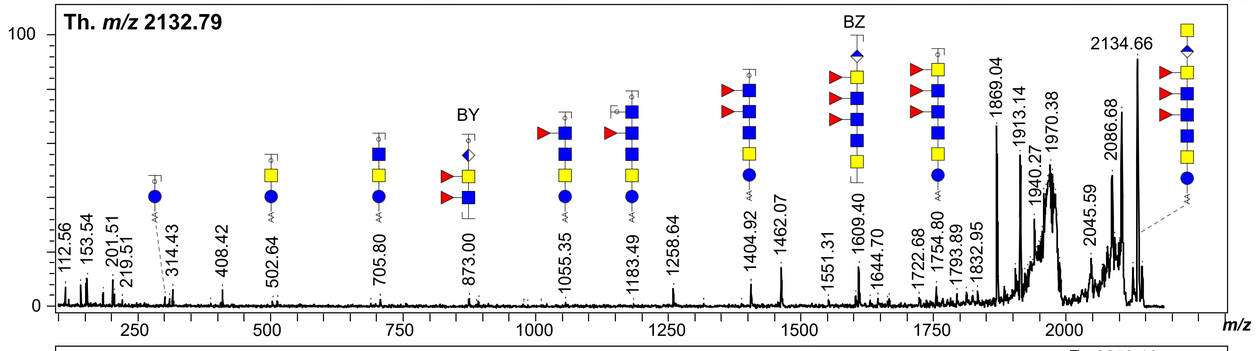


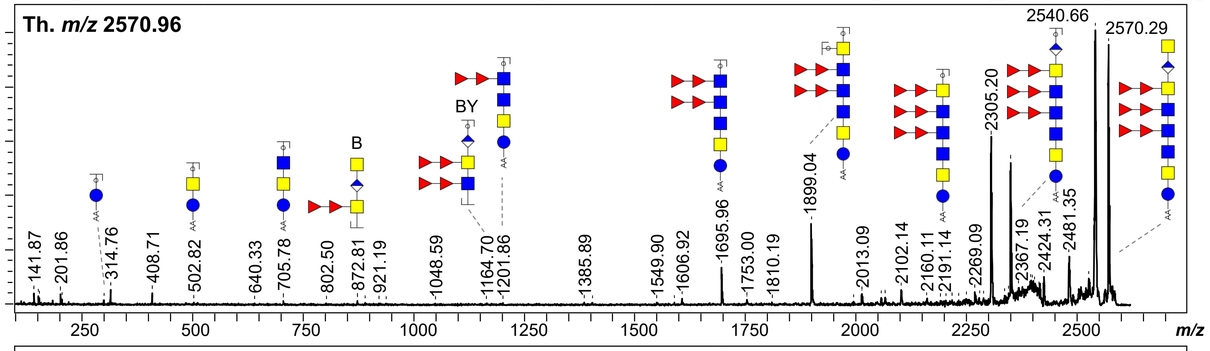


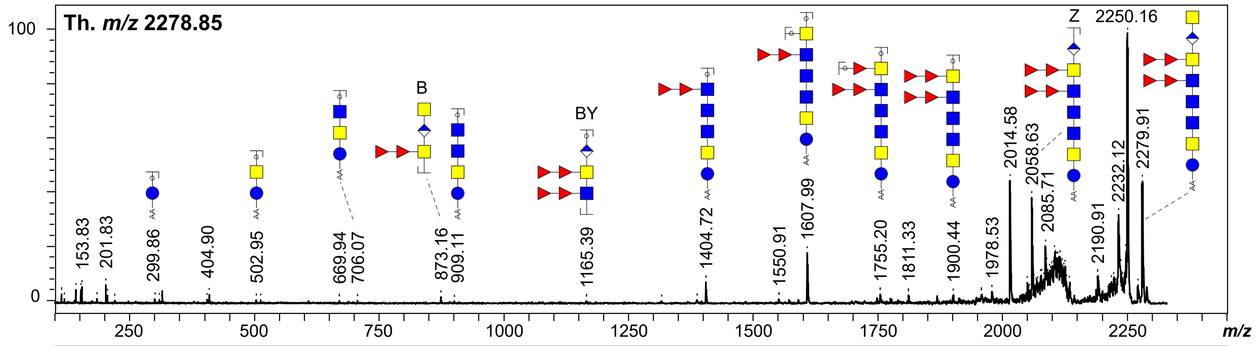
(D) *S. mansoni* eggs

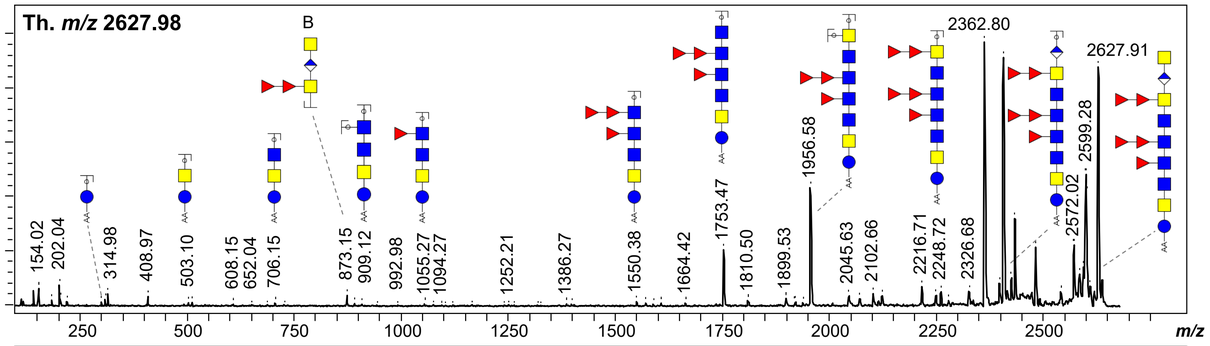


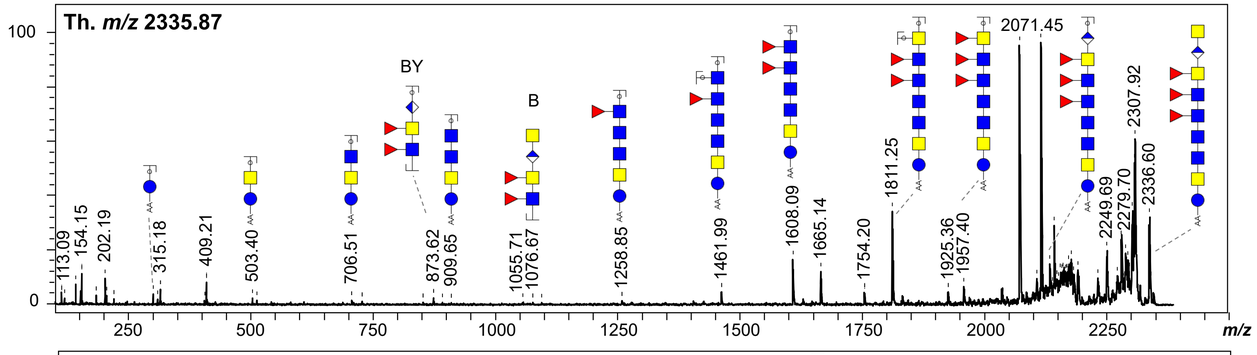


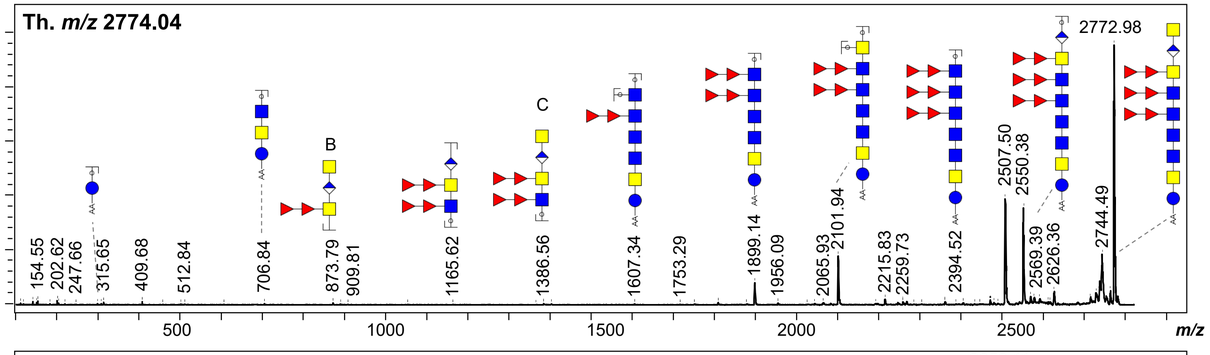


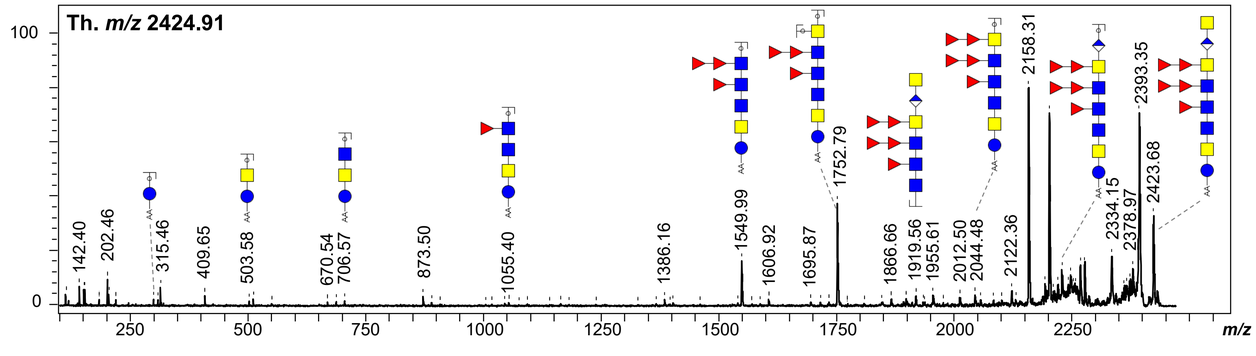












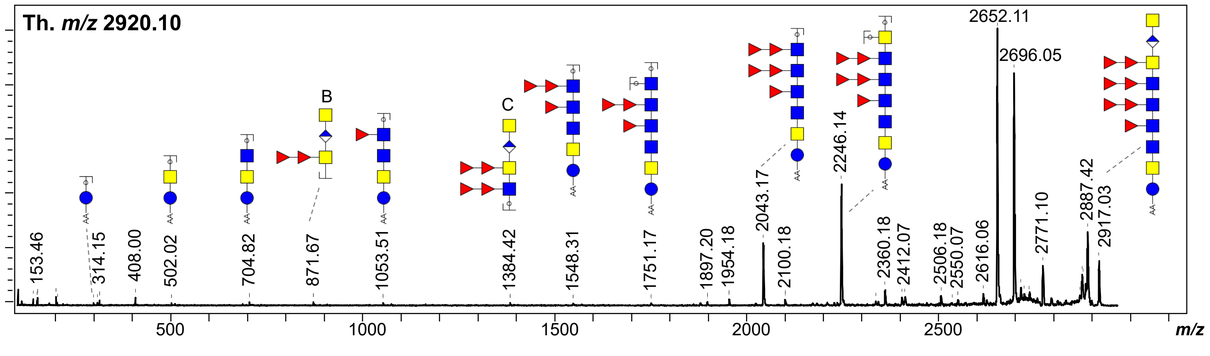
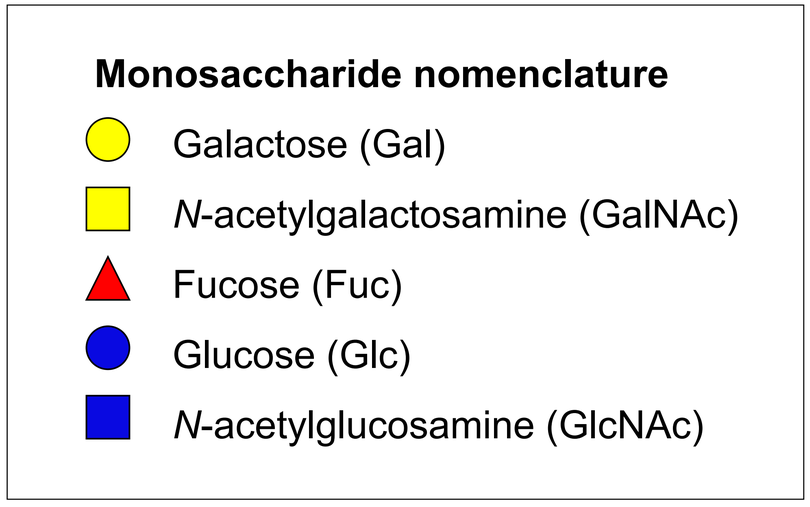


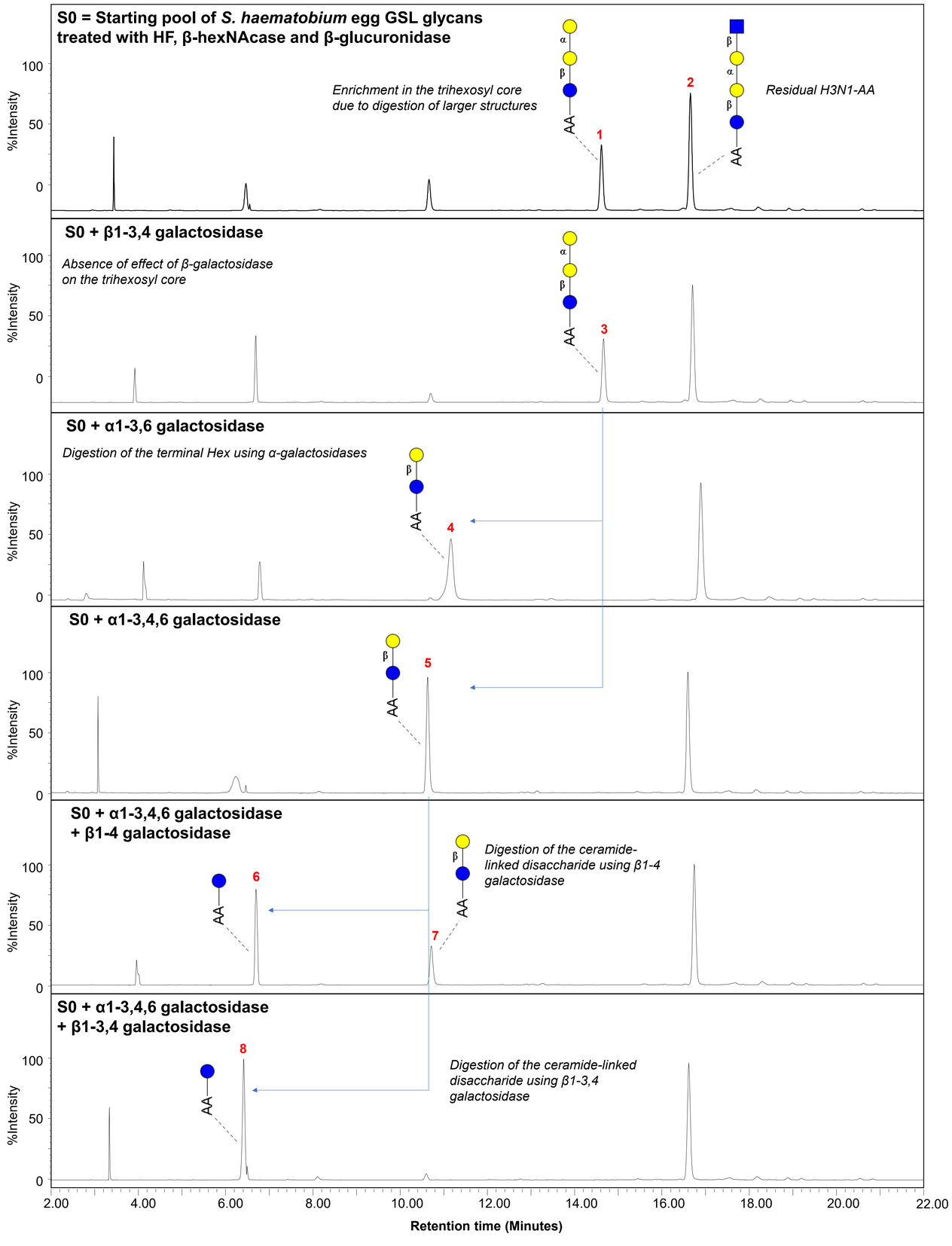
Figure S4 – UHPLC and MS analysis of *S. haematobium* GSL glycans

S. haematobium GSL glycans were released from their lipid carriers using rEGCase II prior to AA-labeling. Egg GSL glycans were treated with HF, β-hexNAcase and β-glucuronidase to digest the structures to their trihexosyl core in the conditions detailed in Figure 2 and in the Methods section. The obtained pool (S0) was further digested using a panel of exoglycosidases (Figure 2) and the resulting digestions were analyzed by UHPLC along with S0 untreated control. Fluorescence (A) and MS data (B) were recorded. Masses were obtained from a QDa mass detector set in positive mode (*m/z,* $[M-H]^+$). Signal intensities in percentages are indicated on the Y-axis. Blue arrows highlight products resulting from the aforementioned digestions.

Glycans are represented using the CFG nomenclature.



(A) UHPLC analysis of exoglycosidase digestions



(B) Corresponding *m/z*

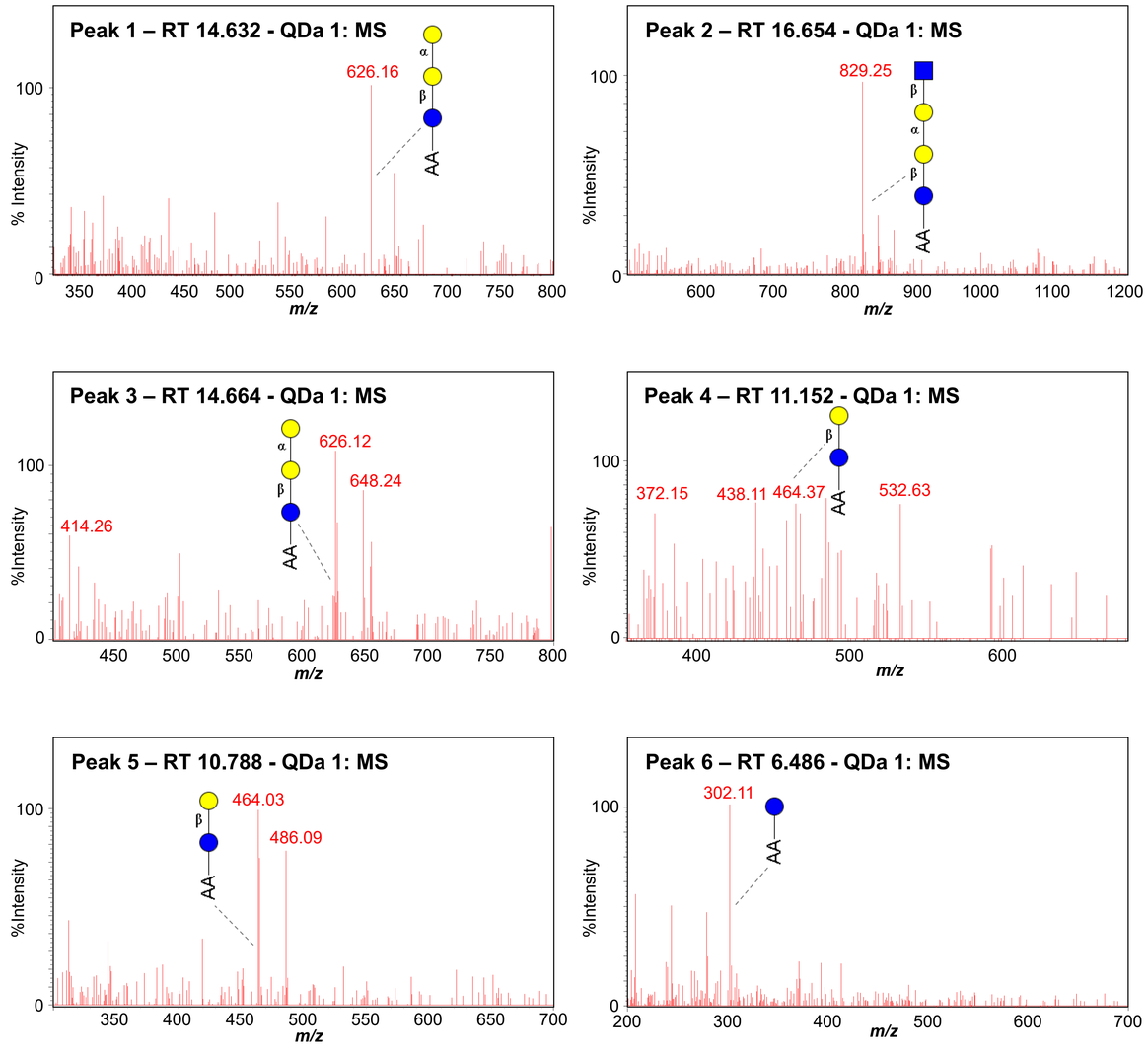


Figure S5 – PGC-nano-LC-MS and MS/MS of *S. haematobium* and *S. mansoni* GSL glycans

GSL glycans were released from their lipid carriers using rEGCase II. Upon purification, acidic and neutral fraction were separated. In specified cases released glycans were treated with HF and/or subjected to exoglycosidase digestions, as detailed in the Methods section. Native glycans and resulting treatment products were all reduced and cleaned using sequential C18 and PGC SPE prior to analysis using PGC-nano-LC-MS.

PGC-nano-LC-MS analysis of ion with *m/z* 505.18 (A) in the neutral GSL glycans of *S. haematobium* eggs either native or treated with HF and β-HexNAcase. Extracted ion chromatogram of *m/z* 505.18 is shown and elution patterns are compared to those of globotriose (Gb3, Galα1-4Galβ1-4Glc) and isoglobotriose (iGb3, Galα1-3Galβ1-4Glc) standards.

PGC-nano-LC-MS/MS analysis of selected GSL glycans (B-F) of *S. haematobium* worms (B), eggs (C-D) and of *S. mansoni* (E-F) eggs. Monoisotopic mass (Monoiso. *m/z*), charge state (Charge), theoretical (Th.) and observed (Exp.) ion *m/z*, the mass deviation between the theoretical and experimental *m/z* as well as the retention time (RT) of the fragmented ions are indicated on the top of each panel. All spectra were acquired in negative-ion reflectron mode and signals are labeled with monoisotopic masses (*m/z*)..

Note that based on glycan sequencing using exoglycosidases (Figure S2), we identified instances where the HexNAc preceding the GlcA residue could be either GlcNAc or GalNAc (Supplementary Data 2). The GalNAc-containing isomer is represented in panels C and E.

All glycans are represented using the CFG nomenclature and ion types are represented using the GlycoWorkBench symbols (see inset below), which follow the nomenclature of fragments of carbohydrates as defined by Domon and Costello (https://doi.org/10.1007/BF01049915).

Diagnostic ions and/or fragmentation patterns indicative of glycan traits of interest are listed in the **Table A** next page, together with the corresponding panels in which they are observed. A summary of all ion species analyzed by PGC-nano-LC-MS is provided in **Table B**.

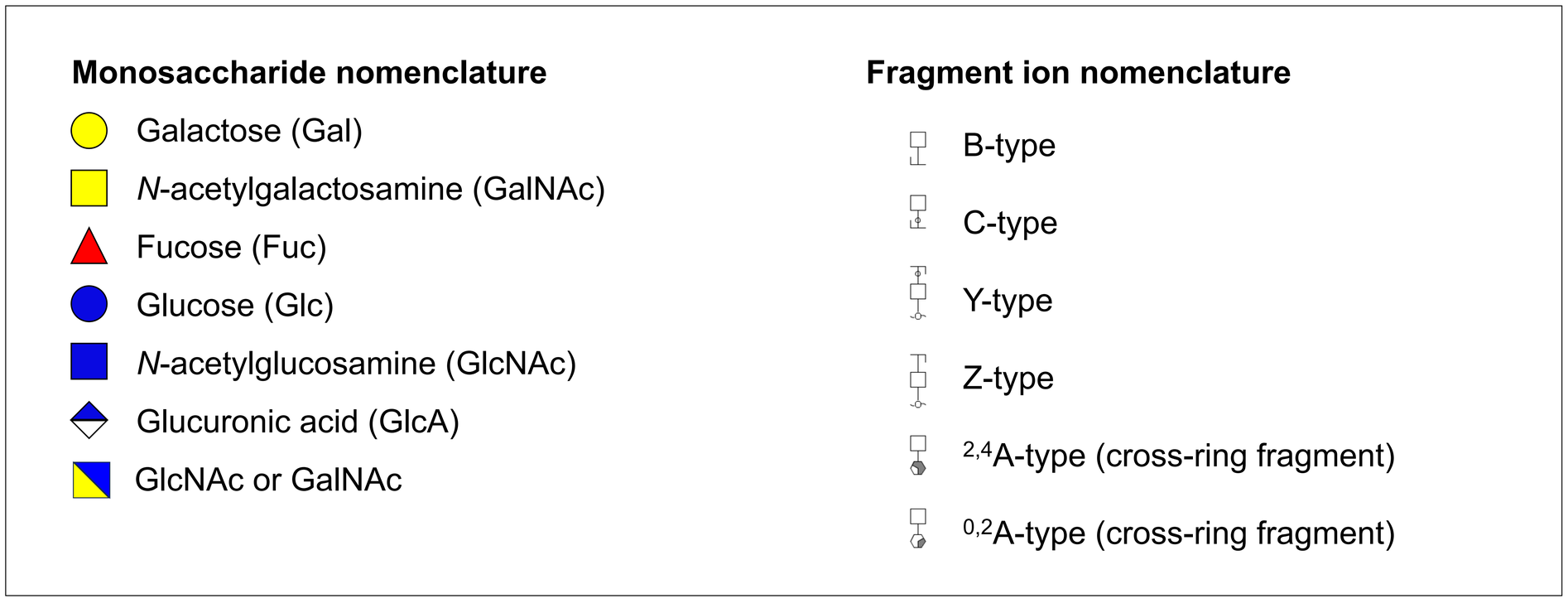


Table A Highlighted diagnostic ions and fragmentation patterns

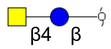
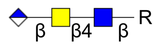
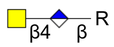
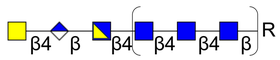
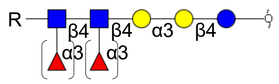
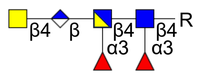
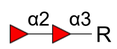
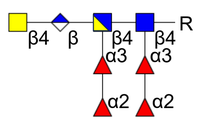
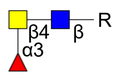
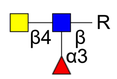
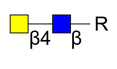
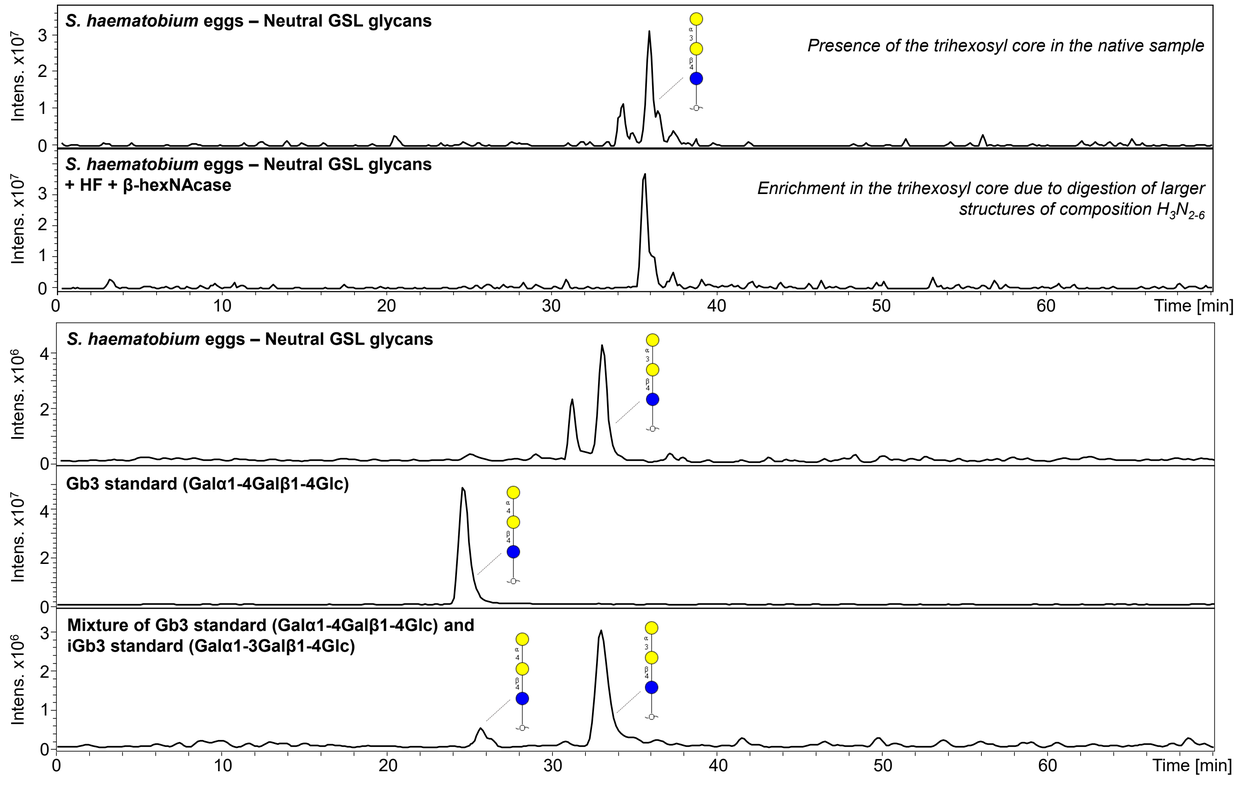
Identified glycan trait		Diagnostic ion *m/z* (theoretical *m/z*) or Δ*m*	Present in	Panel
Trihexosyl GSL core		Y_3-H_3 (*m/z* 505.18) Z_3-H_3 (*m/z* 487.17)	*S. haematobium* worms *S. haematobium* eggs - acidic & neutral	A-D
Disaccharide GSL core		Y_2-H_1N_1 (*m/z* 384.15) Z_2-H_1N_1 (*m/z* 366.14)	*S. mansoni* eggs - acidic	E-F
Terminal GlcA		Δ*m* = 176.03 (A_1) Y_5-H_3N_2 (*m/z* 911.33)	*S. haematobium* adult worms	B
Terminal HexNAc and subterminal GlcA		Δ*m* = 203.08 (N_1) followed by Δ*m* = 176.03 (A_1) B_2-N_1A_1 (*m/z* 378.10) and C_2-N_1A_1 (*m/z* 396.11)	*S. haematobium* eggs - acidic *S. mansoni* eggs - acidic	C E-F
Subterminal GlcA in a chain of 1 to 4 unsubstituted HexNAc		B_{3-6}-$N_{2-5}A_1$ (*m/z* 581.18; 784.26; 987.34;1190.42) C_{4-7}-$N_{3-6}A_1$ (*m/z* 802.27; 1005.35; 1208.43; 1411.51)	*S. haematobium* eggs - acidic	C
Fuc in proximity with trihexosyl core		Z_4-$H_3N_1F_1$ (*m/z* 836.30) and Y_4-$H_3N_1F_1$ (*m/z* 854.31) Z_5-$H_3N_2F_1$ (*m/z* 1039.38) and Y_5-$H_3N_2F_1$ (*m/z* 1057.39) Y_5-$H_3N_2F_2$ (*m/z* 1203.45)	*S. haematobium* eggs - acidic & neutral	C-D
Fuc and F-LDN-F in proximity with subterminal GlcA		C_3Y_{6-8}-$N_1A_1F_1$ (*m/z* 542.17) and C_3-$N_2A_1F_1$ (*m/z* 745.25) C_4-$N_3A_1F_2$ (*m/z* 1094.39)	*S. mansoni* eggs - acidic	F
DF		C_2-F_2 (*m/z* 309.11)	*S. mansoni* eggs - acidic	F
DF-LDN-DF in proximity with subterminal GlcA		B_3Z_{6-8}-$N_1A_1F_2$ (*m/z* 652.20; C_3Y_{6-8}-$N_1A_1F_2$ (*m/z* 688.23) B_3Y_{4-6}-$N_2A_1F_2$ (*m/z* 873.30) C_3-$N_2A_1F_4$ (*m/z* 1165.42) B_4-$N_2A_1F_4$ (*m/z* 1368.50 and C_4-$N_2A_1F_4$ (*m/z* 1386.50)	*S. mansoni* eggs - acidic	F
F-LDN		B_1-N_1F_1 (*m/z* 348.13) and Z_4-H_3N_1 (*m/z* 690.25) and Y_4-H_3N_1 (*m/z* 708.26)	*S. haematobium* eggs - neutral	D
LDN-F		B_2-N_2F_1 (*m/z* 551.21) and Z_4-$H_3N_1F_1$ (*m/z* 836.30) and Y_4-$H_3N_1F_1$ (*m/z* 854.31)	*S. haematobium* eggs - neutral	D
LDN		B_2-N_2 (*m/z* 405.15) C_2-N_2 (*m/z* 423.16) $^{2,4}A_2$-N_2 (*m/z* 465.17)	*S. haematobium* eggs - neutral	D

Table B List of ion species analyzed by PGC-nano-LC-MS

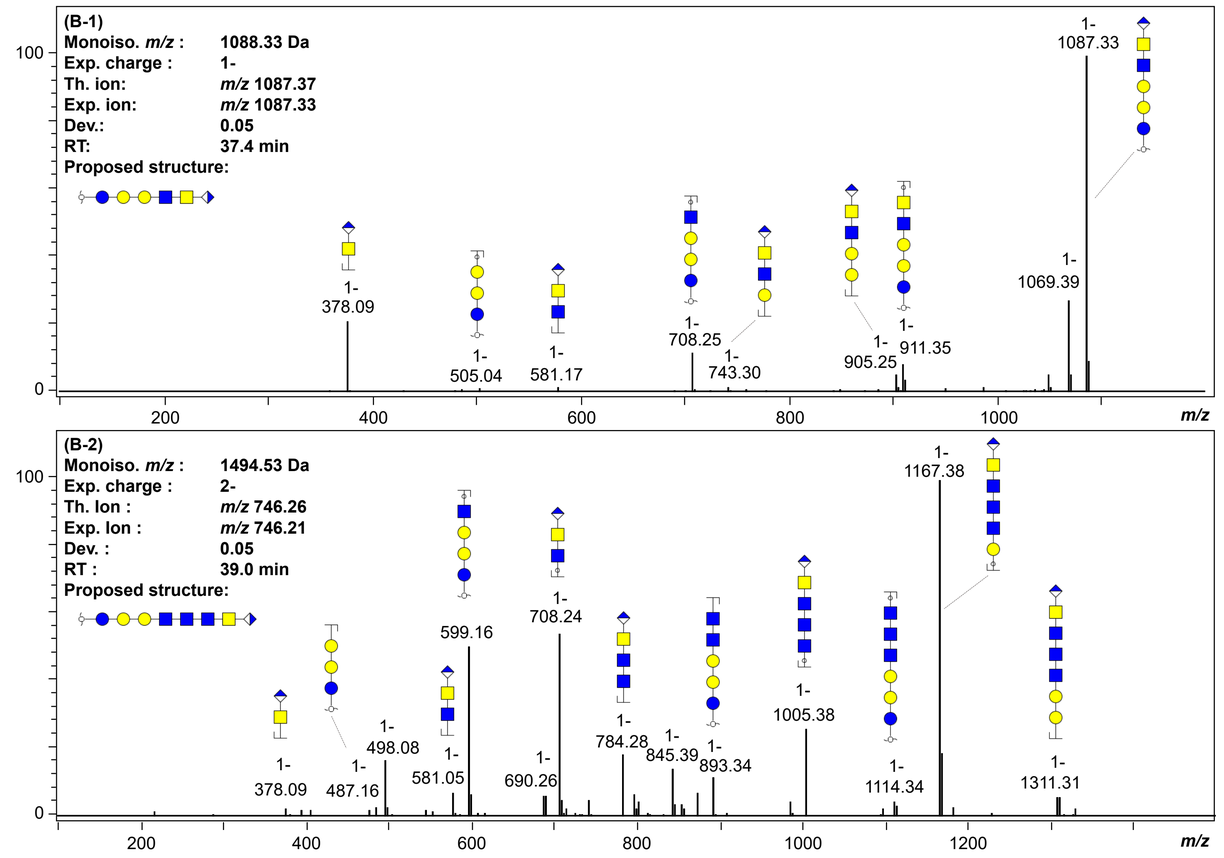
Species Life-stage Acidic/Neutral	Glycan composition	Th. *m/z* MALDI	Th. *m/z* PGC-LC-MS		Glycan ID	Panel
			$[M-H]^{1-}$	$[M-H]^{2-}$		
S. haematobium Eggs - Neutral	H3	624.21	505.18	n.d. (not detected)	-	A
S. haematobium Worms - Acidic	H3N2A1	1206.41	1087.37	n.d.	B1	B
	H3N4A1	1612.56	1493.53	746.26	B2	B
S. haematobium Eggs Acidic	H3N5A1F1	1961.70	1842.66	920.83	C1	C
	H3N6A1	2018.72	1899.69	949.34	C2	C
	H3N6F1A1	2164.78	2045.74	1022.37	C3	C
	H3N7A1	2221.80	2102.77	1050.88	C4	C
	H3N6A1F2	2310.84	2191.80	1095.40	C5	C
	H3N7A1F1	2368.86	2248.82	1123.91	C6	C
	H3N7A1F2	2513.92	2394.88	1196.94	C7	C
	H3N7A1F3	2659.98	2540.94	1269.97	C8	C
S. haematobium Eggs Neutral	H3N2	1030.35	911.34	455.16	D1	D
	H3N2F1	1176.40	1057.39	528.19	D2a-b (2 iso)	D
	H3N4	1436.51	1317.49	658.24	D3	D
	H3N4F1	1582.60	1463.55	731.27	D4a-b (2 iso)	D
	H5N3F2	1849.64	1730.64	864.81	D5	D
	H3N5F2	1931.66	1812.69	905.84	D6	D
S. mansoni Eggs Acidic + HF	H1N5A1	1491.54	1372.50	685.75	E1	E
	H1N6A1	1694.62	1575.58	787.29	E2	E
	H1N7A1	1897.70	1778.66	888.83	E3	E
	H1N8A1	2100.78	1981.74	990.37	E4	E
S. mansoni Eggs Acidic	H1N5A1F4	2075.77	1956.73	977.86	F1a-b (2 iso)	F
	H1N6A1F3	2132.79	2013.75	1006.37	F2	F
	H1N6A1F4	2278.85	2159.81	1079.40	F3a-b (2 iso)	F
	H1N7A1F3	2335.87	2216.83	1107.91	F4	F
	H1N6A1F5	2424.91	2305.87	1152.43	F5	F
	H1N7A1F4	2481.93	2362.89	1180.94	F6	F
	H1N6A1F6	2570.96	2451.92	1225.46	F7a-c (3 iso)	F
	H1N7A1F6	2774.04	2655.01	1327.00	F8	F

Isomeric glycan structure detected (with number of identified isomers)

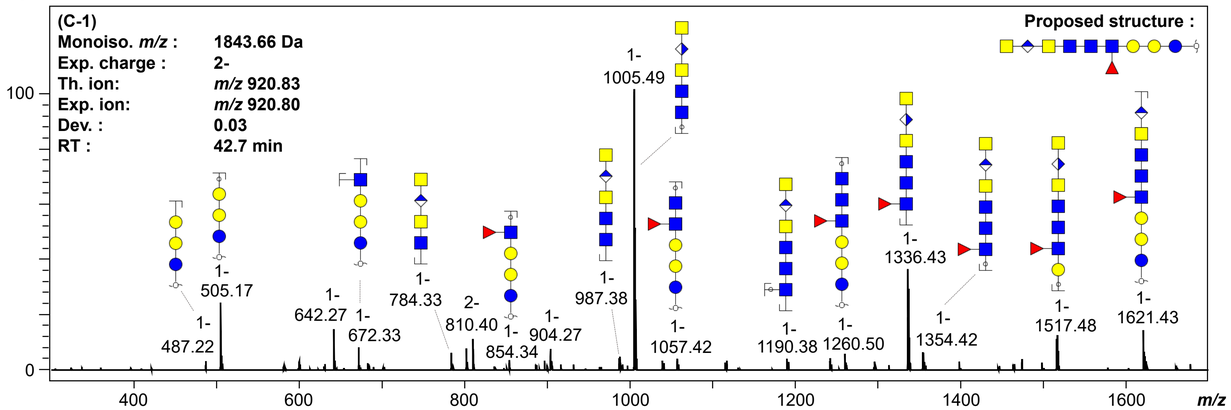
(A) Extracted ion chromatogram of *m/z* 505.18

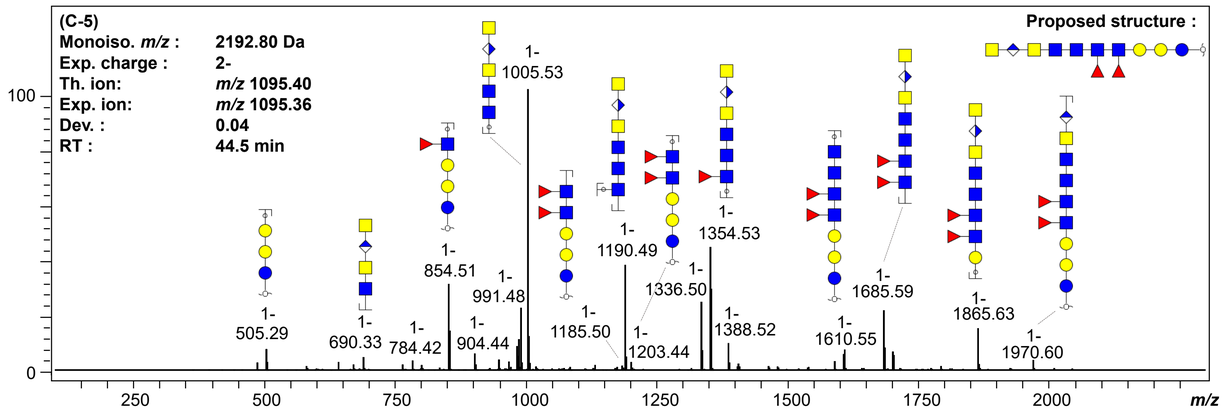


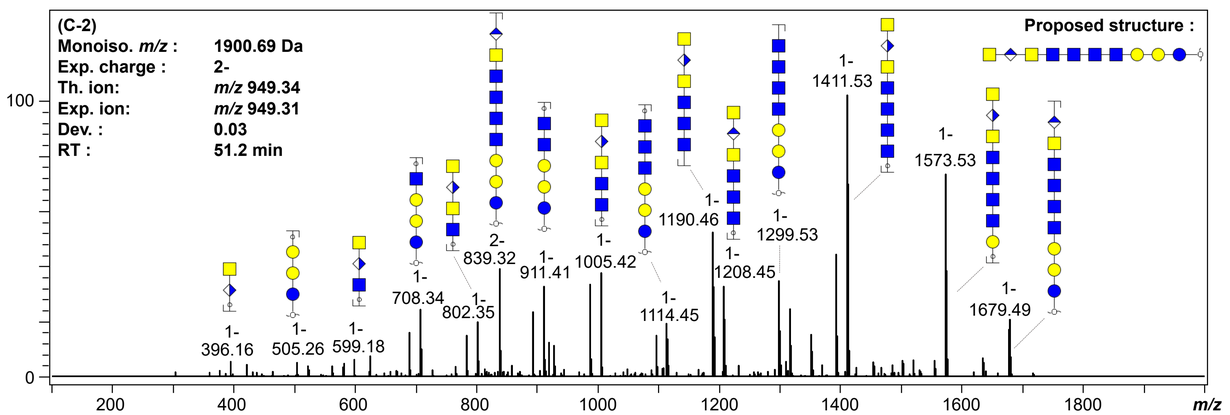
(B) *S. haematobium* worm acidic GSL glycans

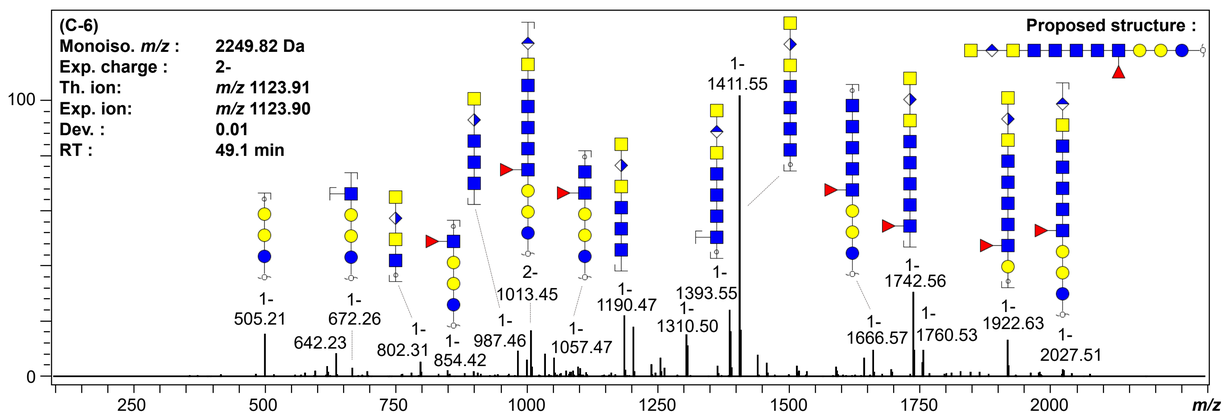


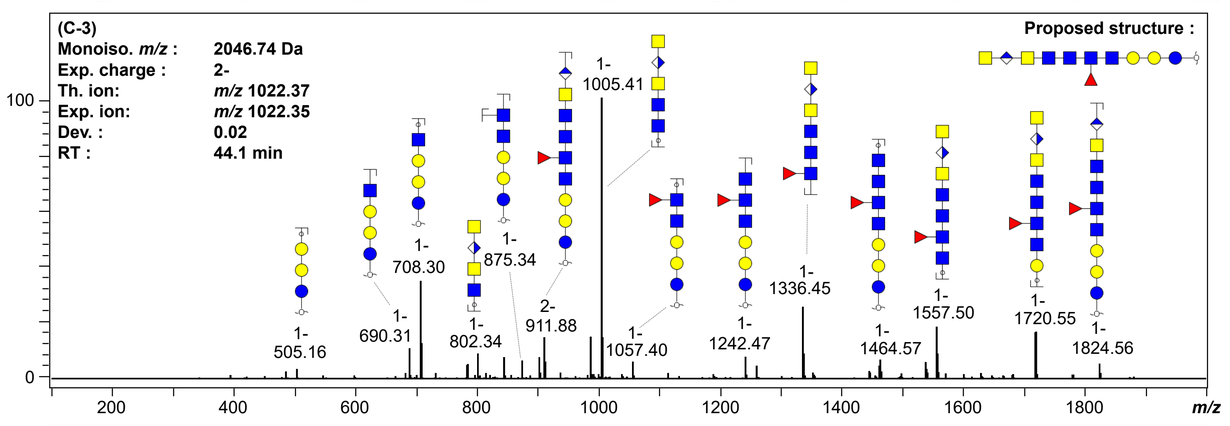
(C) *S. haematobium* egg acidic GSL glycans

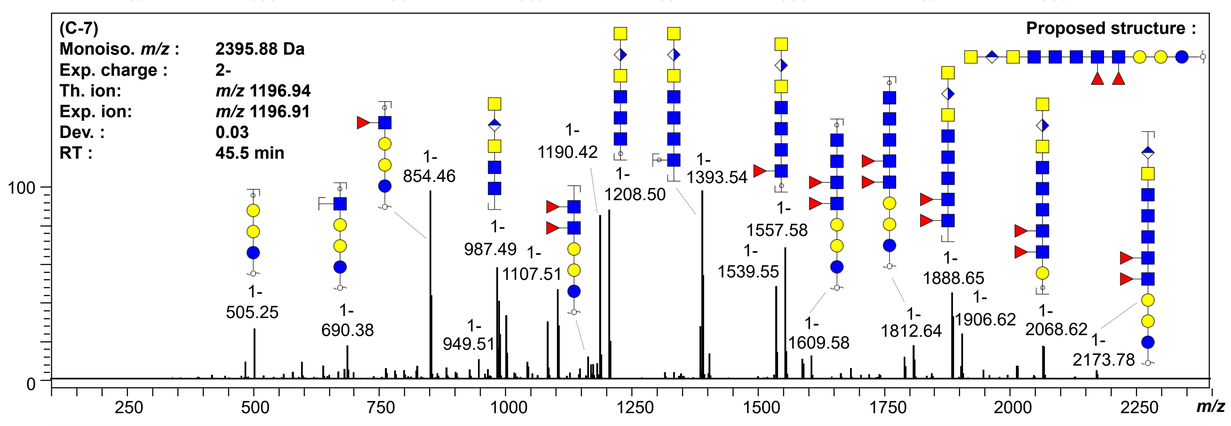


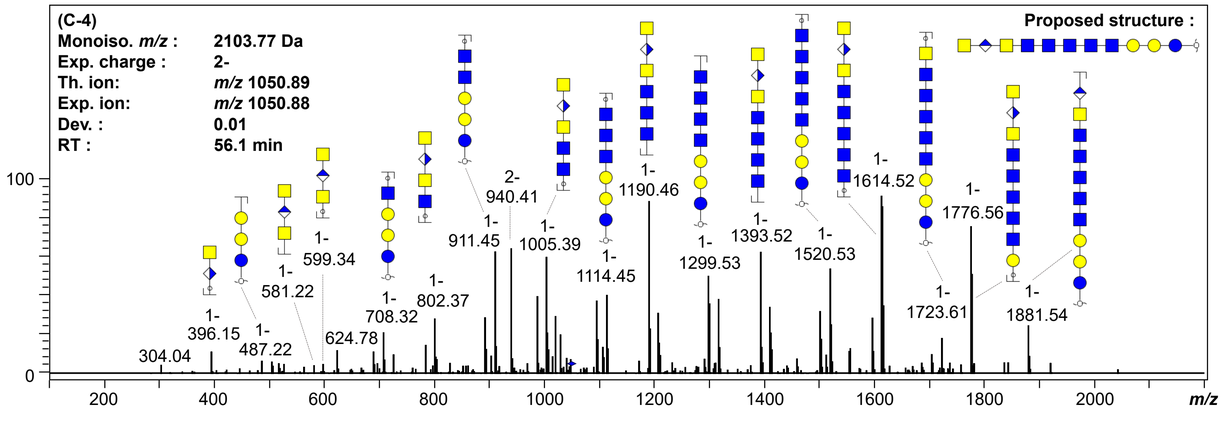


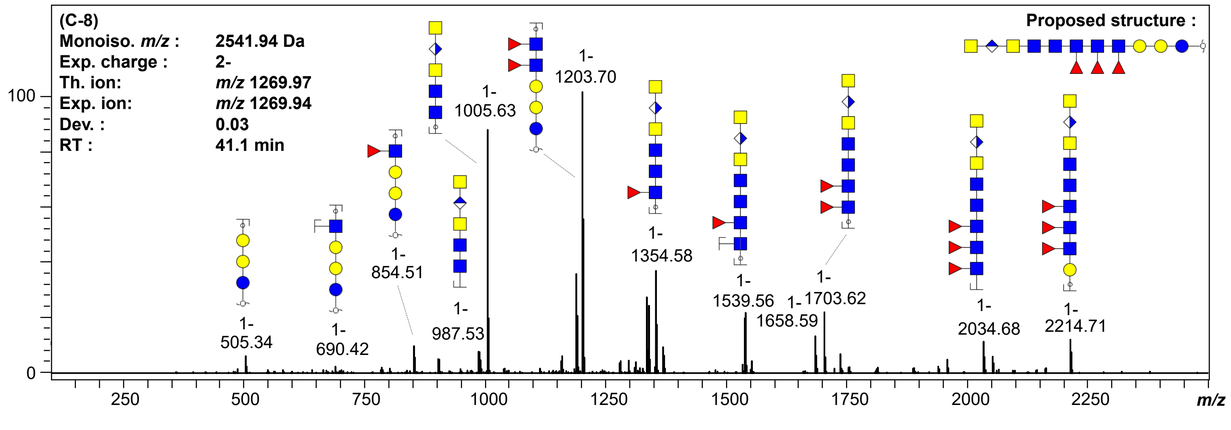




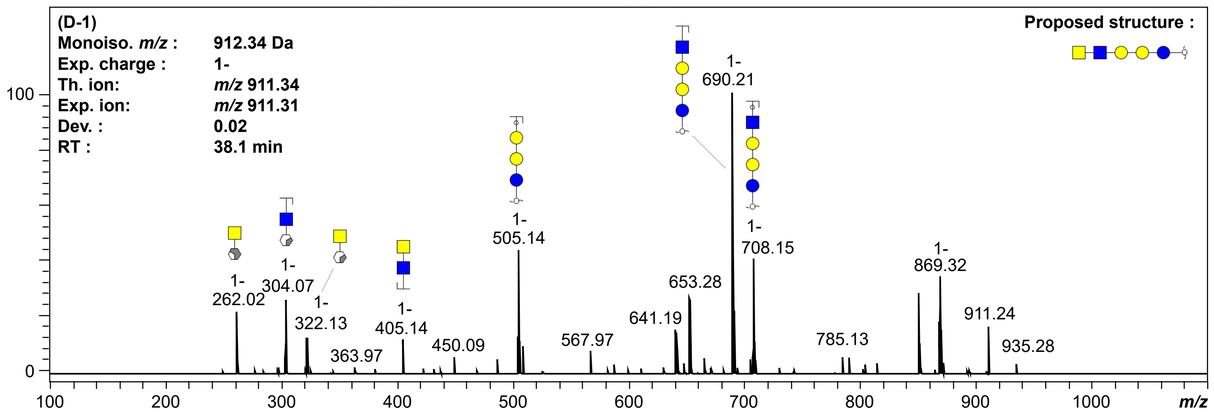


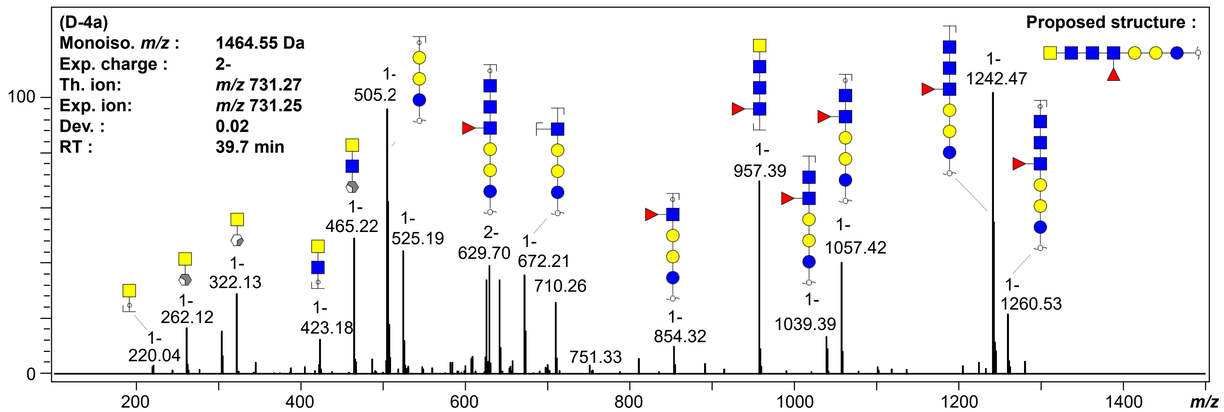


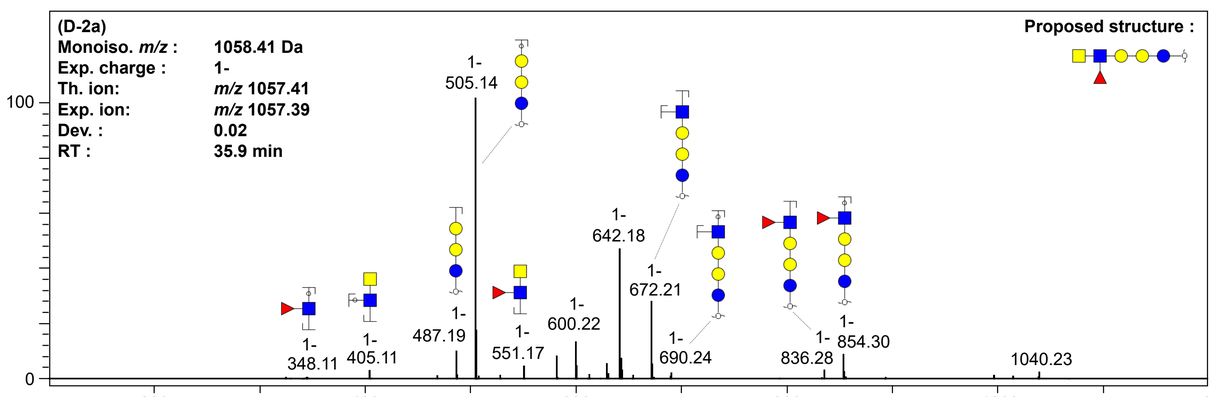


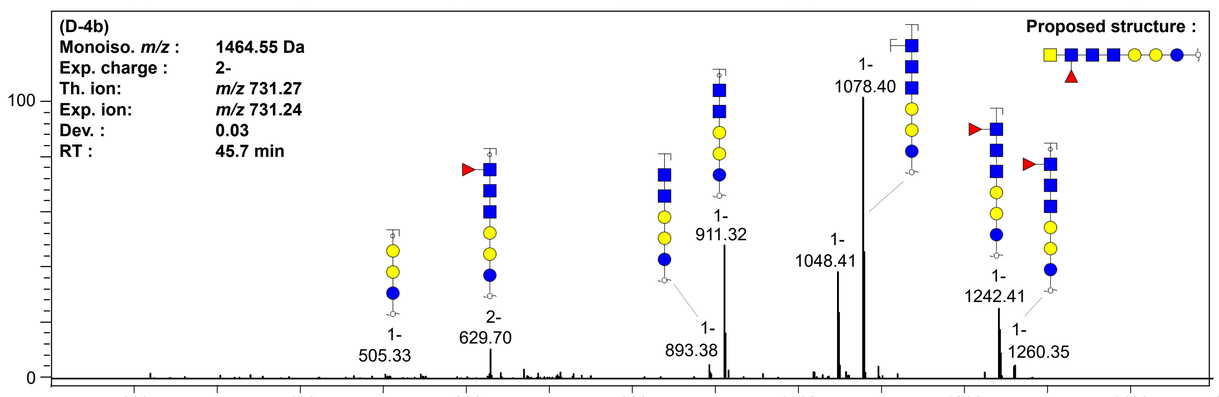


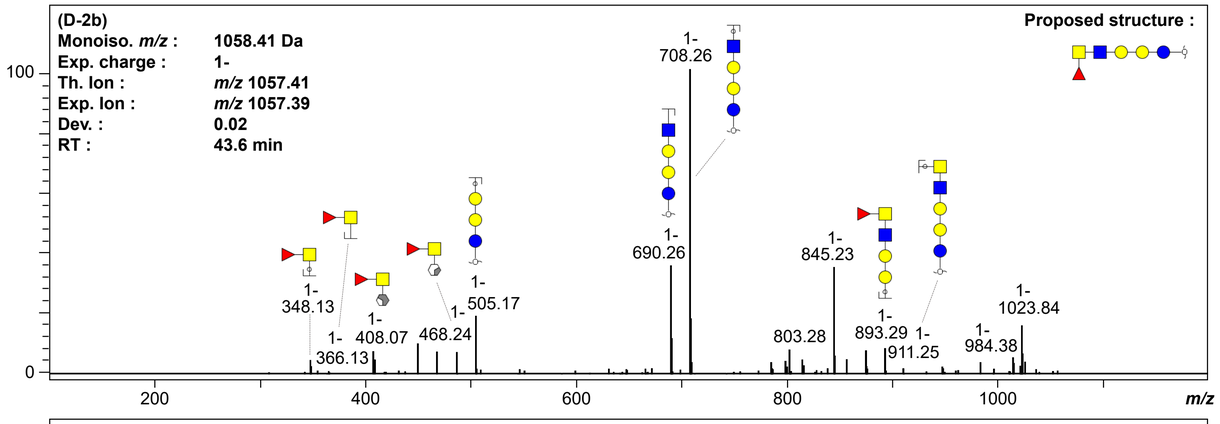
(D) *S. haematobium* egg neutral GSL glycans

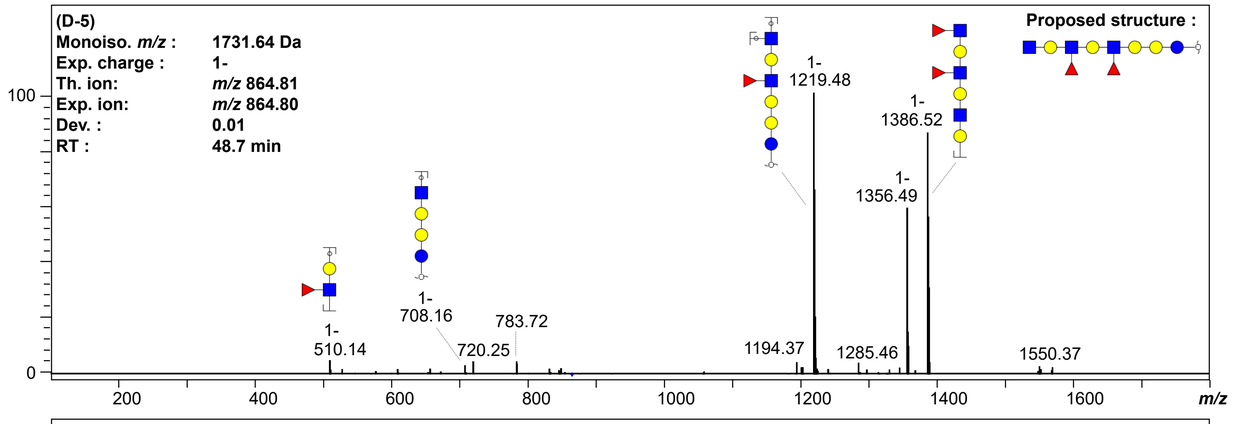


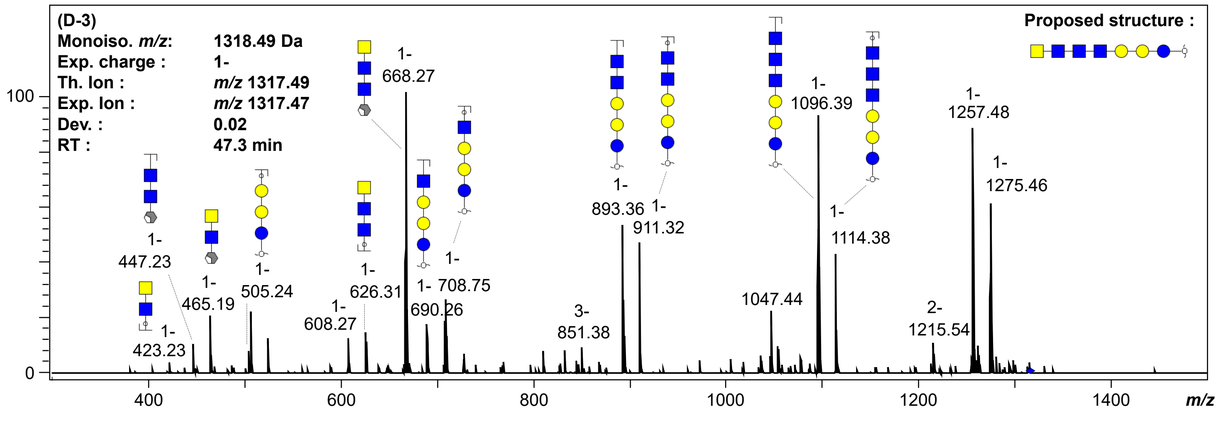


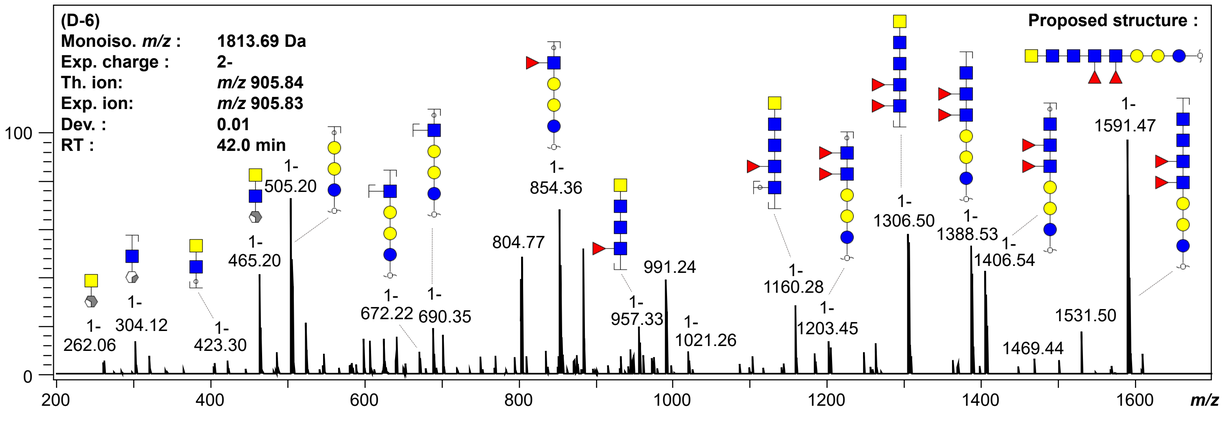




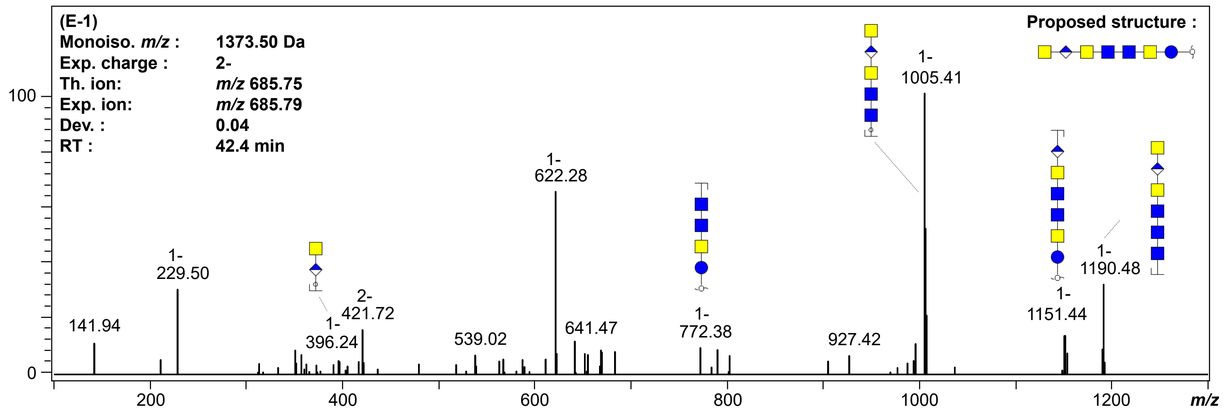


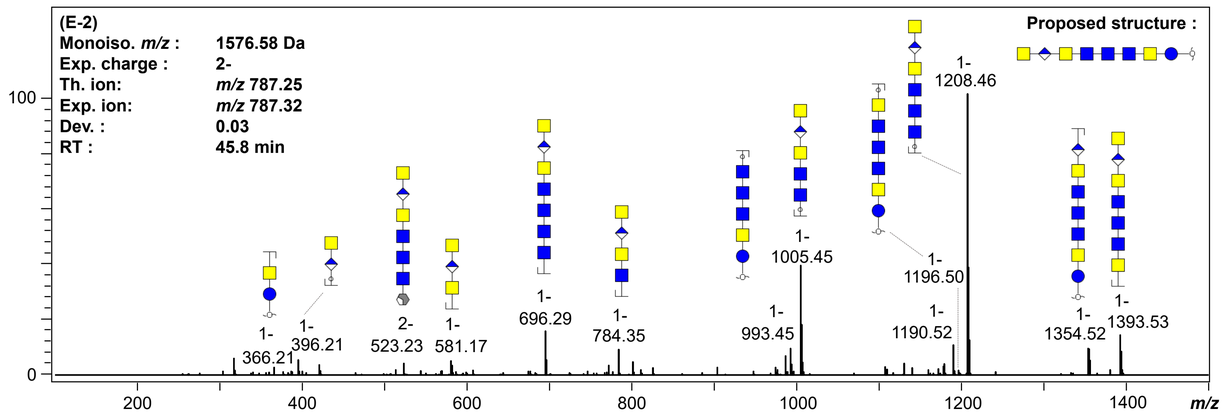


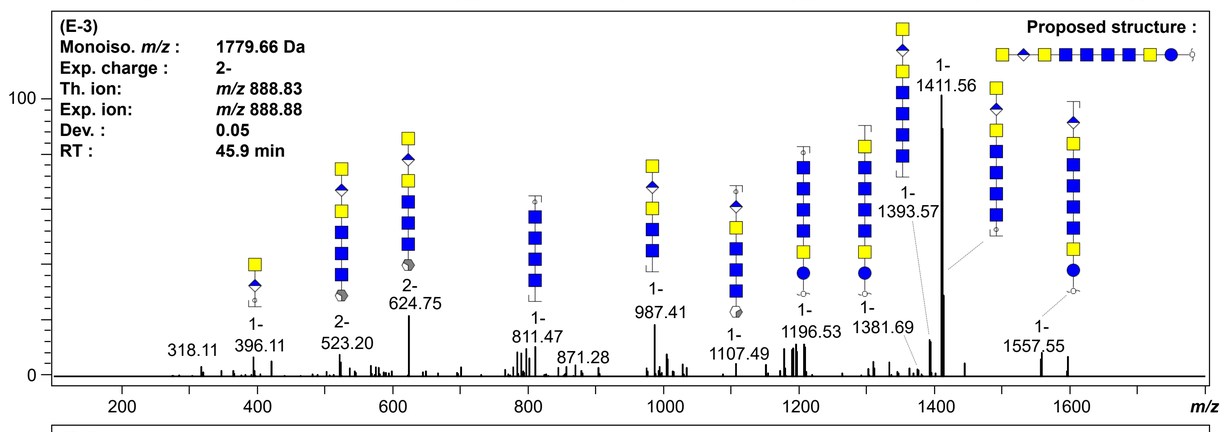


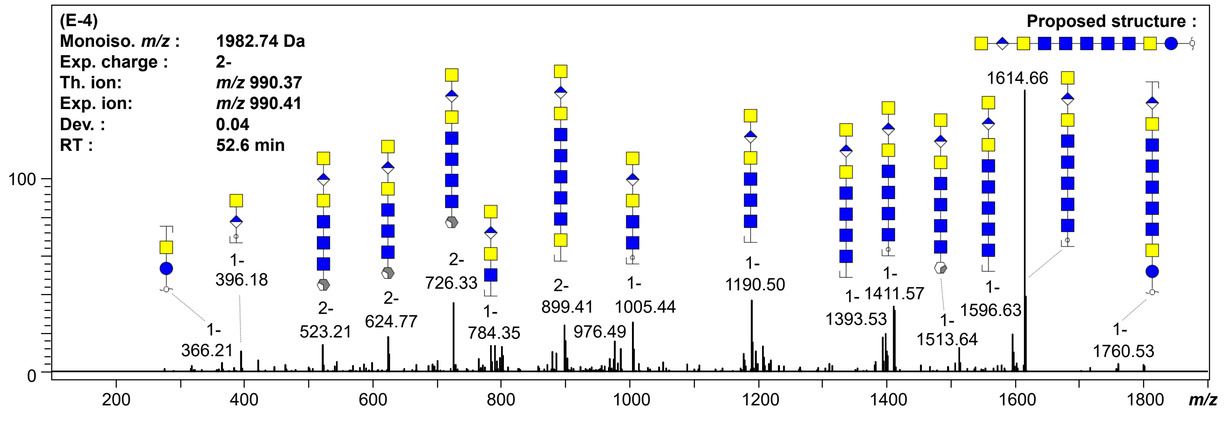


(E) *S. mansoni* egg acidic GSL glycans - HF treated

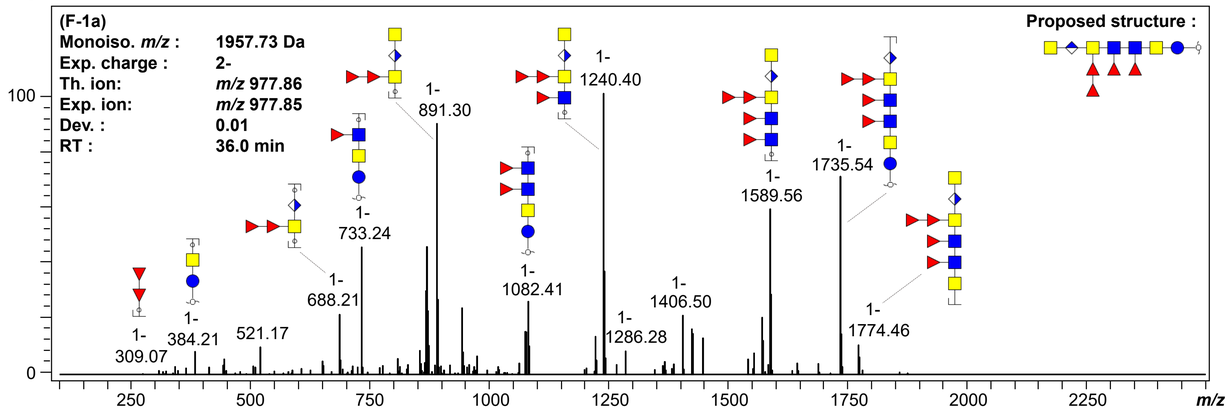


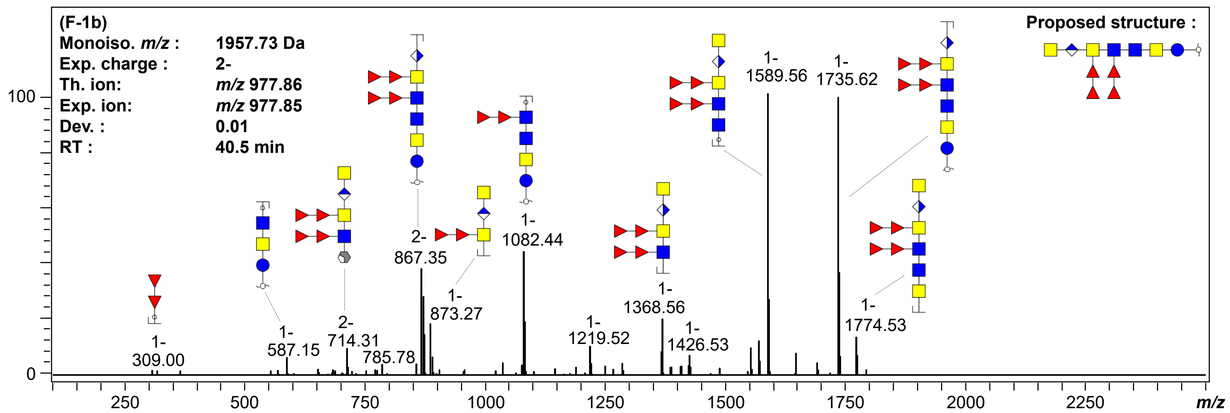




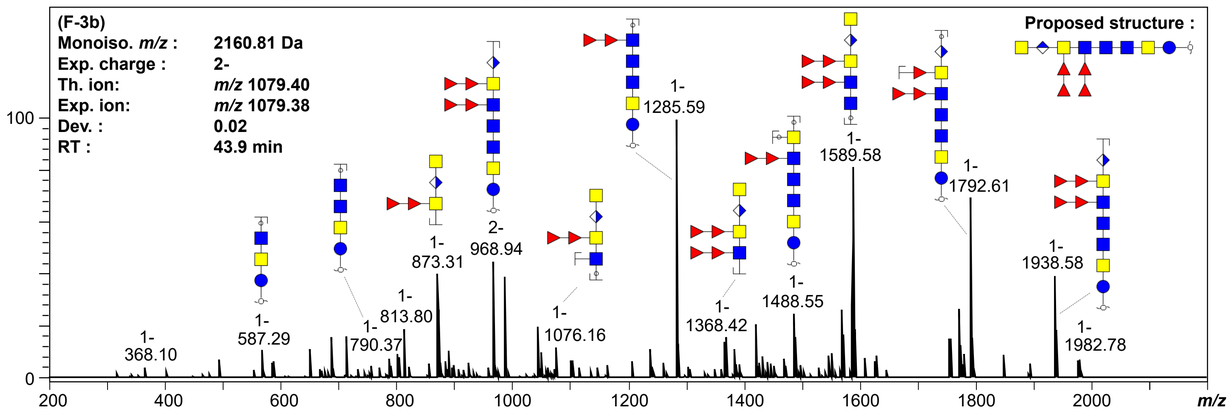


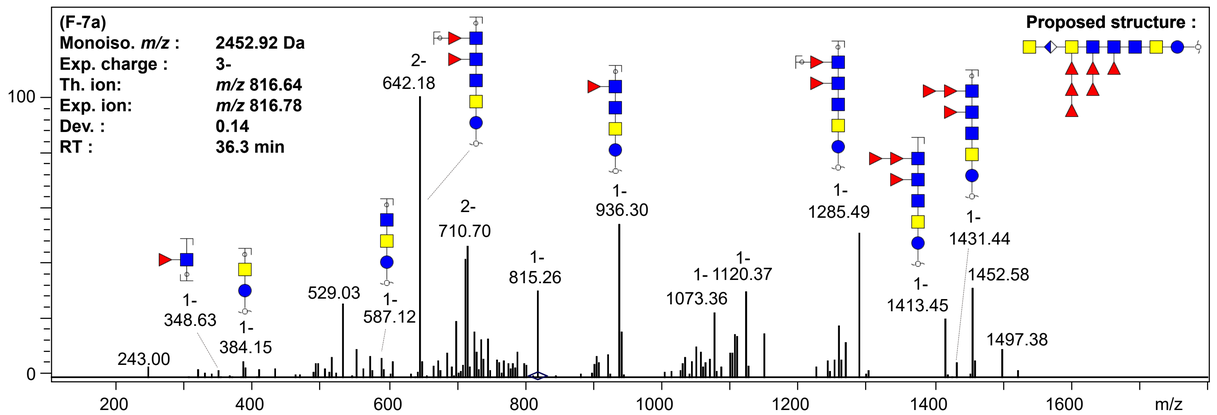
(F) *S. mansoni* egg acidic GSL glycans

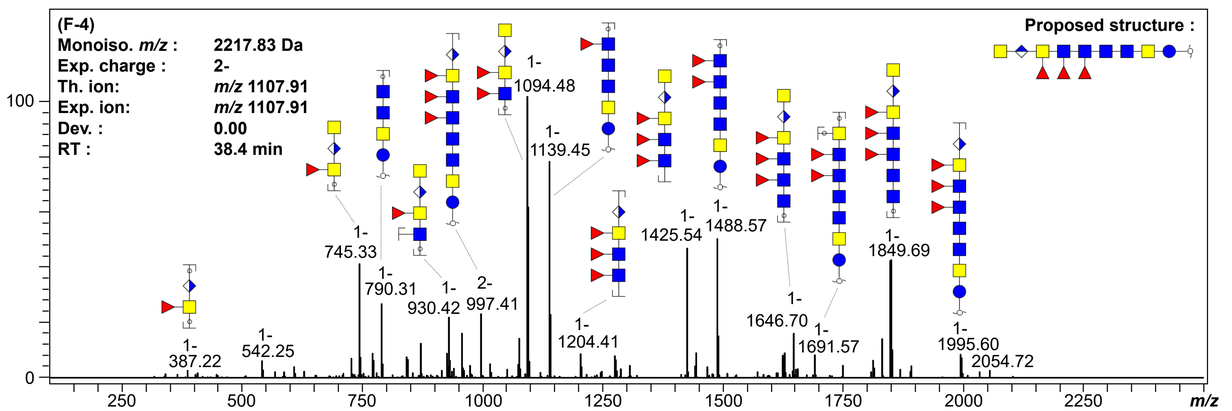


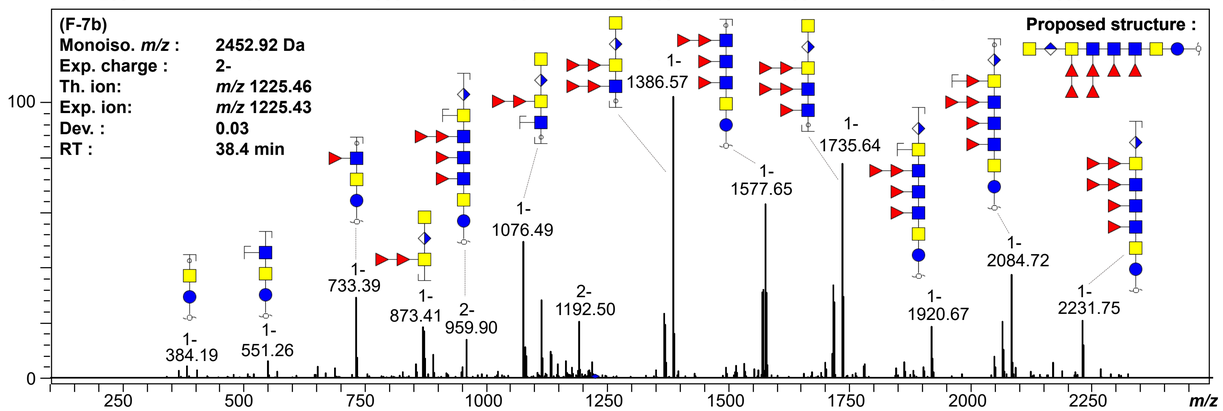


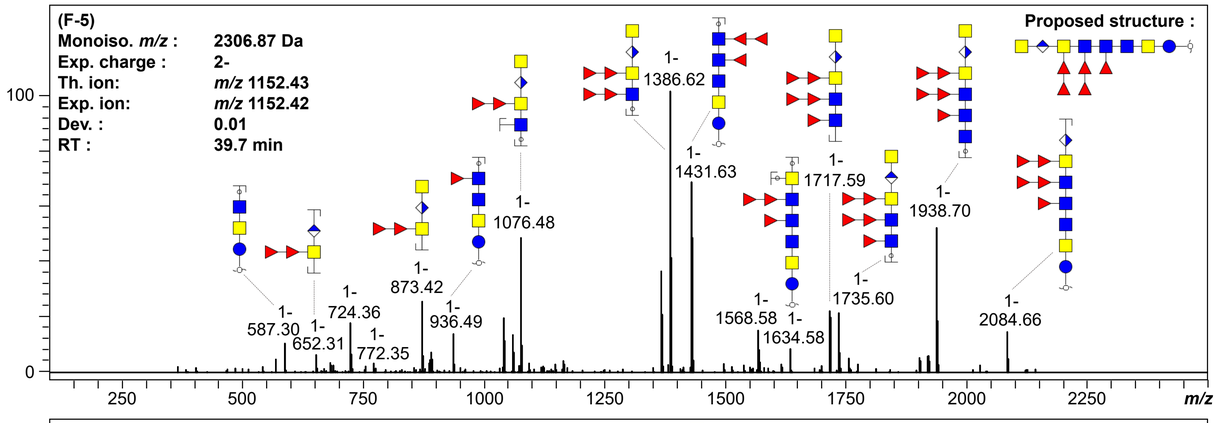
(F) *S. mansoni* egg acidic GSL glycans *(continued)*

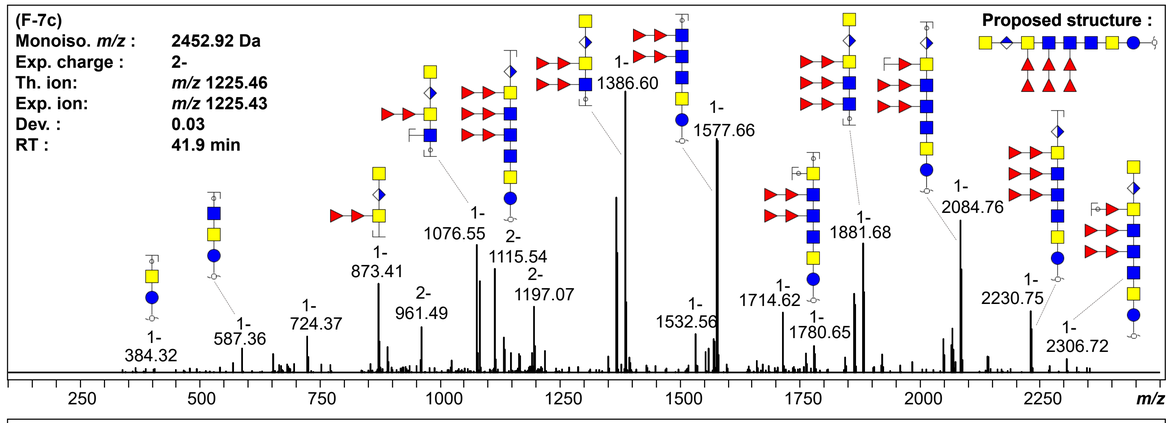


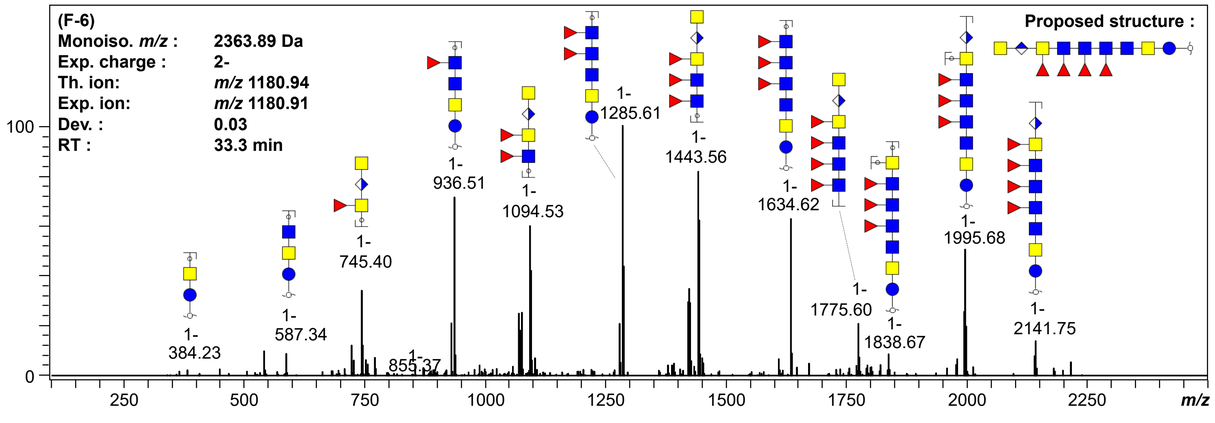












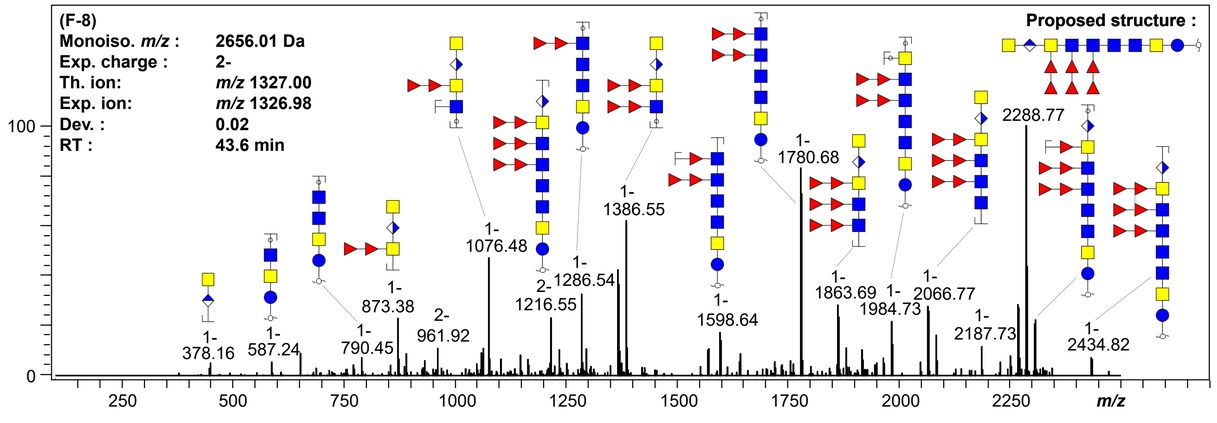


Figure S6 – Structural characterization of *S. haematobium N*-glycans

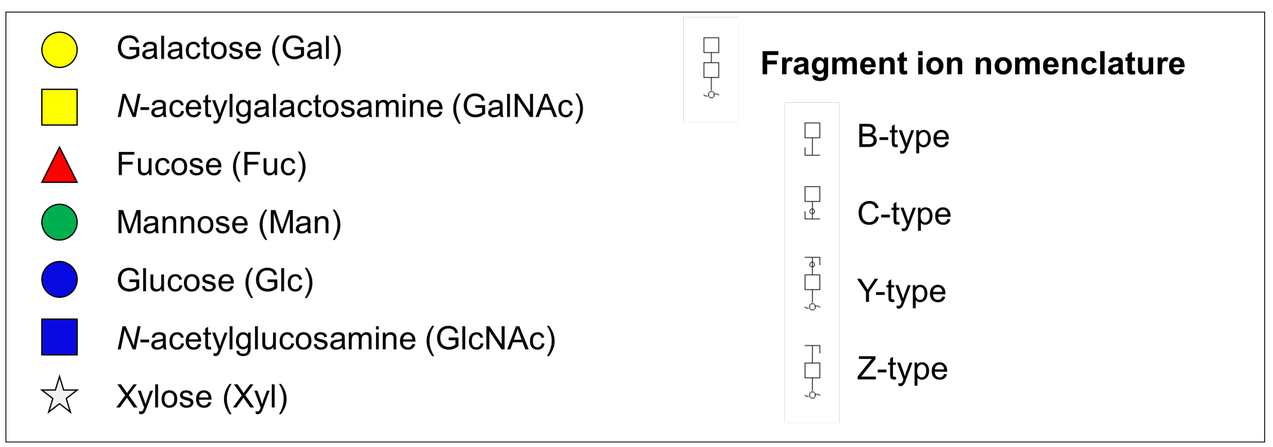
N-linked glycans were released from *S. haematobium* glycoproteins using PNGase F (A-C,F) and PNGase A sequentially (D-G) prior to AA labeling.

Glycan sequencing (A-D) *N*-glycans of *S. haematobium* cercariae (A), adult worms (B) and eggs (C, D) – either a mixture of eggs at different stages of maturity (C) or mature eggs only (E) - were subjected to hydrofluoric acid (HF) treatment and/or to digestion with exoglycosidase(s). Reactions were performed using the conditions detailed in Figure 2 (see Methods section). Treated samples and undigested control were then analyzed using MALDI-TOF-MS. All spectra were acquired in negative-ion reflectron mode and signals are labeled with monoisotopic masses (*m/z,* $[M-H]^-$). Signal intensities in % are indicated on the Y-axis. Sample type (untreated control or treated sample) and parasite life-stage (cercariae, adult worms, total or mature eggs) are indicated at the top of each panel. Blue arrows highlight the products resulting from the aforementioned treatments. Contaminants introduced by the preparation and known non-glycan signals are labeled with the # symbol.

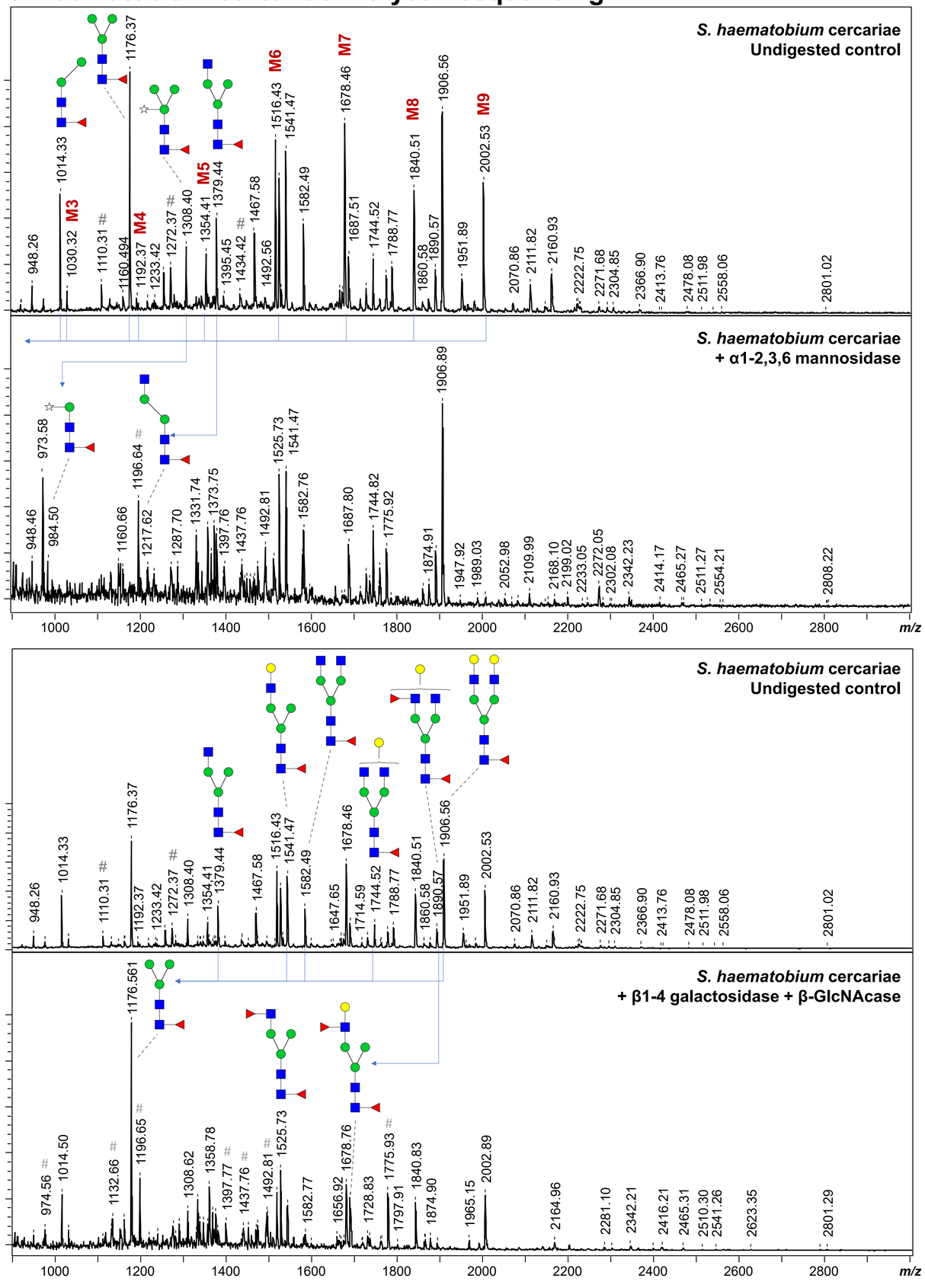
MALDI-TOF-MS/MS (E) of selected PNGase A-specific ions present in the *N*-glycan spectra of mature eggs was performed in negative-ion mode. Theoretical masses of fragmented ions are indicated on the upper left corner of each panel. Y-type fragment-ions, as defined by Domon and Costello (https://doi.org/10.1007/BF01049915) are represented, unless indicated otherwise (B = B type, C = C type, Z = Z type).

MALDI-TOF-MS (F, G) of *N*-glycans derived from immature and mature eggs of *S. haematobium*, separated by Percoll gradient centrifugation (see Methods section). A focus on *m/z* 2500-3500 highlighting the presence of high molecular weight structures is provided (G). Of note, in experiments in which PNGase A was applied to a mixture of eggs at various stages of maturity, minor ions corresponding to *N*-glycan structures up to ~4000 Da were observed (data not shown). The lack of interpretable MS2 spectra prevented verification of whether these carried four antennae or polyLacNAc repeats.

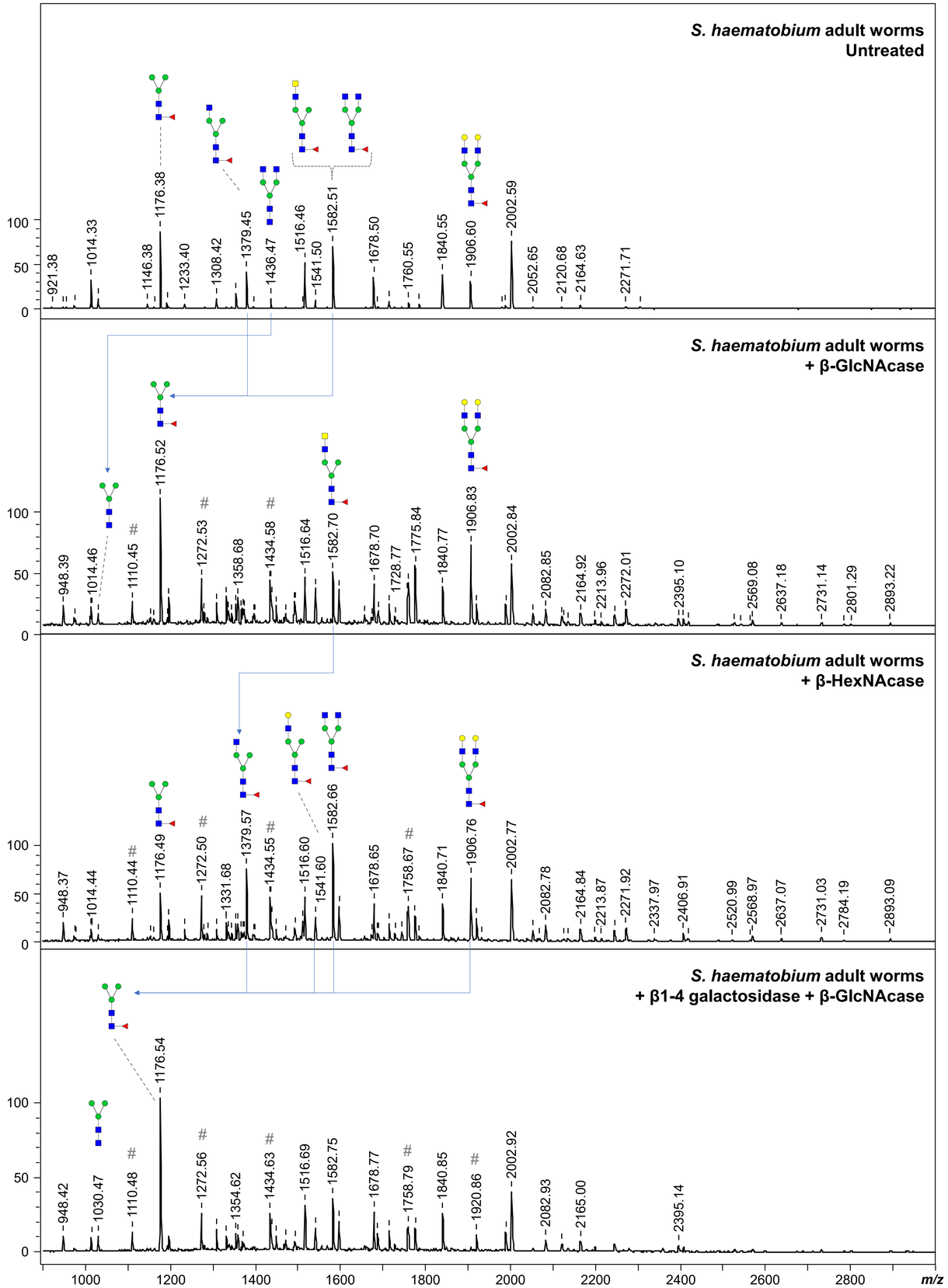
All (putative) glycans are represented using the CFG nomenclature (see inset below).



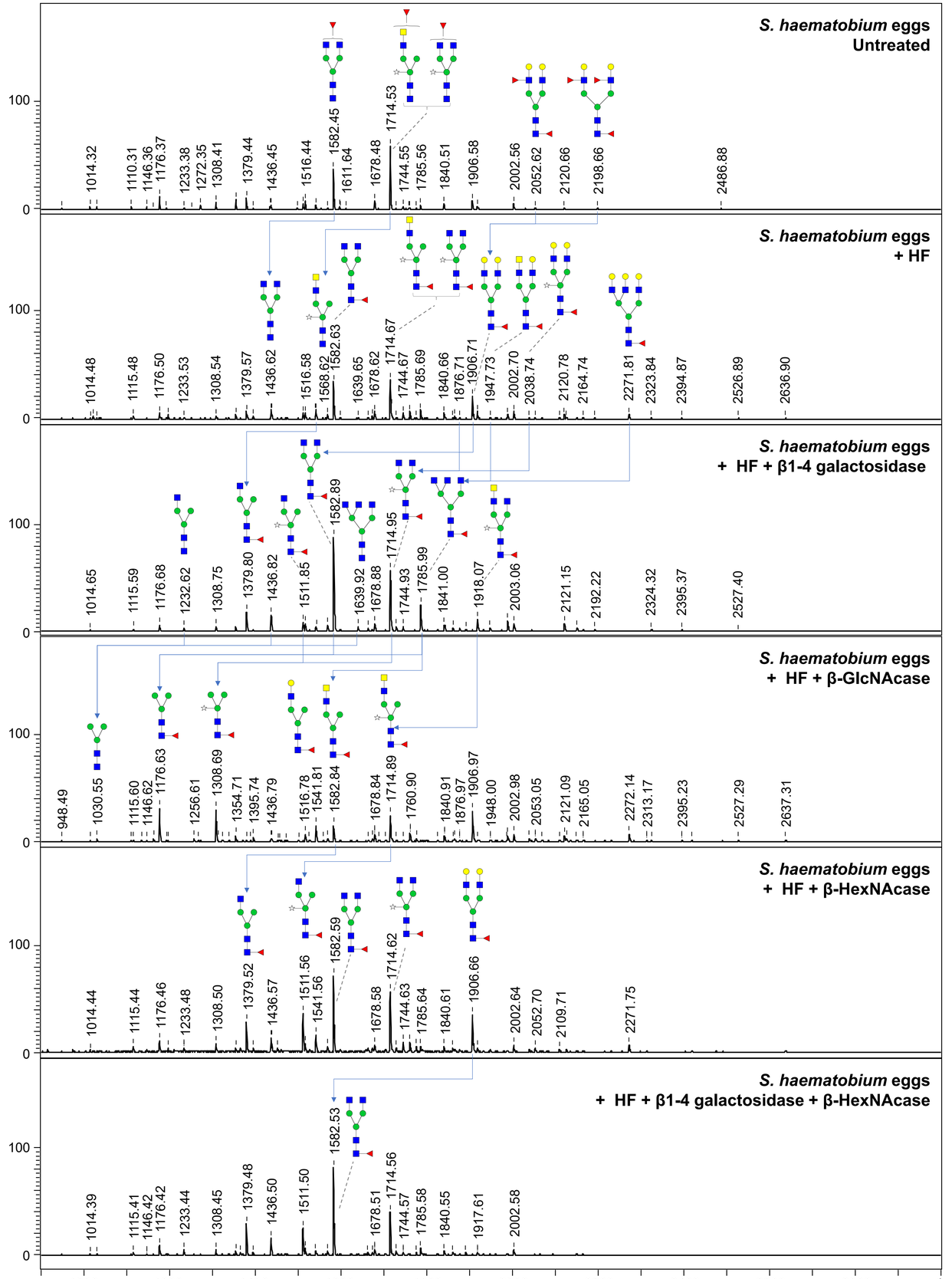
(A) *S. haematobium* cercariae – Glycan sequencing



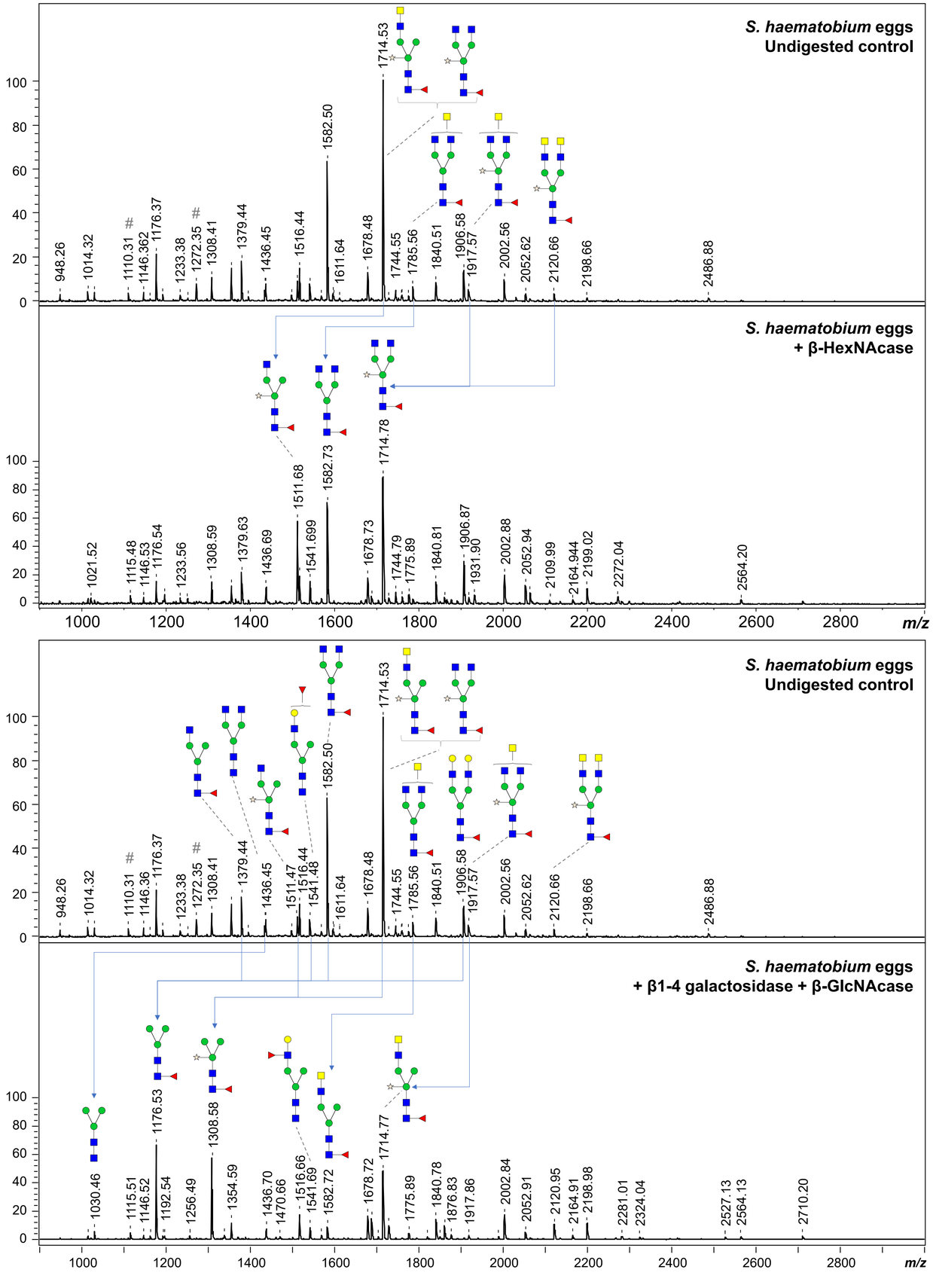
(B) *S. haematobium* adult worms – Glycan sequencing



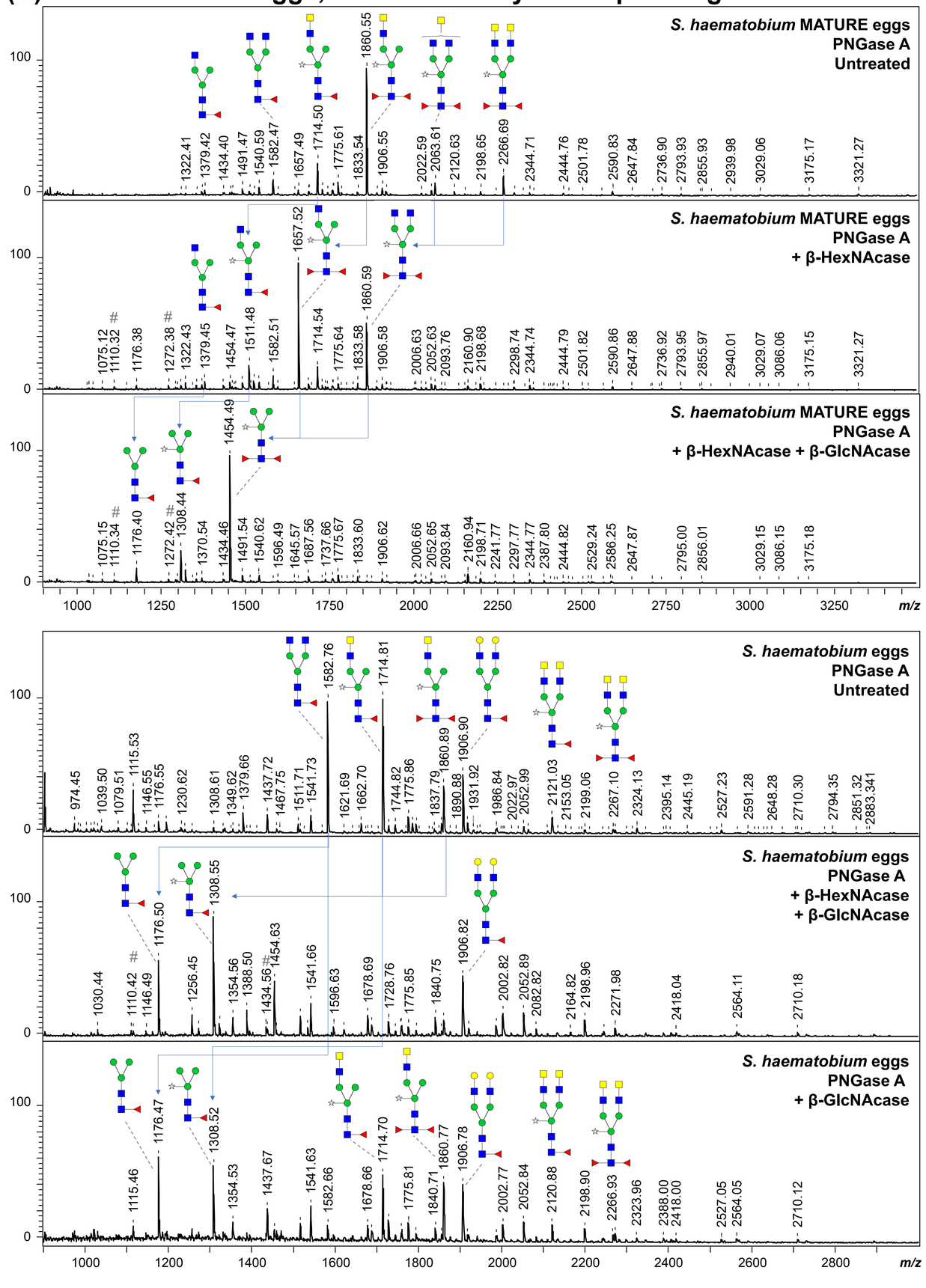
(C) *S. haematobium* eggs – Glycan sequencing



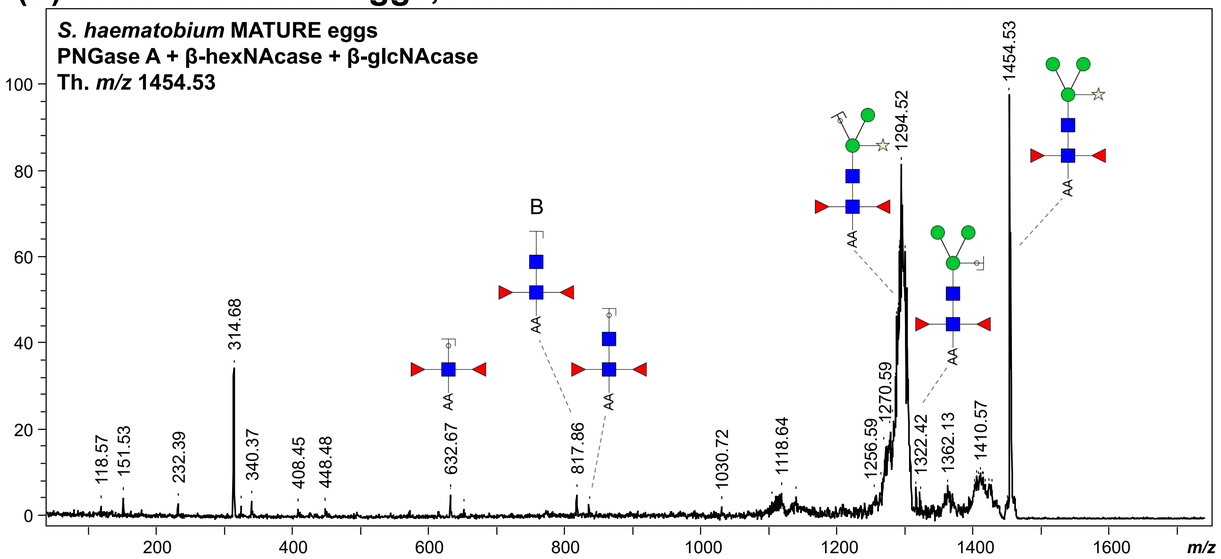
(C) *S. haematobium* eggs – Glycan sequencing *(continued)*

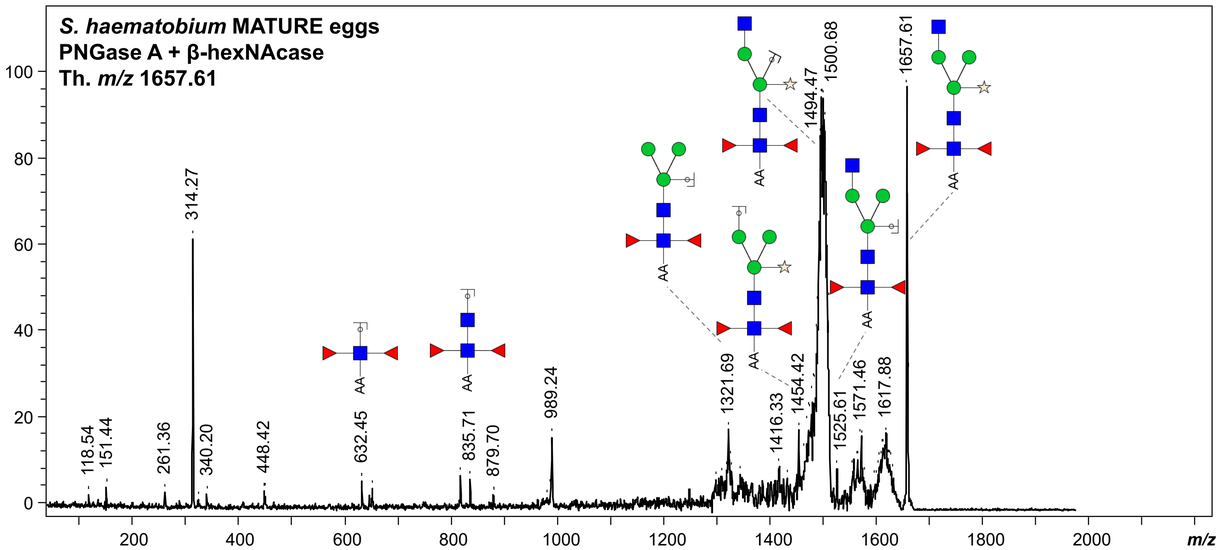


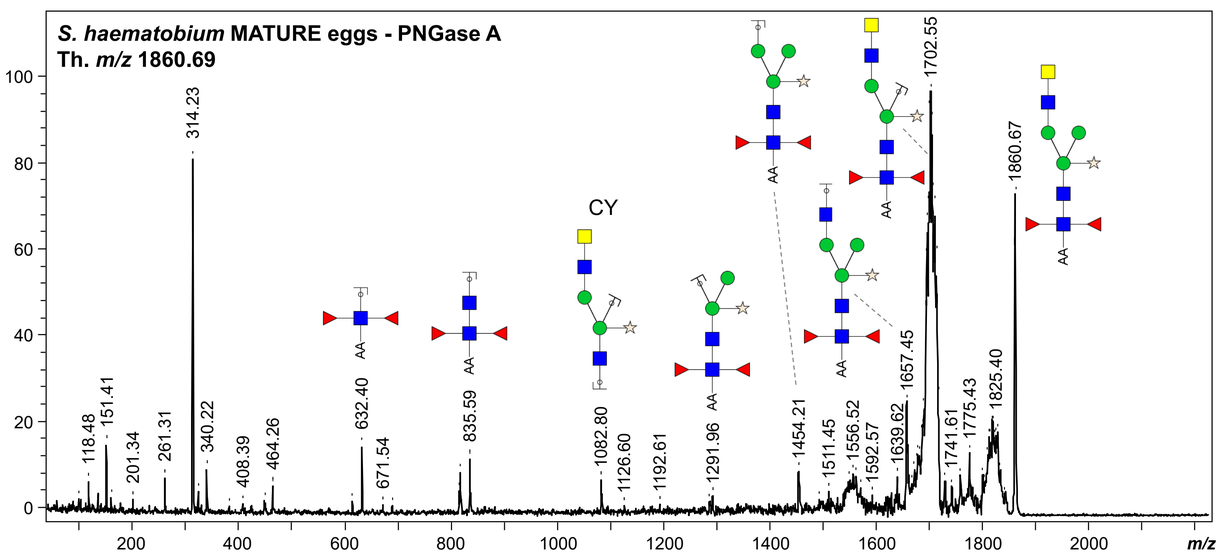
(D) *S. haematobium* eggs, PNGase A – Glycan sequencing



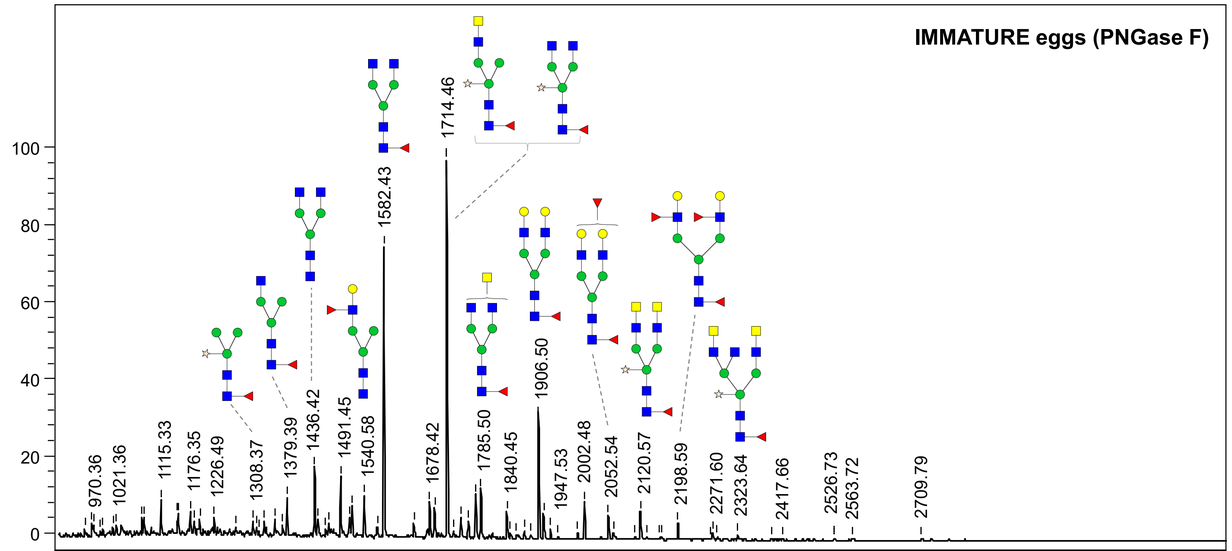
(E) *S. haematobium* eggs, PNGase A – MALDI-TOF-MS/MS

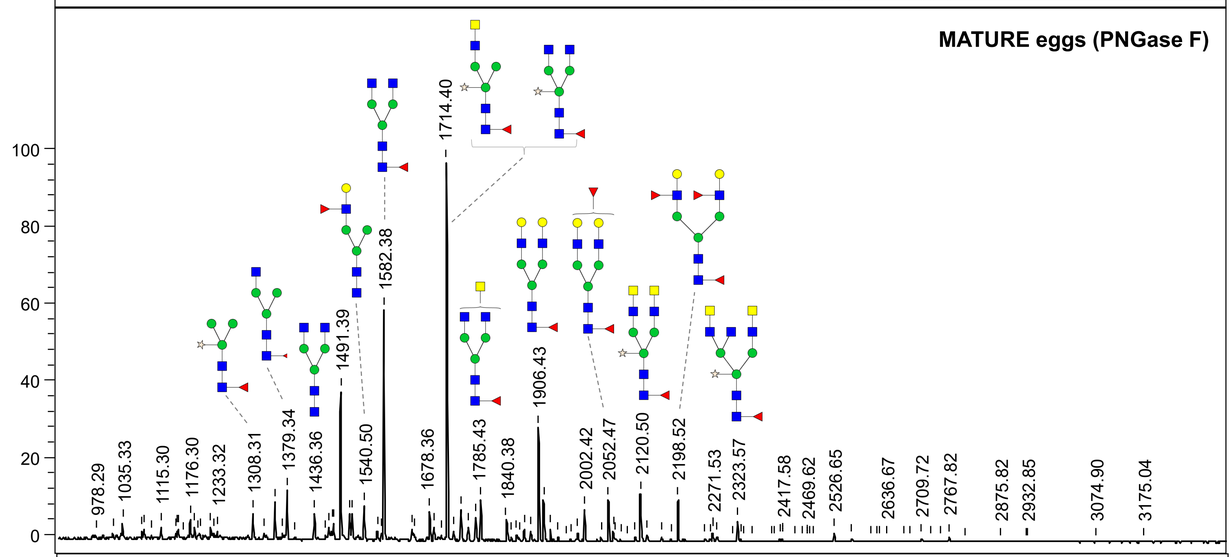


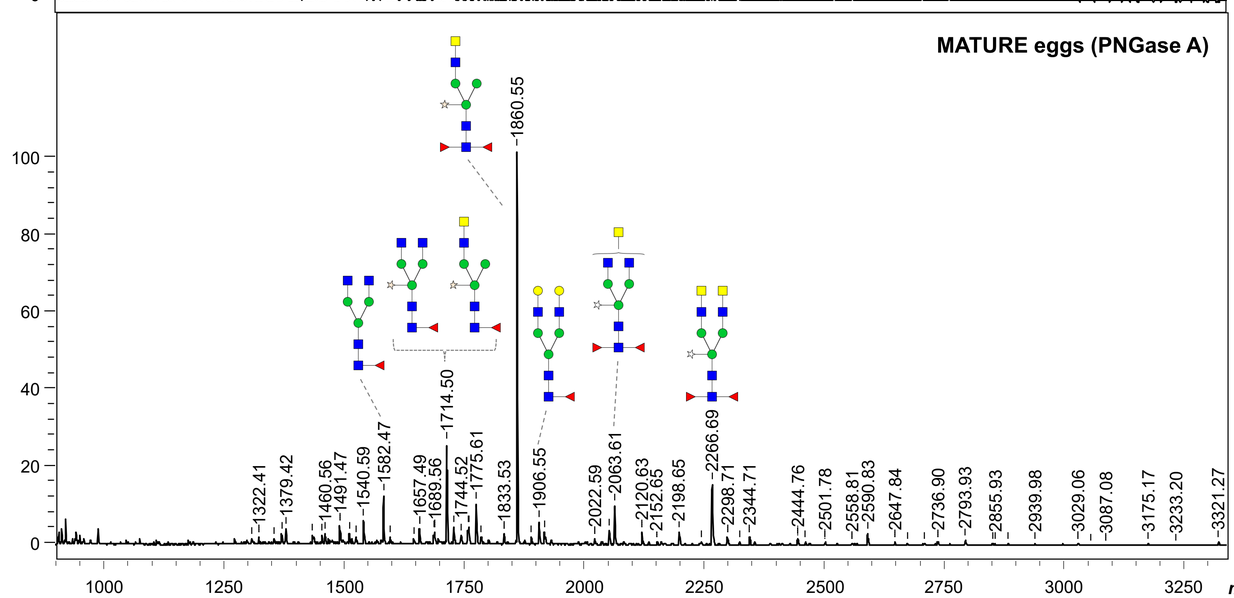




(F) *S. haematobium N*-glycans – Immature and mature eggs

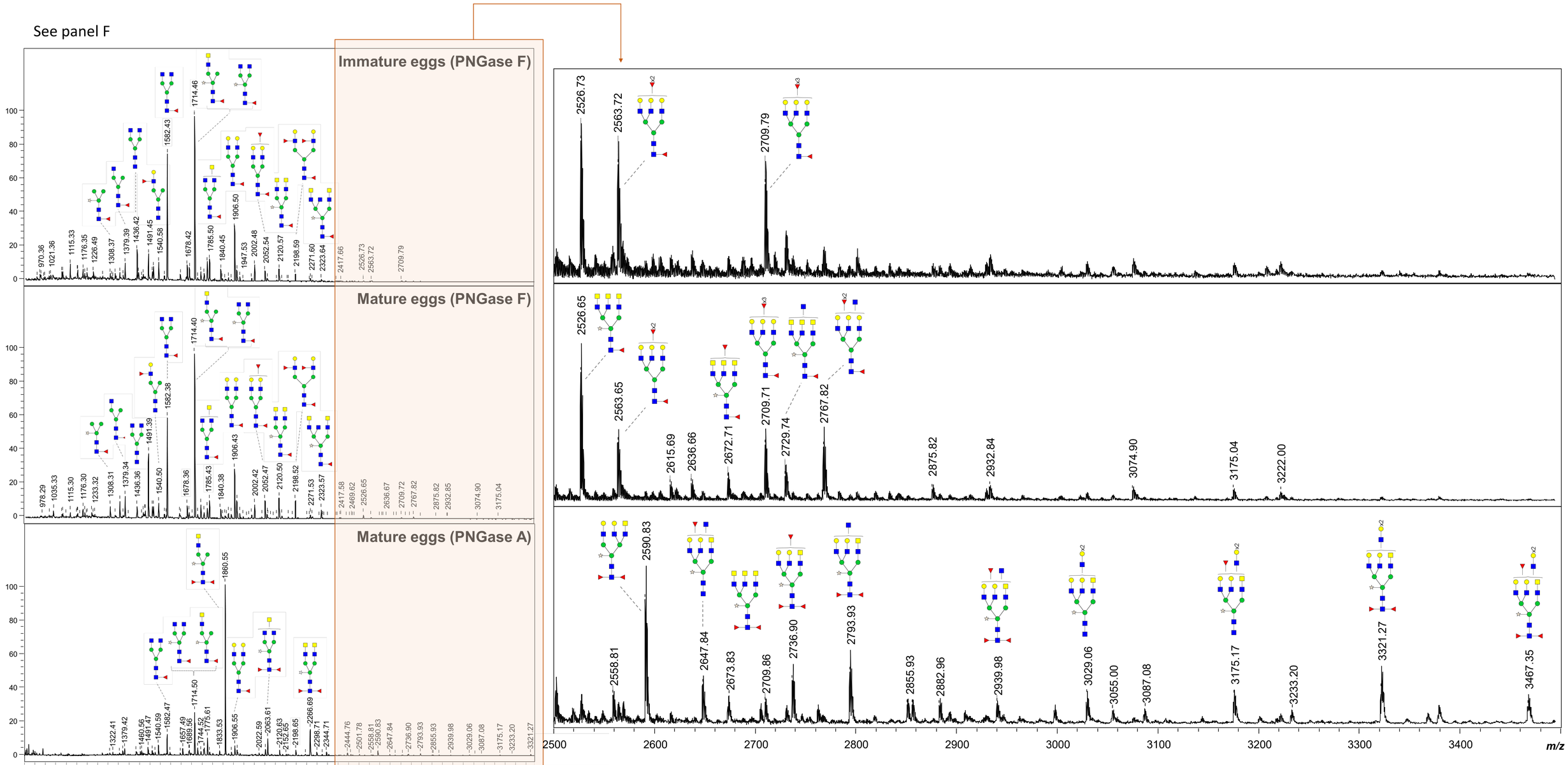






(G) *S. haematobium* eggs, PNGase A-specific N-glycans (focus on higher molecular range structures)

See panel F



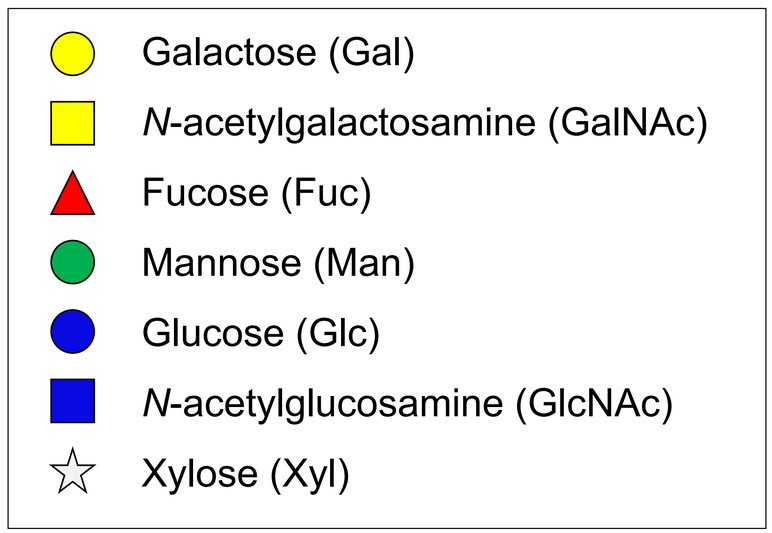
Most probable structures based on monosaccharide compositions are represented using the CFG nomenclature.

Figure S7 – Comparison of *S. haematobium* and *S. mansoni N*-glycans

Proteins were extracted in parallel from the cercariae (A), adult worms (B) and eggs (C-D) of *S. haematobium* (left panels) and *S. mansoni* (right panels). *N*-linked glycans were released from their glycoprotein carriers using PNGase F (A-C) and PNGase A, sequentially (D) prior to AA labeling and analysis by MALDI-TOF-MS. All spectra were acquired in negative-ion reflectron mode and signals are labeled with monoisotopic masses (*m/z,* $[M-H]^-$). Raw data can be found in Supplementary Data 1. The 15 most intense ion species based on % of total signal intensity of the MALDI-TOF-MS spectrum were annotated with corresponding glycan structures, as previously determined using MALDI-TOF-MS in combination with glycan sequencing techniques and MALDI-TOF-MS/MS (see Figure S5 and Supplementary Data 3 for *S. haematobium* and previously published work on *S. mansoni*[15]).

Major differences observed between species are highlighted using boxes. Fucosylation = outer-arm Fuc residues, Xylosylation = β1-2 linked core Xyl, LN = terminal Galβ1-4GlcNAc, LeX = terminal Galβ1-4(Fucα1-3)GlcNAc, Triantennary = *N*-glycans with 3 antennae, Oligomannosidic = *N*-glycans with 5 to 9 mannose residues.

All (putative) glycans are represented using the CFG nomenclature (see inset below). Known non-glycan signals are labeled with the # symbol. M3 to M9 are used to label oligomannosidic *N*-glycans with 3 to 9 mannose residues.

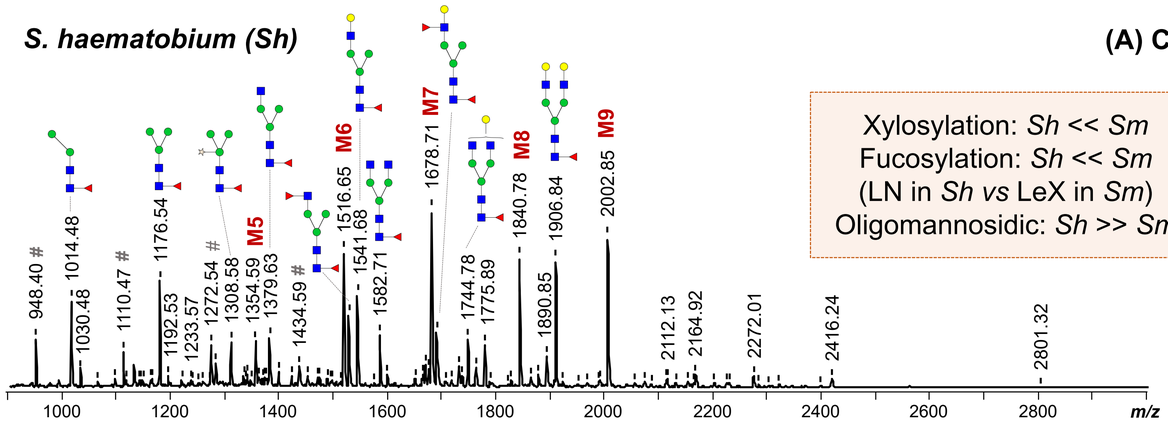


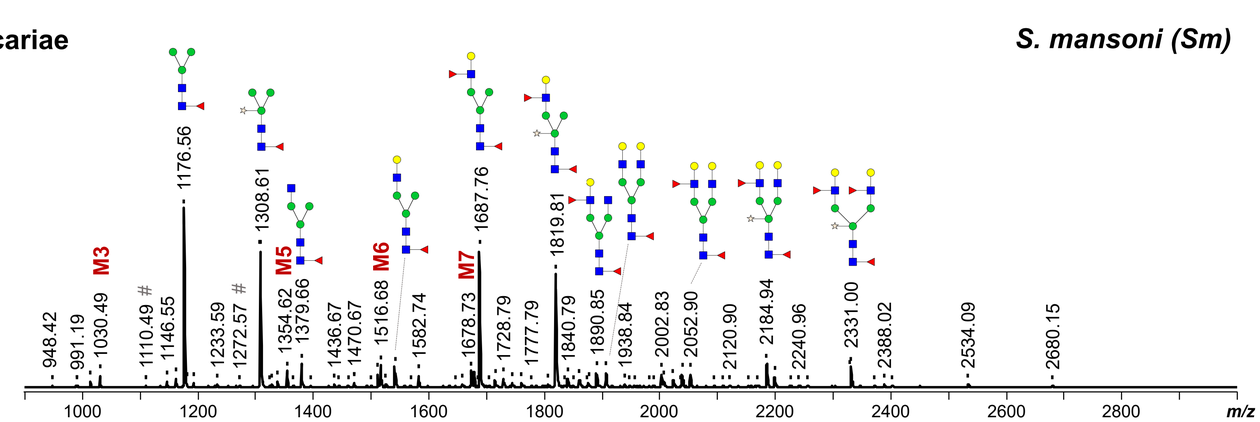
S. haematobium (Sh)

(A) Cercariae

S. mansoni (Sm)

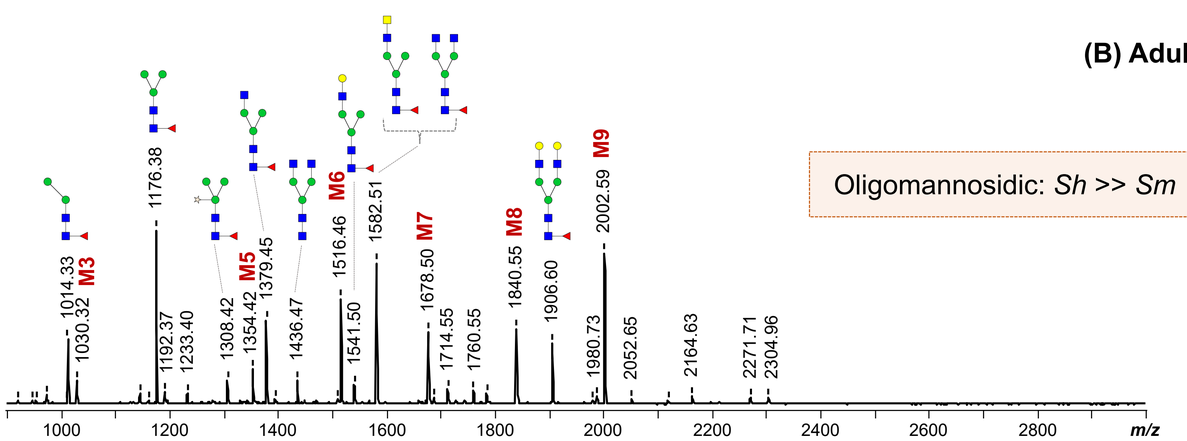
Xylosylation: *Sh* << *Sm*
Fucosylation: *Sh* << *Sm*
(LN in *Sh vs* LeX in *Sm*)
Oligomannosidic: *Sh* >> *Sm*

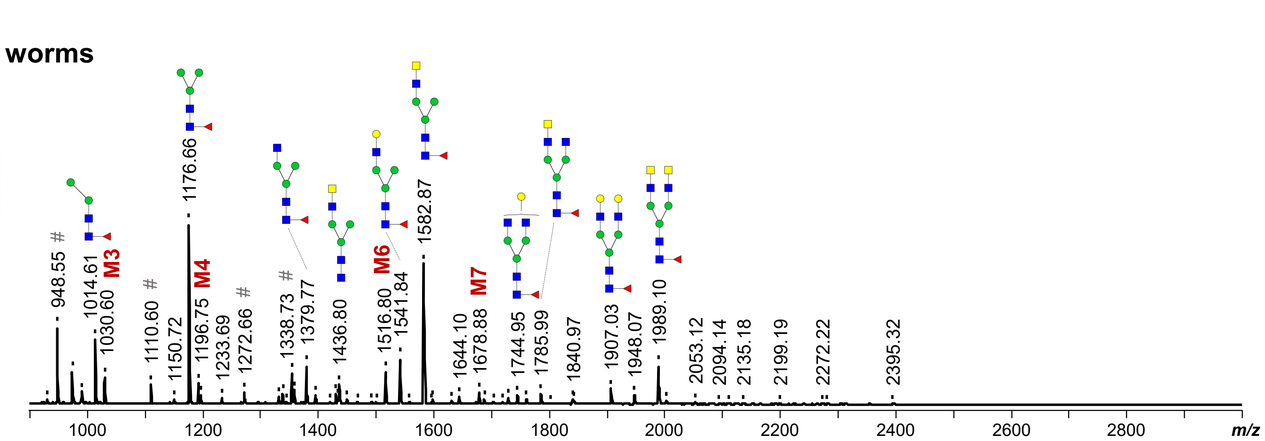




(B) Adult worms

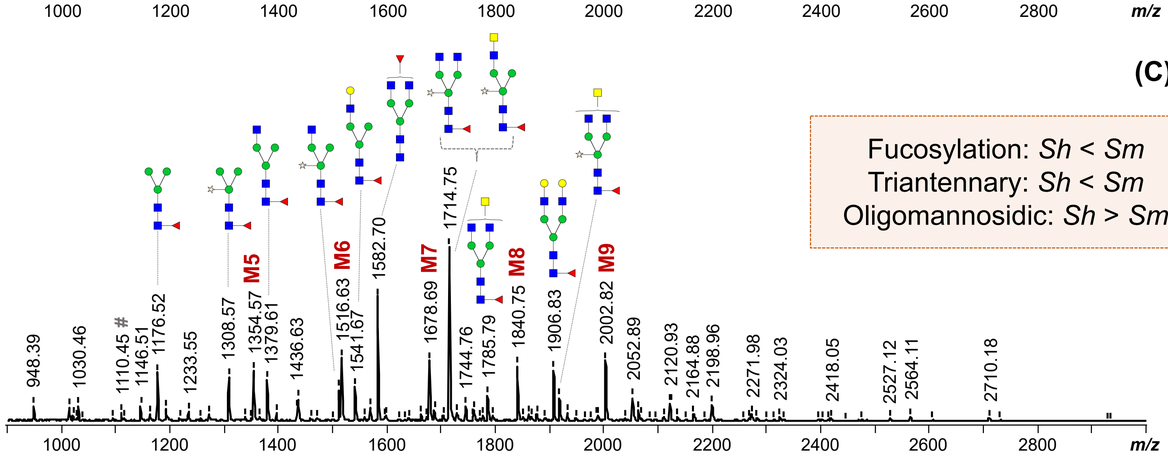
Oligomannosidic: *Sh* >> *Sm*

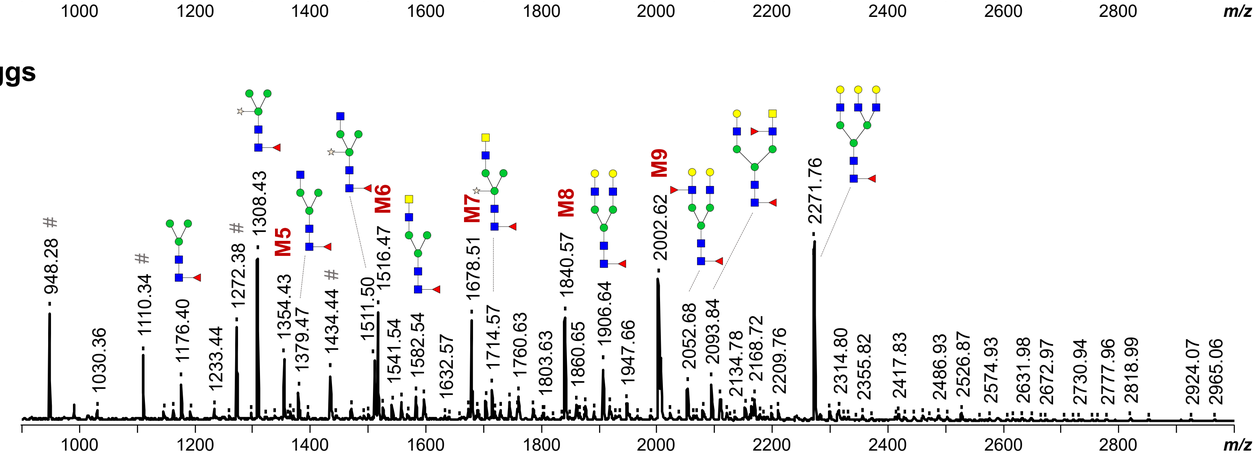




(C) Eggs

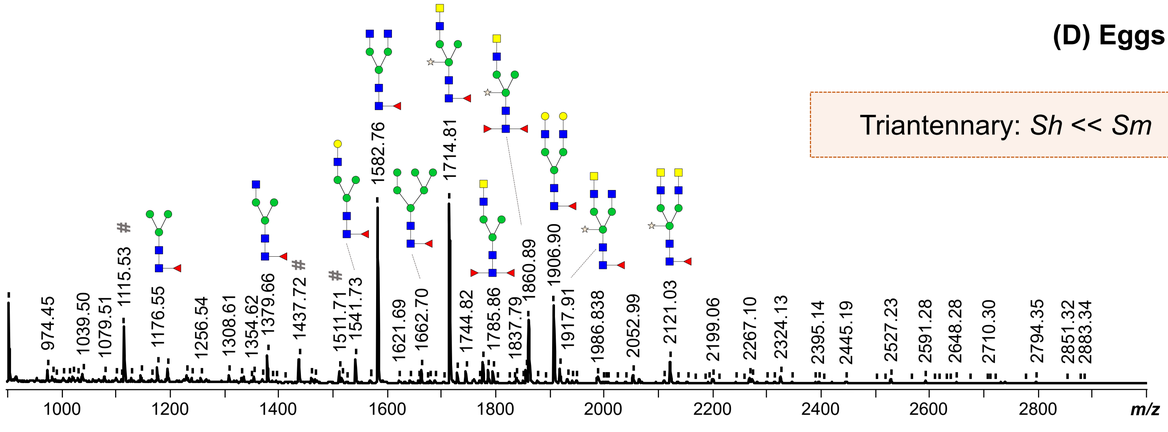
Fucosylation: *Sh* < *Sm*
Triantennary: *Sh* < *Sm*
Oligomannosidic: *Sh* > *Sm*





(D) Eggs – PNGase A

Triantennary: *Sh* << *Sm*



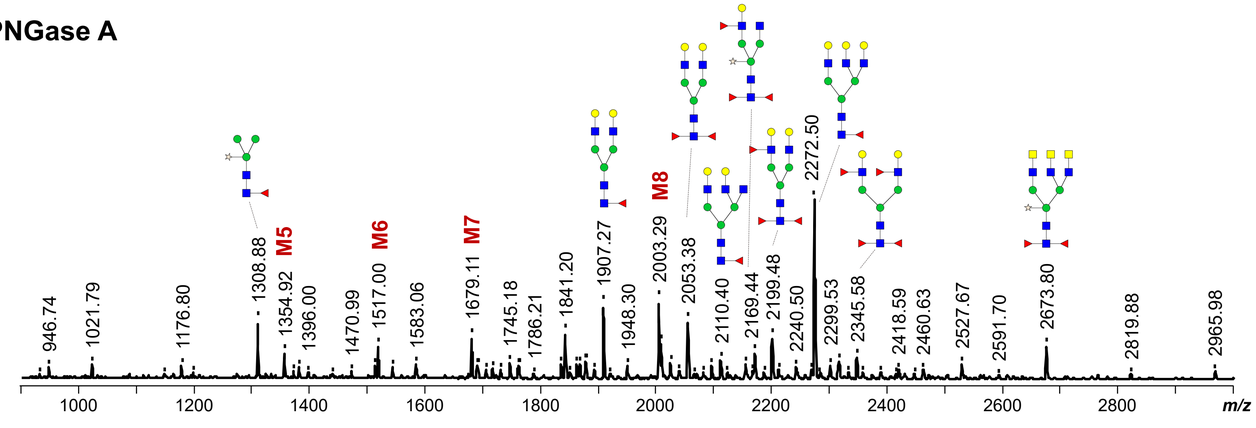
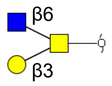
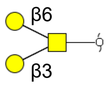
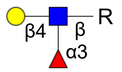
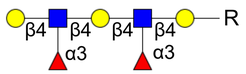
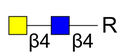
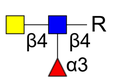
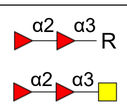
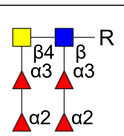


Figure S8 – MALDI-TOF-MS/MS of *S. haematobium O*-glycans

O-glycans were chemically released using β-elimination from the glycoproteins of *S. haematobium* cercariae and eggs and were subsequently permethylated. Selected ions were structurally characterized using MALDI-TOF-MS/MS. All spectra were acquired in positive-ion reflectron mode and signals are labeled with monoisotopic masses (*m/z,* $[M+Na]^+$). Signal intensities in % are indicated on the Y-axis. Fragments were registered as hydrogen or sodium adducts in positive-ion reflectron mode. Diagnostic ions and/or fragmentation patterns indicative of glycan traits of interest are listed in the Table below together with the corresponding panels in which they are observed. Y-type ions, as defined by Domon and Costello (https://doi.org/10.1007/BF01049915) are represented, unless indicated otherwise (B = B type, C = C type, Z = Z type), using the CFG nomenclature (see symbol key insert below).

Identified glycan trait		Diagnostic ion *m/z* (theoretical *m/z*) or Δ*m*	Present in	Panel
Core-2		Z_1-N_1H_1 (*m/z* 502.27) Y_1-N_1H_1 (*m/z* 520.27) Z_1-N_2 (*m/z* 543.29) Y_1-N_2H_1 (*m/z* 765.40)	Cercariae and eggs (more abundant in eggs)	A B
Schisto-core		Z_1-N_1H_1 (*m/z* 502.27) Y_1-N_1H_1 (*m/z* 520.27) Y_2-H_2N_1 (*m/z* 724.37)	Cercariae and eggs (more abundant in cercariae)	A B
LeX		B_2-N_1H_1 (*m/z* 660.32)	Cercariae and eggs	A B
Multiple LeX		C_5-N_1H_1 (*m/z* 1505.74)	Eggs	B
Fuc-GlcNAc		B_1-N_1F_1 (*m/z* 456.22)	Cercariae	A
LDN		B_2-N_2 (*m/z* 527.26)	Eggs	B
LDN-F		B_2-N_2F_1 (*m/z* 701.35)	Eggs	B
DF		C_2-F_2 (*m/z* 403.20) B_1-F_2N_1 (*m/z* 630.31)	Cercariae	A
DF-LDN-DF		B_2-F_4N_2 (*m/z* 1223.62)	Cercariae	A

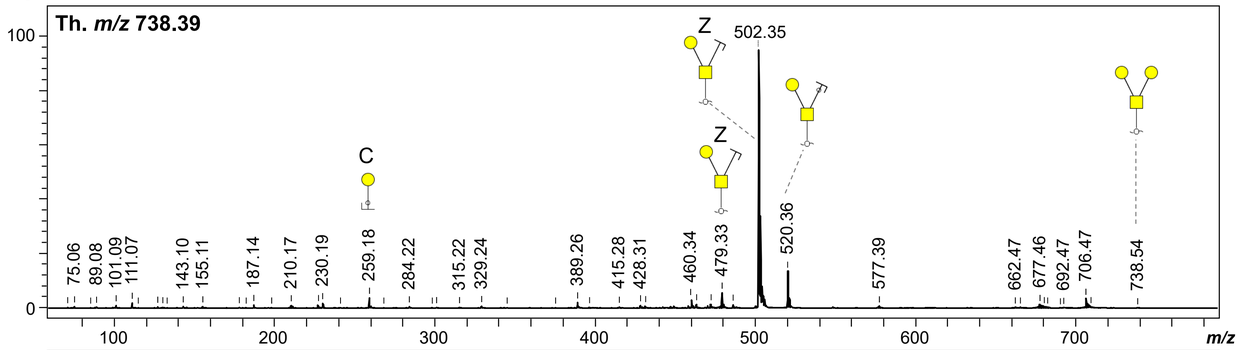
Monosaccharide nomenclature

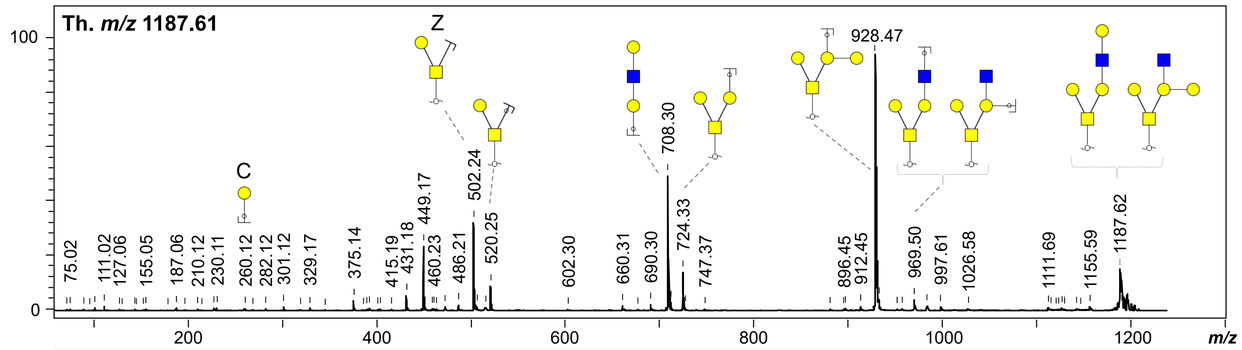
- Galactose (Gal)
- *N*-acetylgalactosamine (GalNAc)
- Fucose (Fuc)
- *N*-acetylglucosamine (GlcNAc)

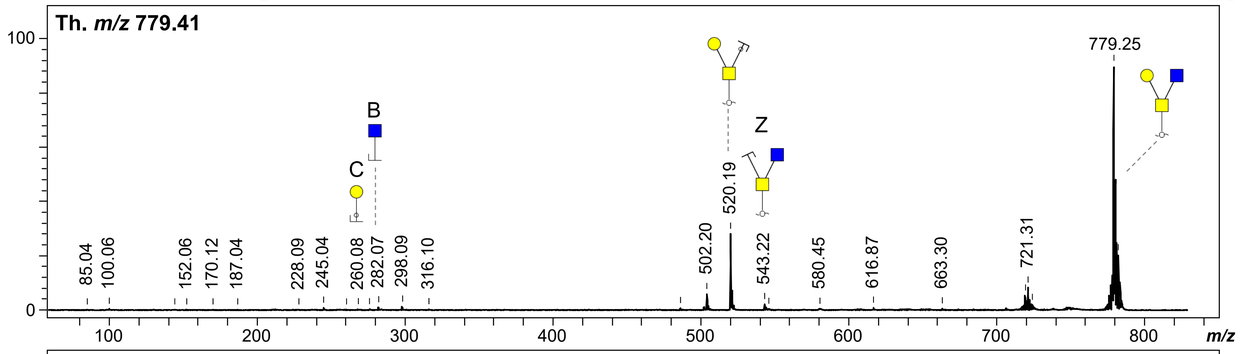
Fragment ion nomenclature

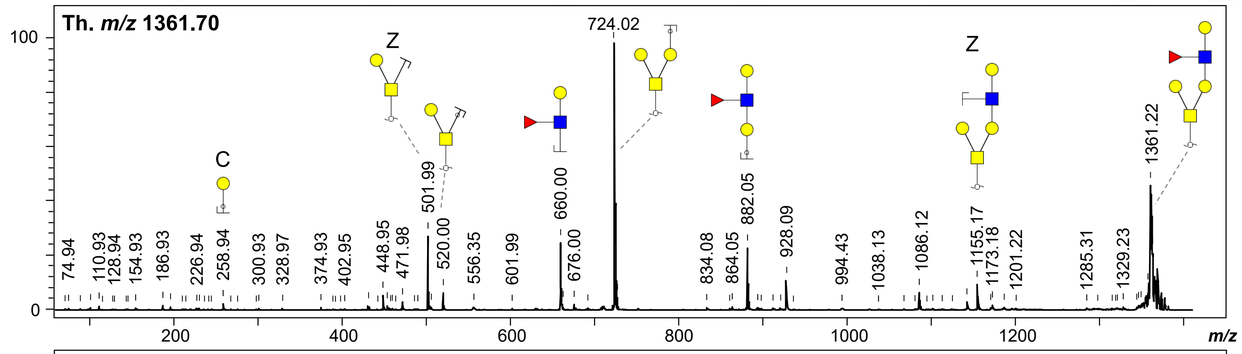
- B-type
- C-type
- Y-type
- Z-type

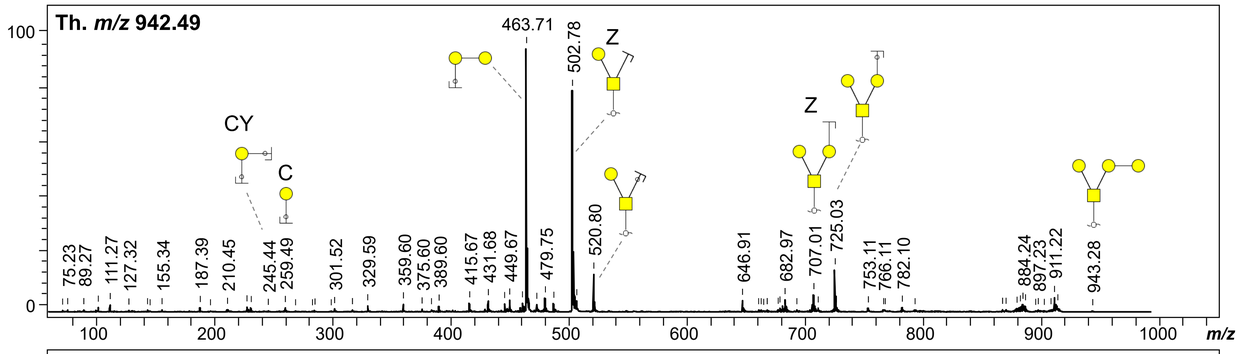
(A) *S. haematobium* cercariae – MALDI-TOF-MS/MS

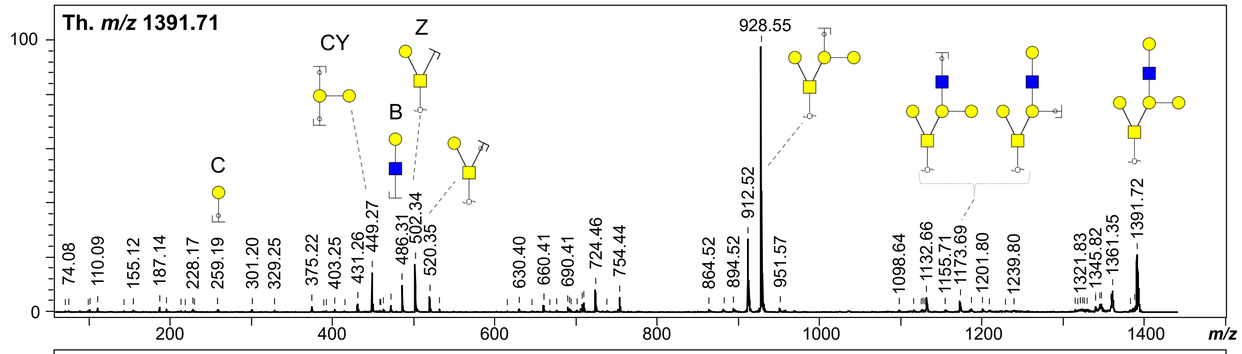


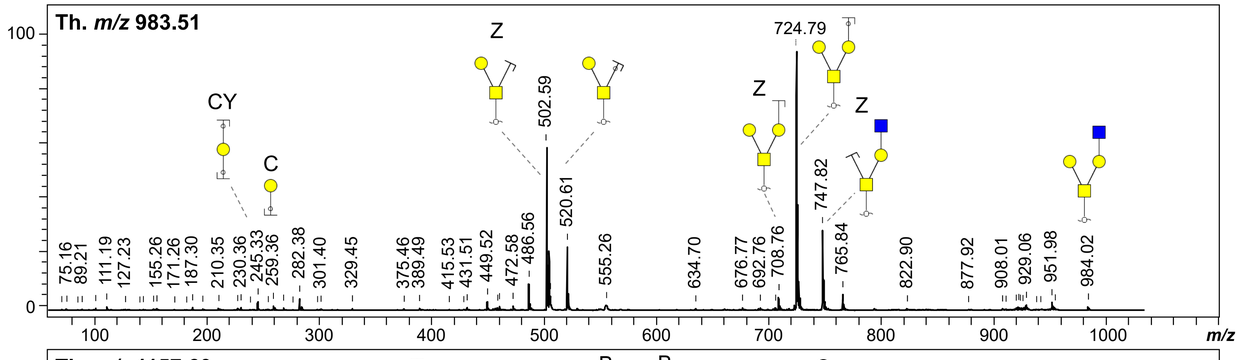


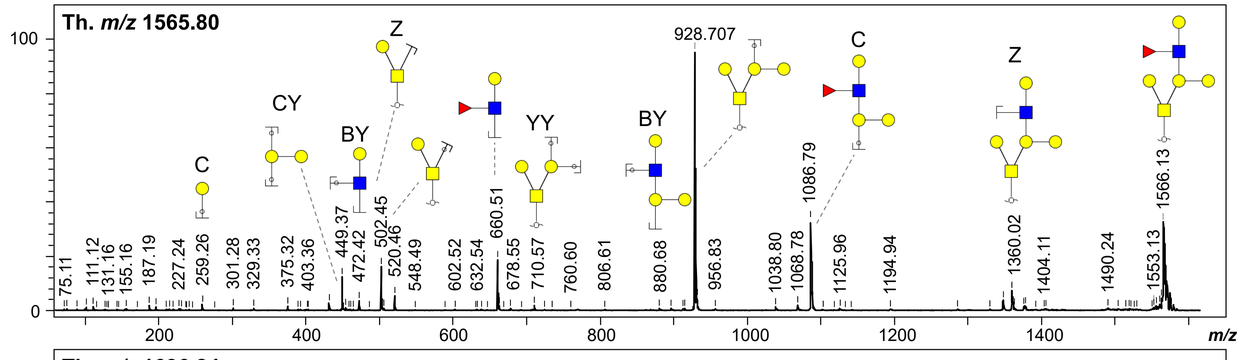


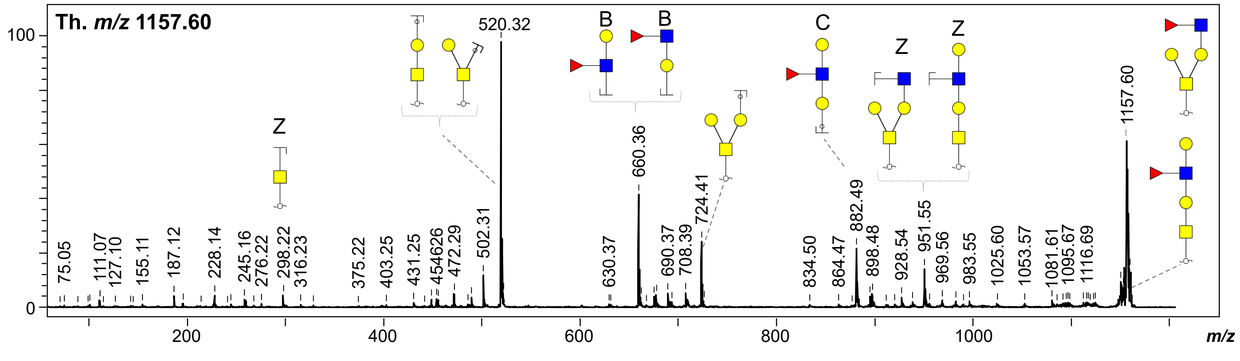


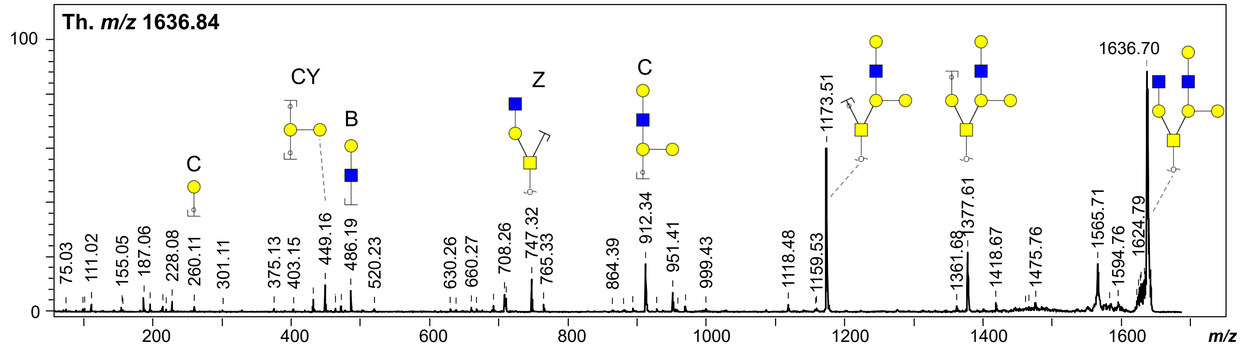




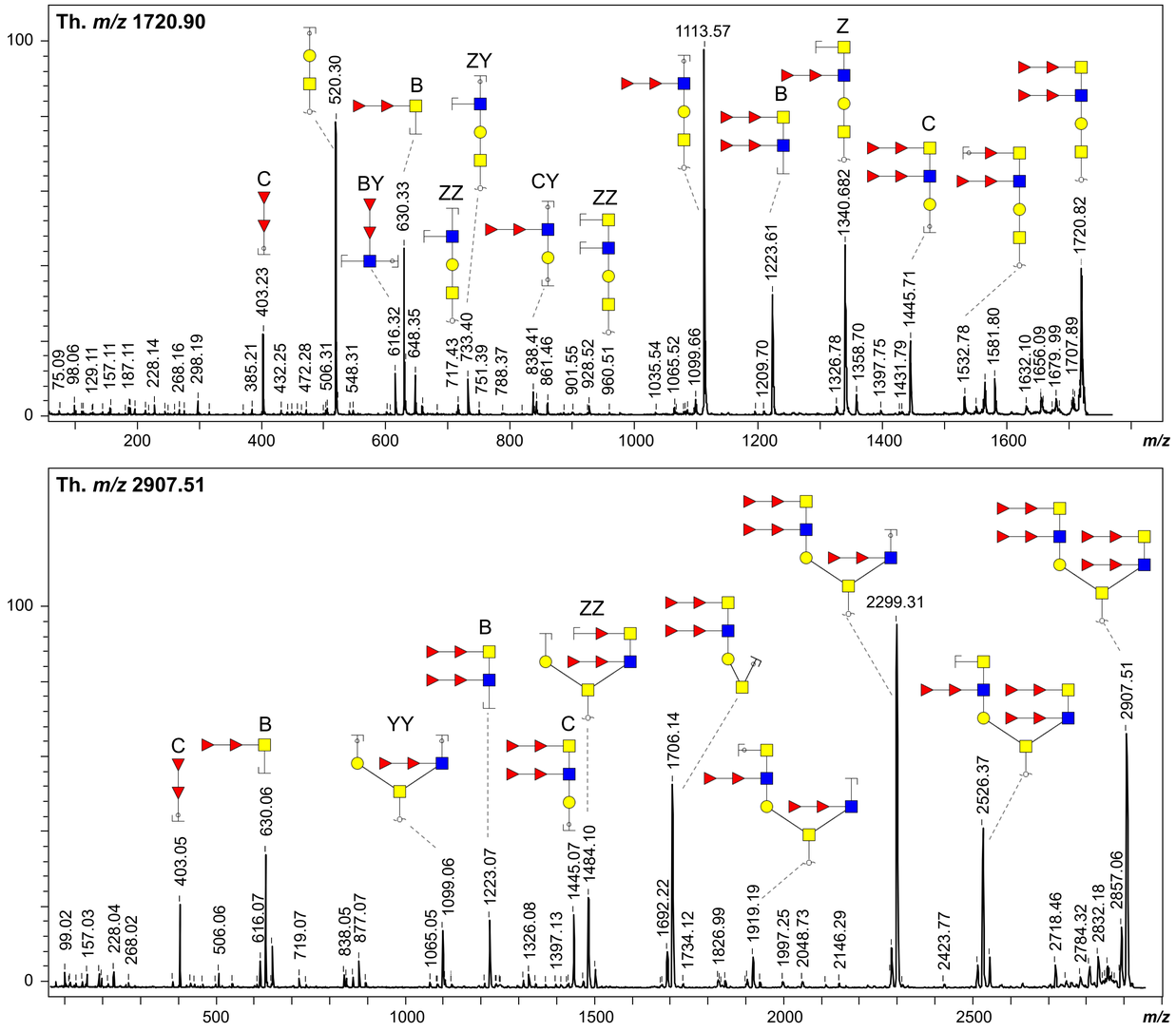




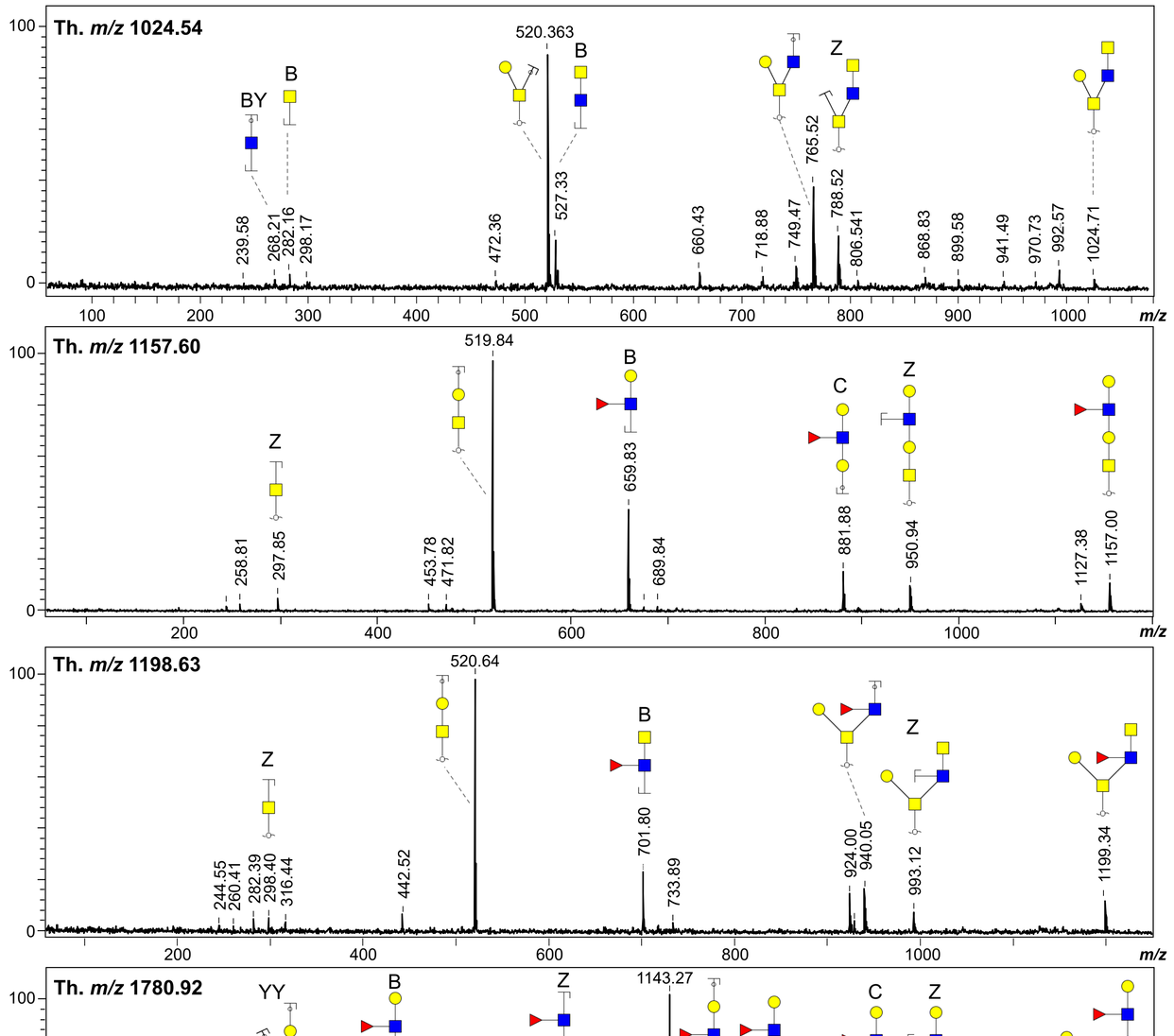




(A) *S. haematobium* cercariae – MALDI-TOF-MS/MS, *CONTINUED*



(B) *S. haematobium* eggs – MALDI-TOF-MS/MS



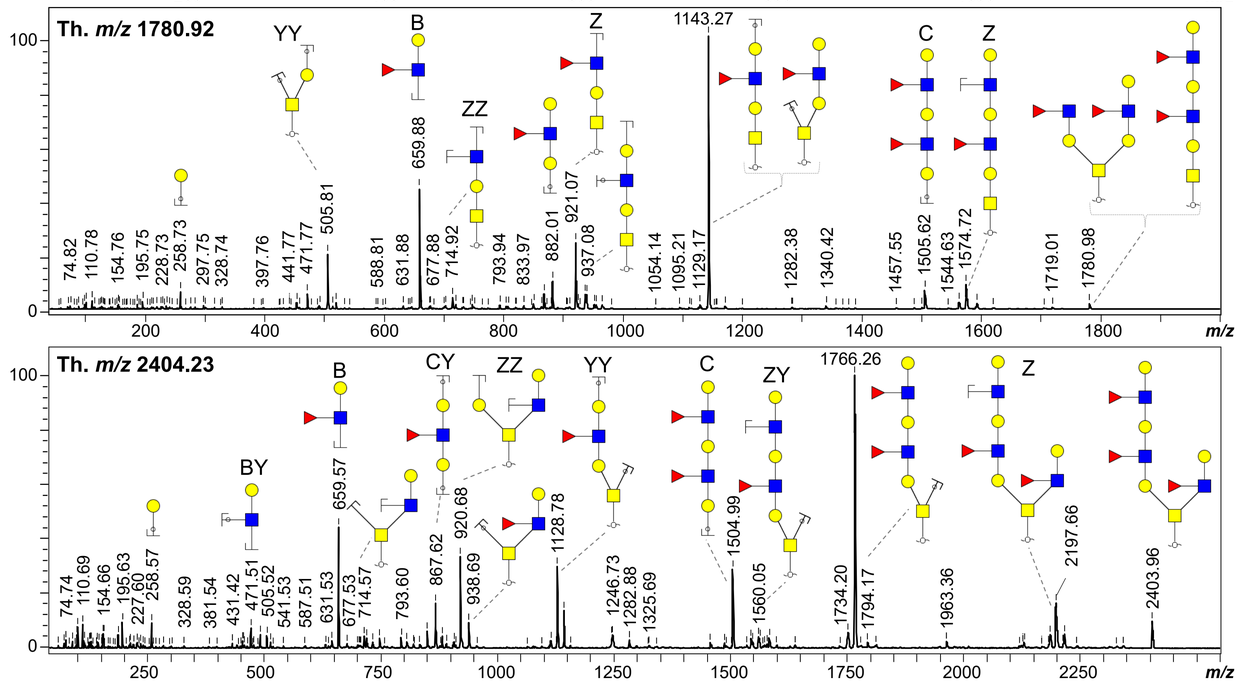


Figure S9 – Glycan microarray construction and validation

I. Preparation of glycan material for microarray printing

1. Glycan purification using UHPLC fractionation

S. haematobium and *S. mansoni N*-linked and GSL glycans were extracted, released and labeled with AA as described in the Methods section. About 2-2.5 million eggs, ~172,000 cercariae and ~1,300 worms were used as starting material for each species to generate sufficient amounts of purified glycans for immobilization on microarray slides. Purification of AA-labeled glycans was performed according to methods previously applied in the construction of parasite glycan microarrays (https://doi.org/10.1016/j.mcpro.2022.100201, https://doi.org/10.1371/journal.pntd.0001922, https://doi.org/10.1016/j.ijpara.2015.02.008).

Briefly, glycans were first separated by normal phase UHPLC on a Dionex UltiMate 3000 system (Thermo Fisher Scientific) using hydrophilic interaction chromatography (HILIC) on a TSKgel Amide-80 column (4.6 mm × 15 cm, particle size 3 µm, #0021867, Tosoh). Eluent A consisted of 50 mmol/L formic acid (pH 4.4) while eluent B consisted of 100% ACN. A linear gradient from 78% to 46% eluent B was applied at a flow rate of 1 mL/min. Fluorescence detection was performed at λ_{ex}-λ_{em} 250-425 nm and data were acquired using Chromeleon™ 7, version 7.1.2.1478 (Thermo Fisher Scientific). Fractions were manually collected in individual tubes, concentrated to dryness using a Speedvac and resuspended in 50 µL MQ. Glycan species present in the obtained fractions were assessed using MALDI-TOF-MS analysis as detailed in the Methods section. Glycan amounts were estimated by comparing peak height in the UHPLC fraction to a known amount of glycan isolated from the glycoprotein RNase B following an established procedure for glycan array building (https://doi.org/10.1021/ac071187g). Fractions with a minimal amount of 20 pmol of glycan material were selected for a second phase of separation using reverse-phase UHPLC using a C18 column (250 x 4mm, particle size 4 µm, Superspher® RP-18 endcapped, # 116858, Millipore Sigma). Eluent A was 0.1% formic acid while eluent B was 95% ACN, 0.1% formic acid. Flowrate was set at 0.5 mL/min and gradient conditions were 5% eluent B for the first 5 minutes followed by a linear gradient ramping from 5 to 50% B for t=5 min to t=30 min. Resulting fractions were further selected for printing based on structural features, assessed using MALDI-TOF-MS, and amount of glycan material, estimated by comparison with the RNase B standard. A comprehensive overview is available in Supplementary Data 5.

2. Preparation of crude antigens

To generate cercarial antigen (CA) and soluble egg antigen (SEA), cercariae and eggs from *S. mansoni* and *S. haematobium* were homogenized in cold PBS by mechanical disruption using a Potter-Elvehjem homogenizer. Homogenates were sonicated and centrifuged at 4°C, 17000g. Soluble fractions were collected, and protein concentrations estimated by BCA assay. Aliquots were diluted with MQ to 100 µg/mL.

Figure S9 – Glycan microarray construction and validation

3. *O*-glycan pools

O-glycans were obtained from a previous study (https://doi.org/10.1016/j.ijpara.2015.02.008) during which they were released from several *S. mansoni* life stages by hydrazinolysis, labelled with AA and purified and fractionated on UHPLC as described for the *N*-linked and GSL glycans. Following MALDI-TOF-MS analysis to determine the glycan compositions, a glycan microarray was constructed and screened with mAbs binding to specific glycan motifs to determine the *O*-glycan epitopes present in the various fractions. For the present work, we pooled the content of five fractions from this previous study containing SEA-derived *O*-glycans to create two pools of *O*-glycans isolated from *S. mansoni* eggs. Similarly, three pools of cercarial *O*-glycans were made from 11 fractions containing CA-derived glycans. Fractions from the original study were pooled based on similarity in their structural content. Glycan amounts were determined previously using height of UHPLC signals. The 5 resulting *O*-glycan pools were included onto the microarray library.

II. Glycan array printing

The fractions listed above were used to generate a glycan library of 288 samples as described in the Methods section. These samples were printed in triplicate to epoxysilane-coated glass slides (Slide E, Nexterion #1066643, Schott AG), leading to a total of 864 spots printed per array. The Microgrid 600 microarrayer (Genomic Solutions) equipped with SMP3 pins that deposit 0.7 nL at each contact was used to print each array eight times per glass slide with a 0.245 mm spacing between spots and 4.60 mm spacing between the printing areas of each array.

III. Array validation with mAbs

1. Binding assay

Upon printing, glycan microarrays were validated using a selection of mAbs binding to specific glycan motifs as determined in previous studies (see Table below). These glycan-binding antibodies were produced by hybridomas generated from mice infected with schistosomes or immunized with schistosome extracts as described previously (https://doi.org/10.1017/s0031182000065045). Slides were incubated following the binding assay procedure detailed in Methods. After blocking, the slides were incubated with mAbs diluted in PBS with 1% BSA and 0.01% Tween 20 in ratios specified in the table below. Binding of mAbs was detected using anti-mouse IgM (heavy chain) Alexa Fluor™ 555 (Invitrogen #A-21426). Fluorescence was measured using a G2565BA scanner (Agilent Technologies) at 10 µm resolution using two laser channels at 532 nm and 633 nm and with the sensitivity level of the photomultiplier tube set at 10%. Image analysis was processed with GenePix Pro 7.0 software (Molecular Devices) according to published methods (https://doi.org/10.1021/pr900515y). The acquired data were exported to Excel and for each glycan-containing fraction, fluorescence intensities of triplicates were averaged and corrected for background using average fluorescence intensity of “blank” spots - i.e. spots printed from wells containing spotting buffer only - as a baseline.

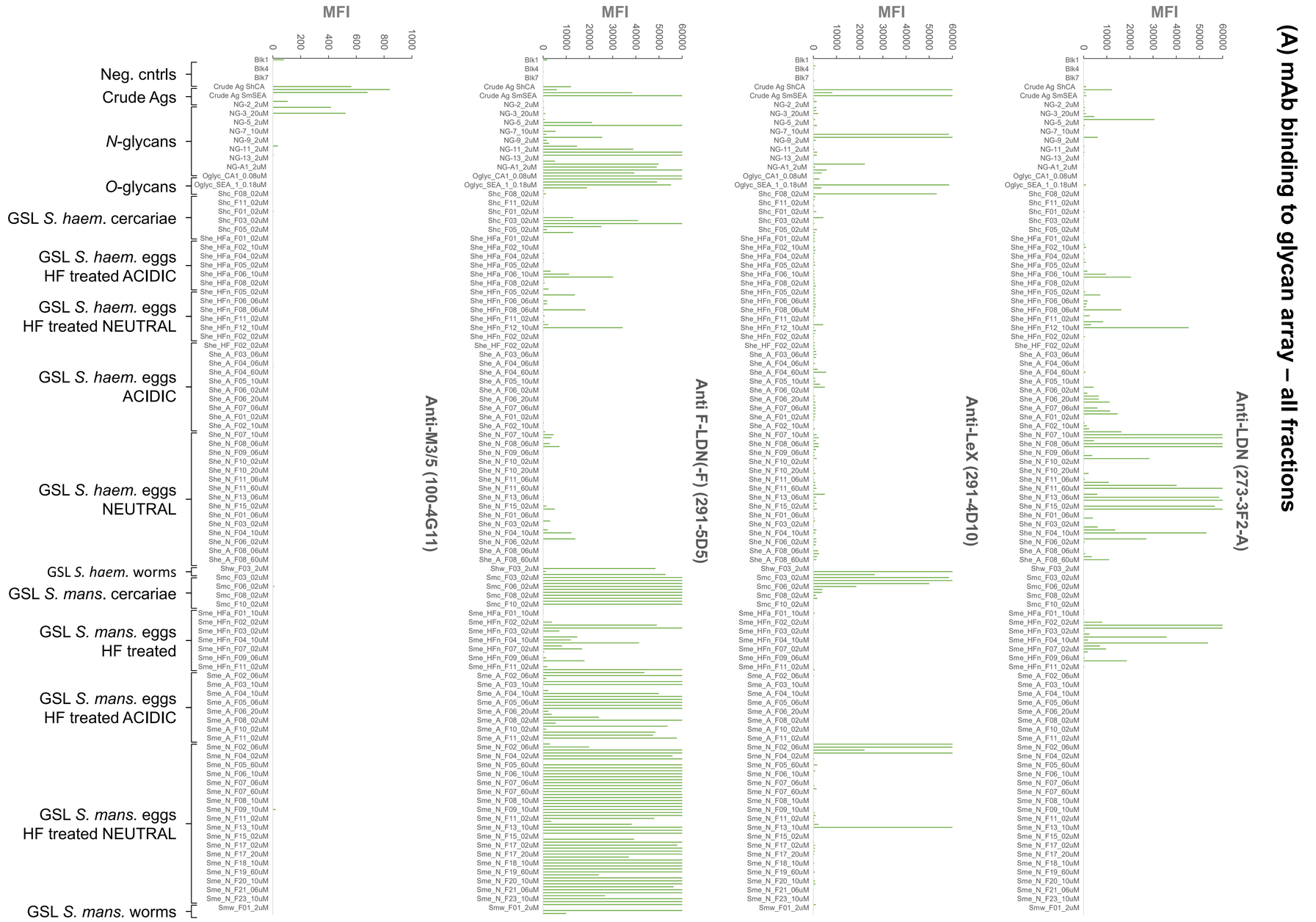
mAb	Dilution ratio	Binding specificity	References
273-3F2-A	1:100	LDN (LacDiNAc, GalNAcβ1-4GlcNAc)	https://doi.org/10.1093/glycob/10.6.601
291-4D10-A	1:200	LeX (Galβ1-4(Fucα1-3)GlcNAc)	https://doi.org/10.1093/glycob/10.6.601
291-5D5-A	1:200	F-LDN(-F) (Fucα1-3GalNAcβ1-4(Fucα1-3)GlcNAc)	https://10.1017/s0031182004006390; https://10.1016/j.ab.2010.07.008
100-4G11	1:250	Trimannosyl branching (R-α1-6Man(α1-6Man)α1-3Man)	https://10.1093/glycob/cwg025

2. Results (see figure next slide)

Background-corrected median fluorescence intensities (MFIs, y-axis) are shown for each glycan fraction printed on the array (x-axis) in panel A. Fraction contents are described on the x-axis: Neg. controls = negative controls without glycan content (print buffer only); *S. haem.* = *S. haematobium*, *S. mans.* = *S. mansoni*. Raw data can be found in Supplementary Data 6 and fraction content is detailed in Supplementary Data 5. Panel B shows a closer view of mAb binding to crude antigens, *N*- and *O*-linked glycan-containing fractions. Compositions of the major glycan structures present in the fractions are provided below the x-axis: H = hexose, N = N-acetylhexosamine, F = fucose, X = core-xylose, M3-9 = mannosylated *N*-glycans with 3 to 9 mannose residues.

In line with its specificity, the mAb 273-3F2-A was found to mainly bind fractions containing structures from either *S. haematobium* or *S. mansoni* egg GSL glycans that were defucosylated using HF treatment; or structures from *S. haematobium* egg GSL glycans carrying little or no fucose(s) in their native stage, all leaving the terminal LDN epitope accessible (panel A). Expectedly, the mAb 291-5D5, showed an opposite pattern of binding, with high MFI values measured for GSL from *S. haematobium* cercariae and from *S. mansoni* (native) egg and cercarial glycans as well as for several *N*-glycan fractions (panel A), all of them containing fucosylated LDN motifs (panel B). Binding of 291-4D10 was observed to a few GSL glycan-containing fractions from *S. mansoni* eggs and cercariae, and a subset of *N*-glycan fractions, in accordance with their content in LeX motifs. Finally, very low fluorescence levels were obtained from incubation with 100-4G11 consistent with the Man3/5 epitope only being present in low amounts in a couple of *N*-glycan-containing fractions (panel B). These findings confirm both the successful covalent immobilization of glycans onto the microarray and the structural assignment of terminal LeX, LDN, and F-LDN(−F) motifs in *S. haematobium* GSL and *N*-linked glycans. Thus, binding of mAb to relevant fractions is also shown in Supplementary Data 2 and 3, side by side with the glycan structures present in the microarray fractions. In addition to Supplementary Data 6, the raw data used to generate the graphs in (A) and (B) are also provided in the Source Data file.

(A) mAb binding to glycan array – all fractions



(B) mAb binding to crude antigens and *N*/*O*-glycan-containing fractions

Figure S10 – Serum IgM responses to schistosome glycans

Microarrays were incubated with sera from non-endemic individuals (n = 9) and with sera from children infected with *S. haematobium* (n = 18) or *S. mansoni* (n = 21). IgM binding to the microarray was measured for each individual. Raw data is available in Supplementary Data 8. Graphs display the averaged background corrected median fluorescence intensity (MFI) values obtained for each group while antibody binding to the microarray fractions (x-axis) for each individual can be visualized on the heatmap. The type of fraction content is indicated along the x-axis: Blanks (negative controls), *Schistosoma* crude antigens, *N*-glycans (PNGase F or PNGase A released), *O*-glycans or GSL glycans from either *S. haematobium* or *S. mansoni* cercariae, worms or eggs (treated with HF or native). The raw data used to generate the graphs are provided as a Source Data file.

N-E = non-endemic (uninfected)

S. haematobium infection

S. mansoni infection

MFI

50000
40000
30000
20000
10000
0

Blanks

Crude antigens

N-glycans

PNG-F

PNG-A

O-glycans

GSL glycans

S. haematobium

Cerca.

Eggs + HF

Acidic

Neutral

Eggs

Acidic

Neutral

Worms

S. mansoni

Cerca.

Eggs + HF

A

Neutral

Eggs

Acidic

Neutral

N-E

Sh inf.

Sm inf.

MFI

0

7500

25000

40000

50000

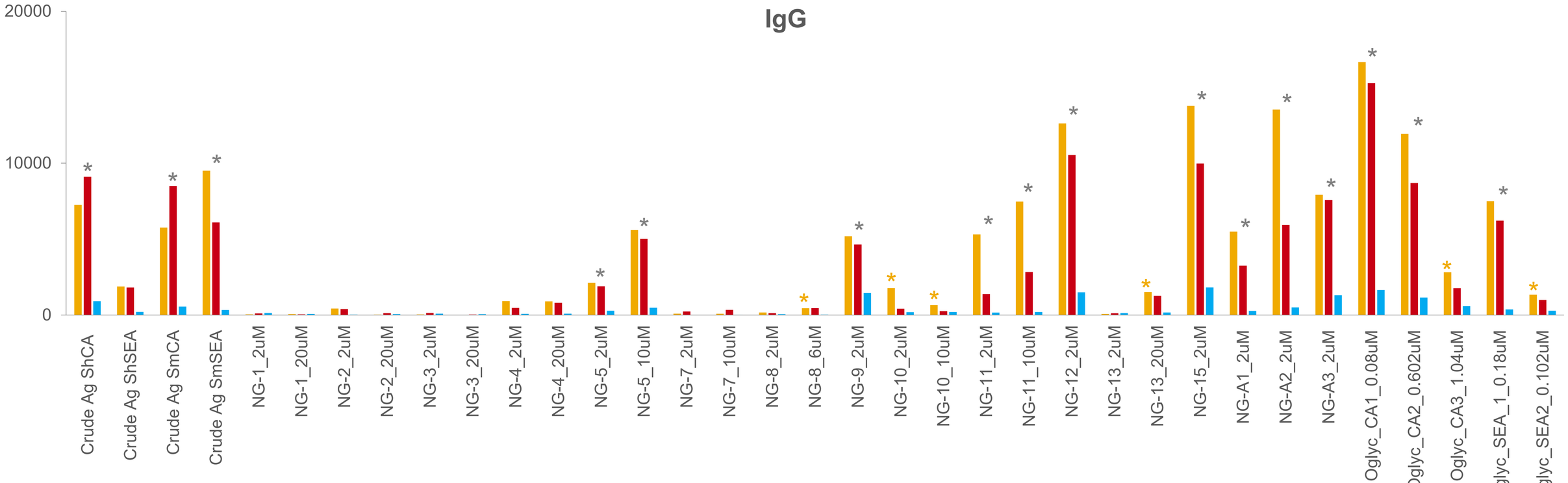
Figure S11 – IgG and IgM responses to crude antigens, *N*-linked and *O*-linked glycans

Microarrays were incubated with sera from non-endemic individuals (n = 9) and with sera from children infected with *S. haematobium* (n = 18) or *S. mansoni* (n = 21). IgG (upper graph) and IgM (bottom graph) binding to the microarray were measured for each individual (Supplementary Data 7 (IgG) and Supplementary Data 8 (IgM)). Averaged MFIs obtained for each group are displayed on the graph for microarray fractions containing *Schistosoma* crude antigens, *N*-glycans (PNG-F or PNG-A released) and *O*-glycans. Details on major glycan structures present in fractions are provided below the x-axis: H = hexose, N = N-acetylhexosamine, F = fucose, X = core-xylose, M3-9 = mannosylated *N*-glycans with 3 to 9 mannose residues.

Significant differences between groups were assessed using Bayesian statistics implemented in the *limma* R package. All tests were conducted as two-sided. P-values were adjusted for multiple comparisons using limma's built-in procedures (Benjamini–Hochberg method), yielding adjusted false-discovery rates. Resulting p-values are shown in Supplementary Data 7 (IgG) and Supplementary Data 8 (IgM) and in the Source Data File. P-values < 0.05 indicating a significant difference in MFI values between groups are labeled using stars. A grey star indicates that both infection groups differ from the non-endemic group, while an orange star indicates that only the *S. haematobium*-infected group differs from the non-endemic group. The raw data used to generate the graphs are provided as a Source Data file.

S. haematobium infection *S. mansoni* infection N-E = non-endemic (uninfected)

IgG



IgM