

# Identification of species-specific glycan antigens expressed by *Schistosoma haematobium*

## Supplementary figures

- Figure S1** – MALDI-TOF-MS spectra of *S. haematobium* GSL derived, N-linked and O-linked glycans ..... Slides 2-5
- Figure S2** – Structural characterization of *S. haematobium* and *S. mansoni* GSL glycans ..... Slides 6-10
- Figure S3** – MALDI-TOF-MS/MS of *S. haematobium* and *S. mansoni* GSL glycans ..... Slides 11-15
- Figure S4** – PGC-nano-LC-MS and MS/MS of *S. haematobium* and *S. mansoni* GSL glycans ..... Slides 16-21
- Figure S5** – Structural characterization of *S. haematobium* N-glycans ..... Slides 22-26
- Figure S6** – Comparison of *S. haematobium* and *S. mansoni* N-glycans ..... Slides 27-28
- Figure S7** – MALDI-TOF-MS/MS of *S. haematobium* O-glycans ..... Slides 29-31
- Figure S8** – Glycan microarray validation using mAbs ..... Slides 32-33
- Figure S9** – Serum IgM responses to schistosome glycans ..... Slides 34-35
- Figure S10** – IgG and IgM responses to crude antigens, N-linked and O-linked glycans ..... Slides 36-37

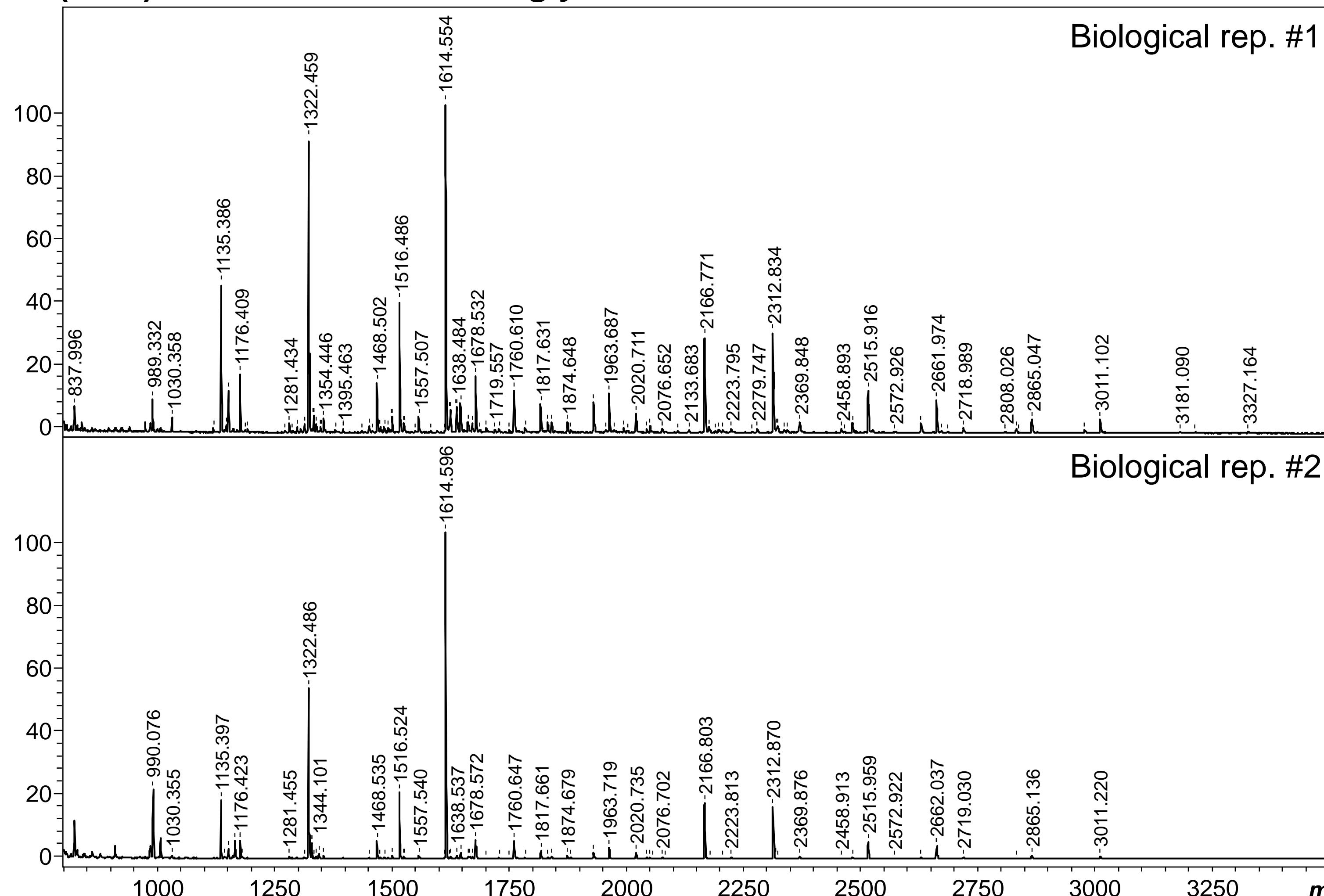
## Figure S1 – MALDI-TOF-MS spectra of *S. haematobium* GSL derived, *N*-linked and *O*-linked glycans

*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).

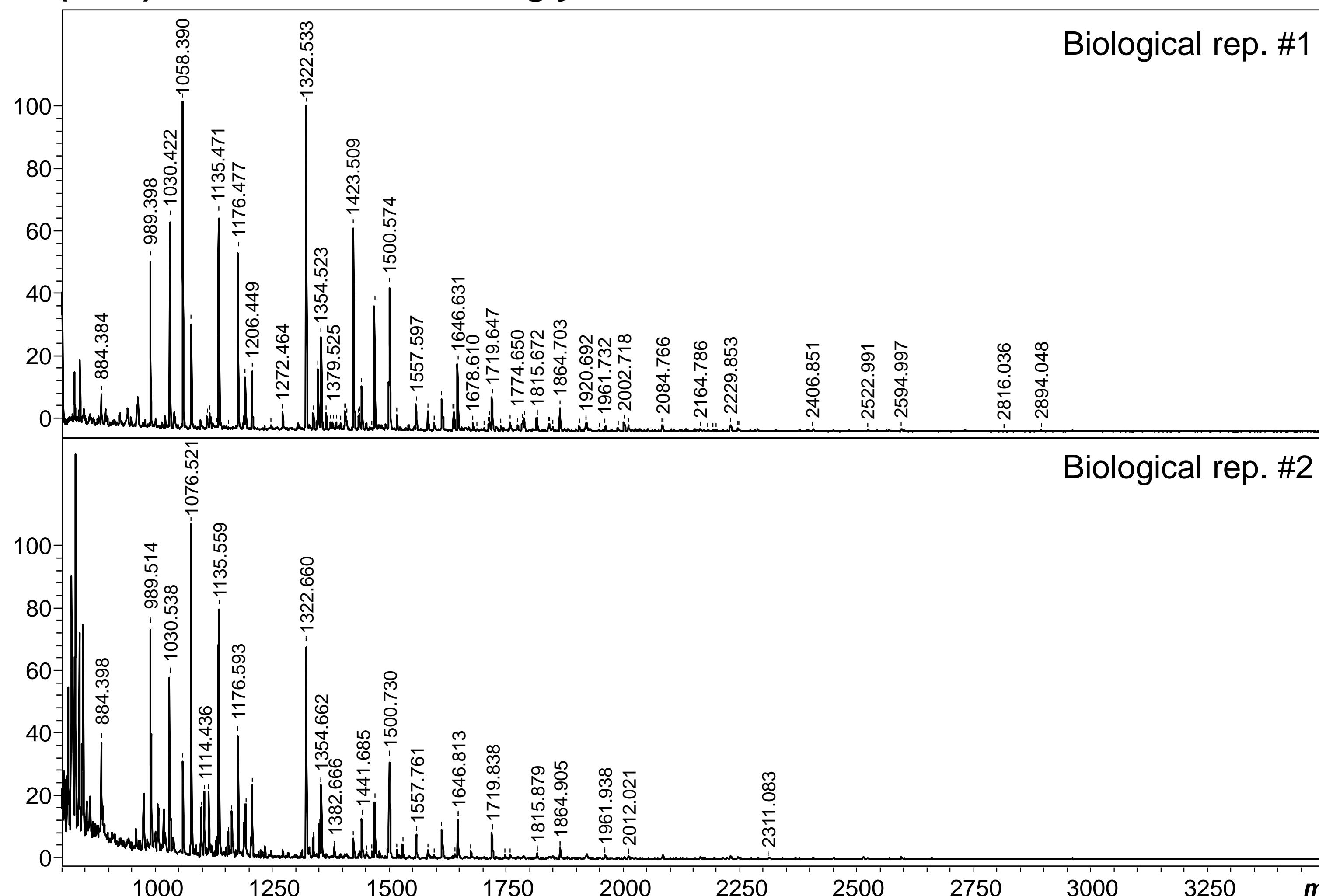
*O*-glycans were released from *S. haematobium* glycoproteins by  $\beta$ -elimination and permethylated prior to MALDI-TOF-MS analysis (C). Spectra were recorded in positive-ion reflectron mode and monoisotopic masses of measured signals are indicated ([M+Na] $^+$ ). Signal intensities in % are indicated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

Glycan class and parasite life-stage from which glycans were extracted is 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 **Table S1**.

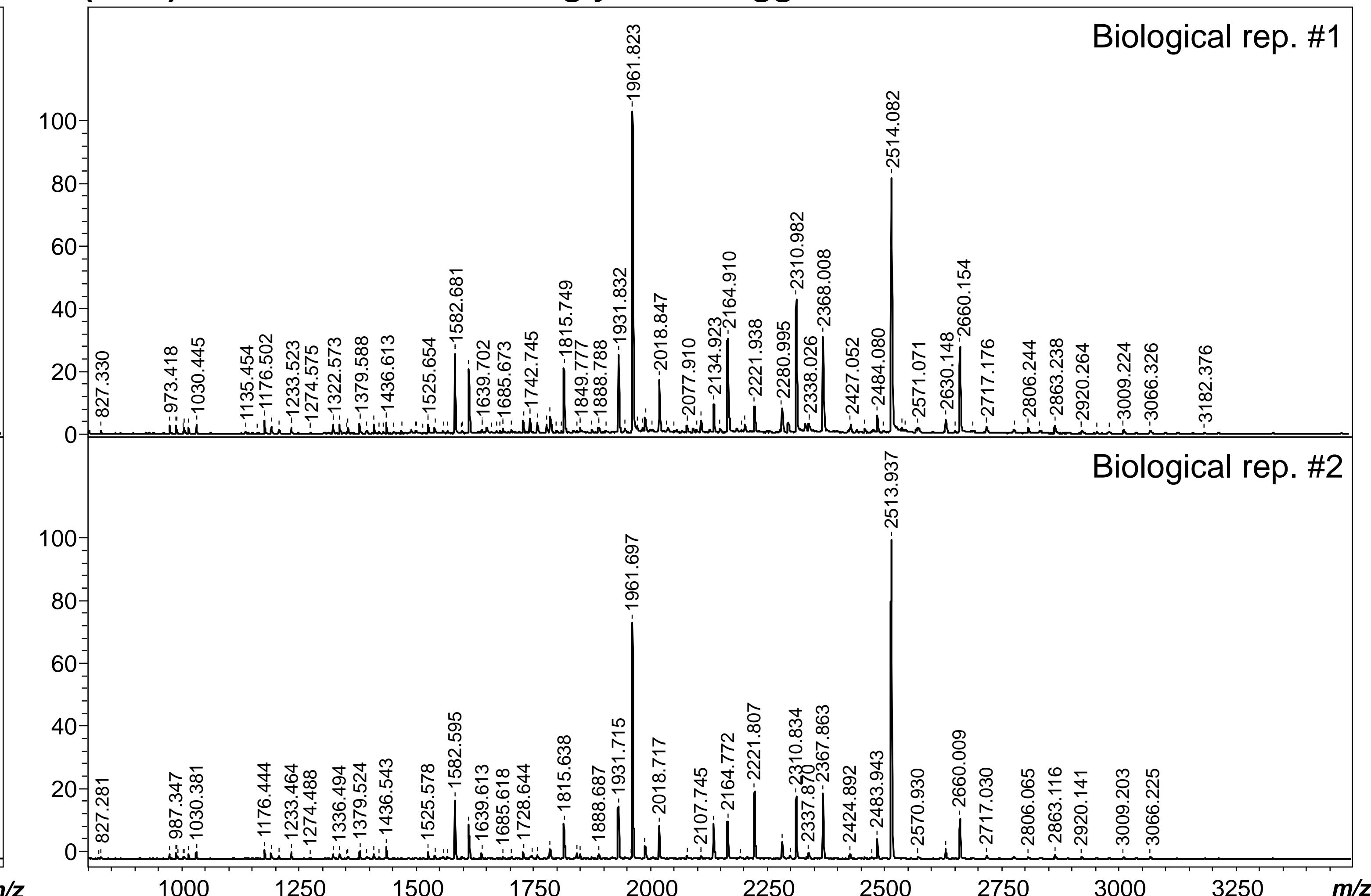
## (A - 1) *S. haematobium* GSL glycans – cercariae



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

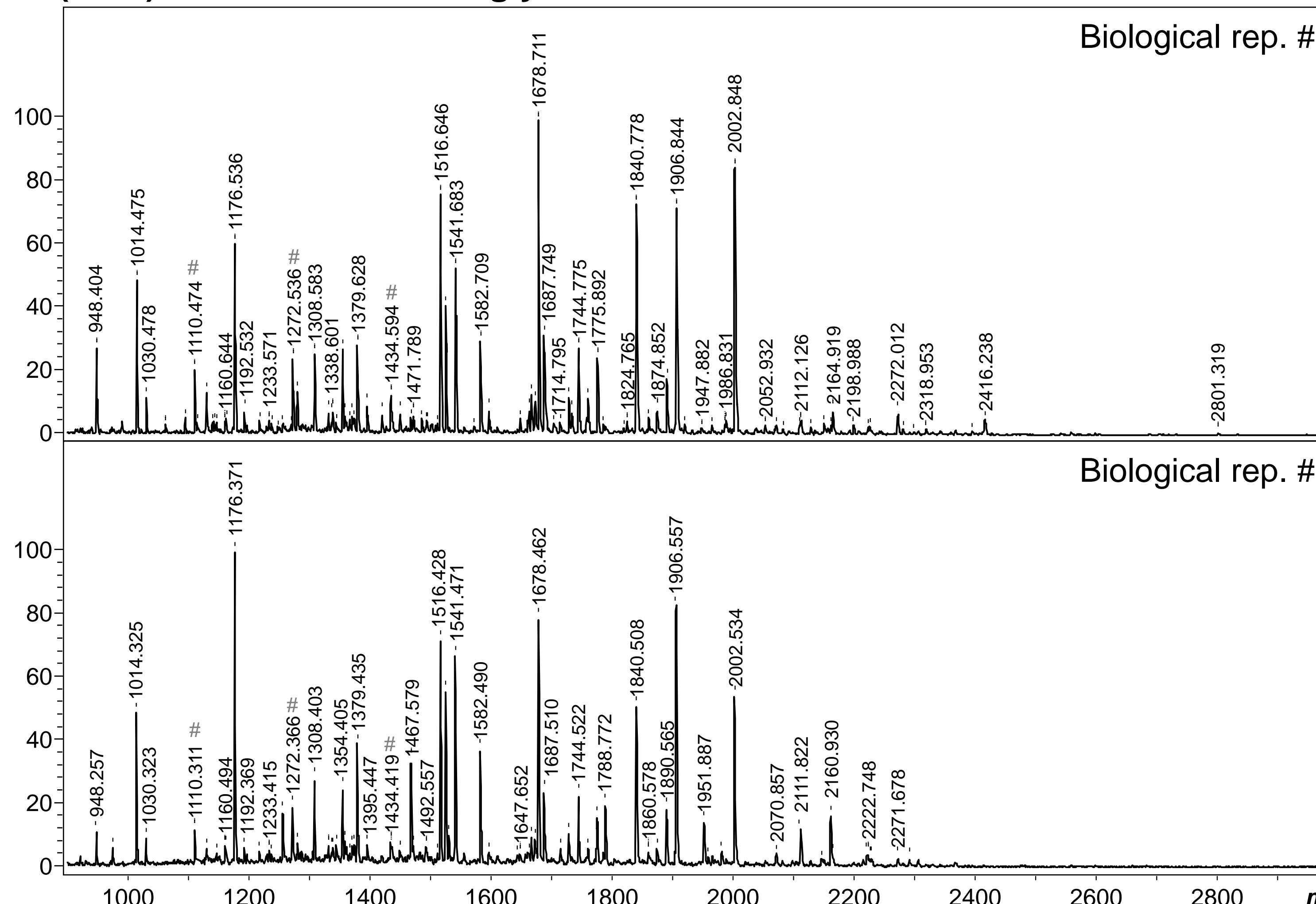


## (A - 3) *S. haematobium* GSL glycans – eggs

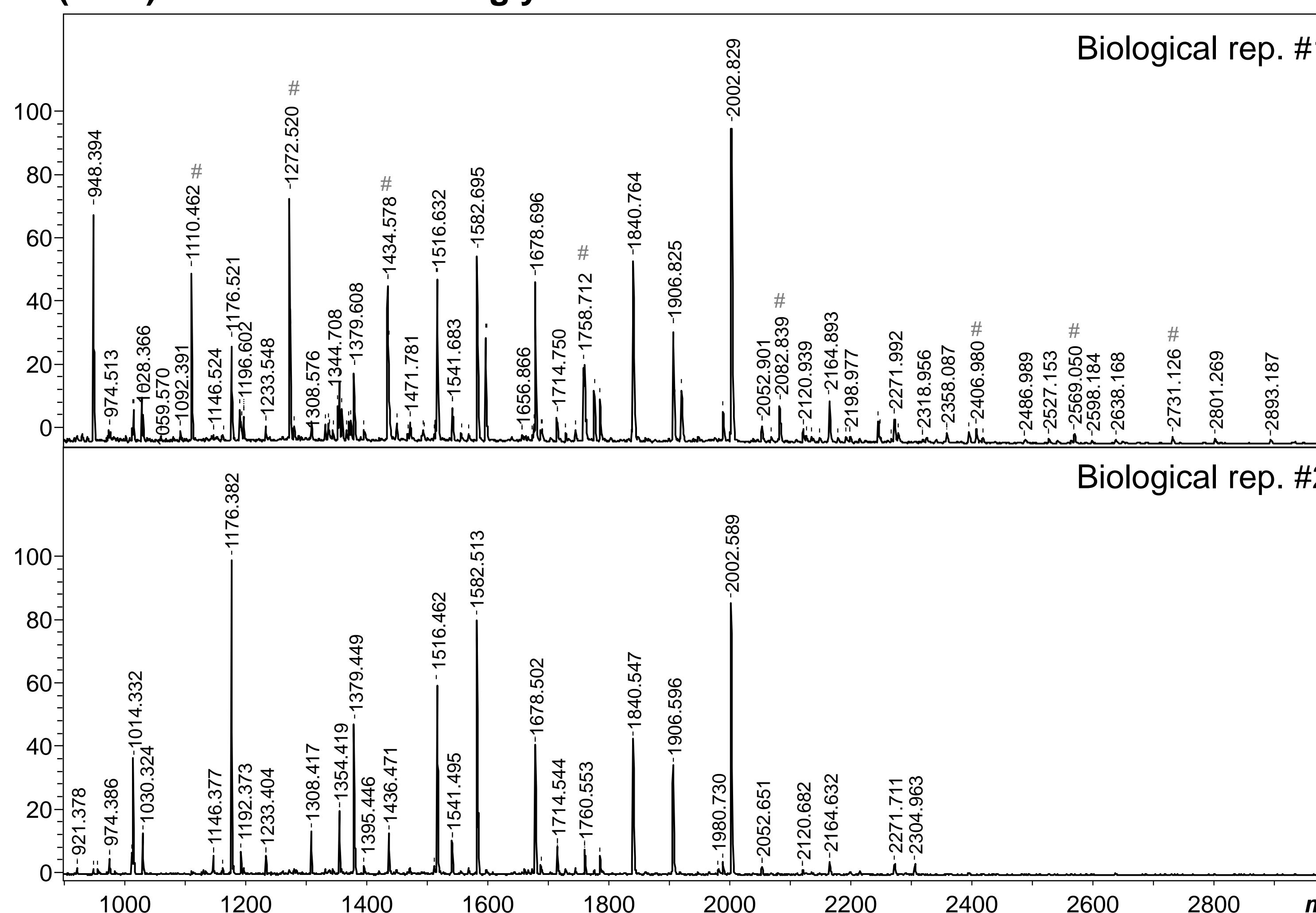


# Biological rep. #1

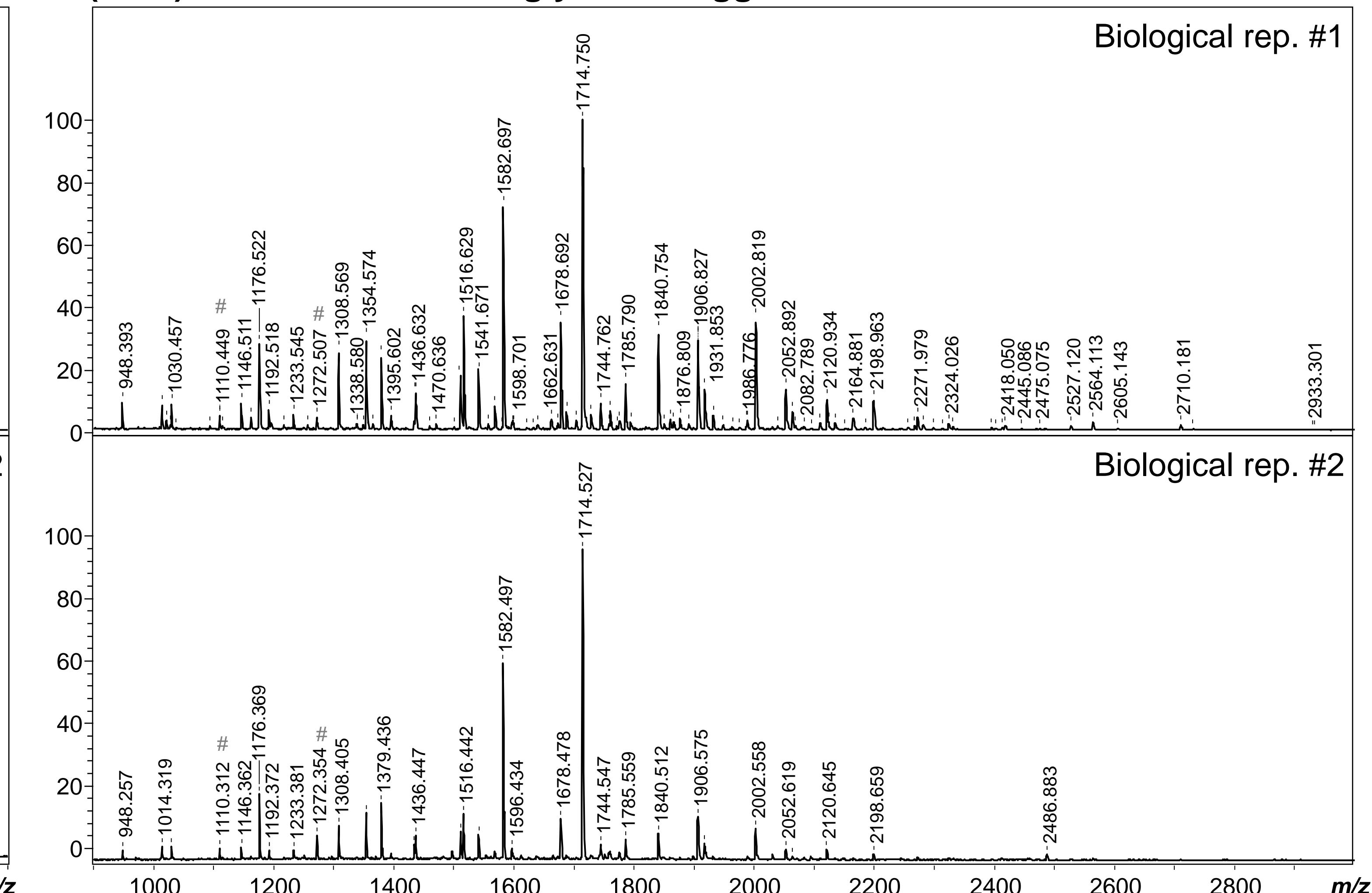
## (B - 1) *S. haematobium* N-glycans – cercariae



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

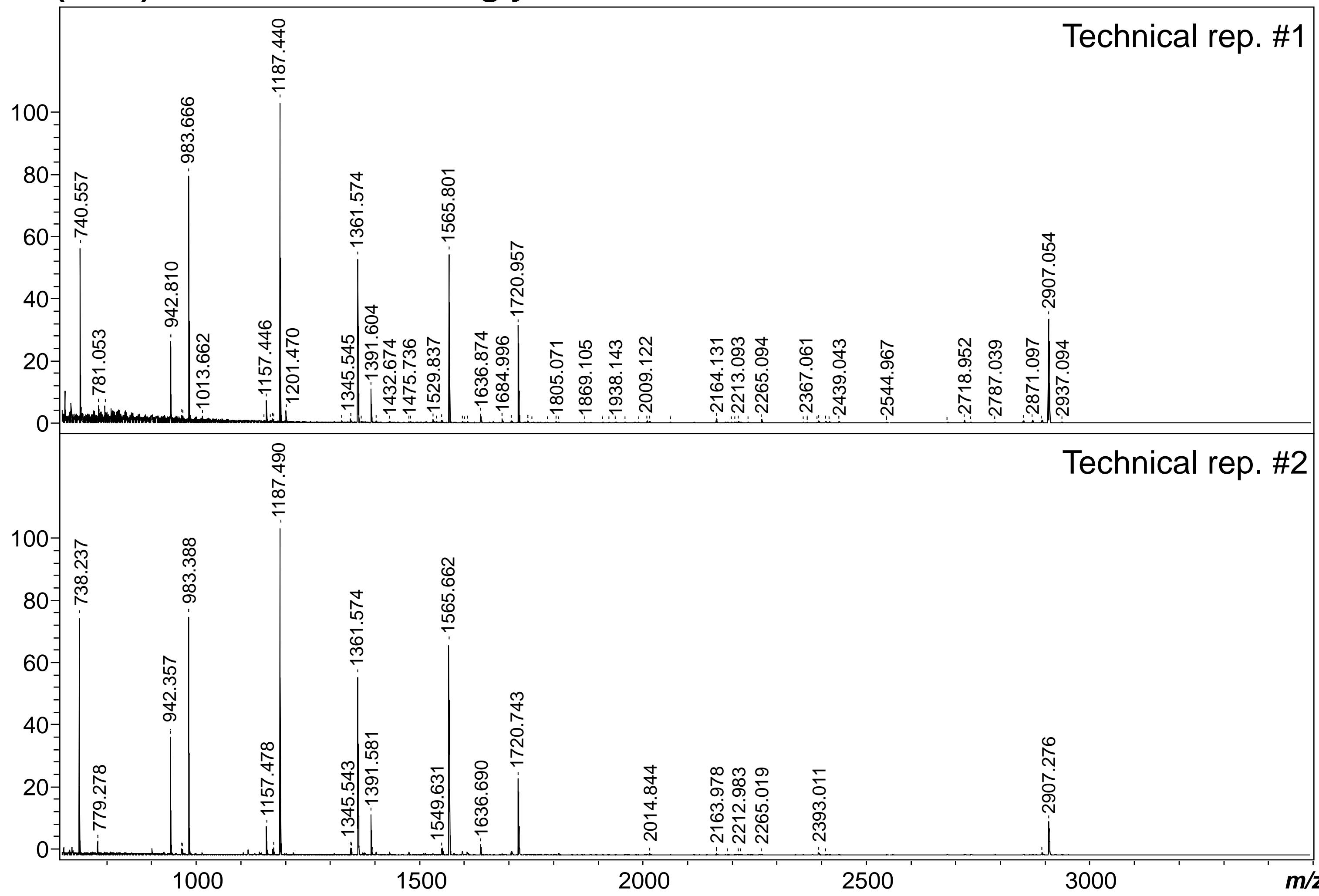


## (B - 3) *S. haematobium* N-glycans – eggs

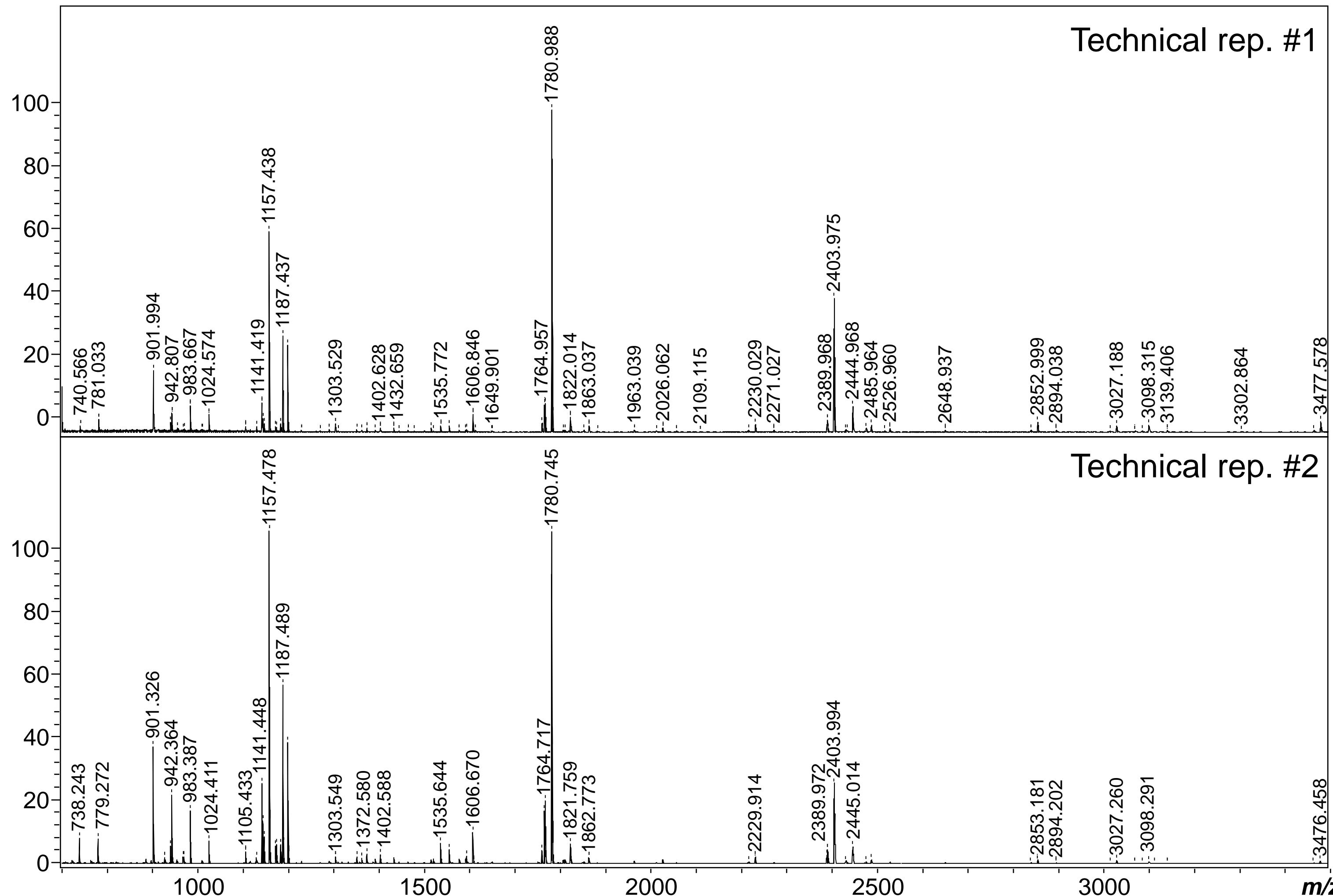


Biological rep. #1

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



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



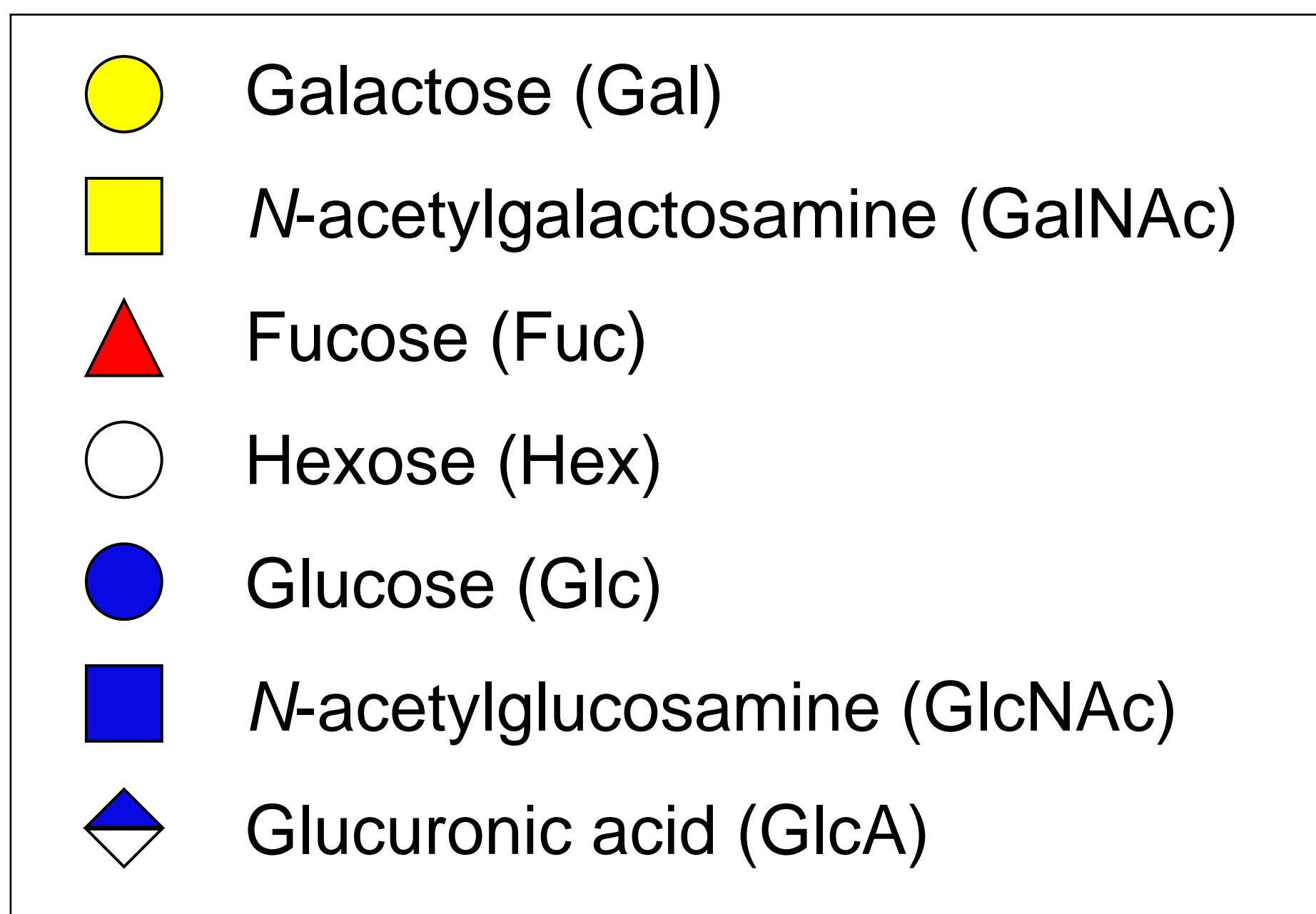
## Figure S2 – Structural characterization of *S. haematobium* and *S. mansoni* GSL glycans

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

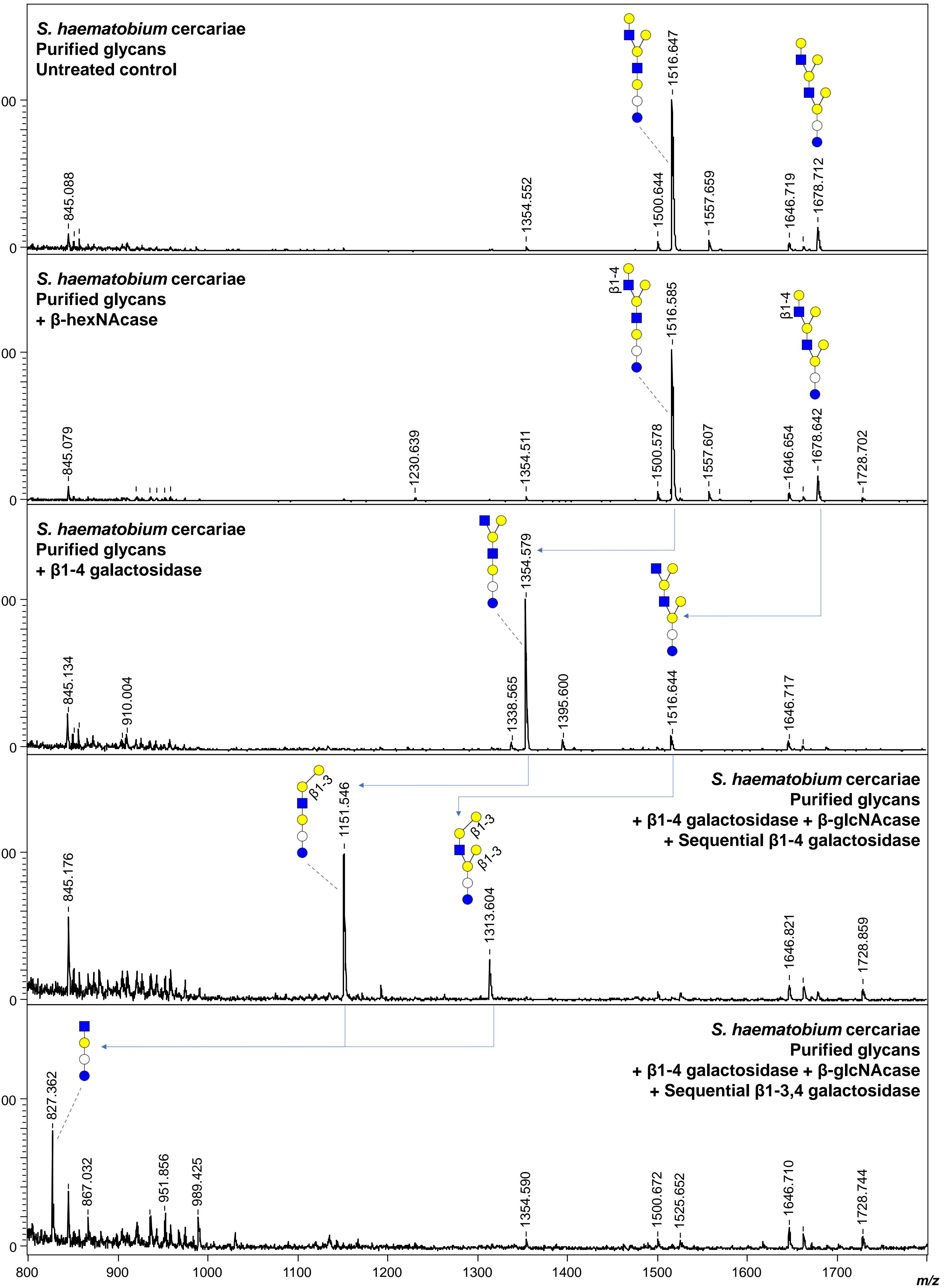
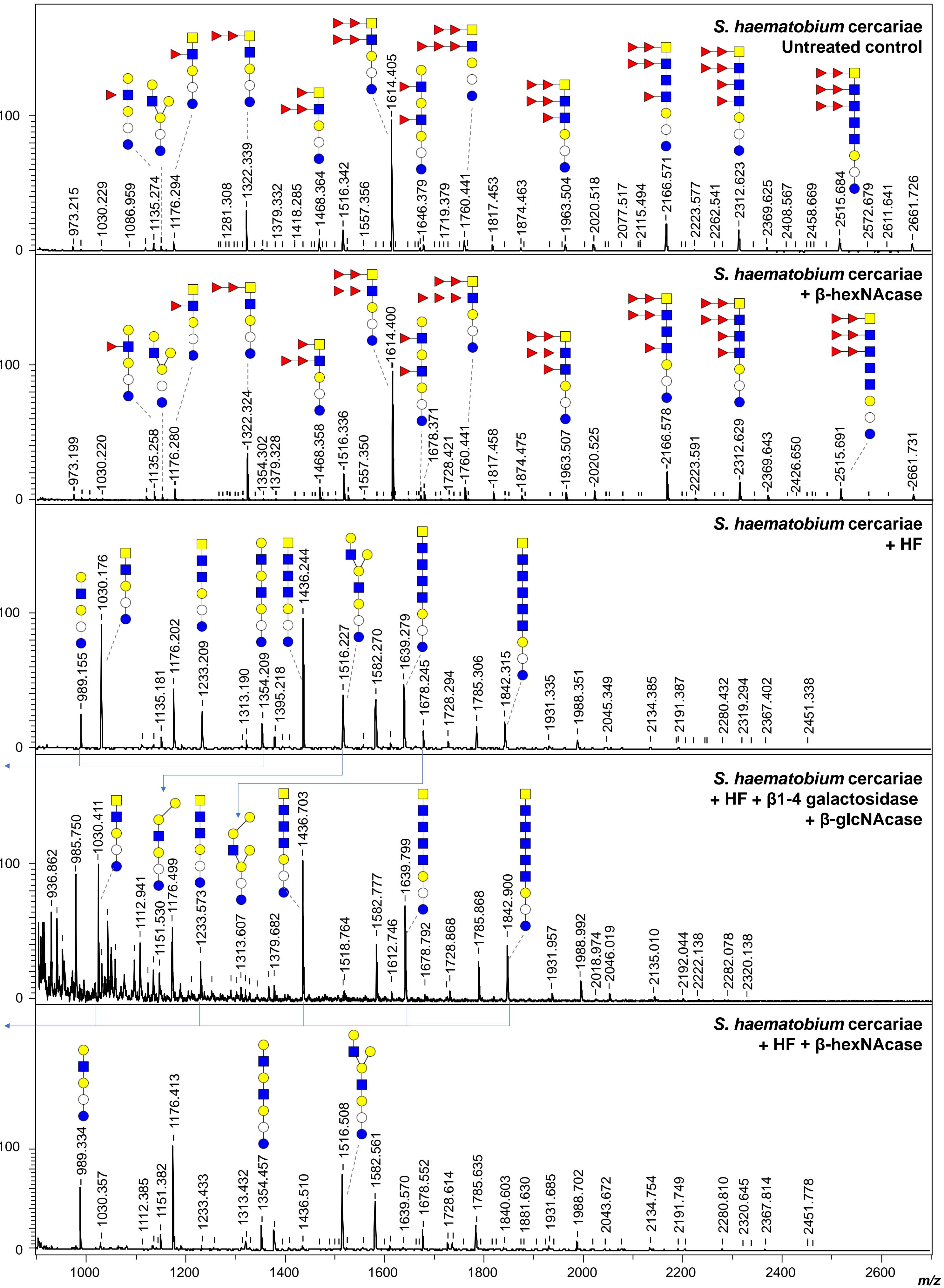
**Glycan sequencing** AA-labeled GSL glycans of *S. haematobium* cercariae (A), adult worms (B), eggs (C) and of *S. mansoni* eggs (acidic fraction only, E) were subjected to hydrofluoric acid (HF) treatment and/or to digestion with exoglycosidase(s). Enzymes were either applied simultaneously to the glycan sample, as a cocktail, or separately (“Sequential”) as indicated. Reactions were performed in the conditions detailed in **Table 1** (see M&M). 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) is 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 & 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 M&M.

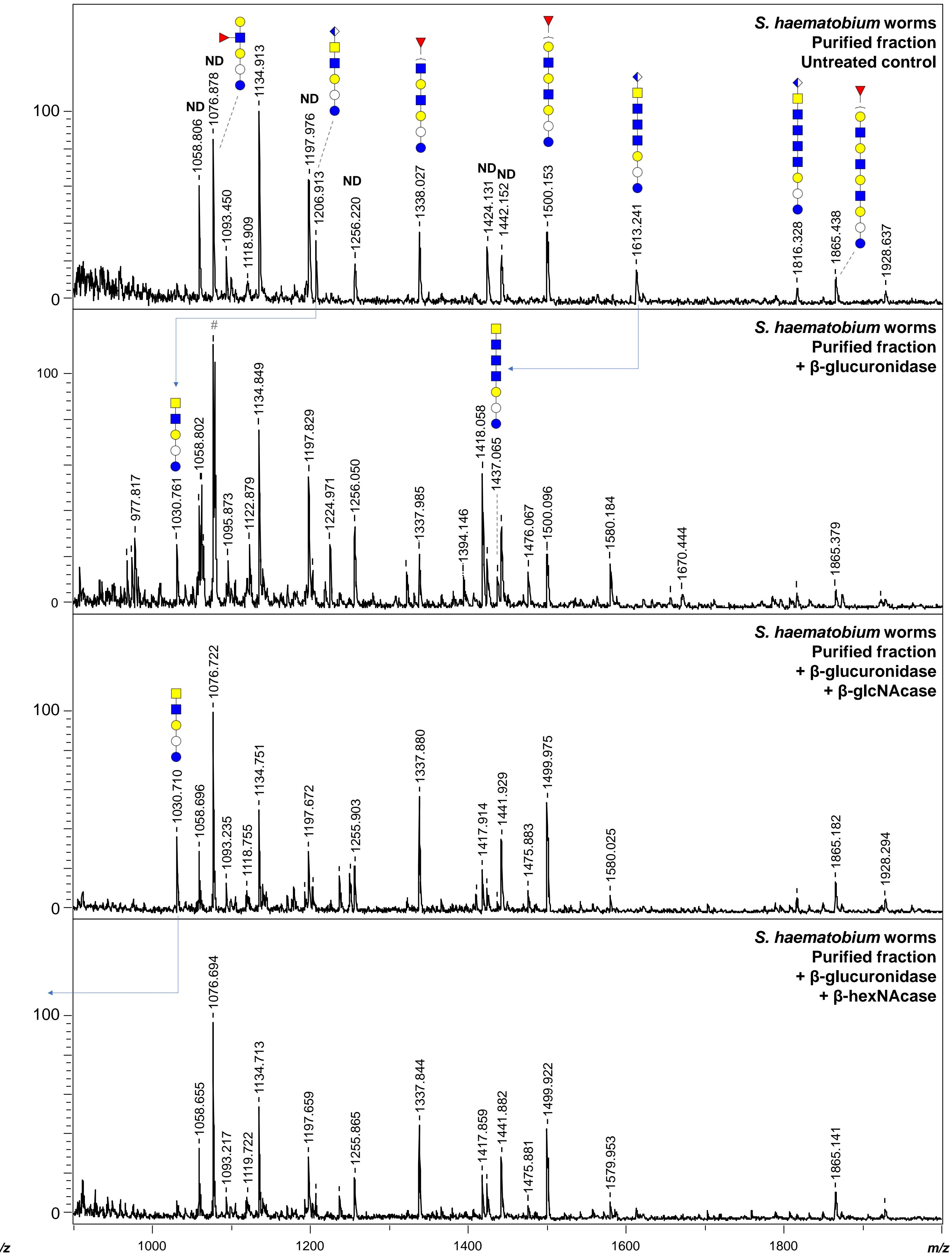
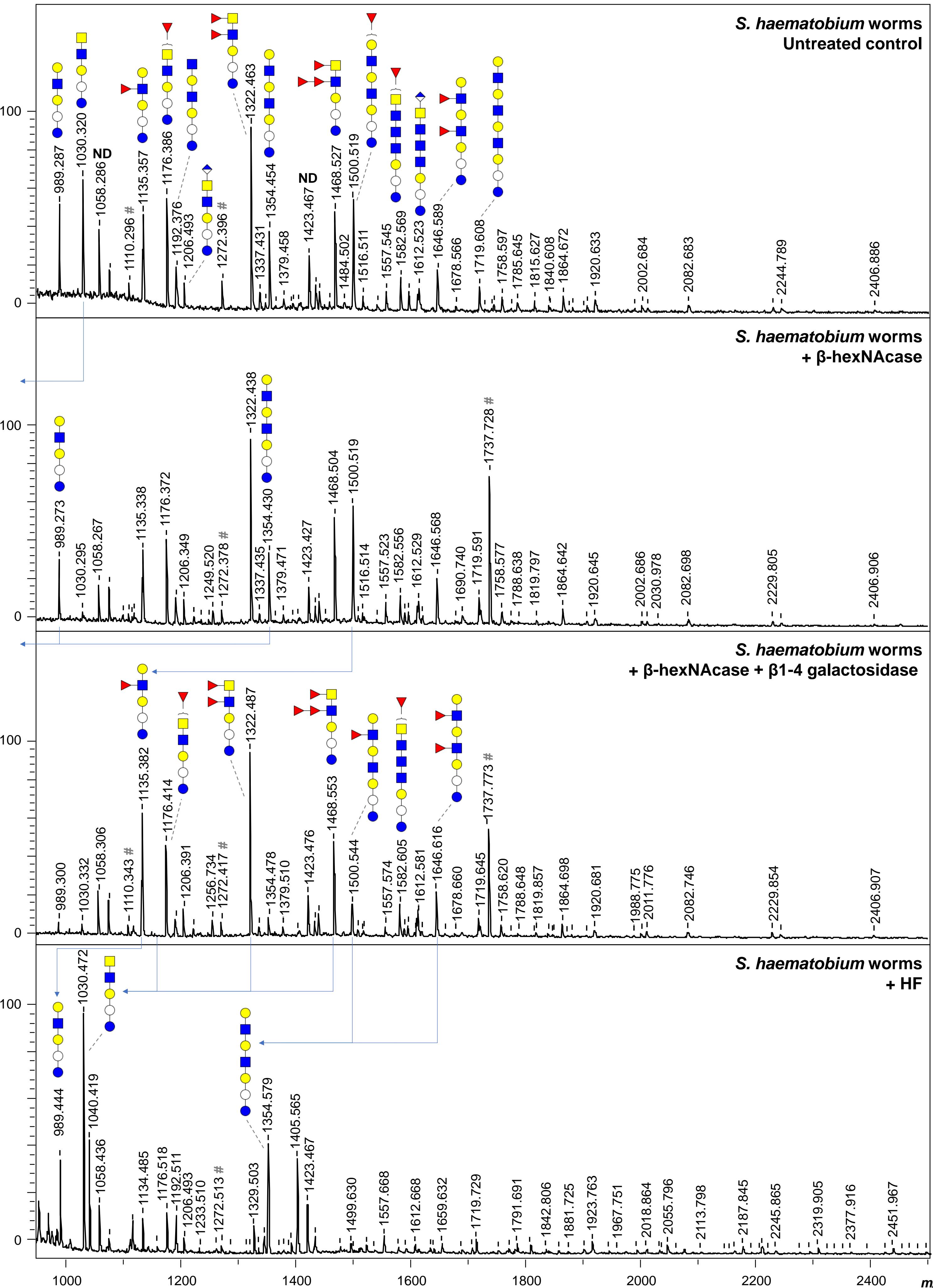
All (putative) glycans are represented using the CFG nomenclature below.



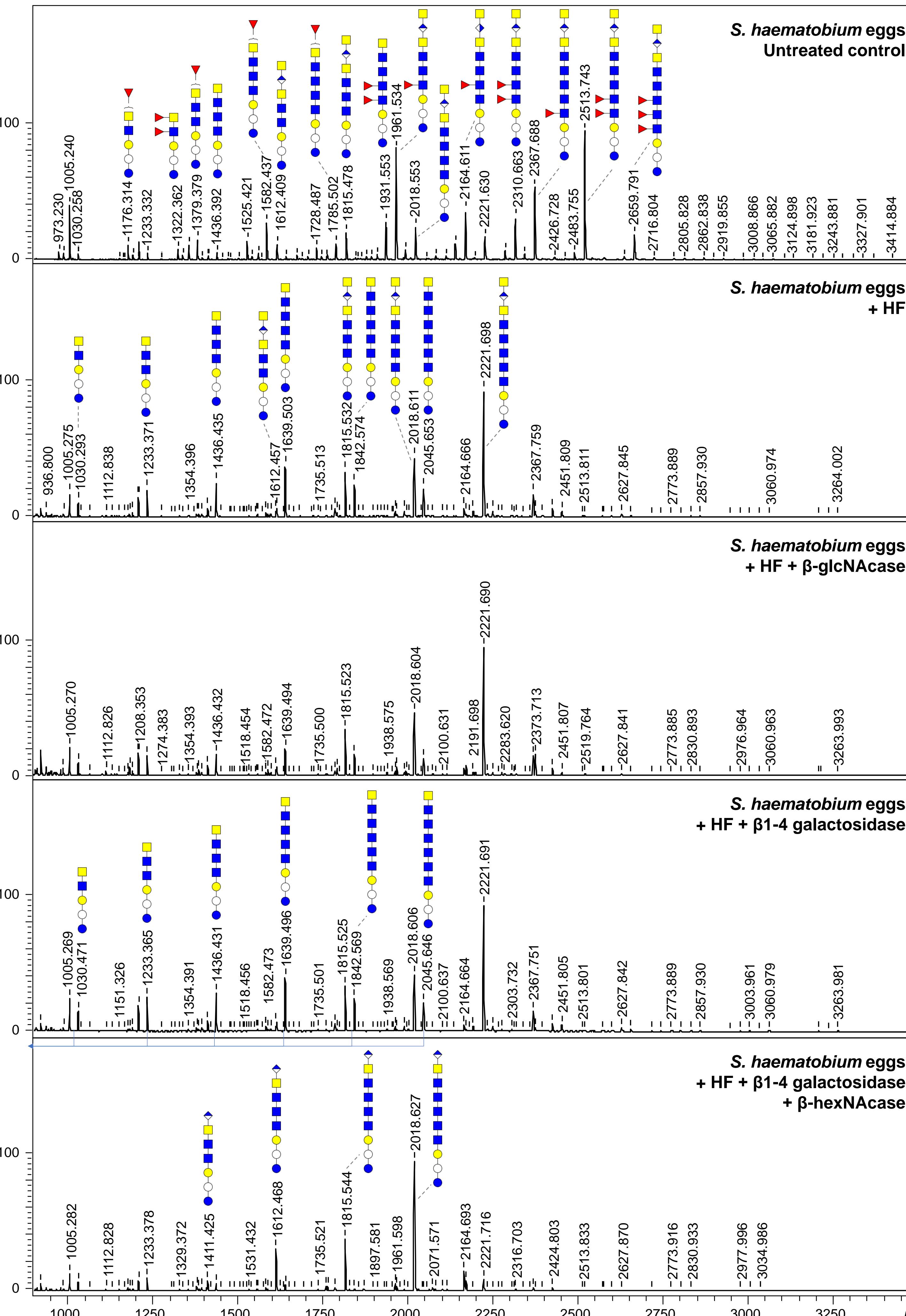
**(A) *S. haematobium* cercariae – glycan sequencing**



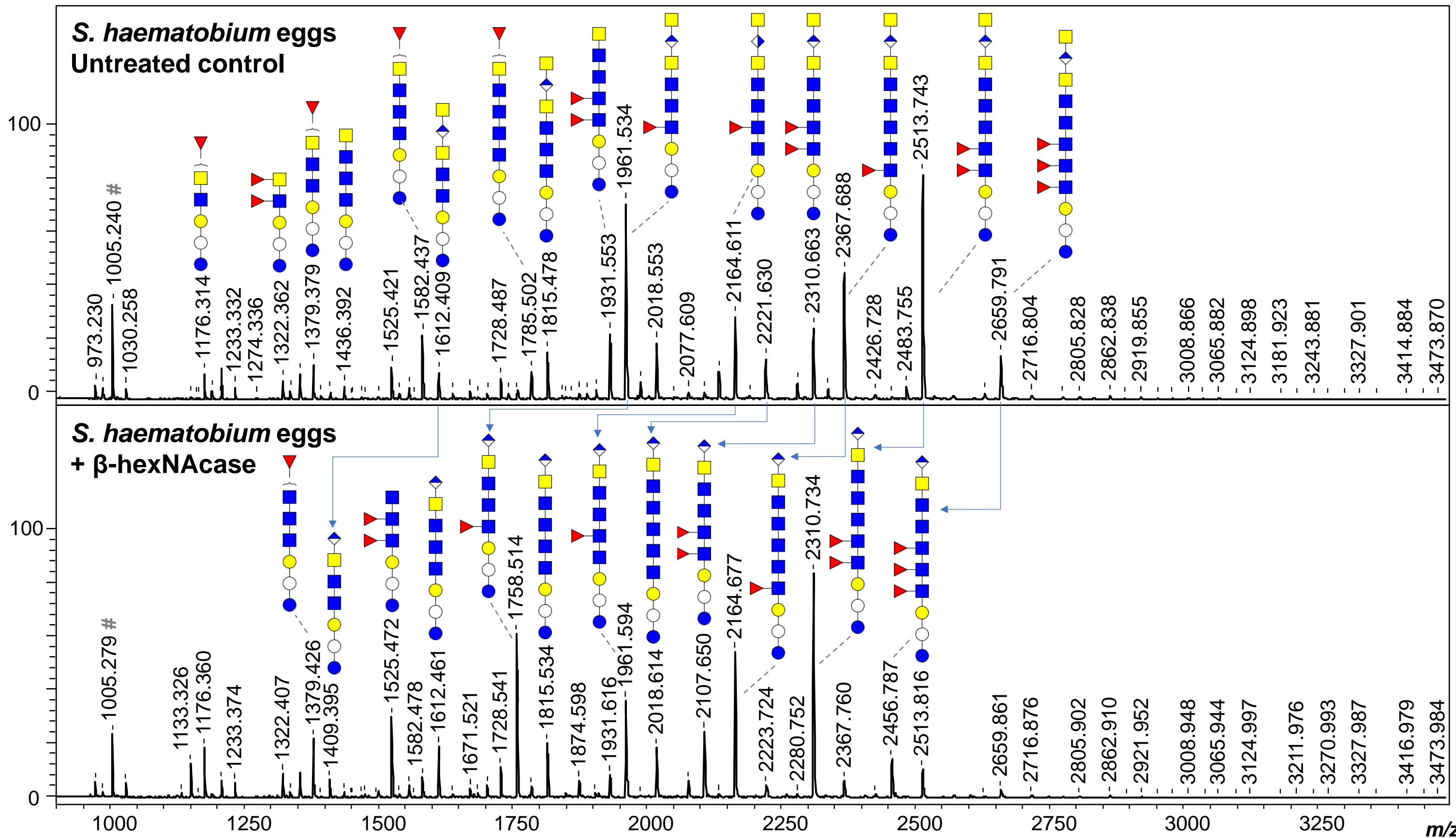
## (B) *S. haematobium* adult worms – glycan sequencing



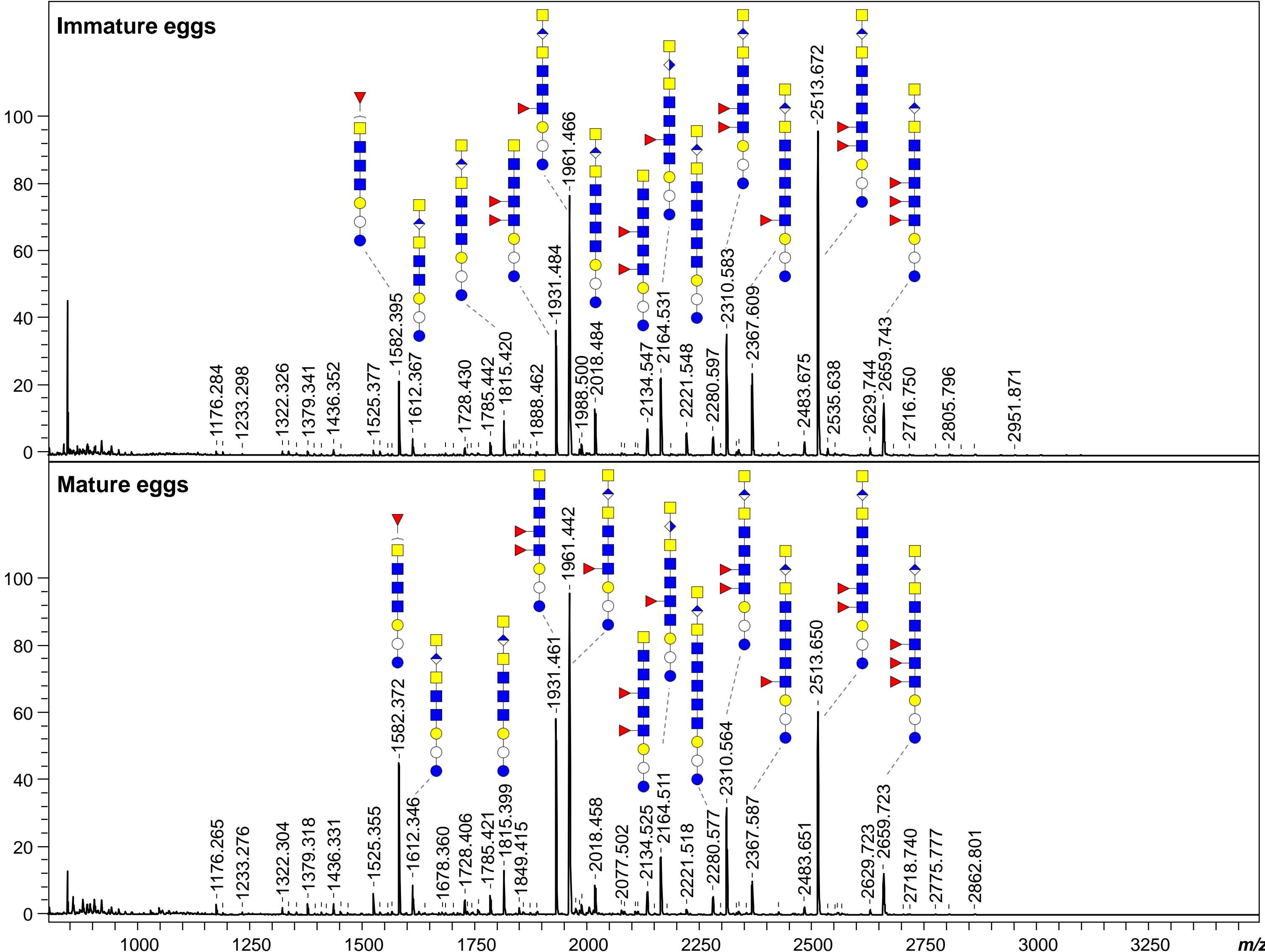
**(C) *S. haematobium* eggs – glycan sequencing**



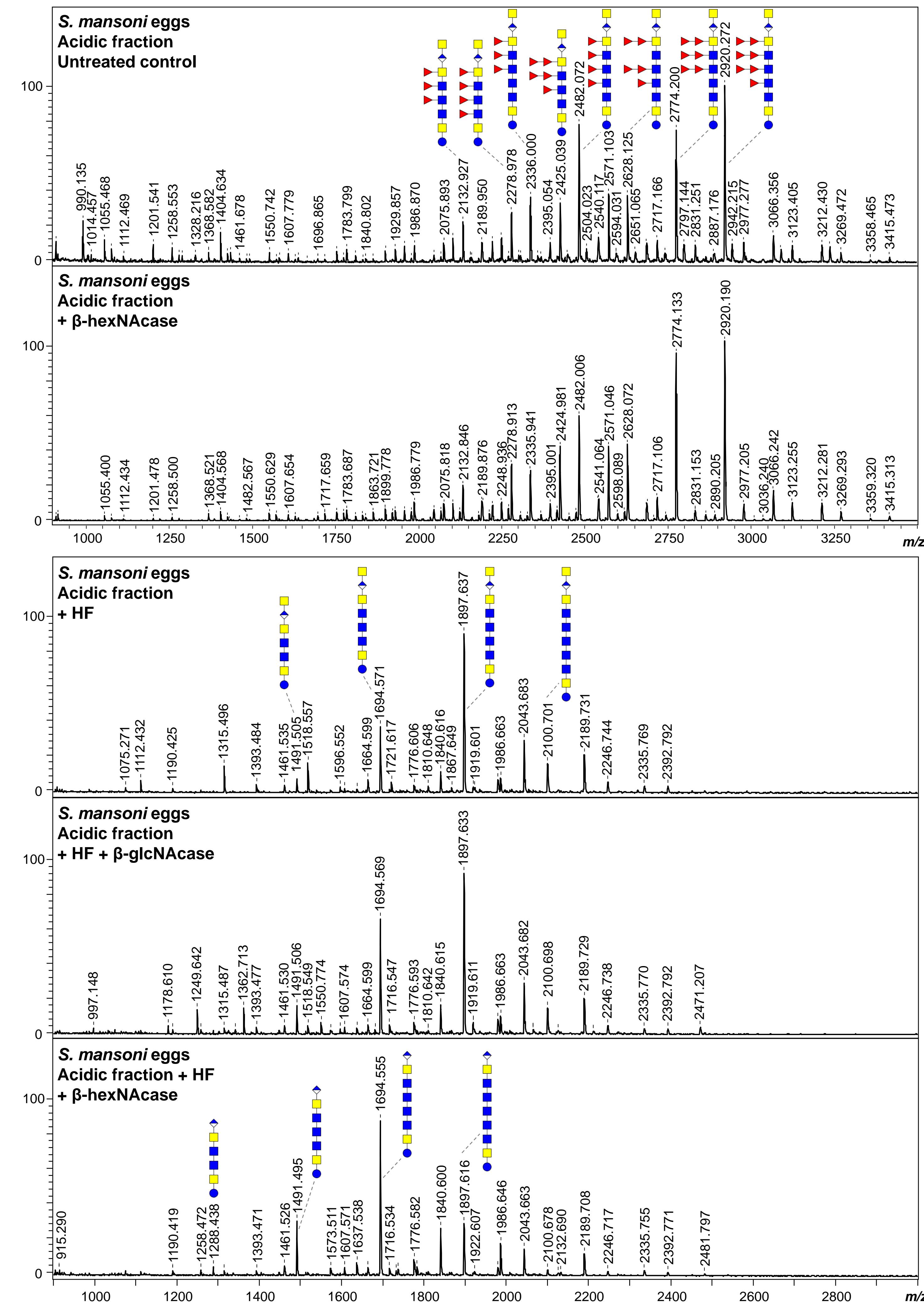
**(C) *S. haematobium* eggs – glycan sequencing (continued)**



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



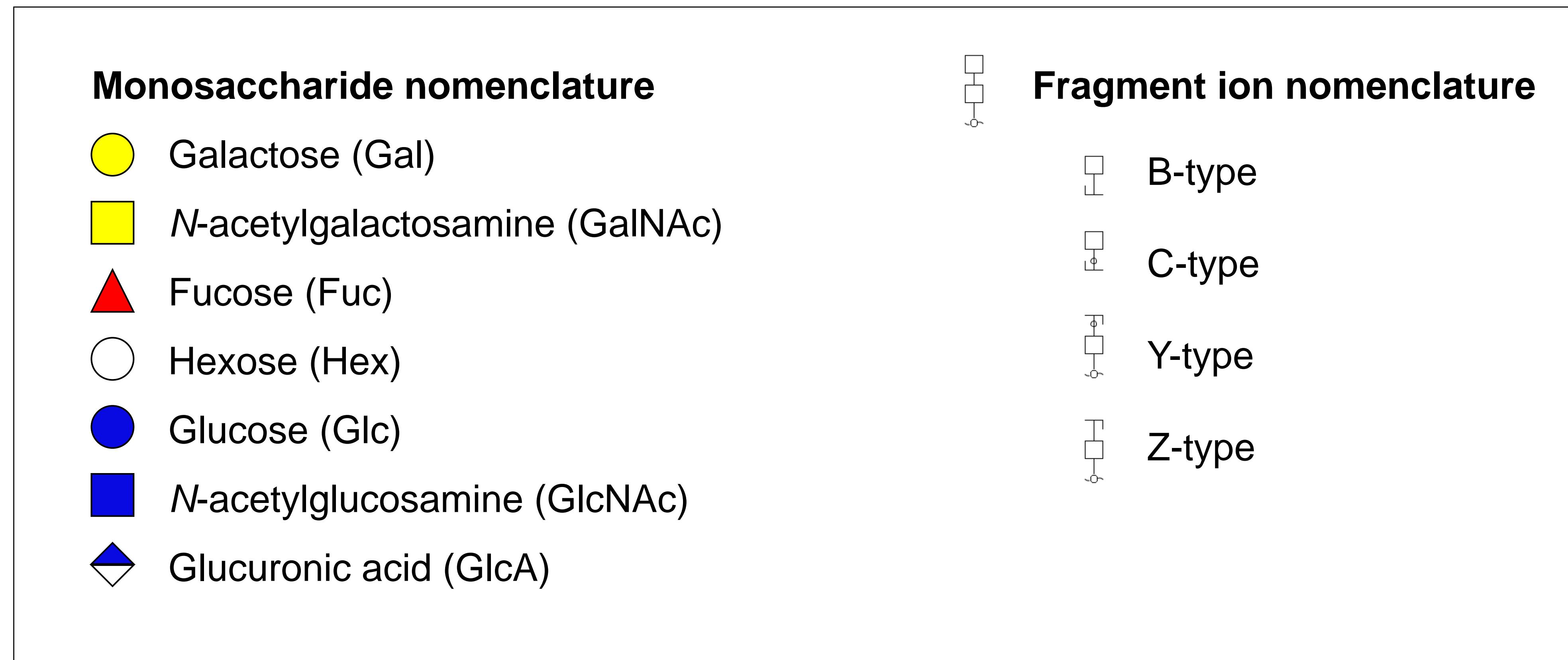
**(E) Acidic GSL glycans of *S. mansoni* eggs – glycan sequencing**



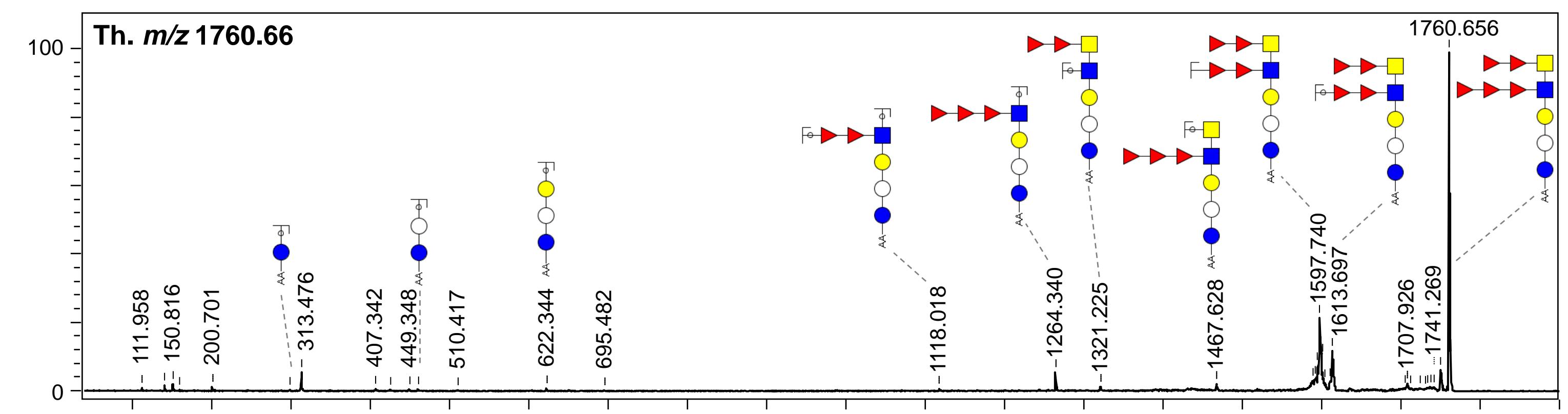
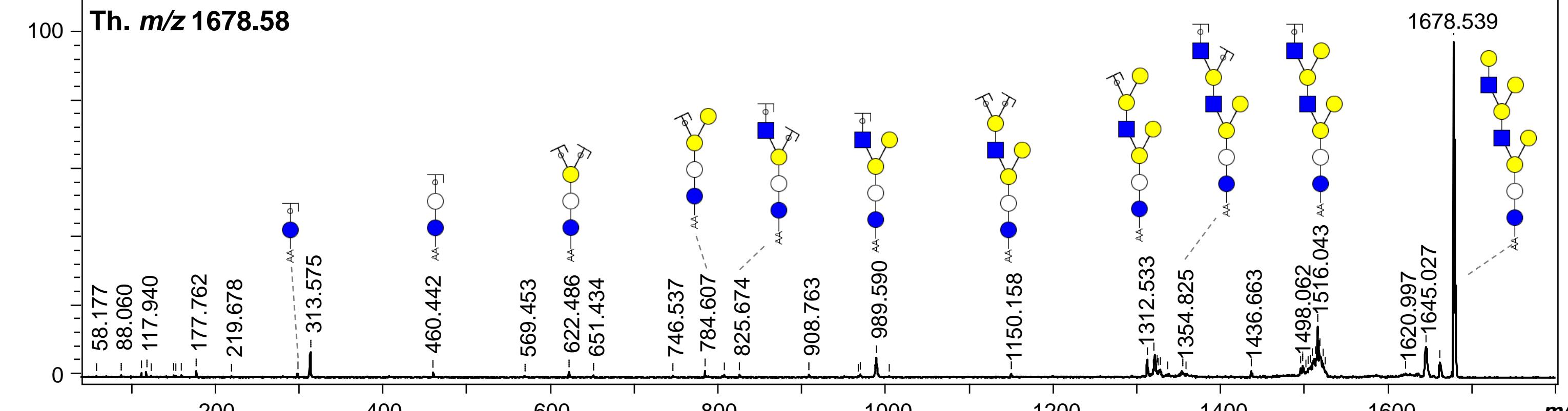
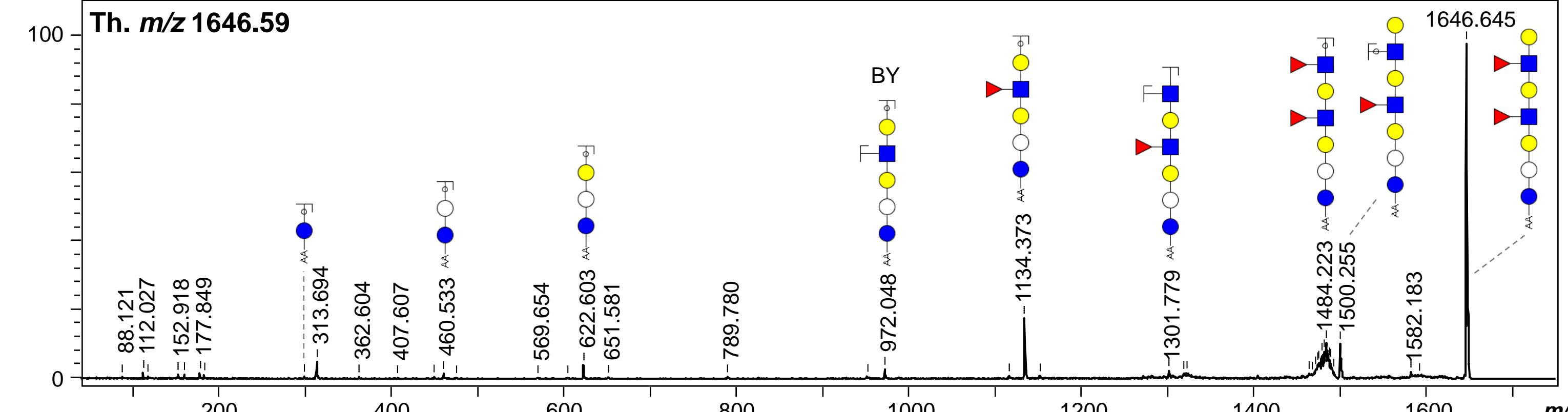
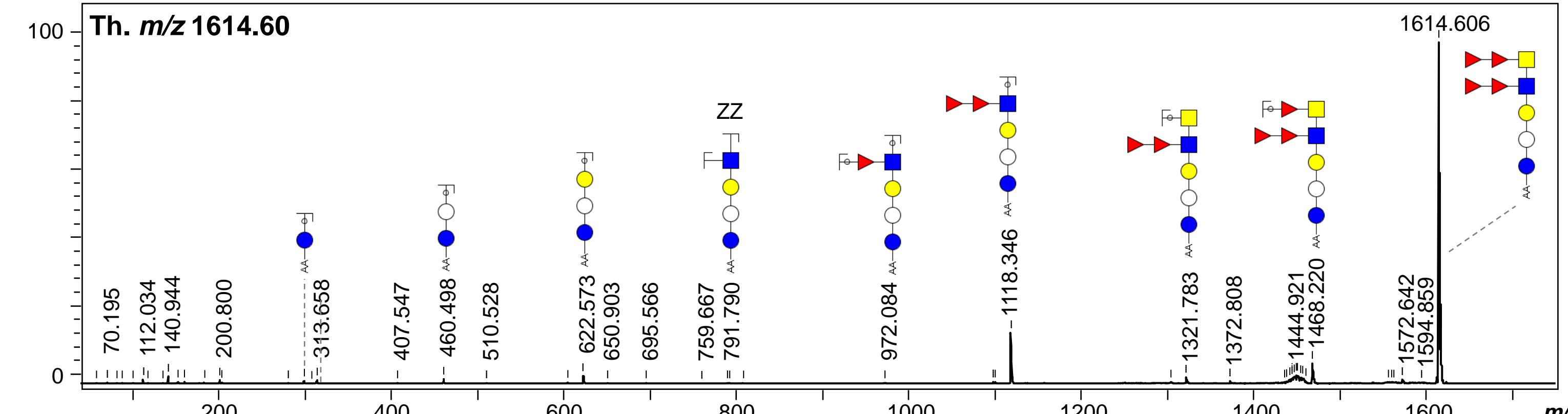
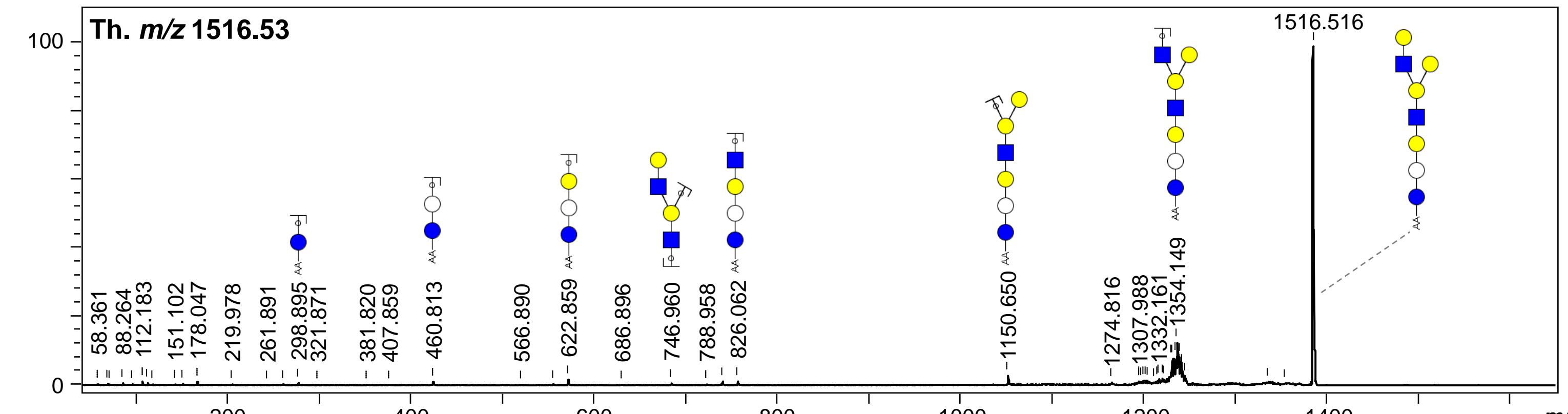
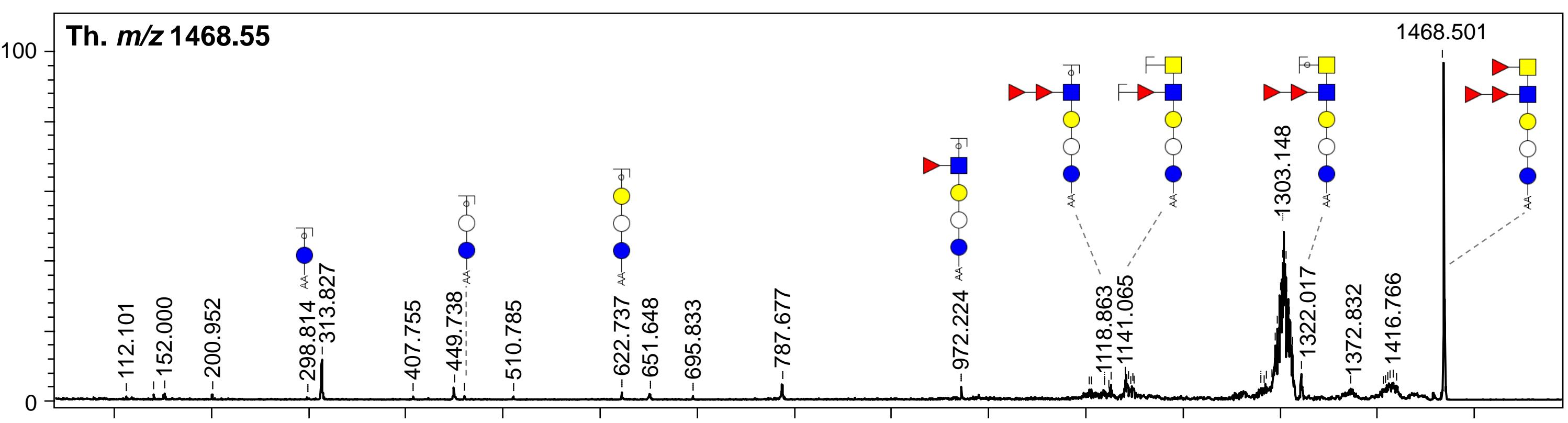
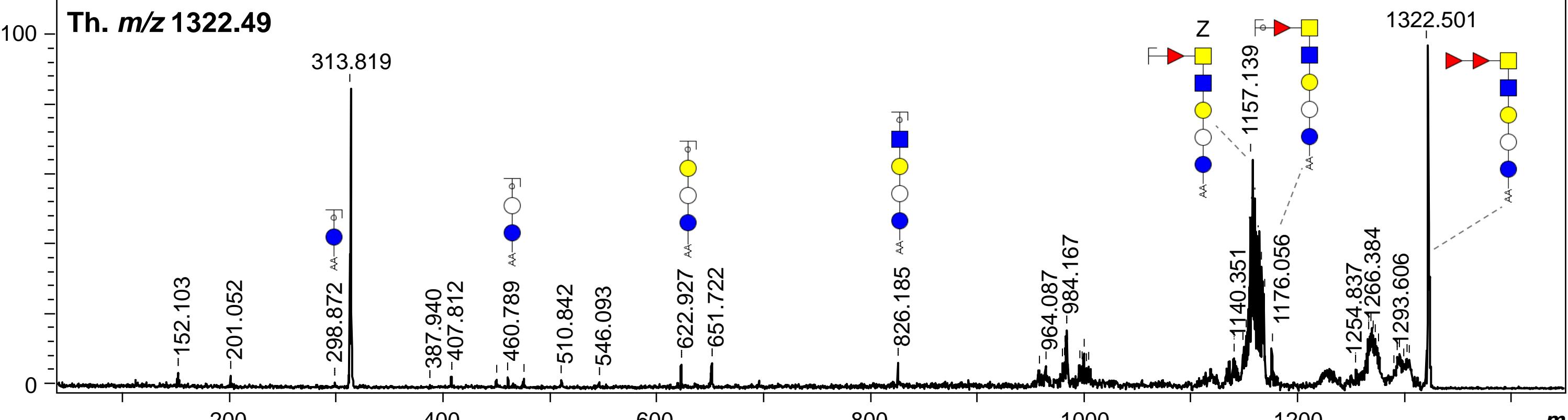
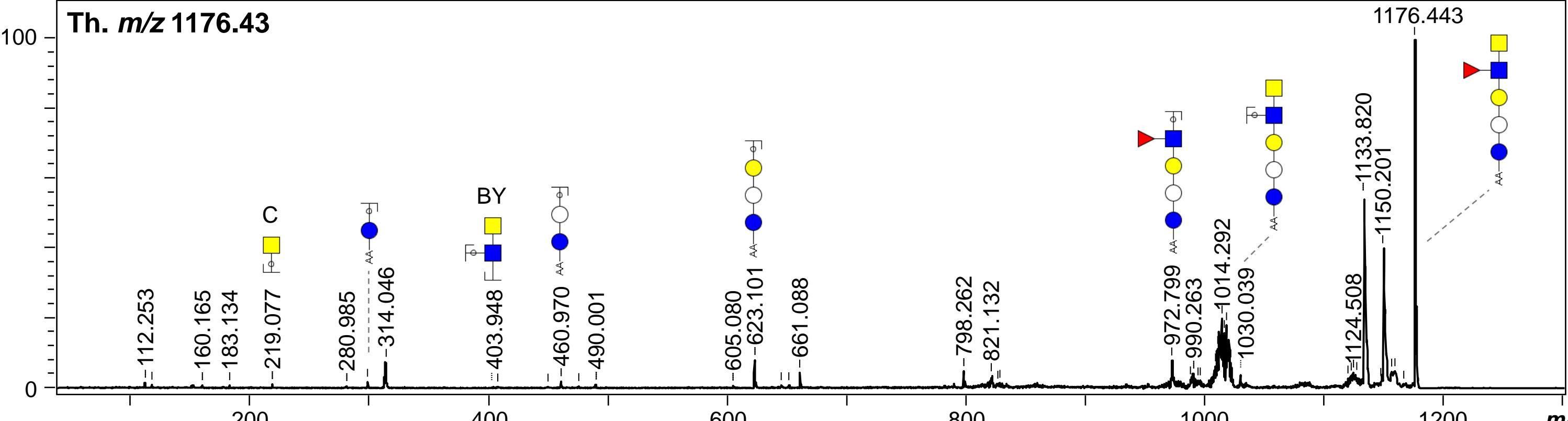
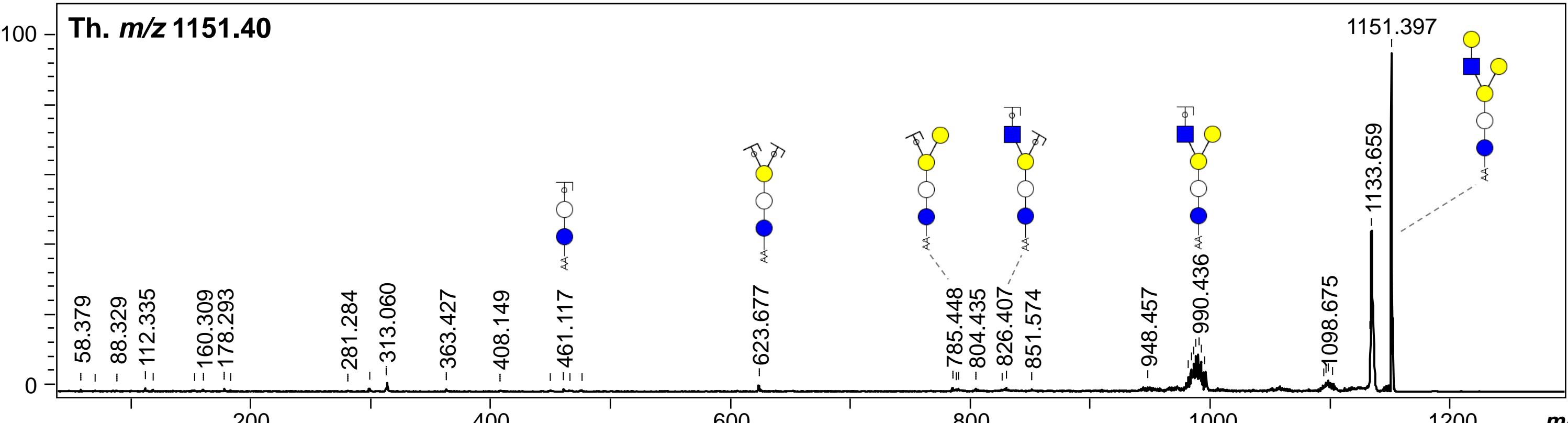
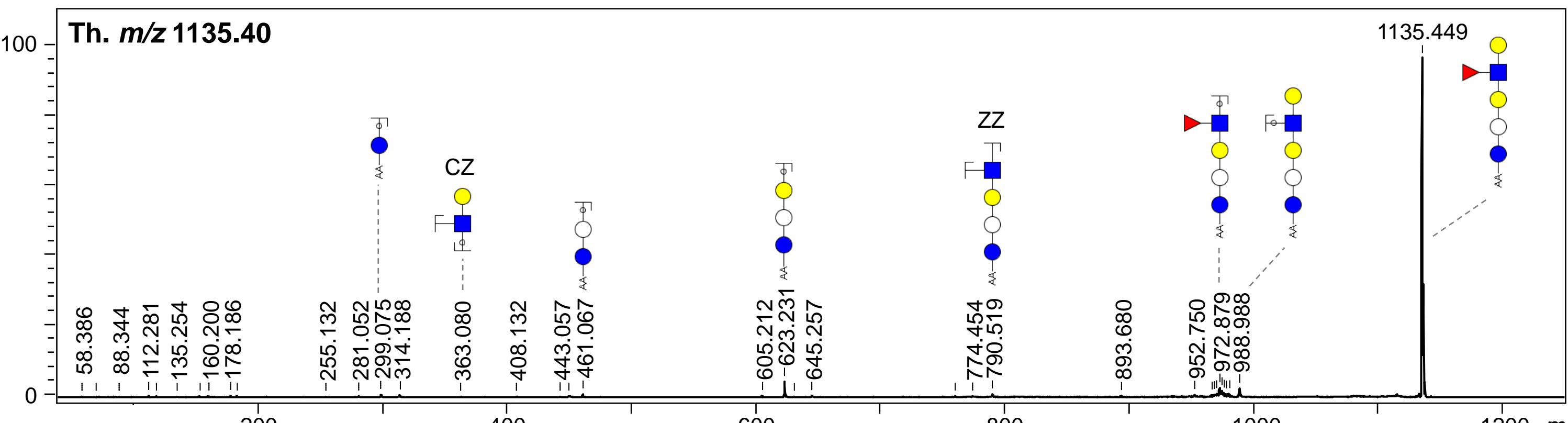
## Figure S3 – MALDI-TOF-MS/MS of *S. haematobium* and *S. mansoni* GSL glycans

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]<sup>-</sup>) 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).

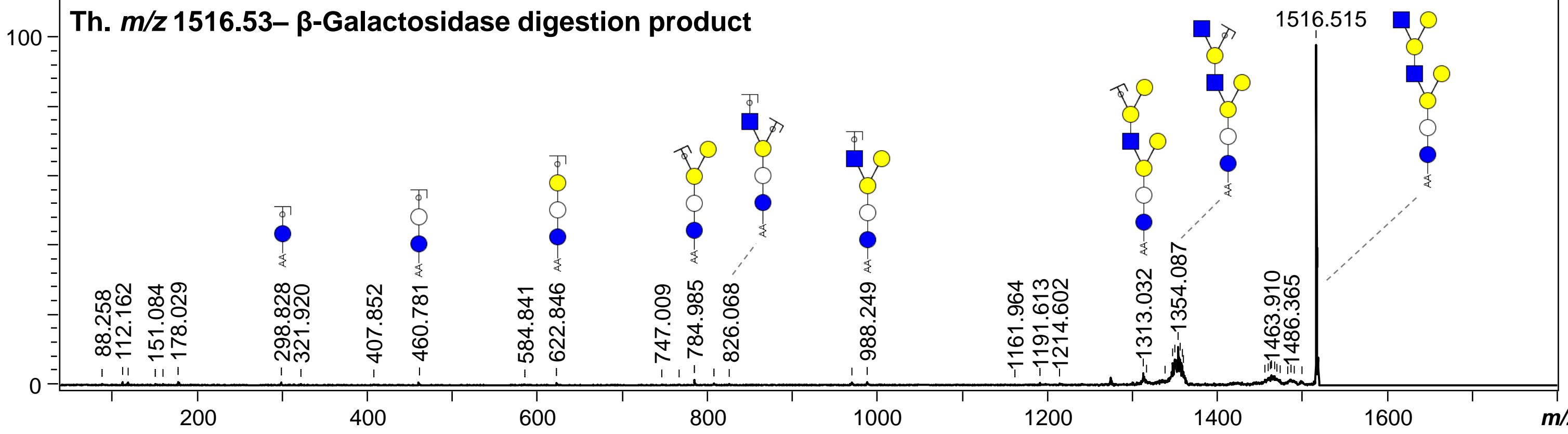
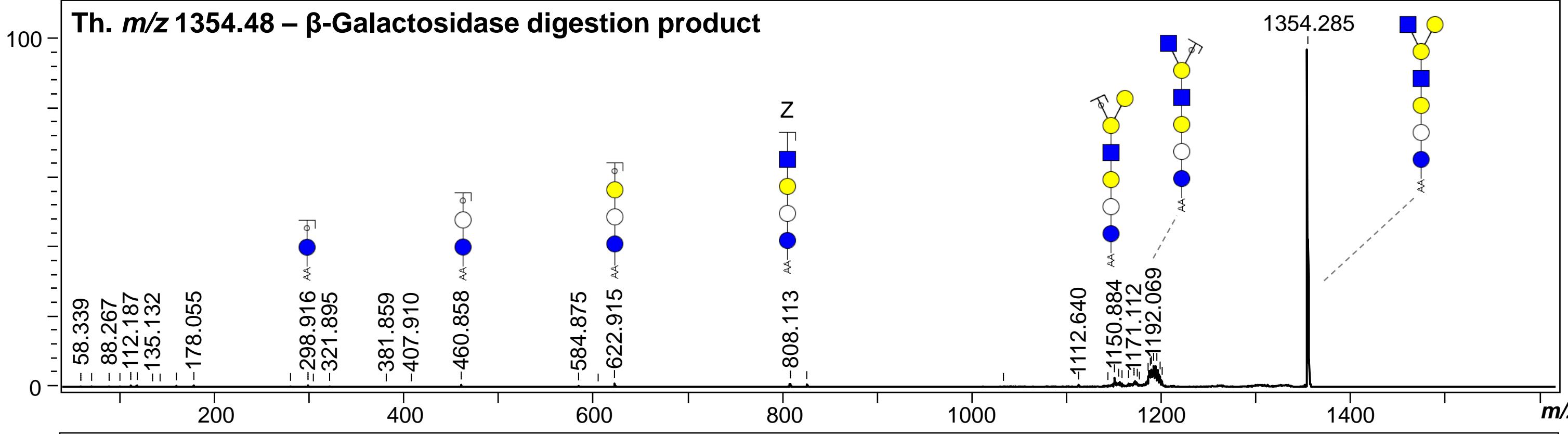
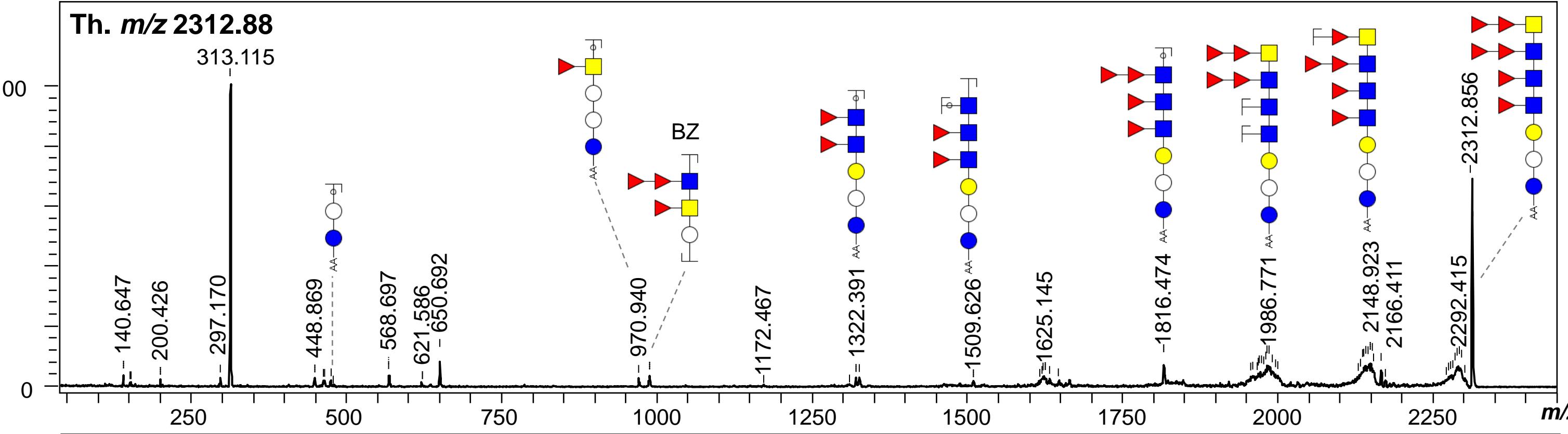
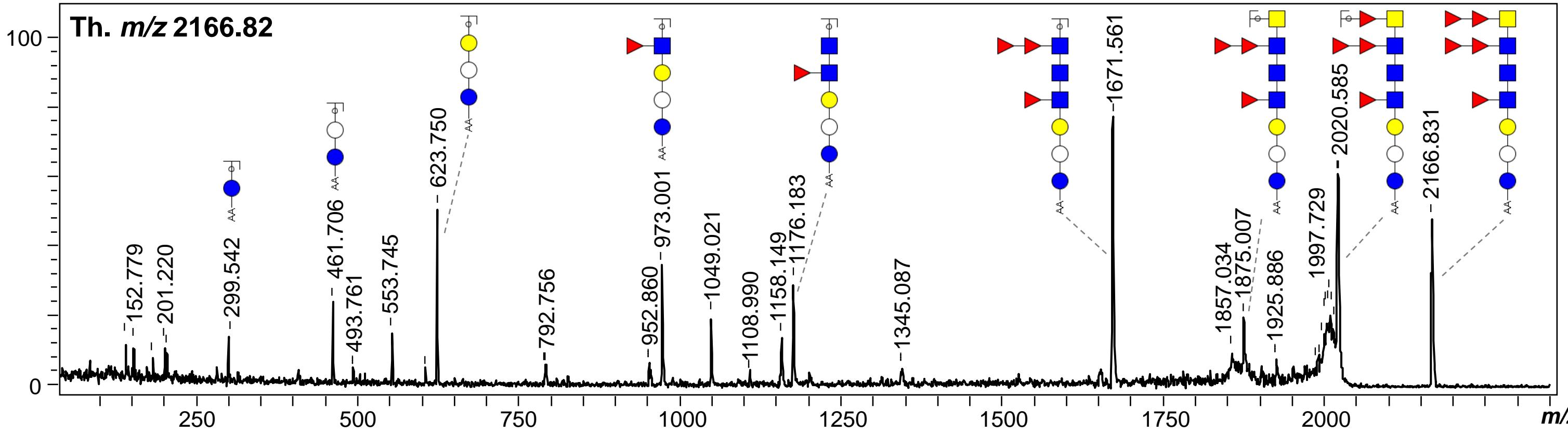
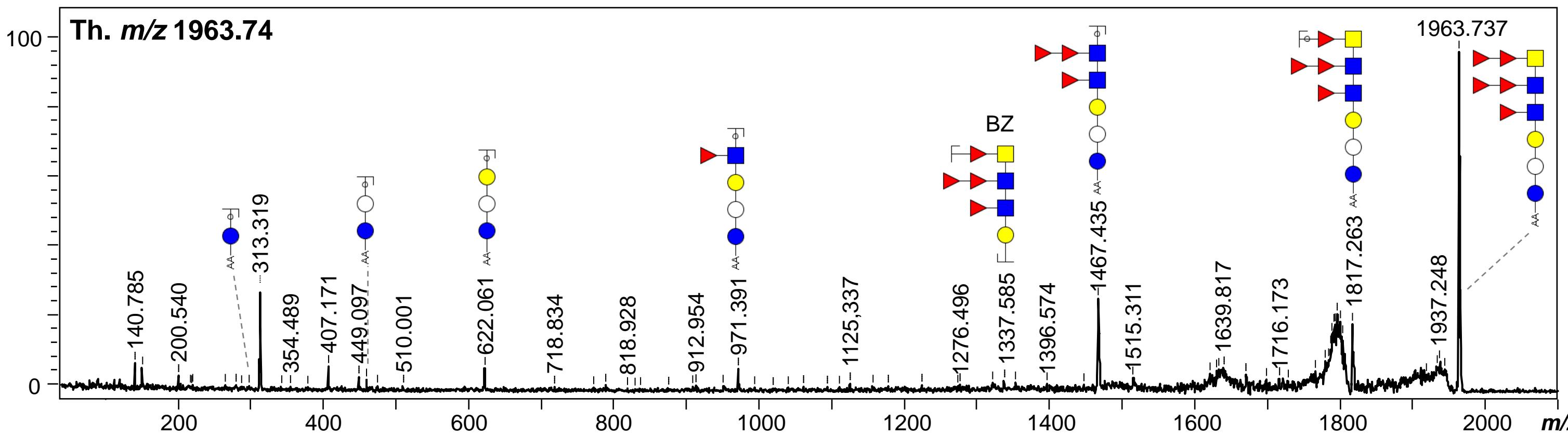
All (putative) glycans are represented using the CFG nomenclature (see inset below).



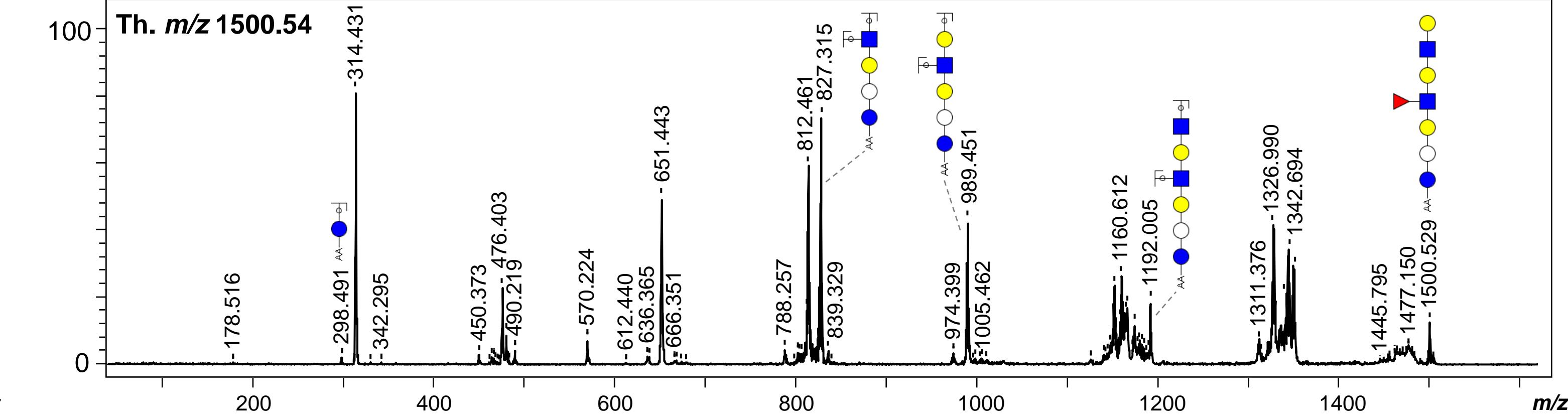
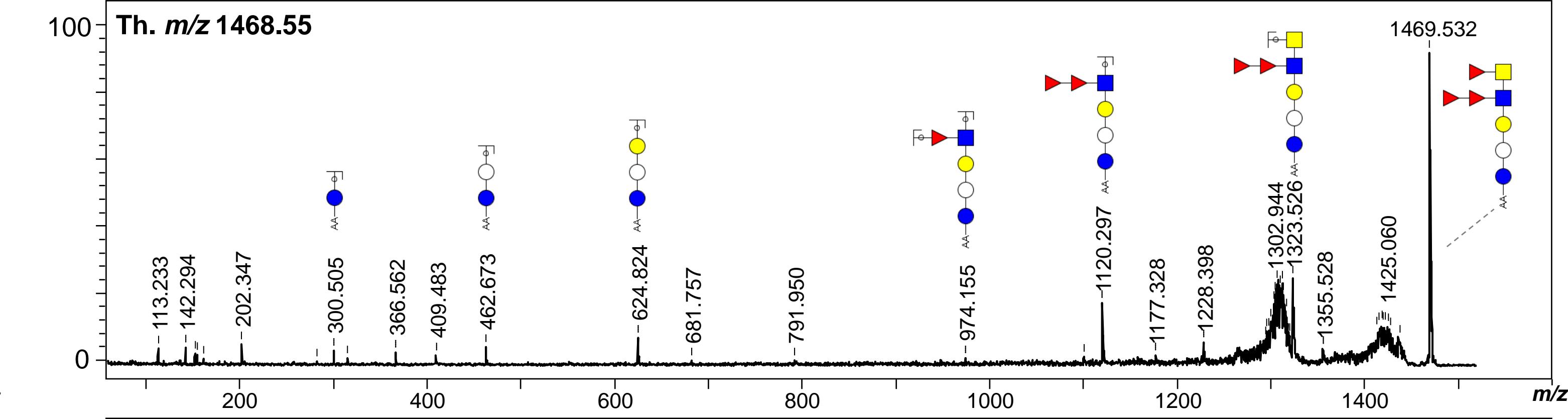
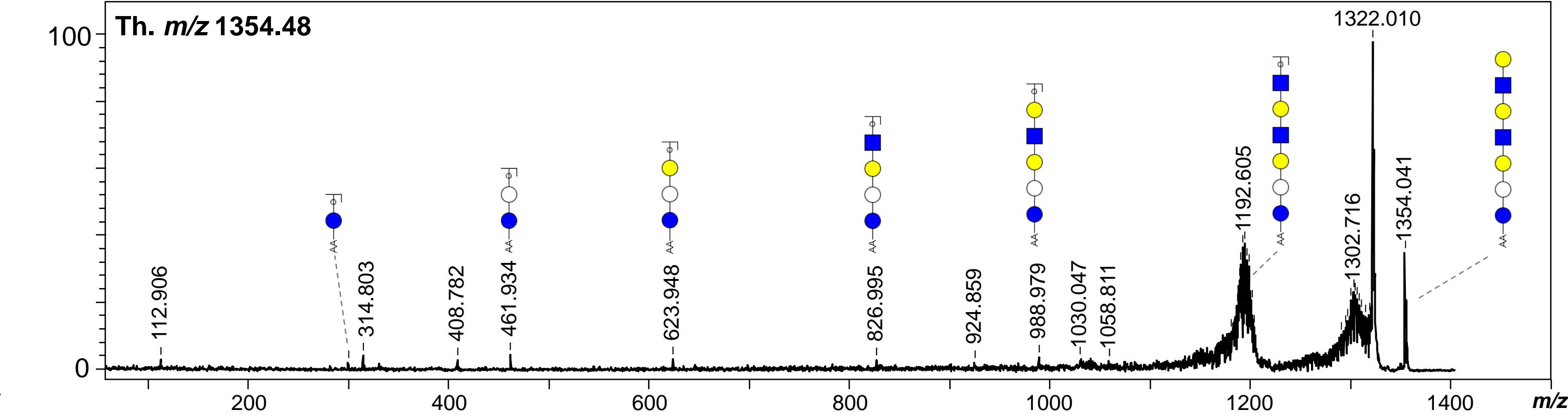
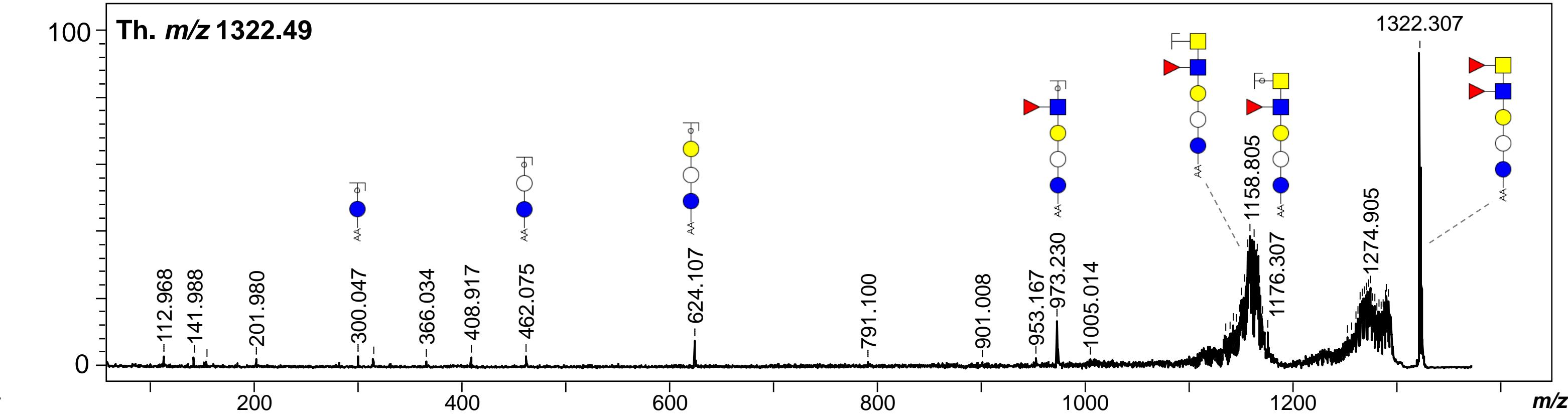
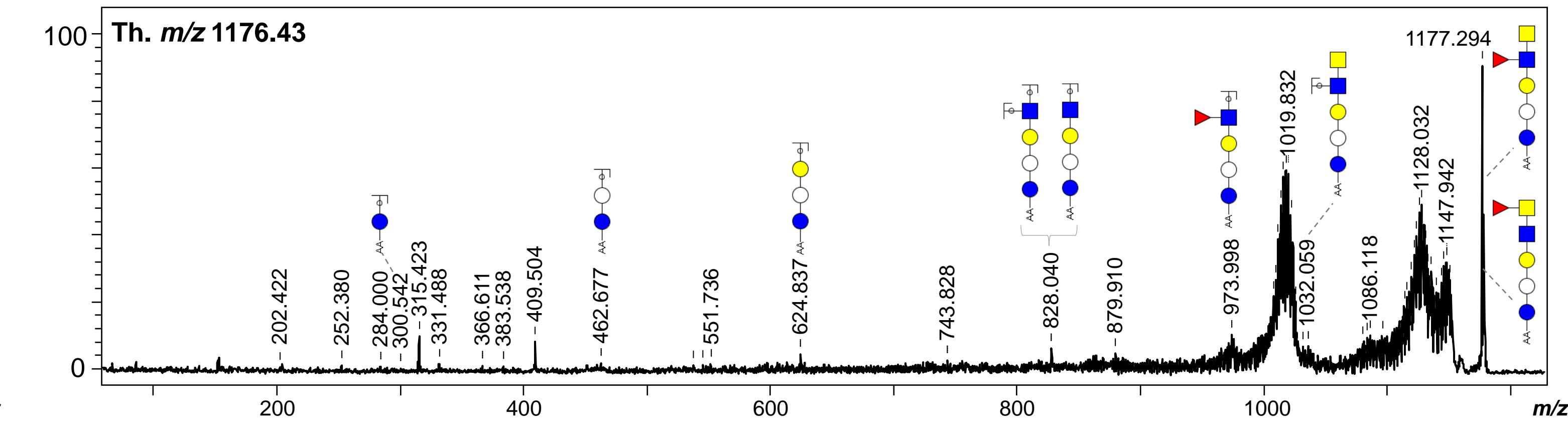
## (A) *S. haematobium* cercariae



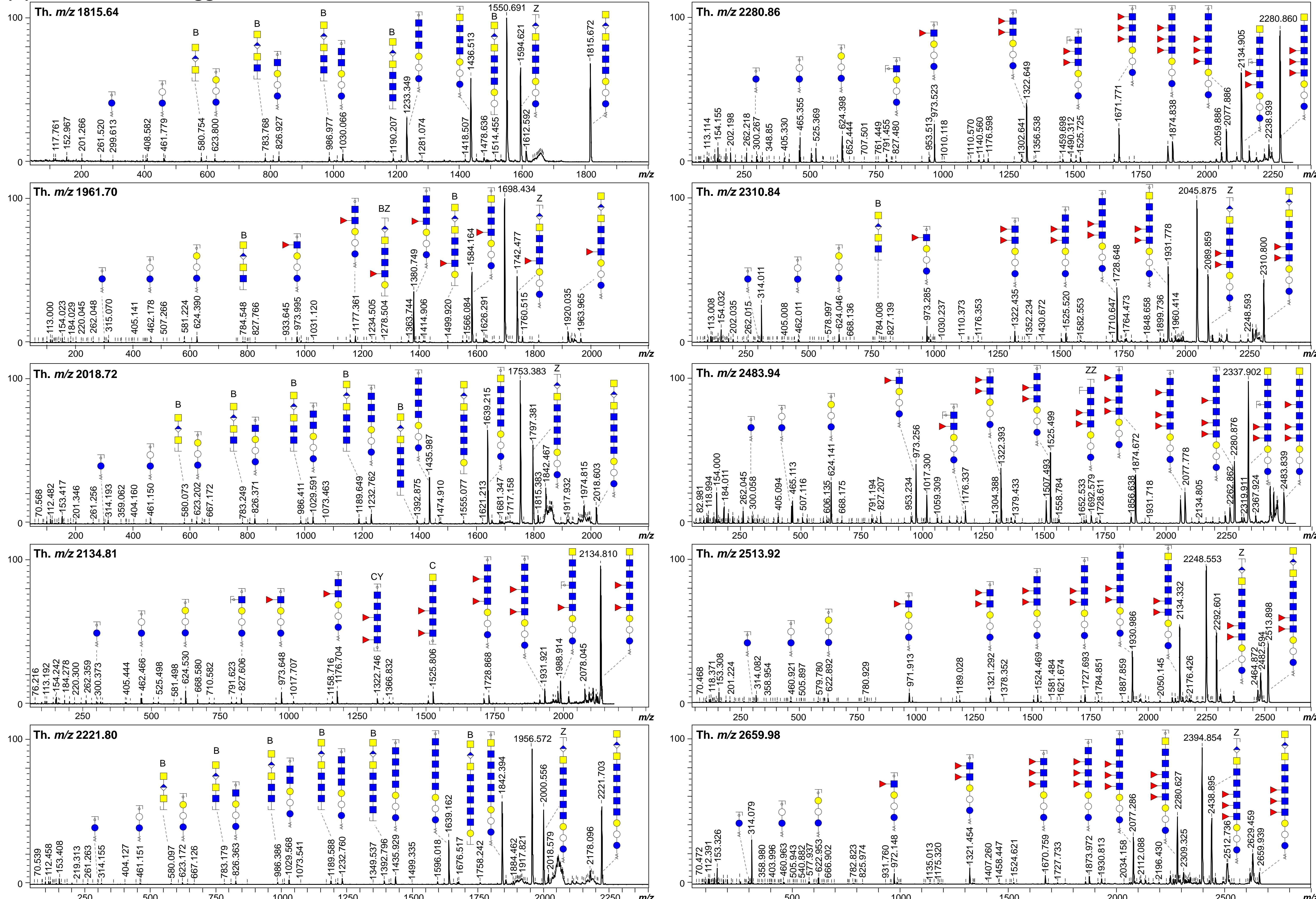
### (A) *S. haematobium* cercariae (continued)



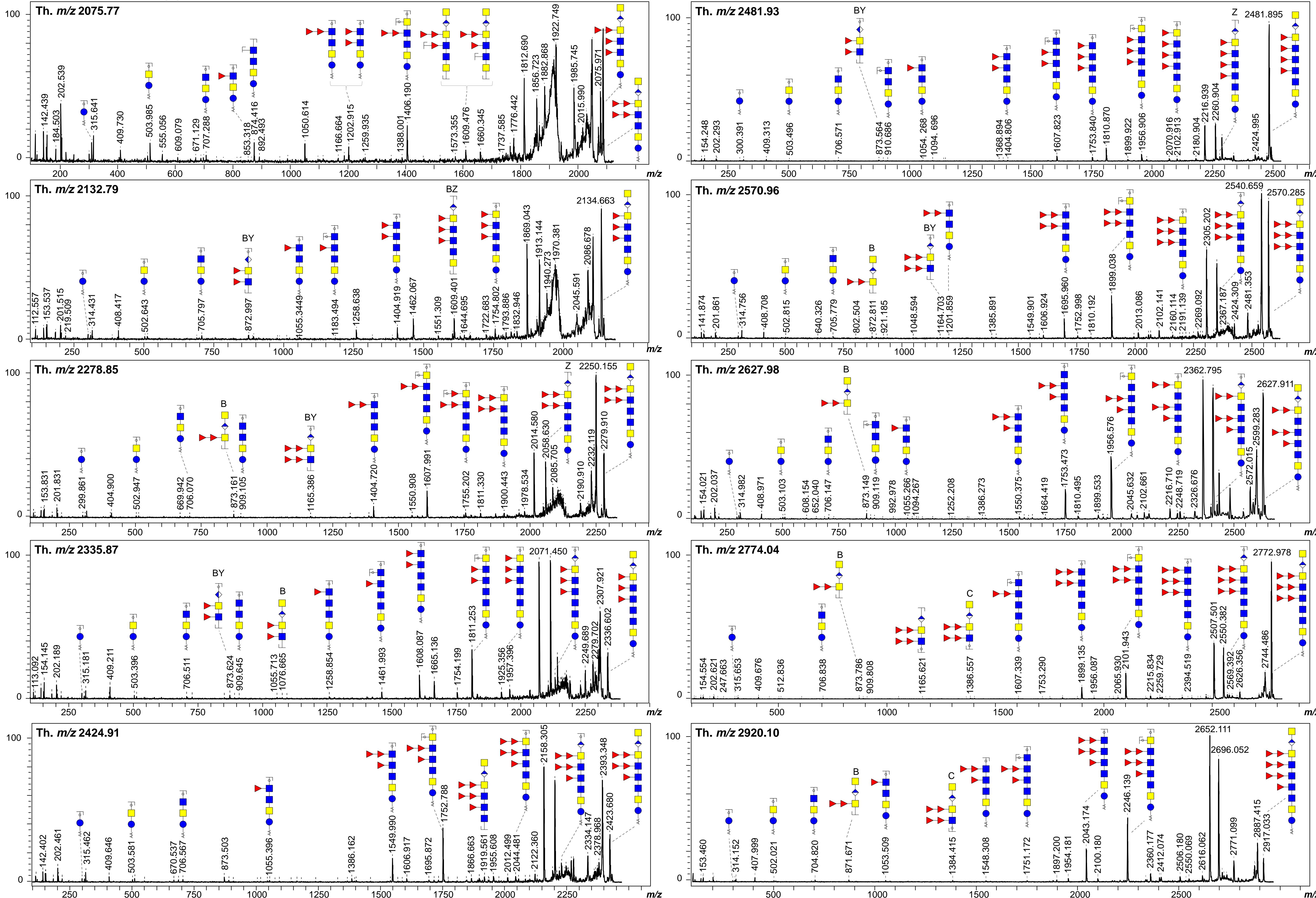
### (B) *S. haematobium* adult worms



**(C) *S. haematobium* eggs**



# (D) *S. mansoni* eggs



## Figure S4 – 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 M&M. 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-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 exoglycosidase(s) sequentially. Extract ion chromatogram of  $m/z$  505.18 is shown for the different conditions specified in the upper left corners. Treatment effects on this ion species can be observed and are highlighted in italic font.

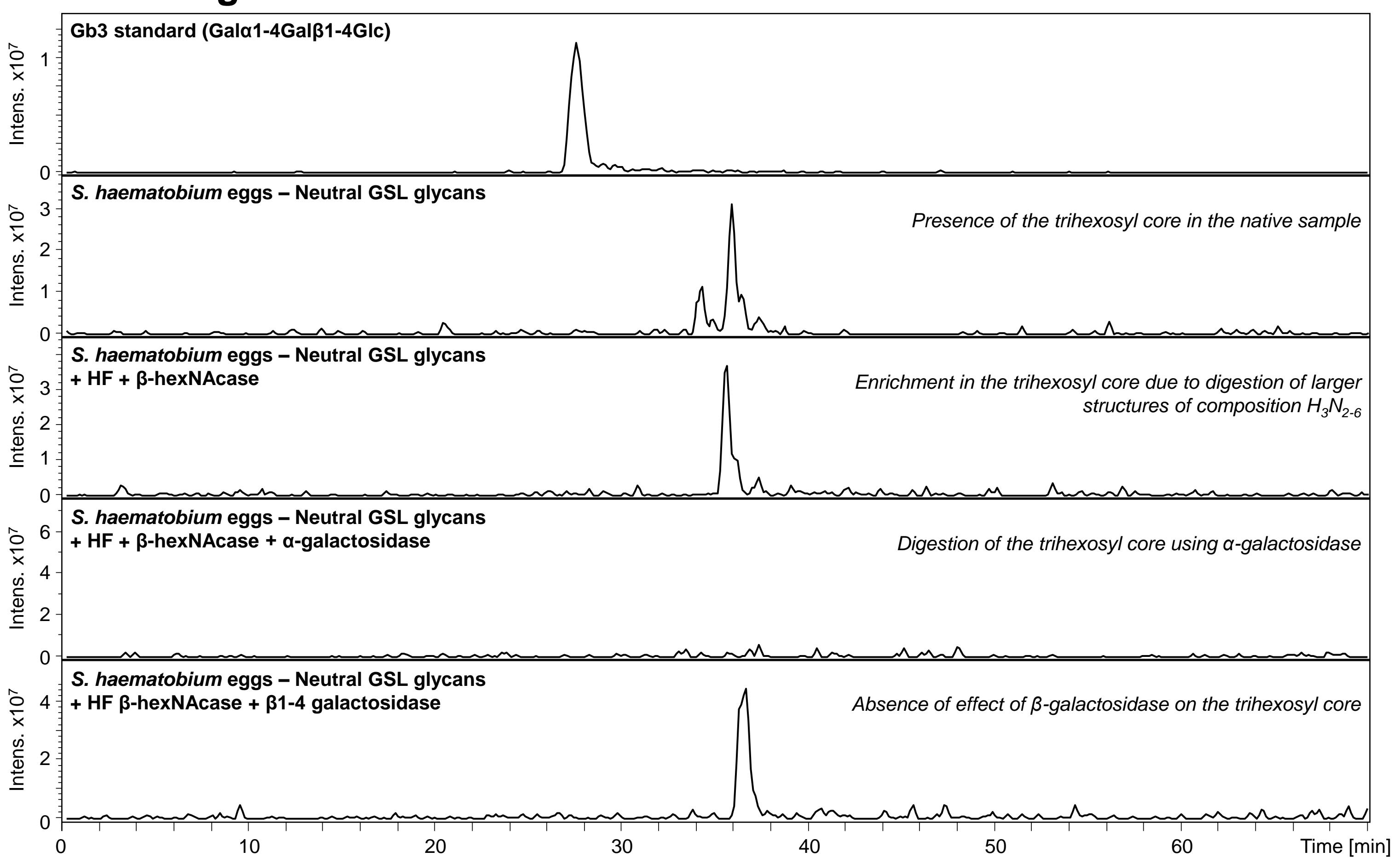
**PGC-LC-MS/MS analysis of selected GSL glycans (B-F)** of *S. haematobium* worms (B), eggs (C-D) and of *S. mansoni* (E-F) eggs. Monoisotopic mass (Monoiso.  $m/z$ ), charge state (Charge), theoretical (Th.) and observed (Exp.) ion  $m/z$ , the mass deviation between the theoretical and experimental  $m/z$  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 ( $m/z$ ). A summary of all ion species analyzed is provided in next page's table.

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>).

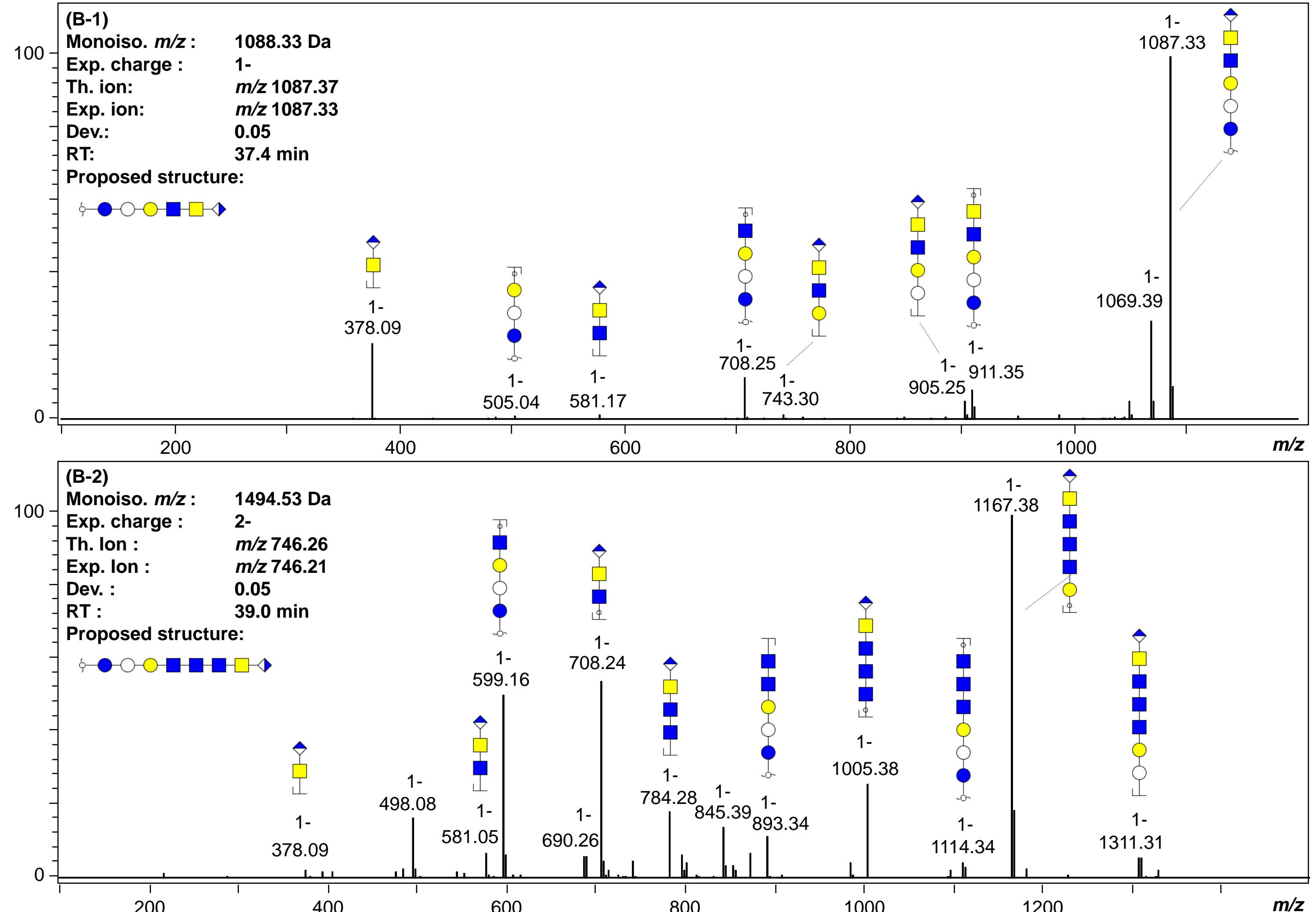
Monosaccharide nomenclature	Fragment ion nomenclature
● Galactose (Gal)	□ B-type
■ N-acetylgalactosamine (GalNAc)	□ C-type
▲ Fucose (Fuc)	□ Y-type
○ Hexose (Hex)	□ Z-type
● Glucose (Glc)	□ 2,4A-type (cross-ring fragment)
■ N-acetylglucosamine (GlcNAc)	□ 0,2A-type (cross-ring fragment)
◆ Glucuronic acid (GlcA)	

Species Life-stage Acidic/Neutral	Glycan composition	Th. m/z MALDI	Th. m/z PGC-LC-MS		Glycan ID	Panel
			[M-H] <sup>1-</sup>	[M-H] <sup>2-</sup>		
<i>S. haematobium</i> Eggs Neutral	H3	624.21	505.18	n.d. (not detected)	-	A
<i>S. haematobium</i> Worms Acidic	H3N2A1	1206.41	1087.37	n.d.	B1	B
	H3N4A1	1612.56	1493.53	746.26	B2	B
<i>S. haematobium</i> 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
<i>S. haematobium</i> Eggs Neutral	H3N2	1030.35	911.34	455.16	D1	D
	H3N2F1	1176.40	1057.39	528.19	D2a-b	D
	H3N4	1436.51	1317.49	658.24	D3	D
	H3N4F1	1582.60	1463.55	731.27	D4a-b	D
	H5N3F2	1849.64	1730.64	864.81	D5	D
	H3N5F2	1931.66	1812.69	905.84	D6	D
<i>S. mansoni</i> 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
<i>S. mansoni</i> Eggs Acidic	H1N5A1F4	2075.77	1956.73	977.86	F1a-b	F
	H1N6A1F3	2132.79	2013.75	1006.37	F2	F
	H1N6A1F4	2278.85	2159.81	1079.40	F3a-b	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	F
	H1N7A1F6	2774.04	2655.01	1327.00	F8	F

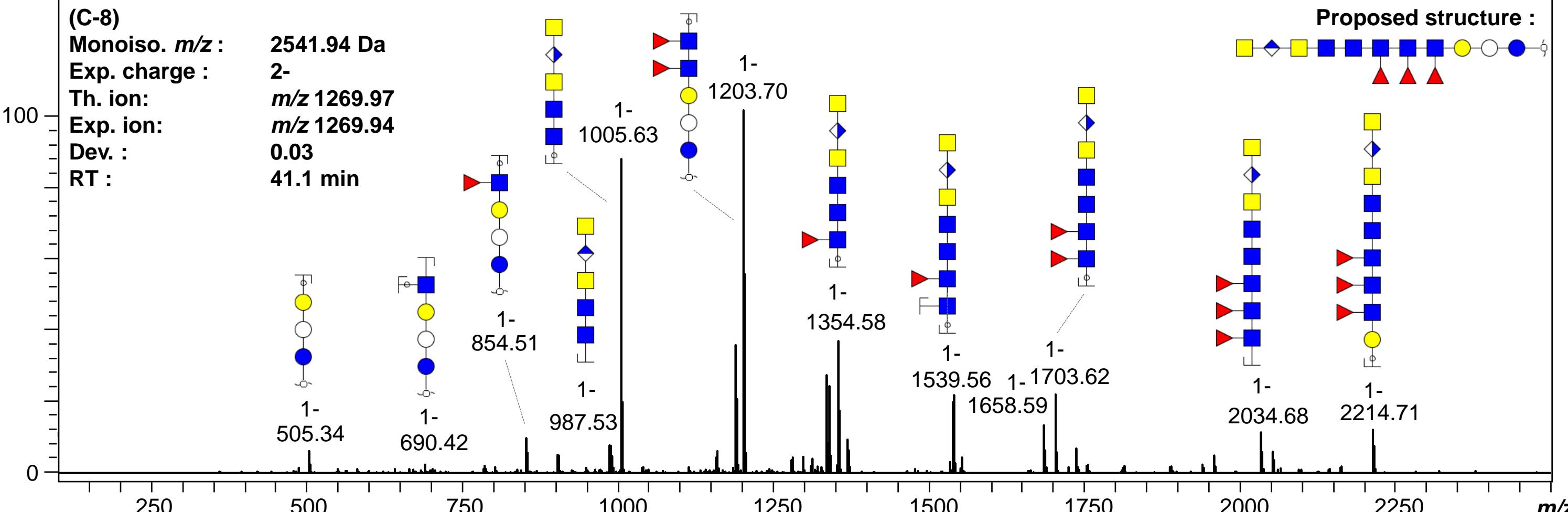
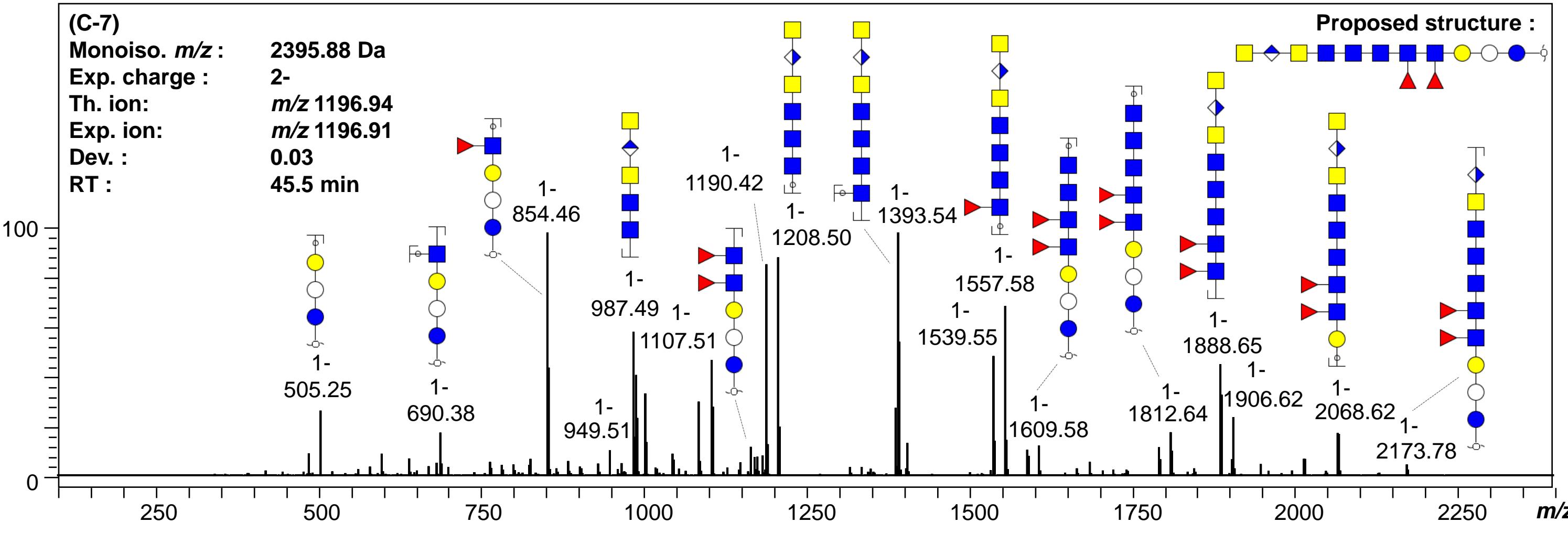
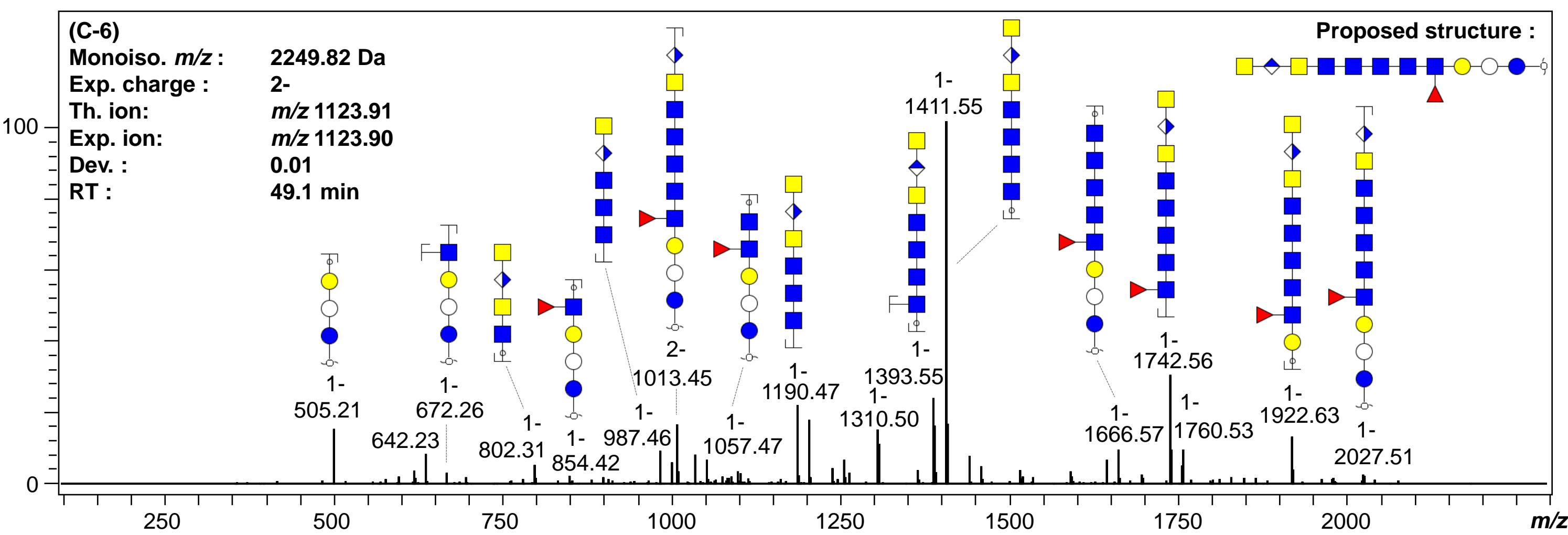
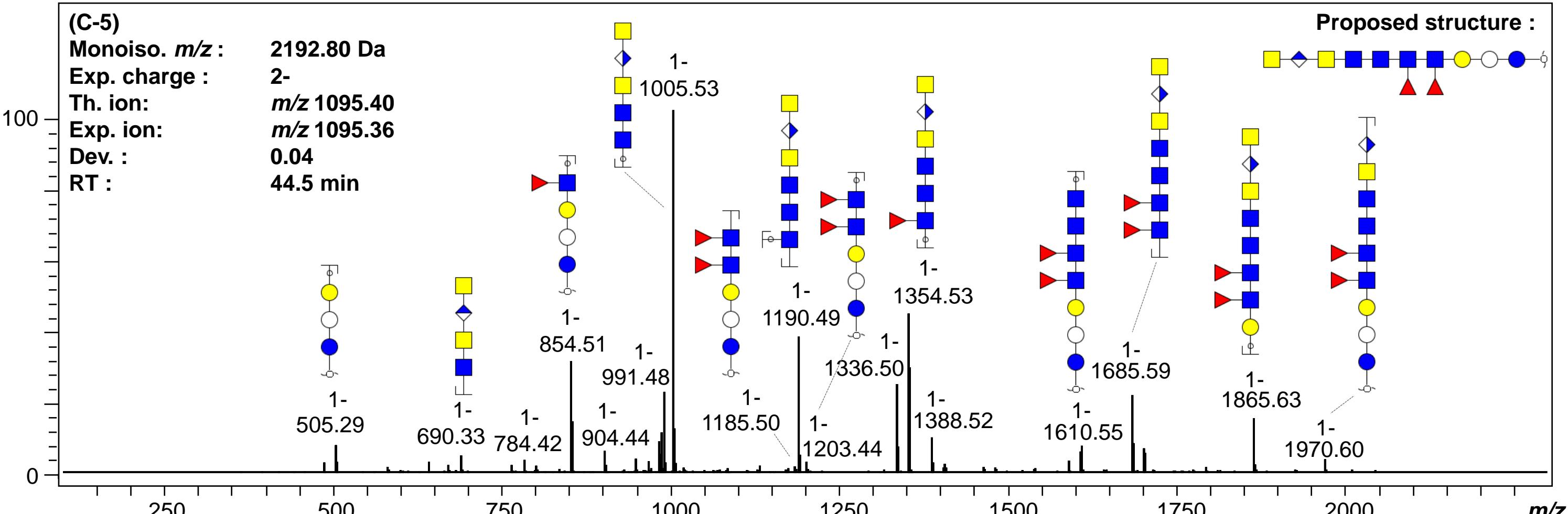
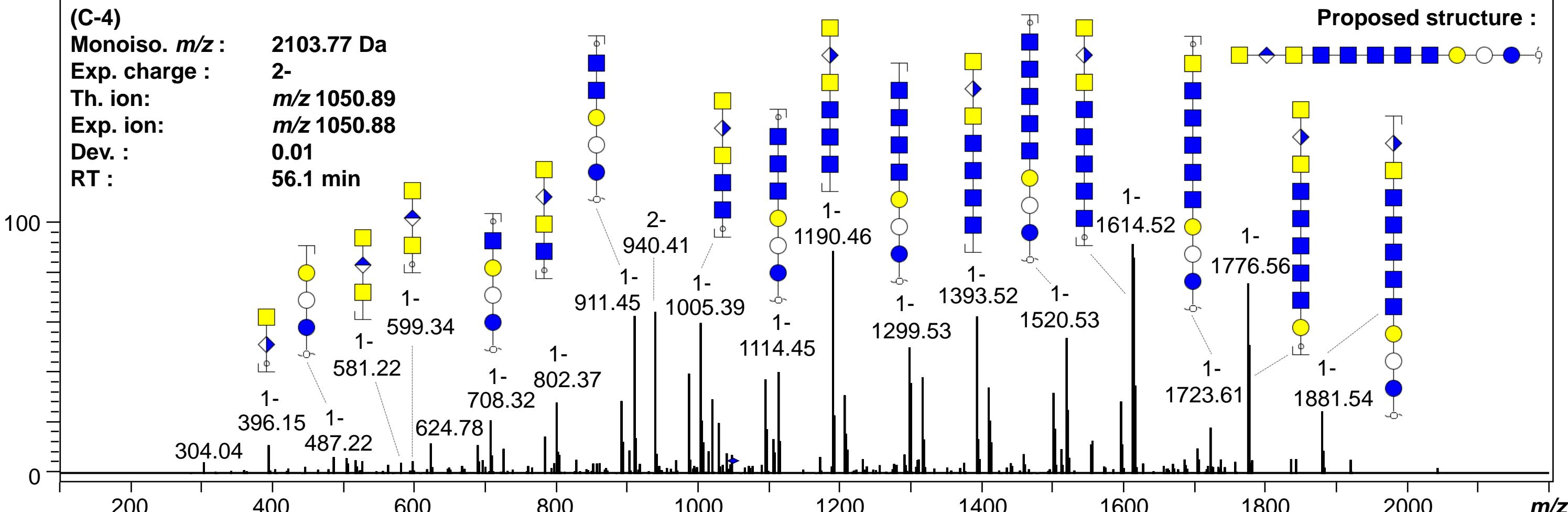
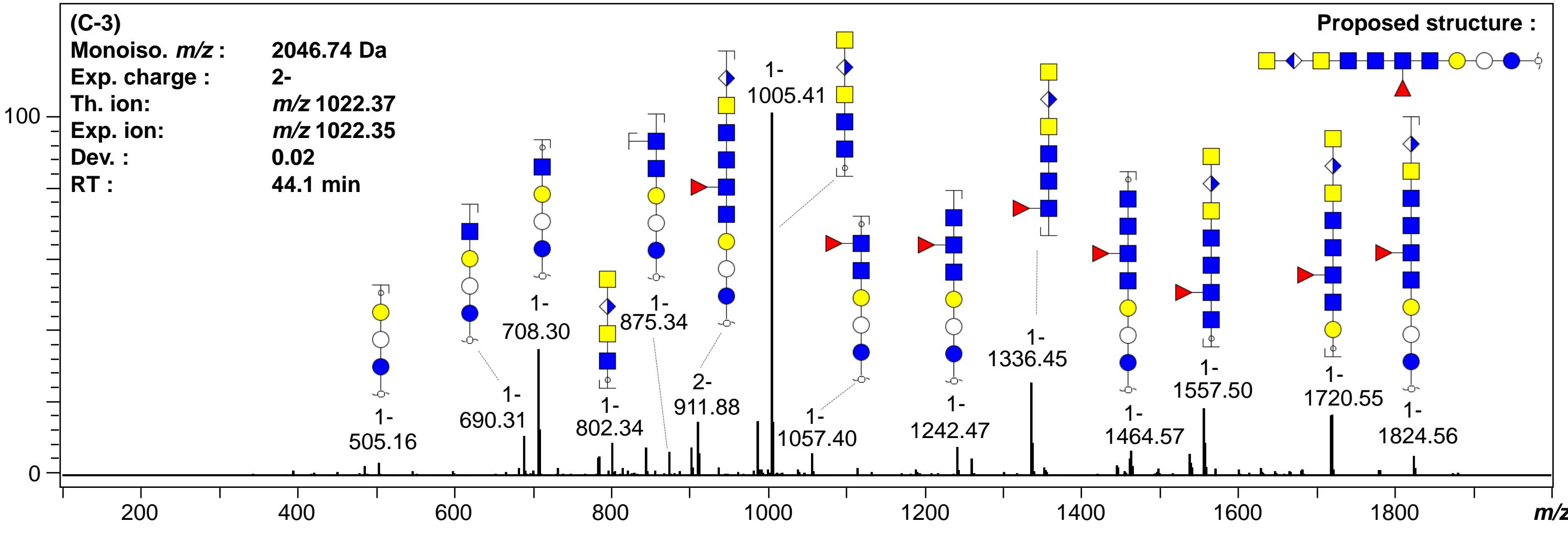
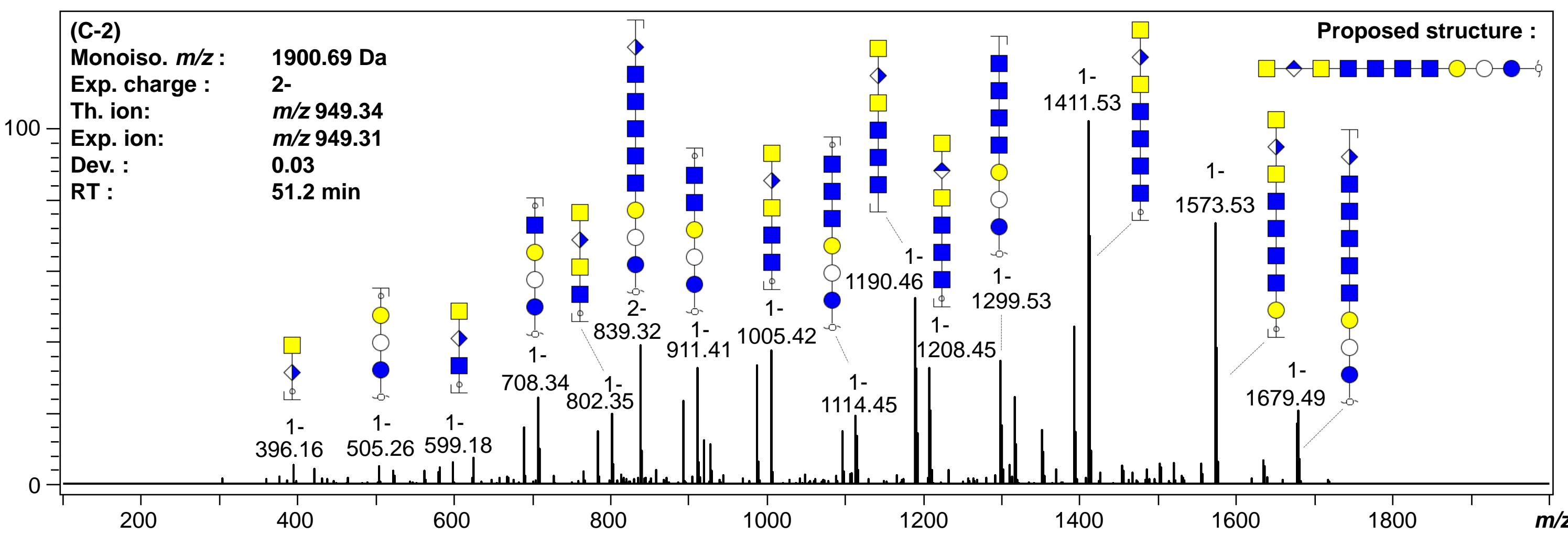
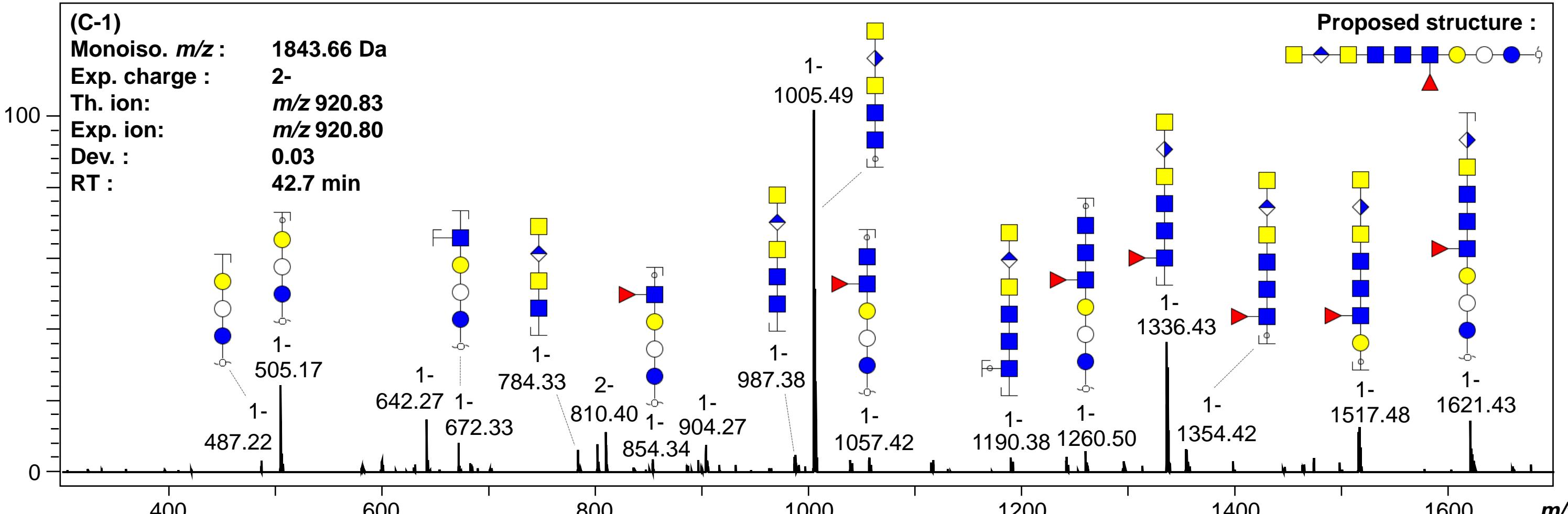
### (A) *S. haematobium* egg neutral GSL glycans - Extract ion chromatogram of m/z 505.18



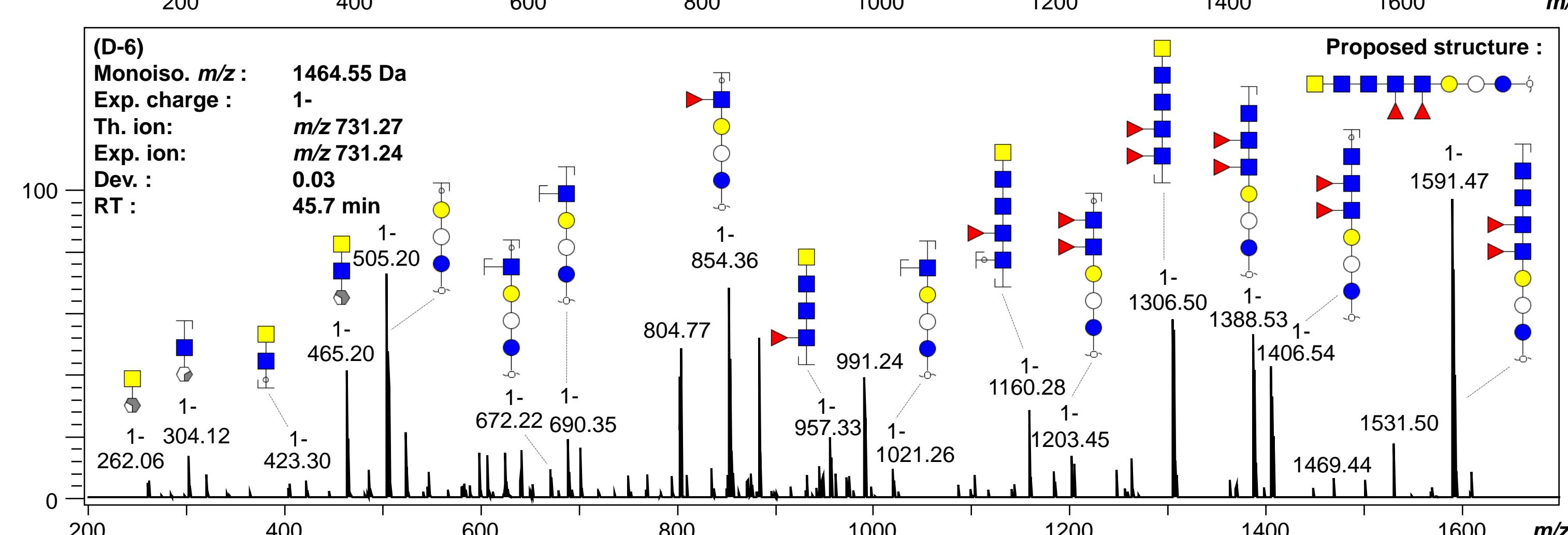
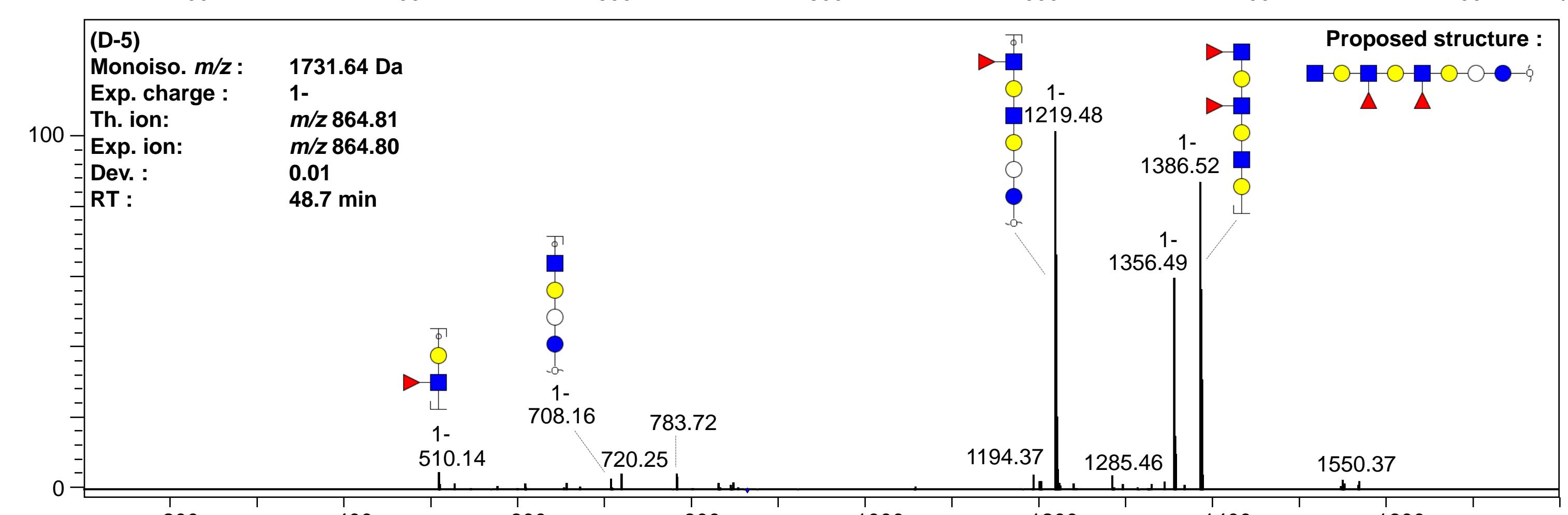
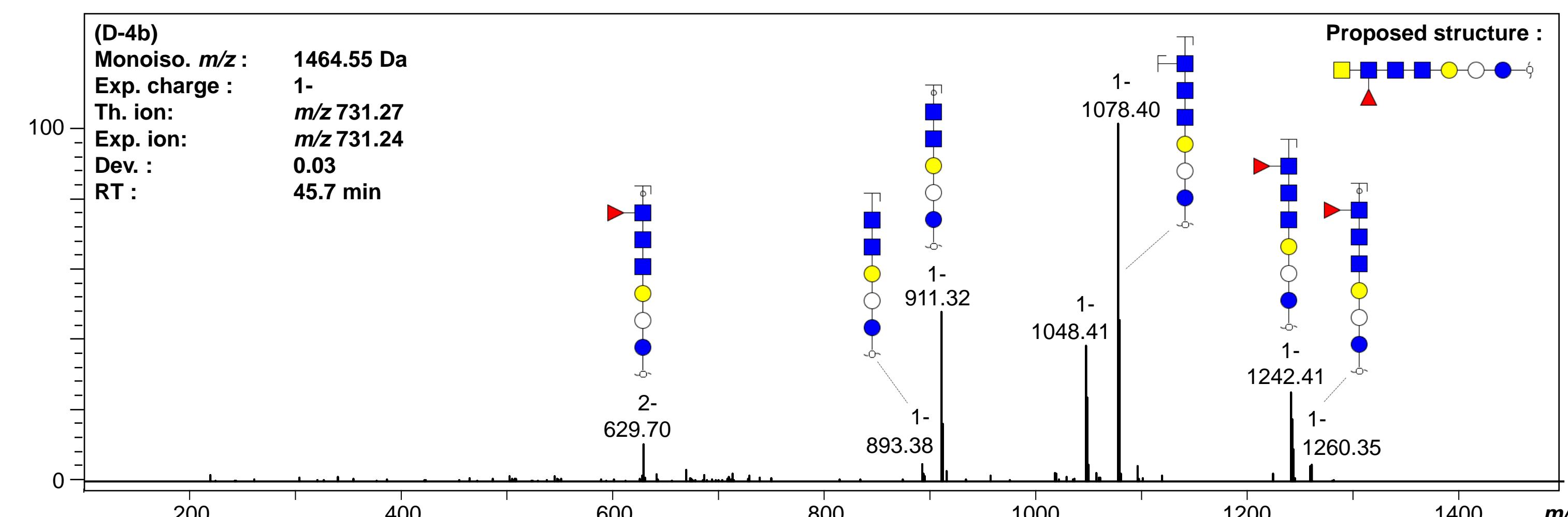
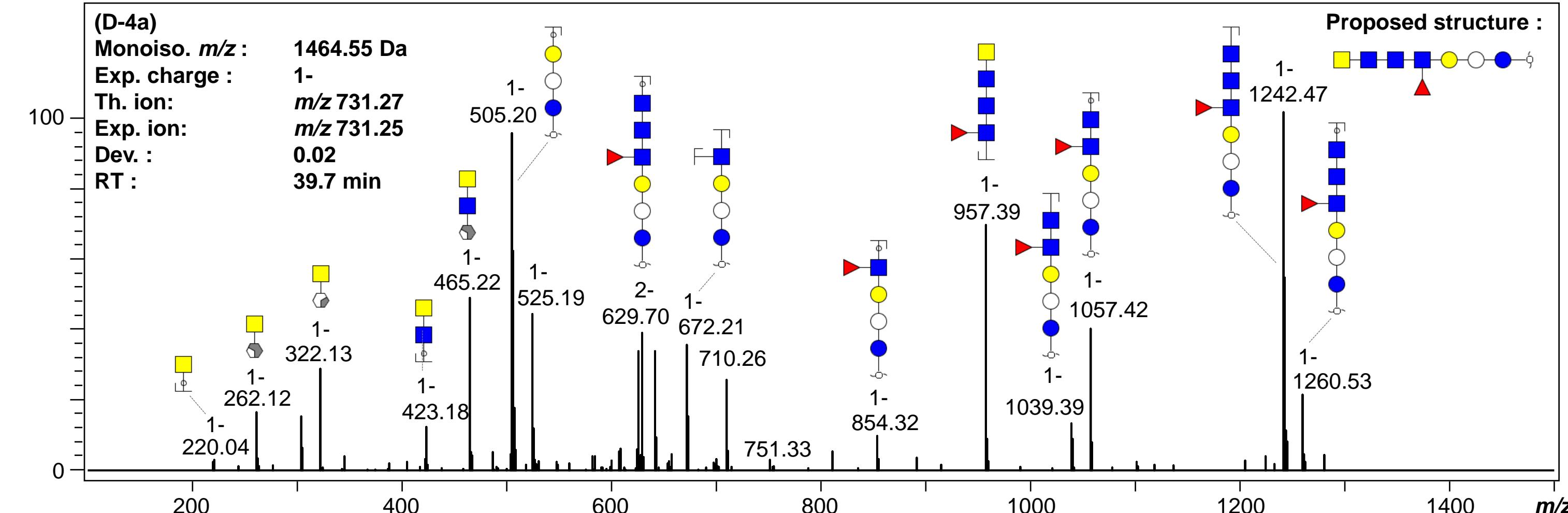
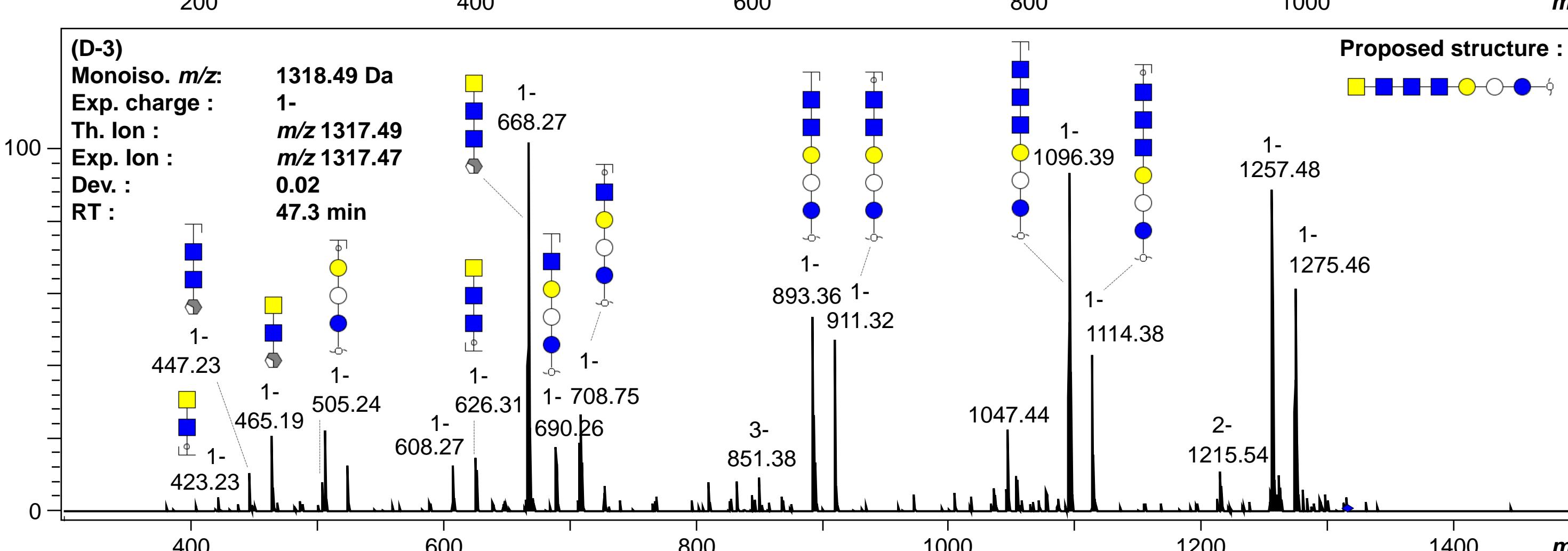
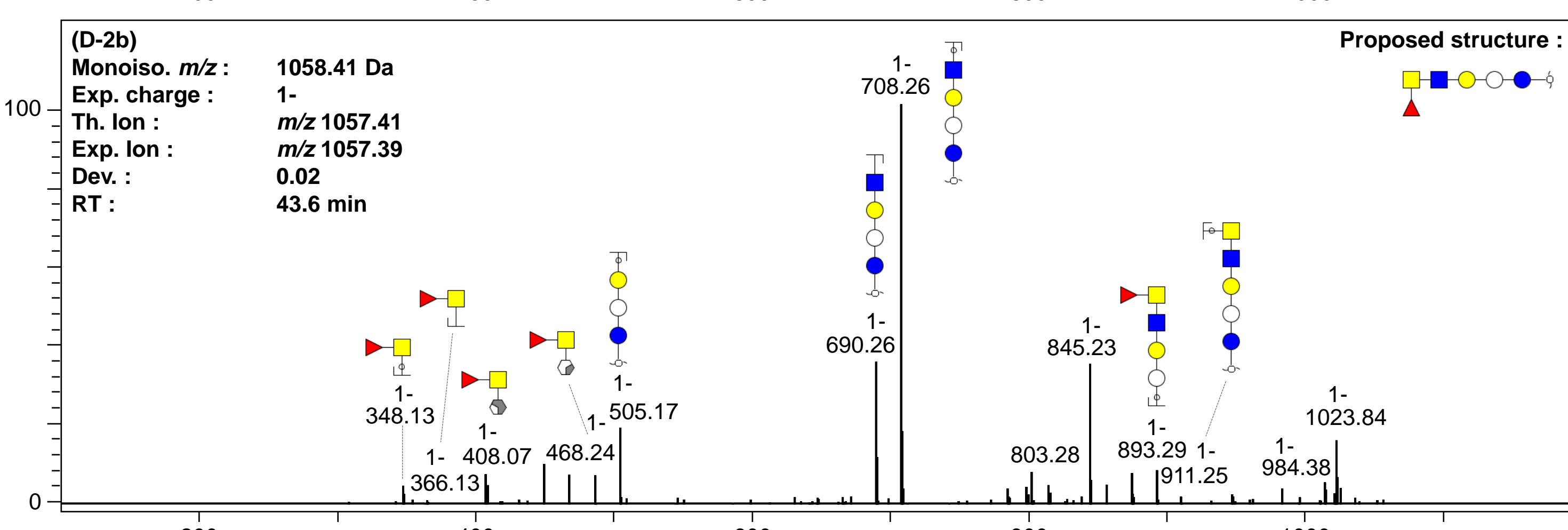
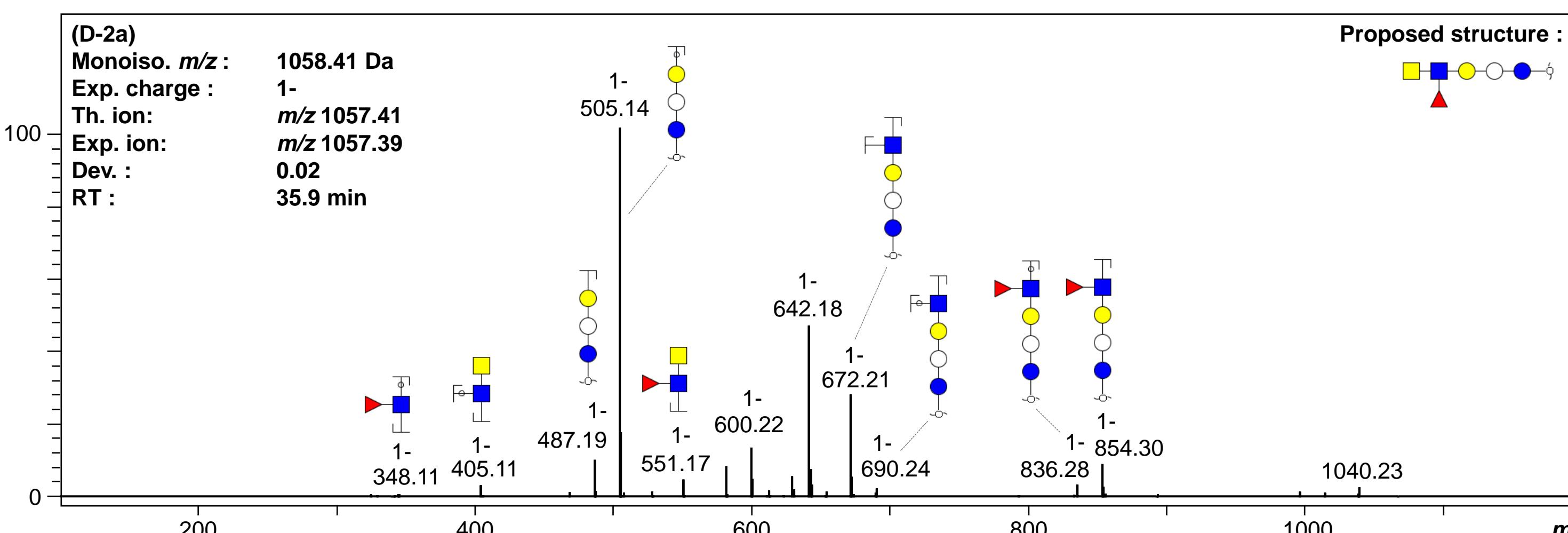
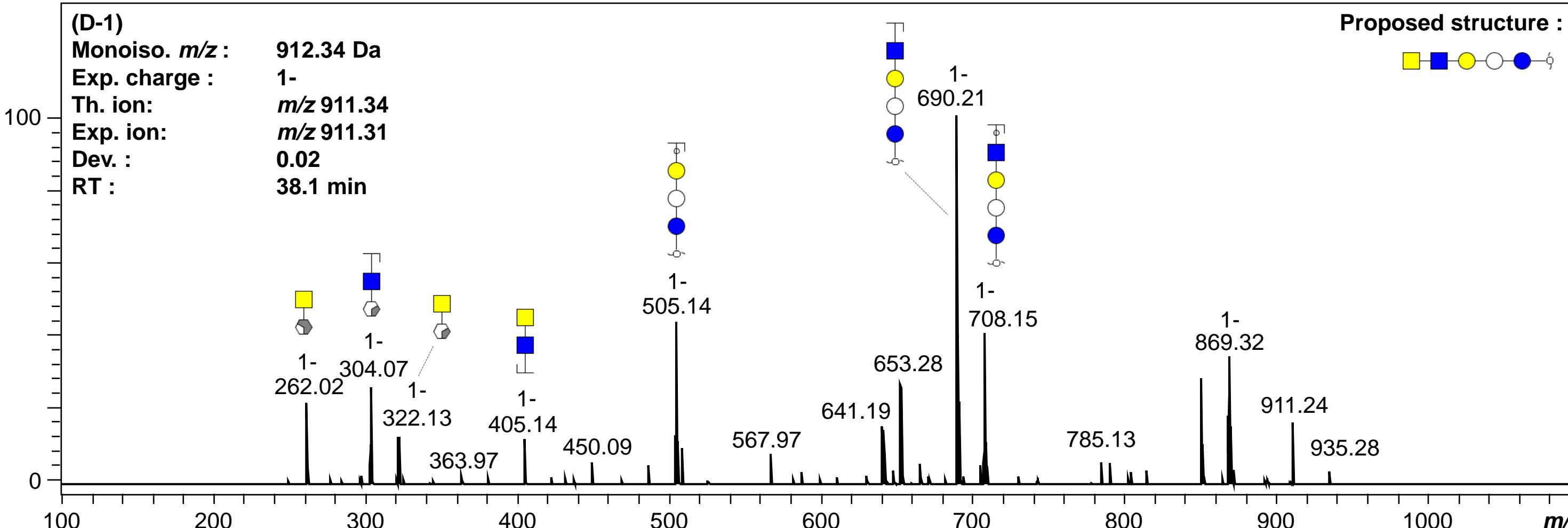
### (B) *S. haematobium* worm acidic GSL glycans



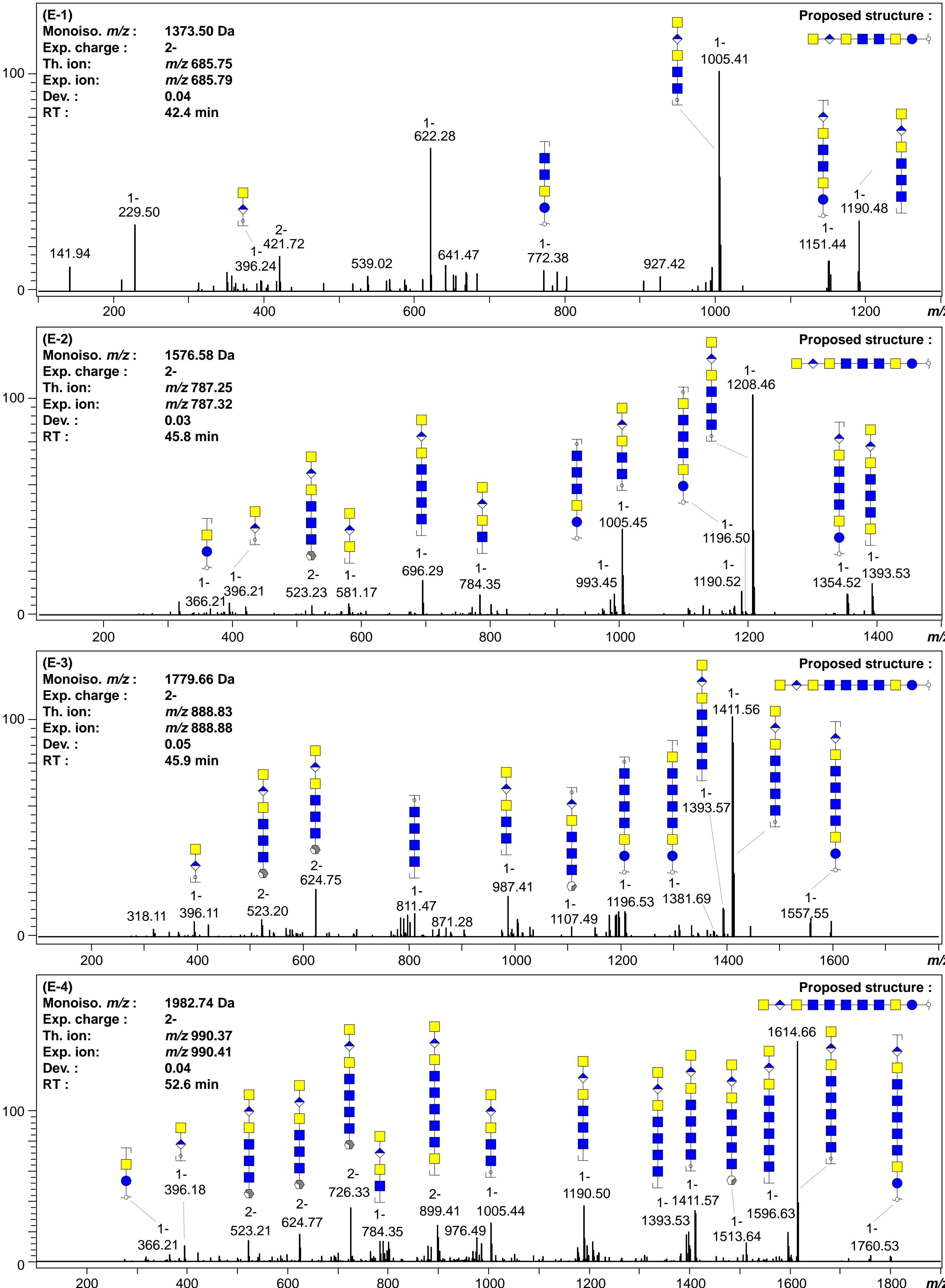
### (C) *S. haematobium* egg acidic GSL glycans



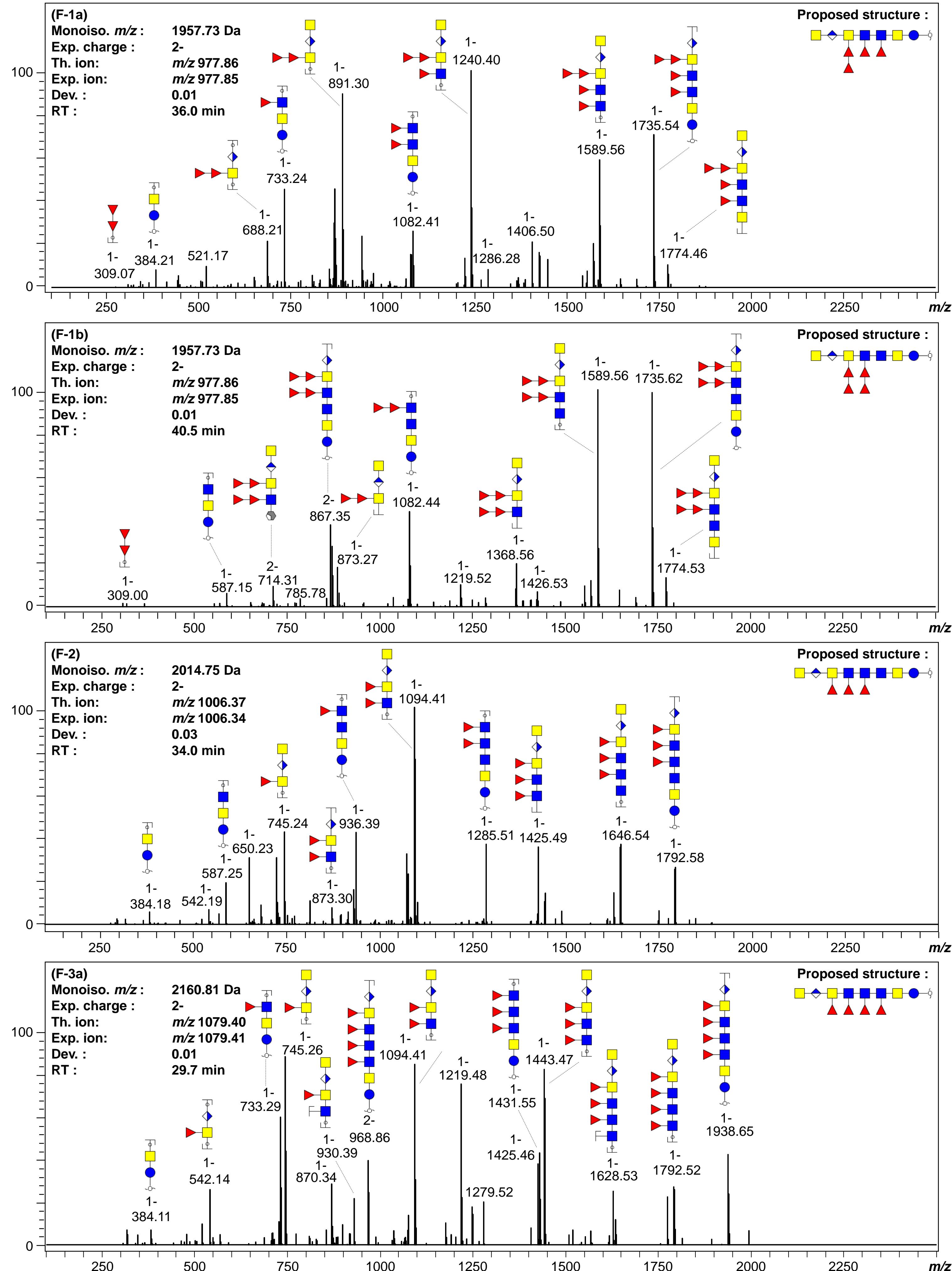
## (D) *S. haematobium* egg neutral GSL glycans



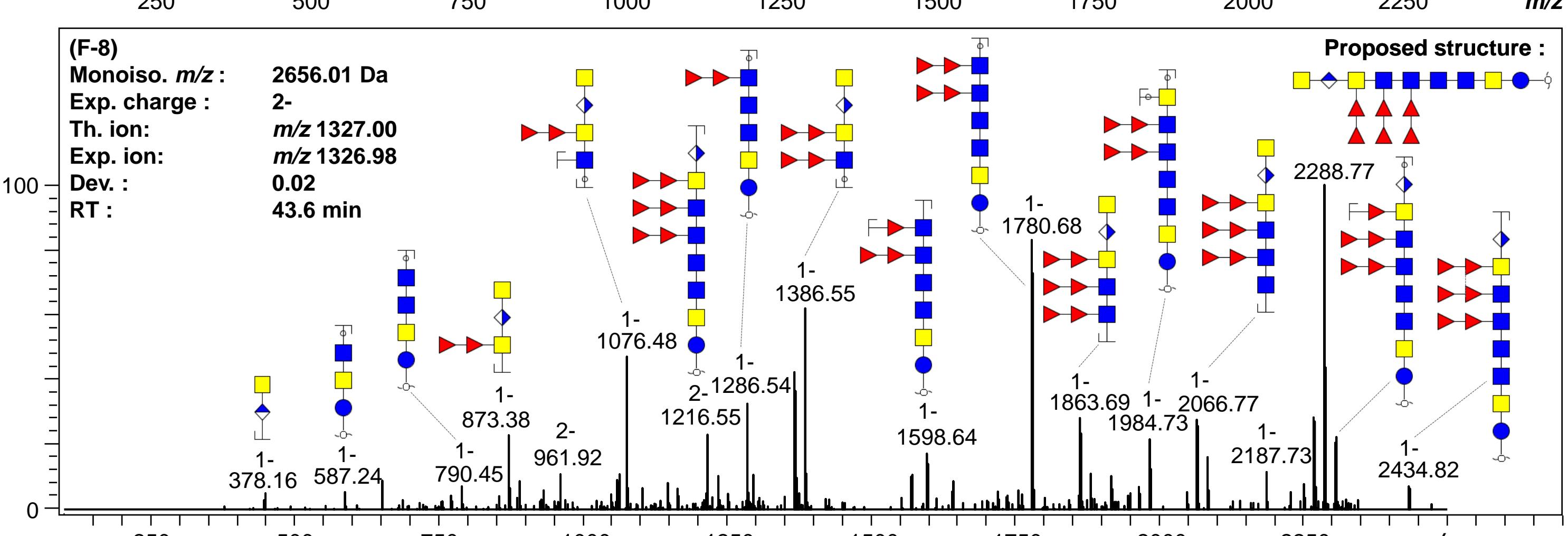
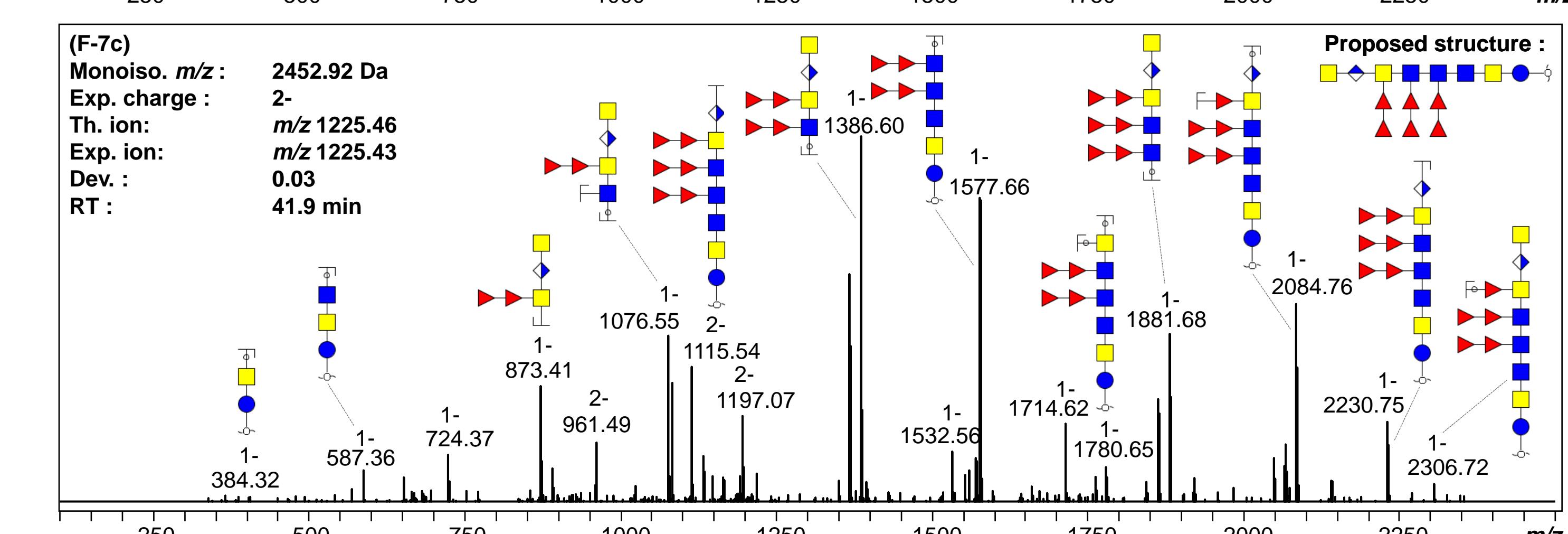
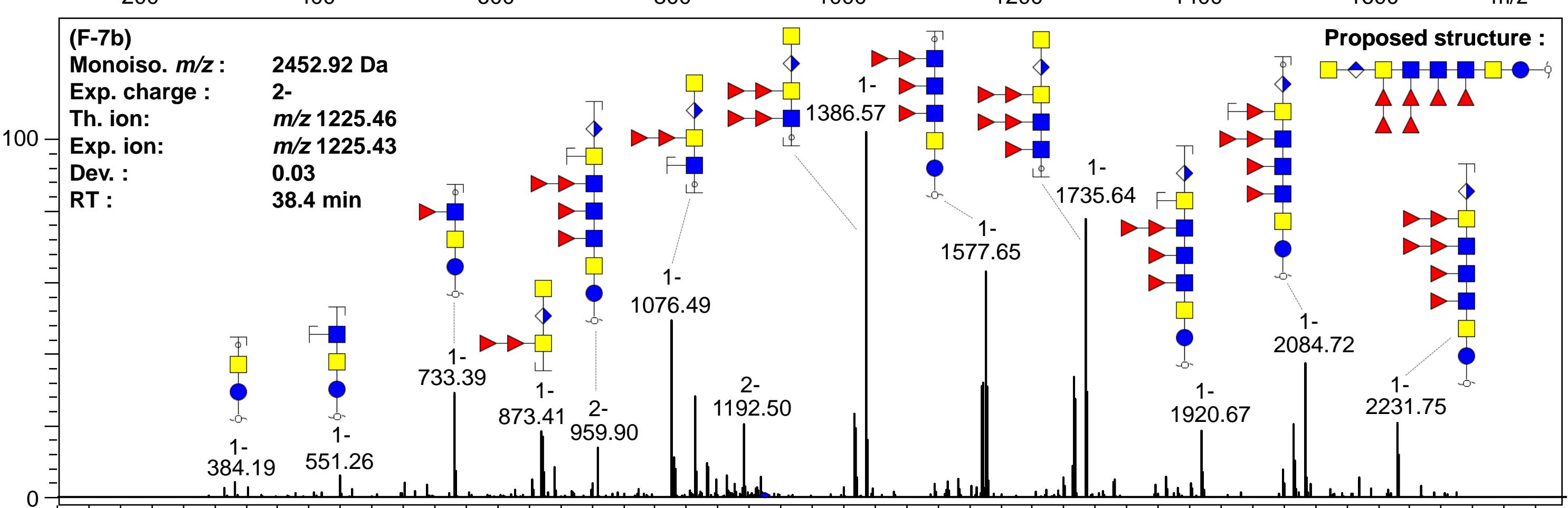
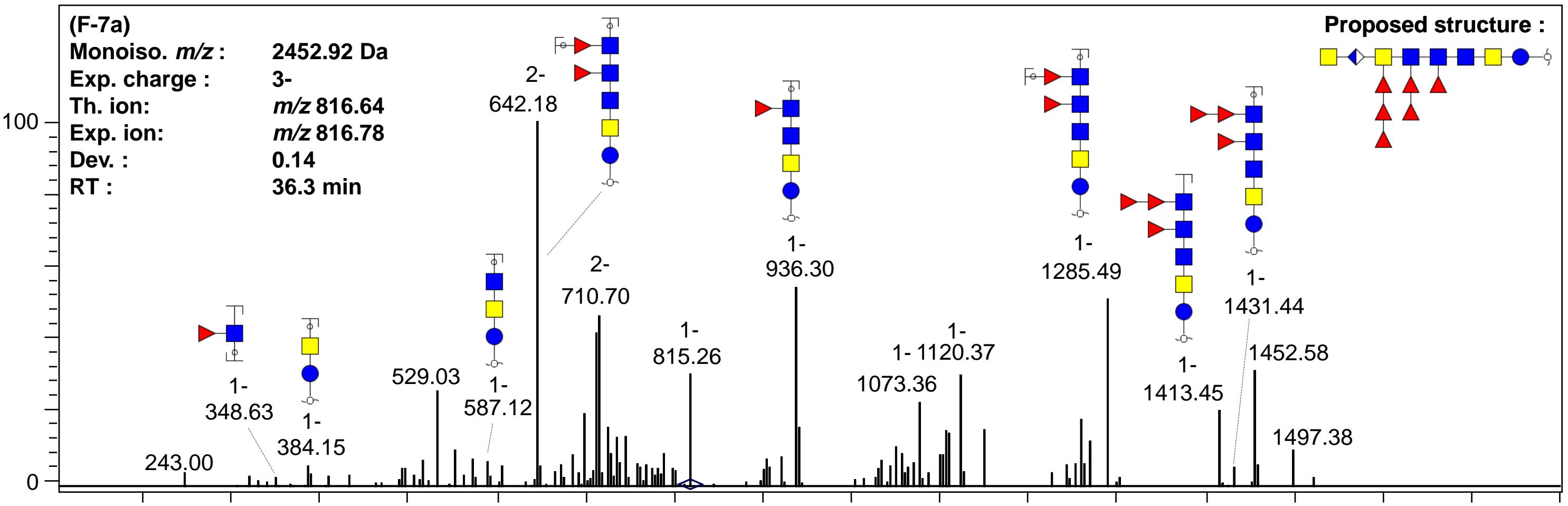
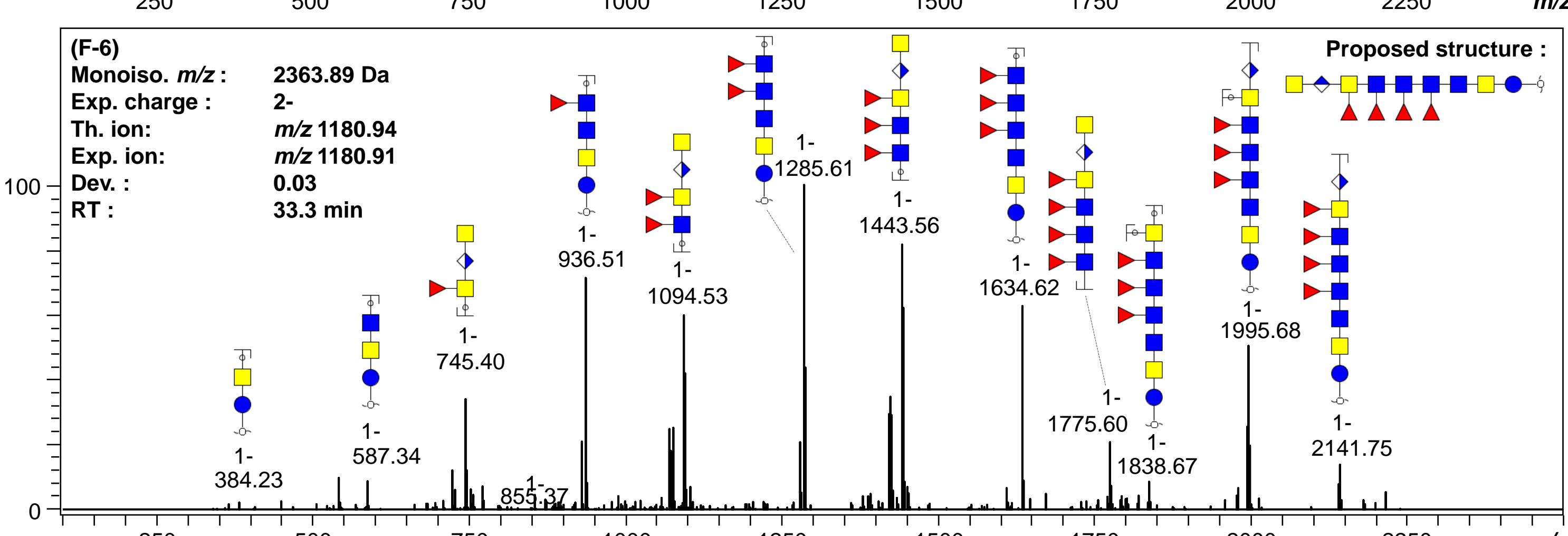
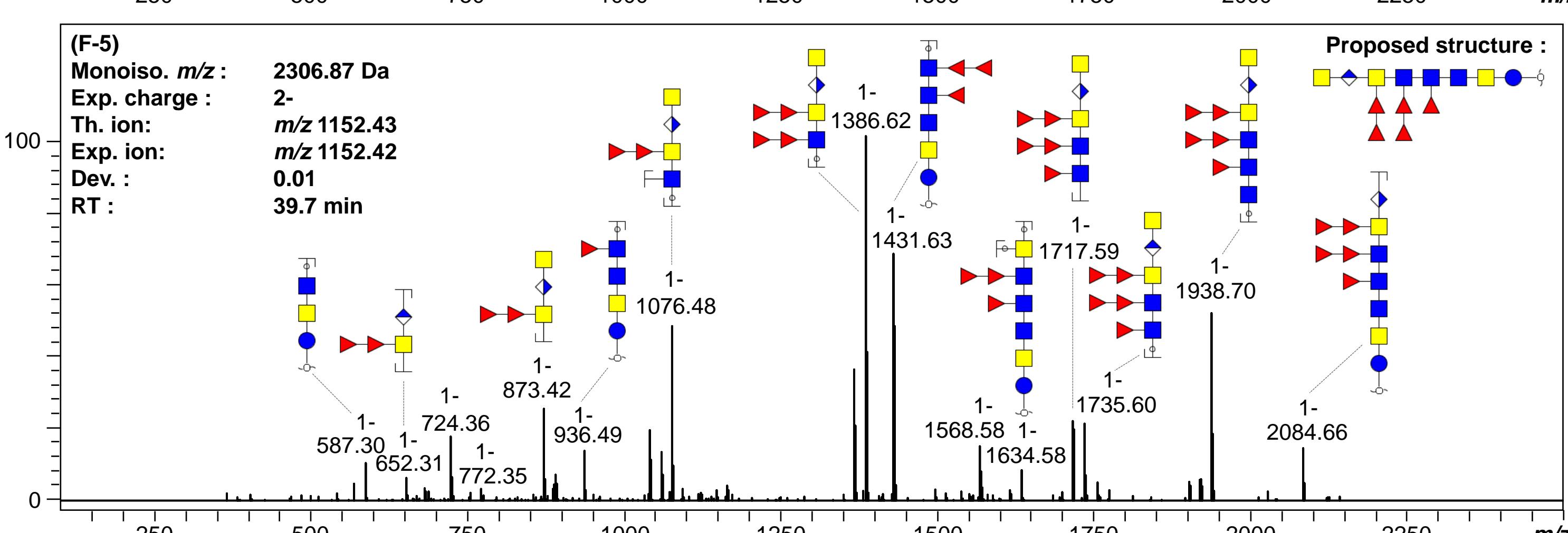
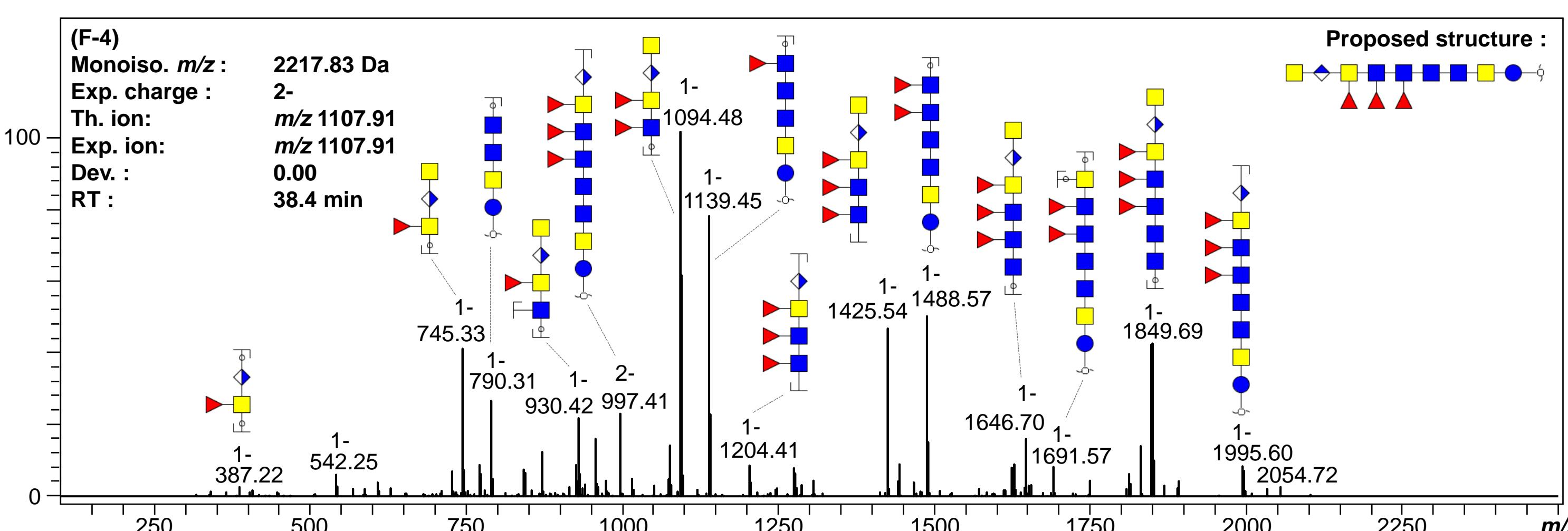
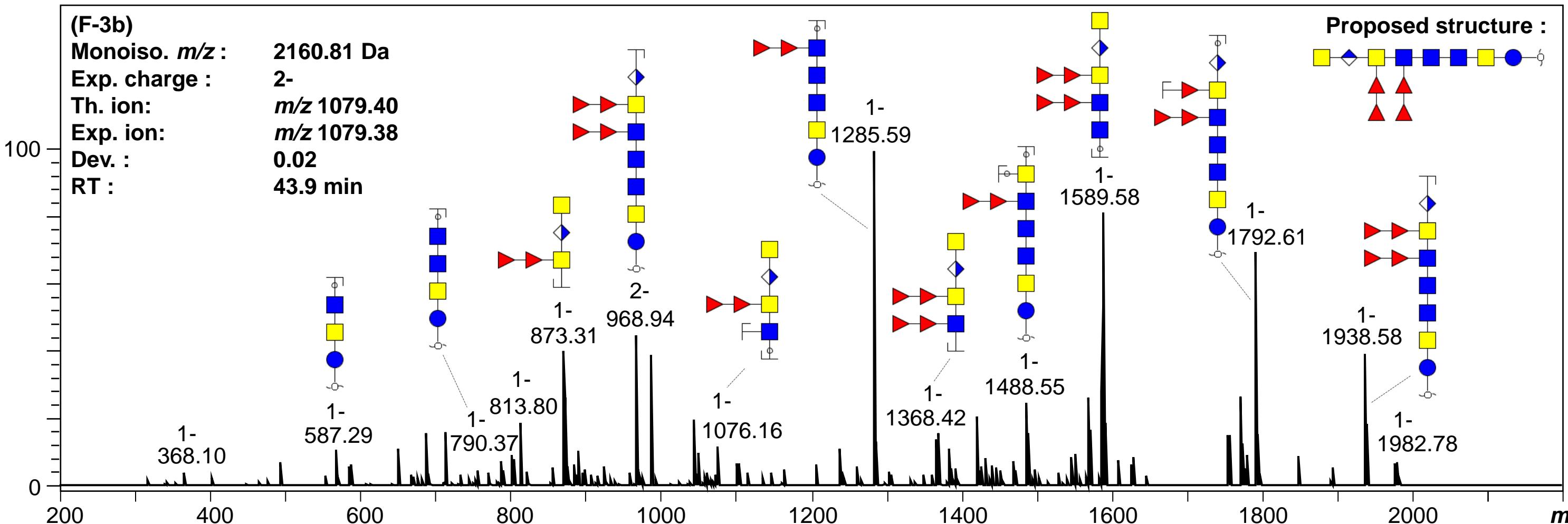
**(E) S. mansoni egg acidic GSL glycans - HF treated**



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



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



## Figure S5 – Structural characterization of *S. haematobium* N-glycans

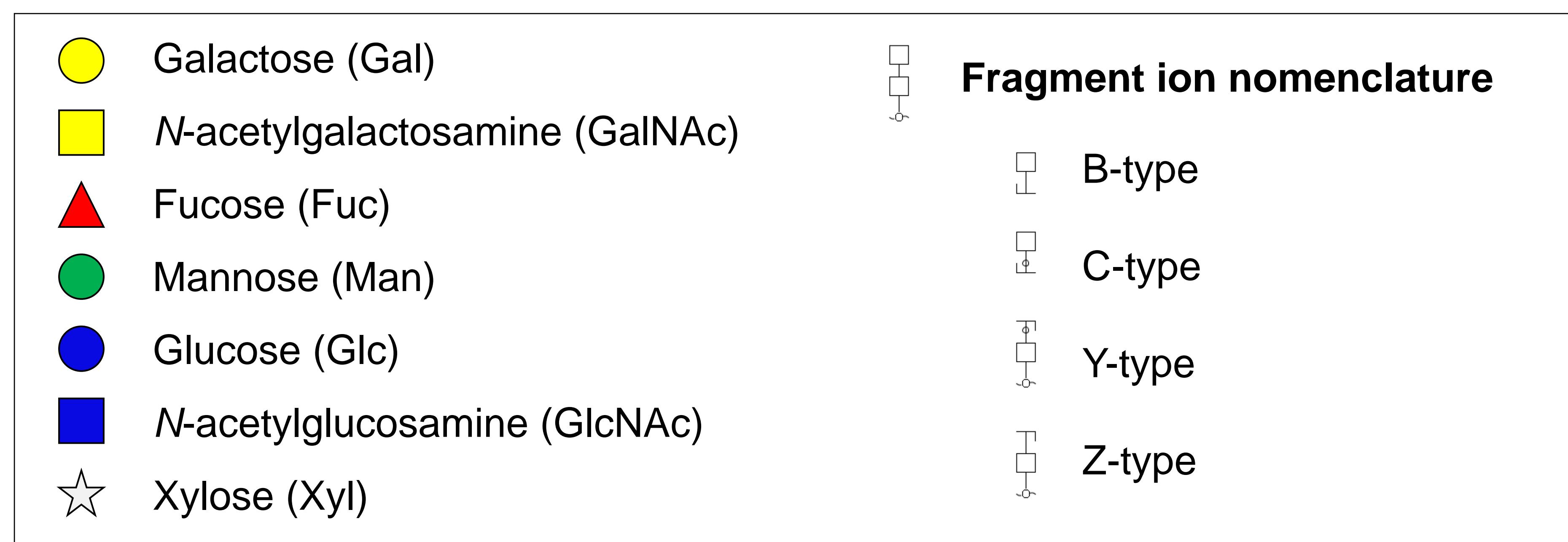
N-linked glycans were released from *S. haematobium* glycoproteins using PNGase F (A-D) and PNGase A sequentially (E-F) prior to AA labeling.

**Glycan sequencing (A-C, E)** N-glycans of *S. haematobium* cercariae (A), adult worms (B) and eggs (C, E) – 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 in the conditions detailed in **Table 1** (see M&M). 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.

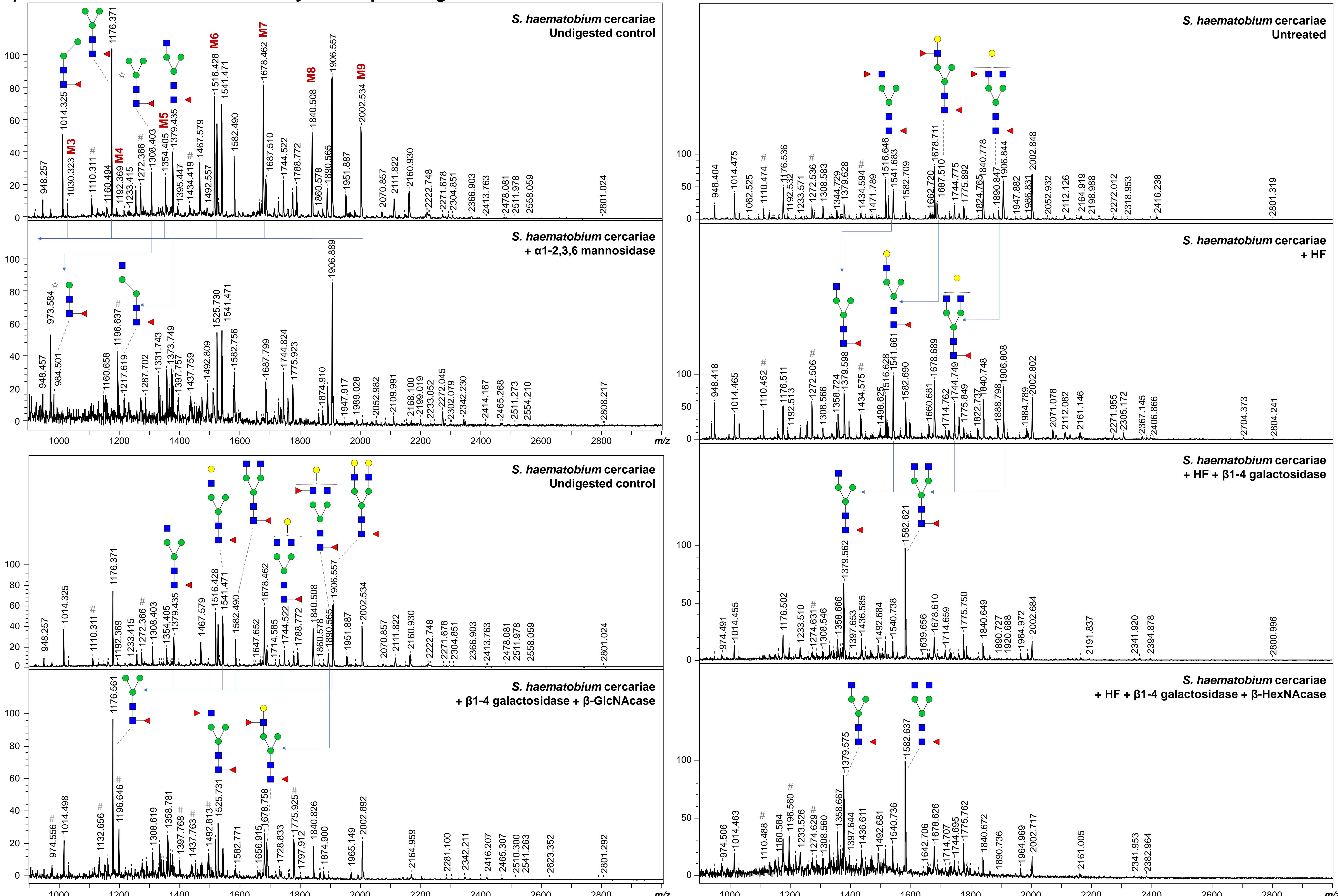
**MALDI-TOF-MS (D)** of N-glycans derived from immature and mature eggs of *S. haematobium*, separated by Percoll gradient centrifugation (see M&M).

**MALDI-TOF-MS/MS (F)** 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).

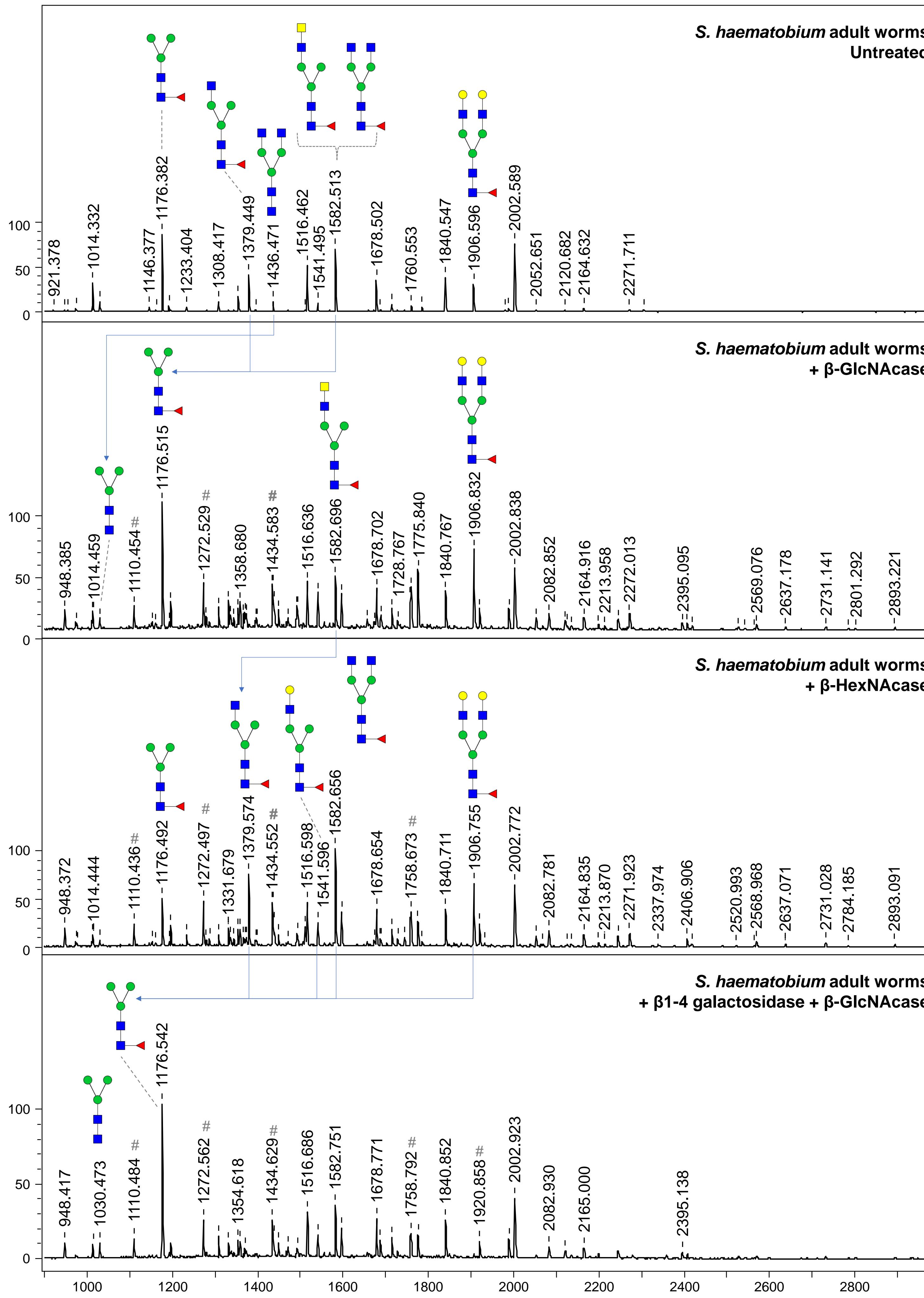
All (putative) glycans are represented using the CFG nomenclature (see inset below).



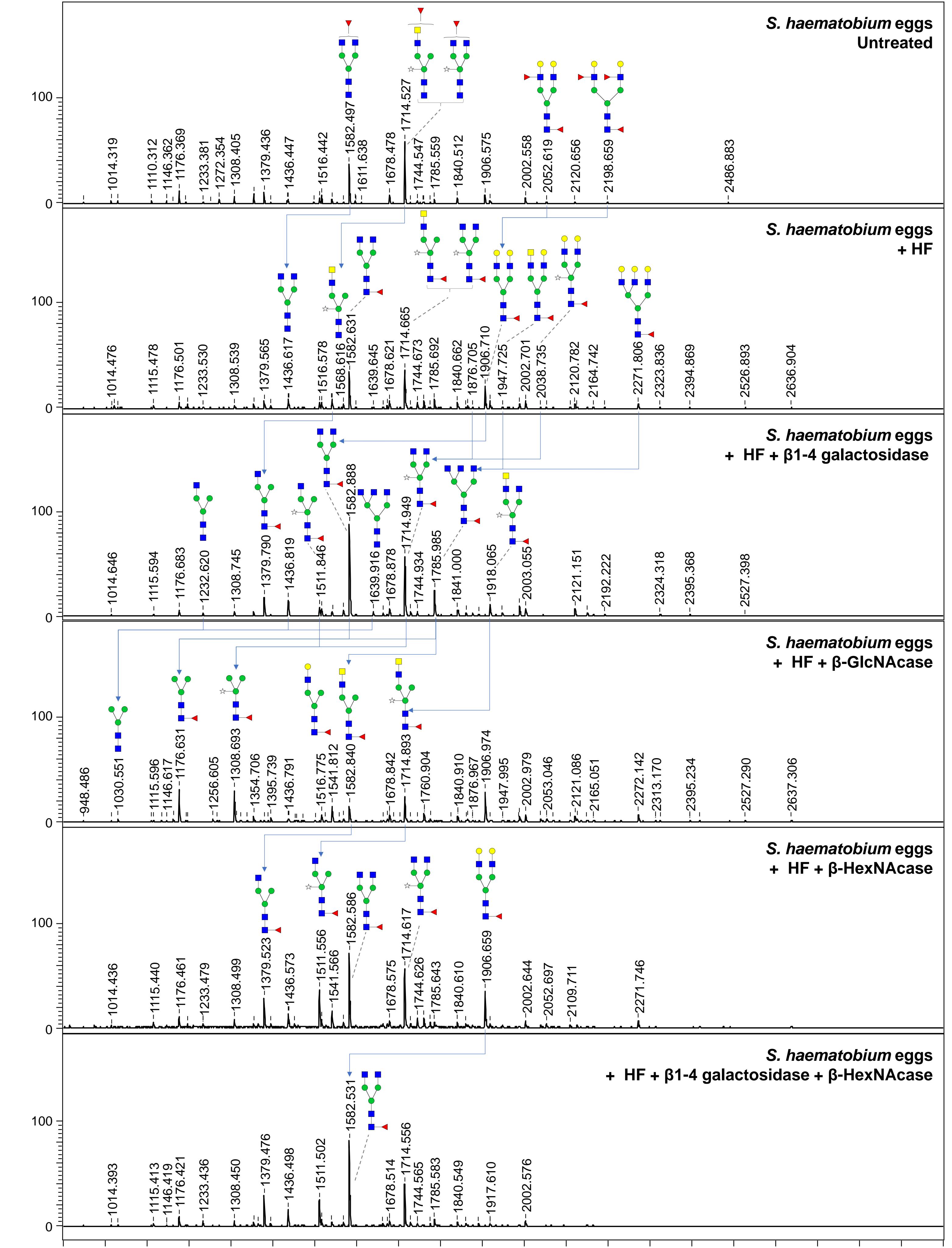
### (A) *S. haematobium* cercariae – Glycan sequencing



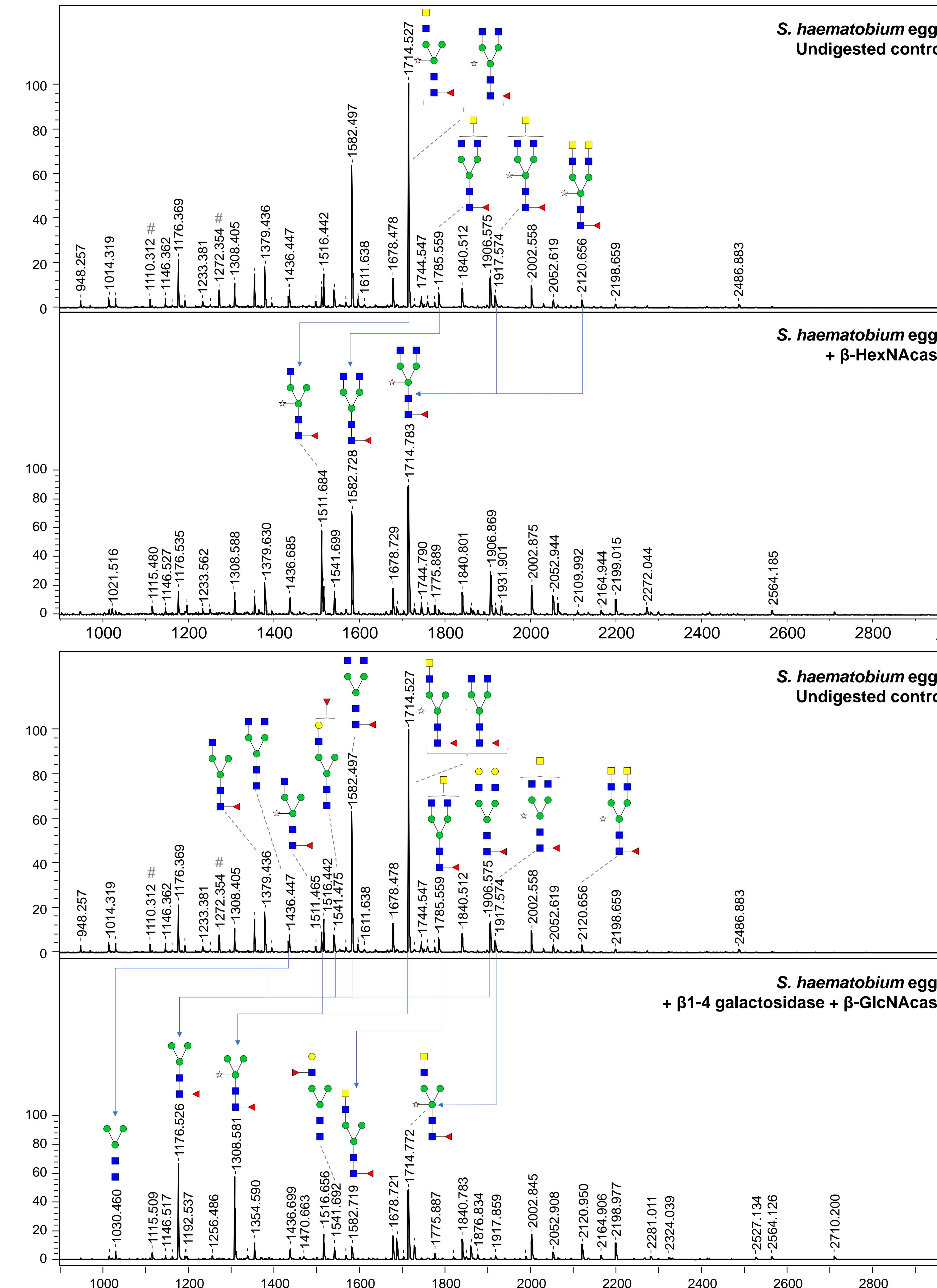
## (B) *S. haematobium* adult worms – Glycan sequencing



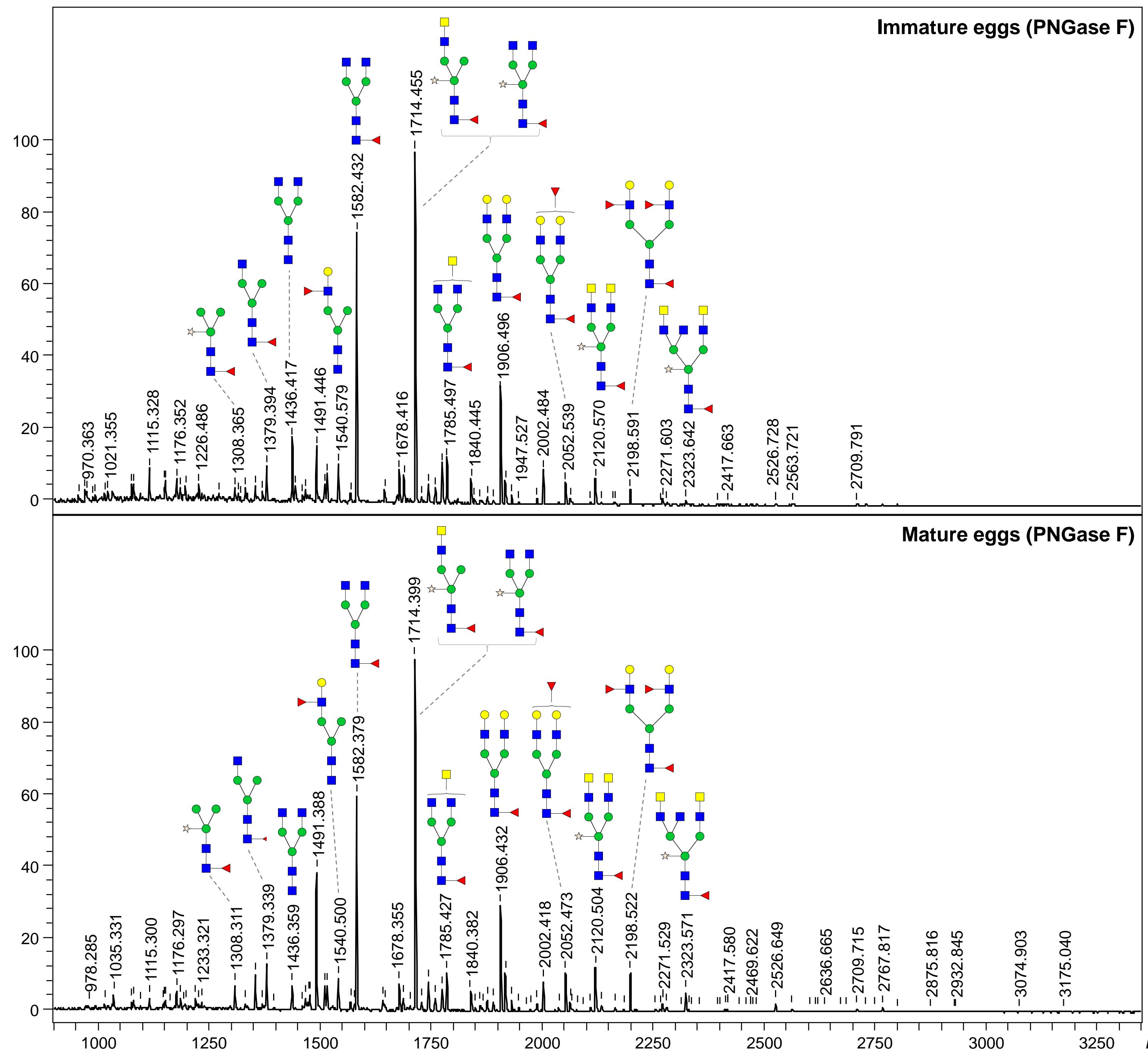
## (C) *S. haematobium* eggs – Glycan sequencing



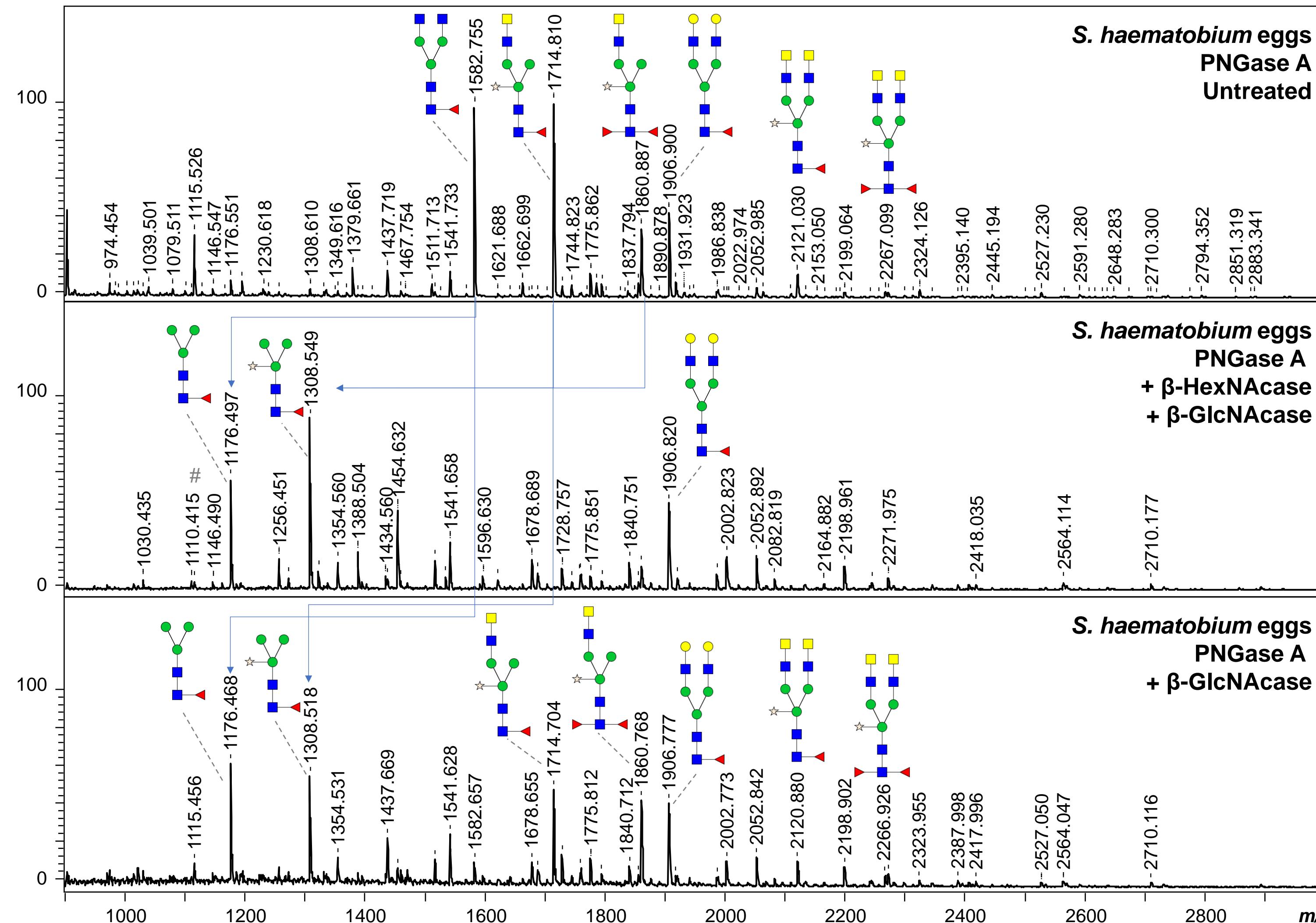
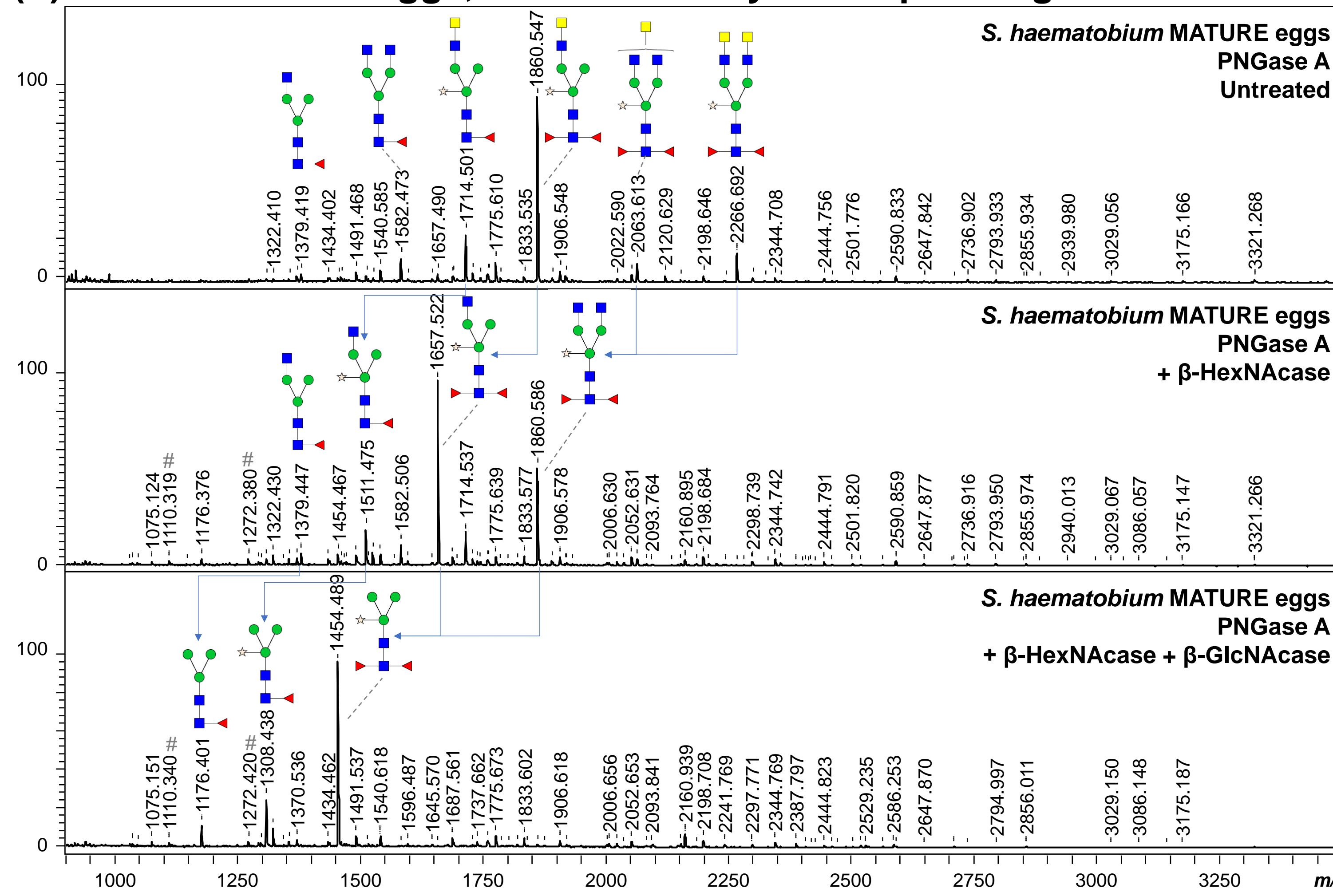
**(C) *S. haematobium* eggs – Glycan sequencing (continued)**



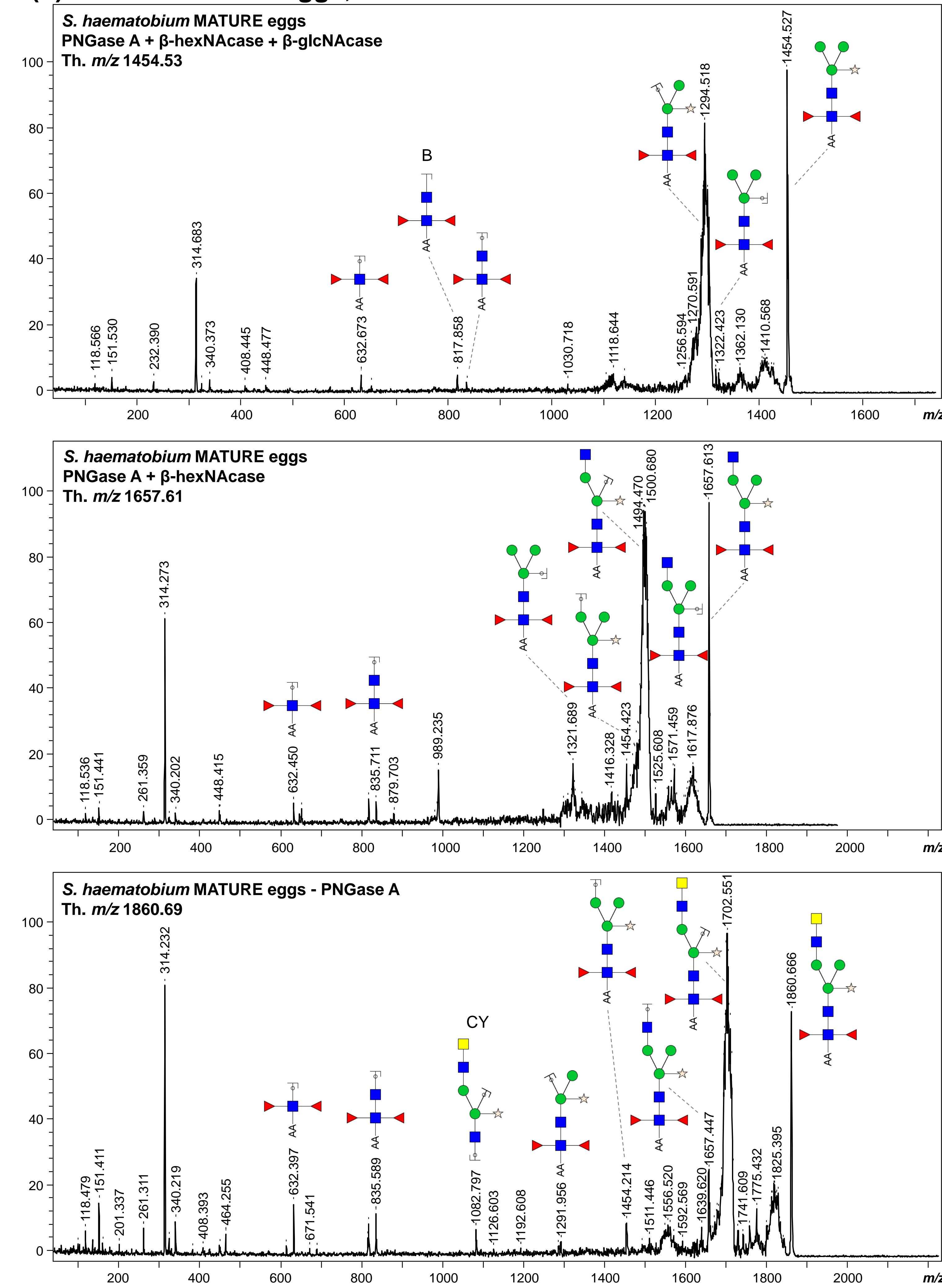
**(D) *S. haematobium* N-glycans – Immature and mature eggs**



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



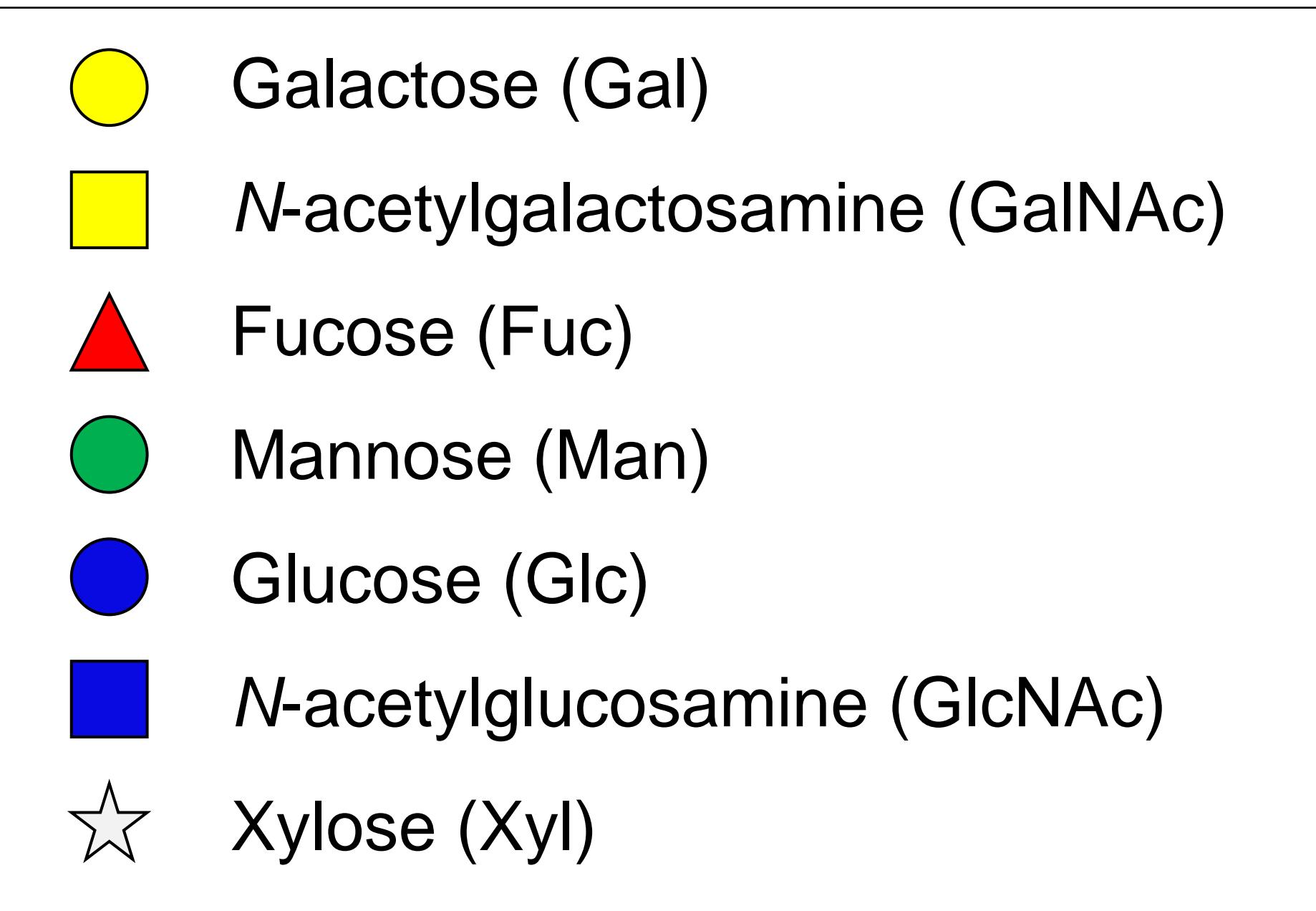
**(F) *S. haematobium* eggs, PNGase A – MALDI-TOF-MS/MS**

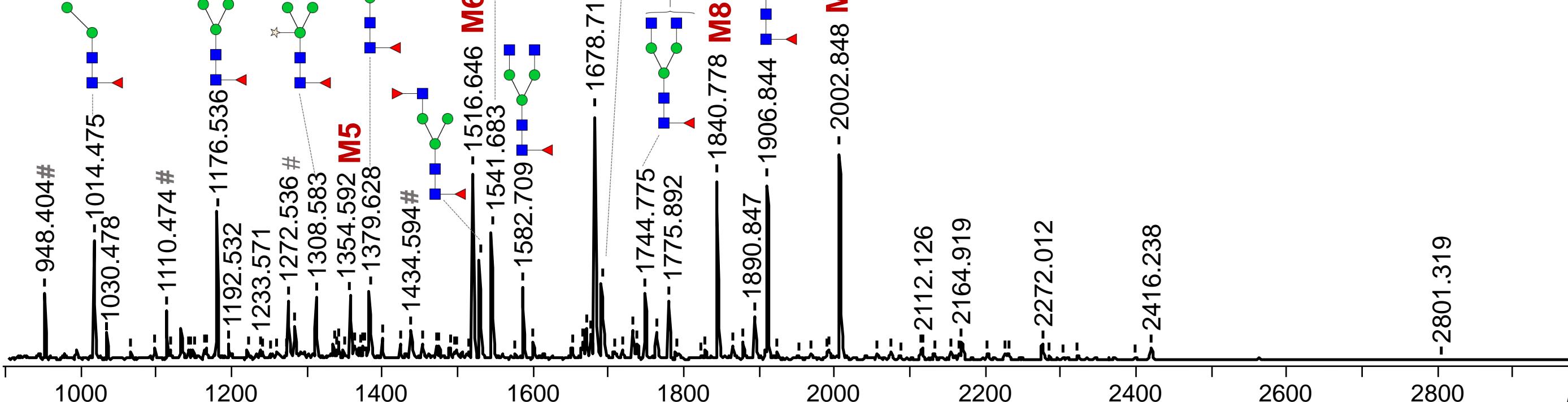
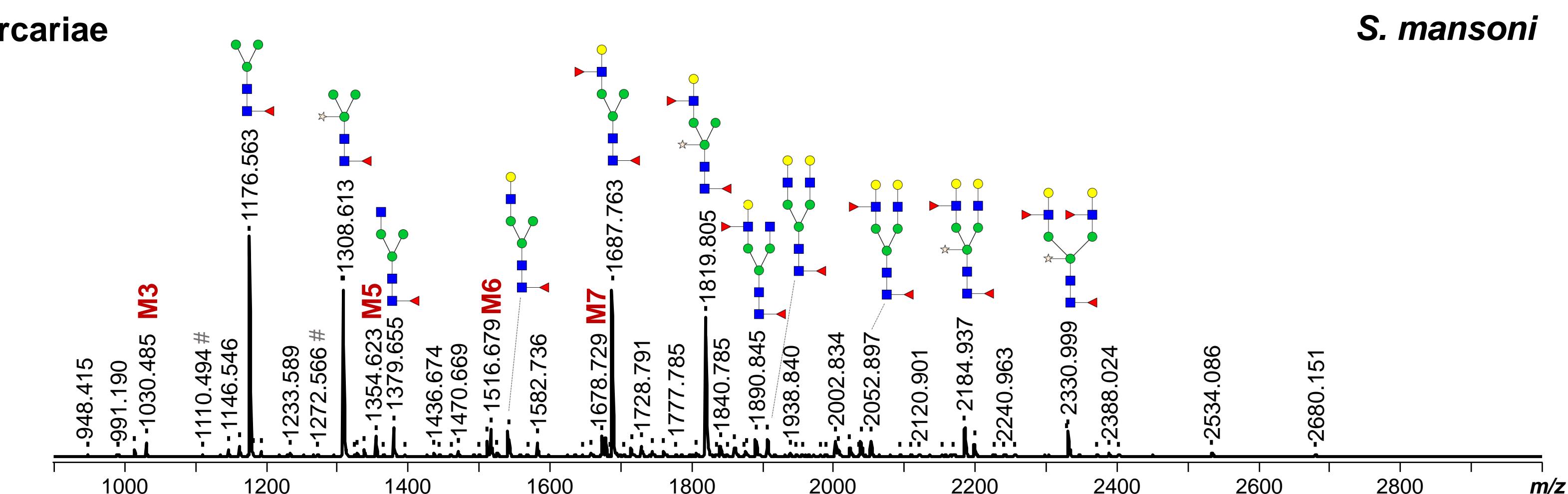
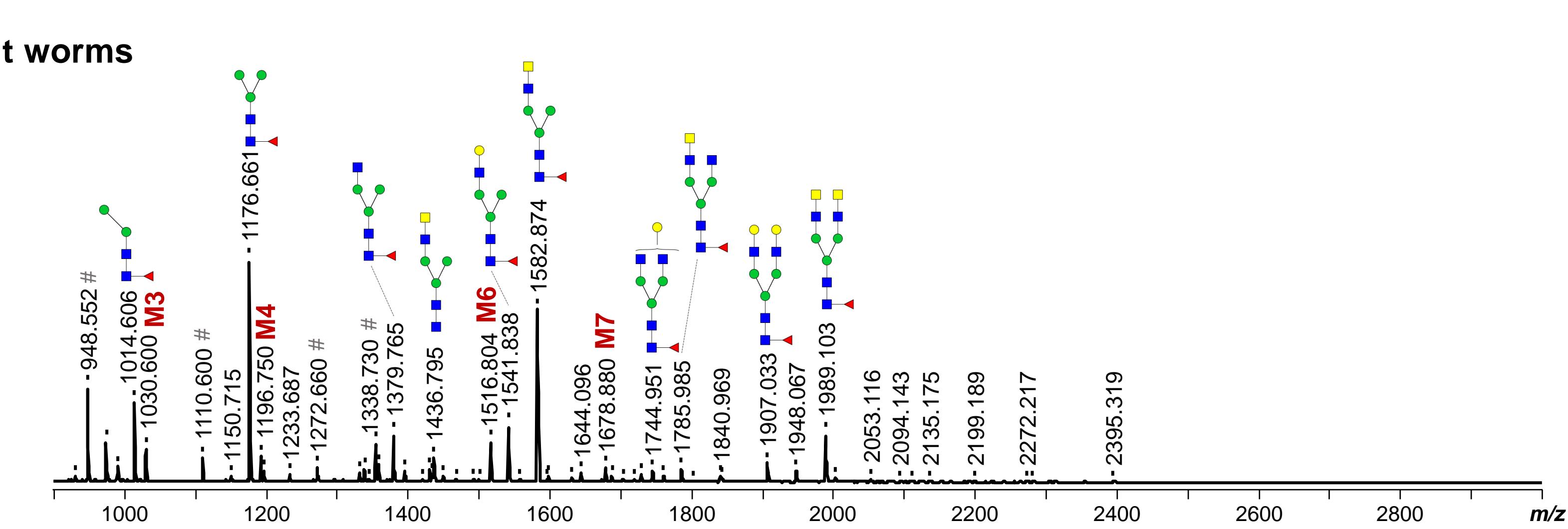
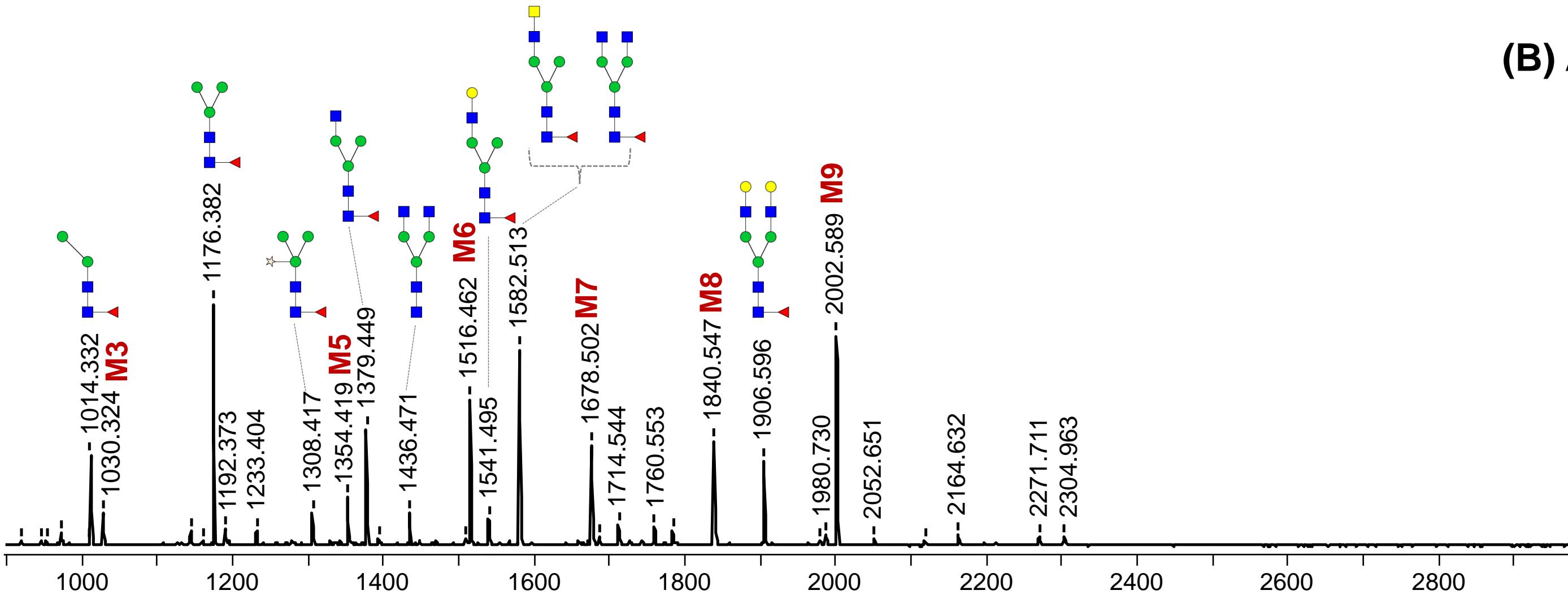
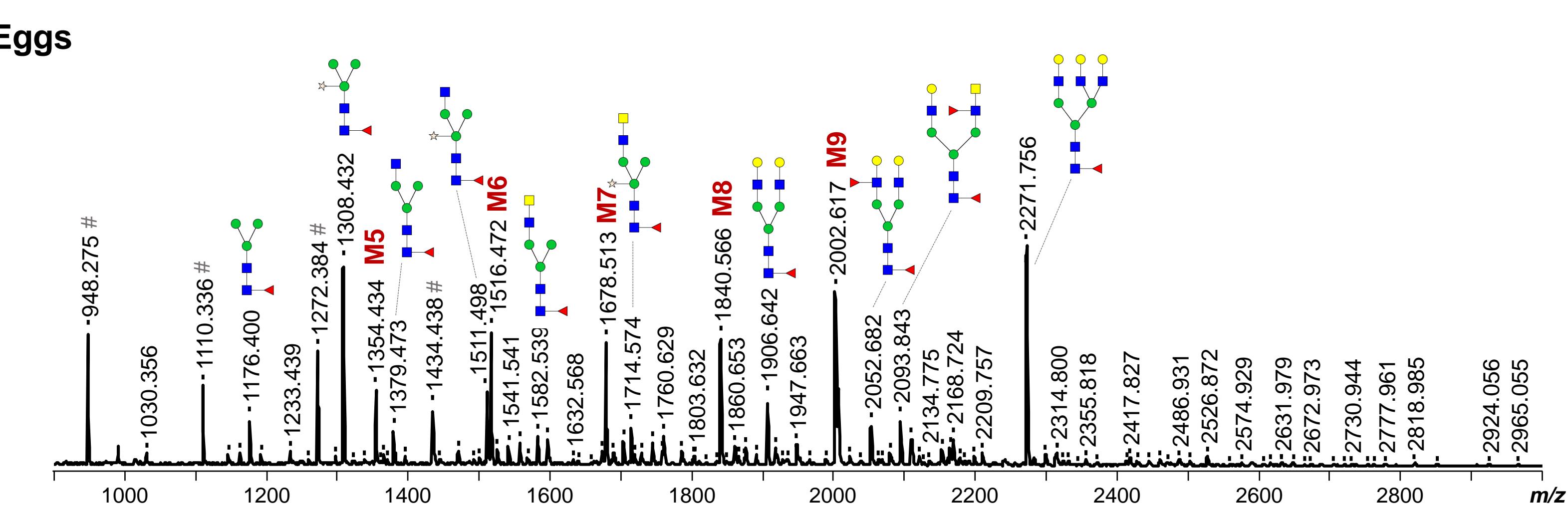
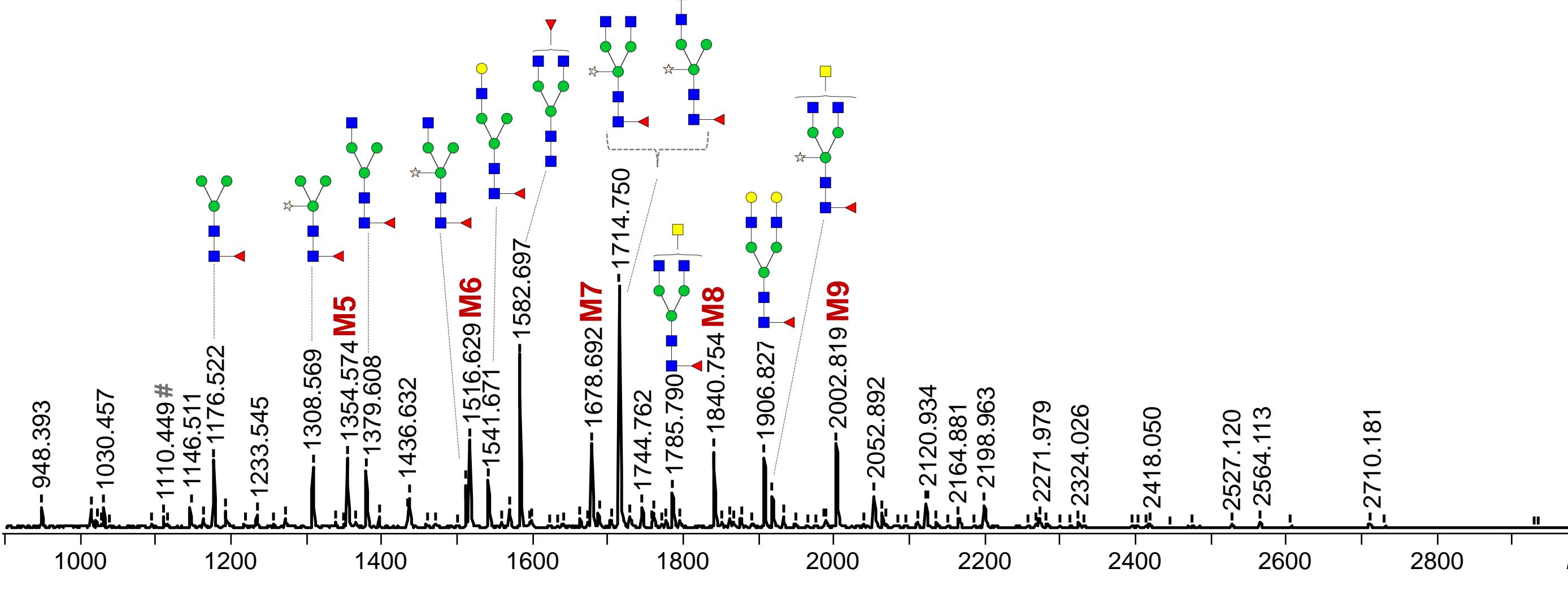
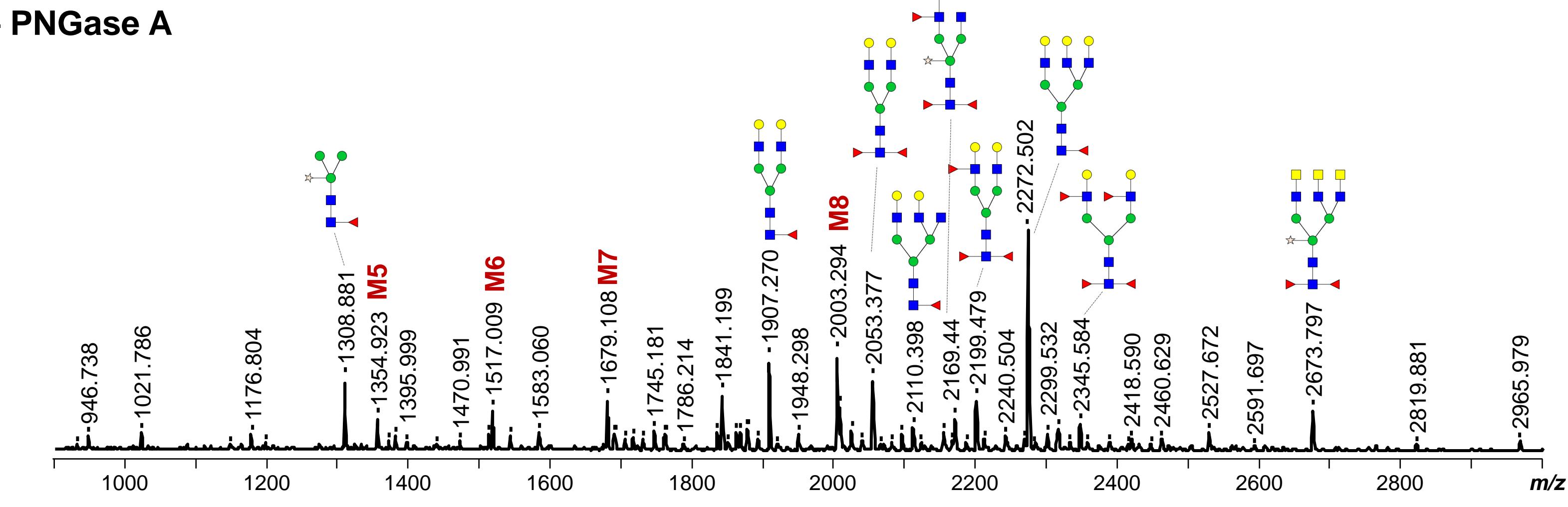
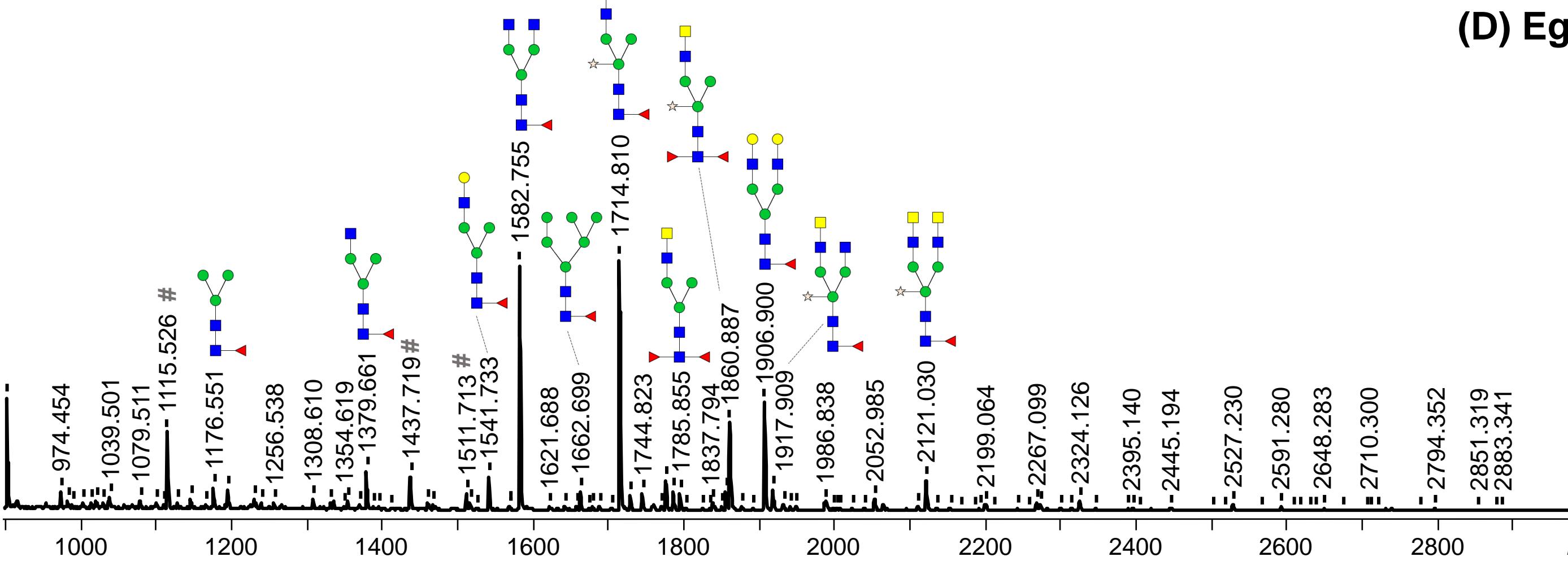


## Figure S6 – Comparison of *S. haematobium* and *S. mansoni* N-glycans

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. Released and labeled glycans were next 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] $^-$ ). Raw data can be found in **Table S1**. 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 **Table S3** for *S. haematobium* and previously published work on *S. mansoni*<sup>15</sup>).

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.

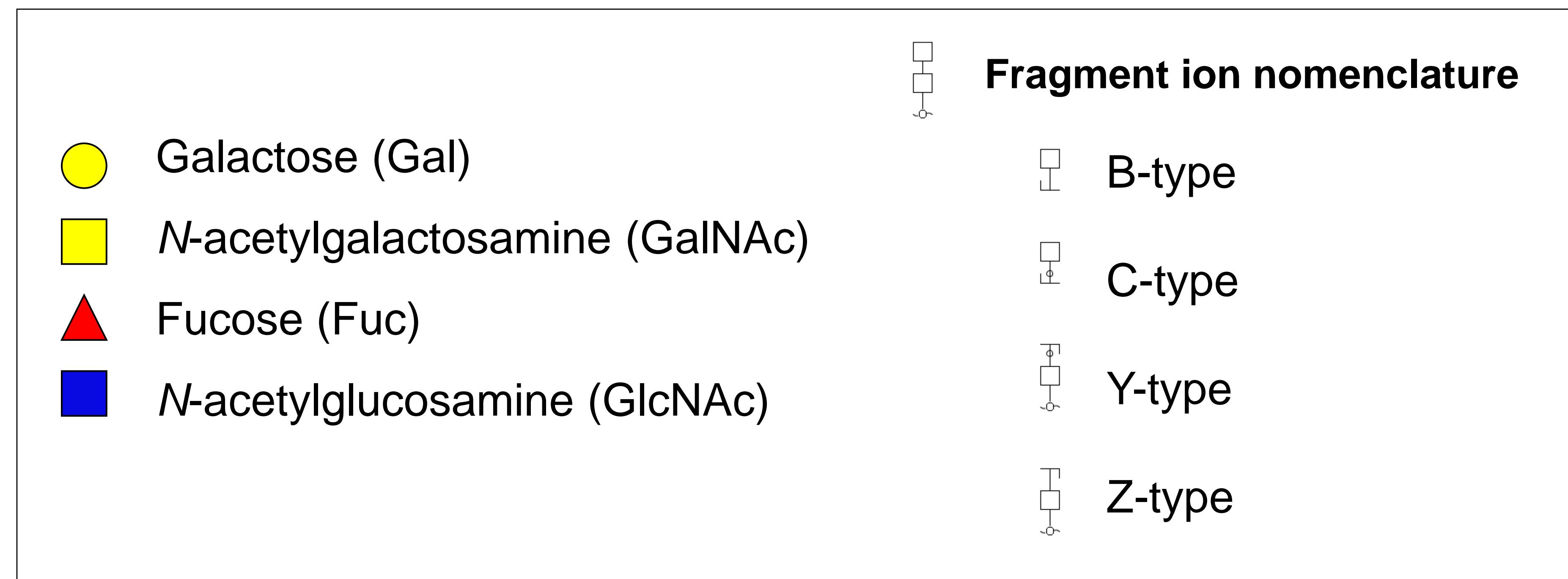


***S. haematobium*****(A) Cercariae****(B) Adult worms****(C) Eggs****(D) Eggs – PNGase A**

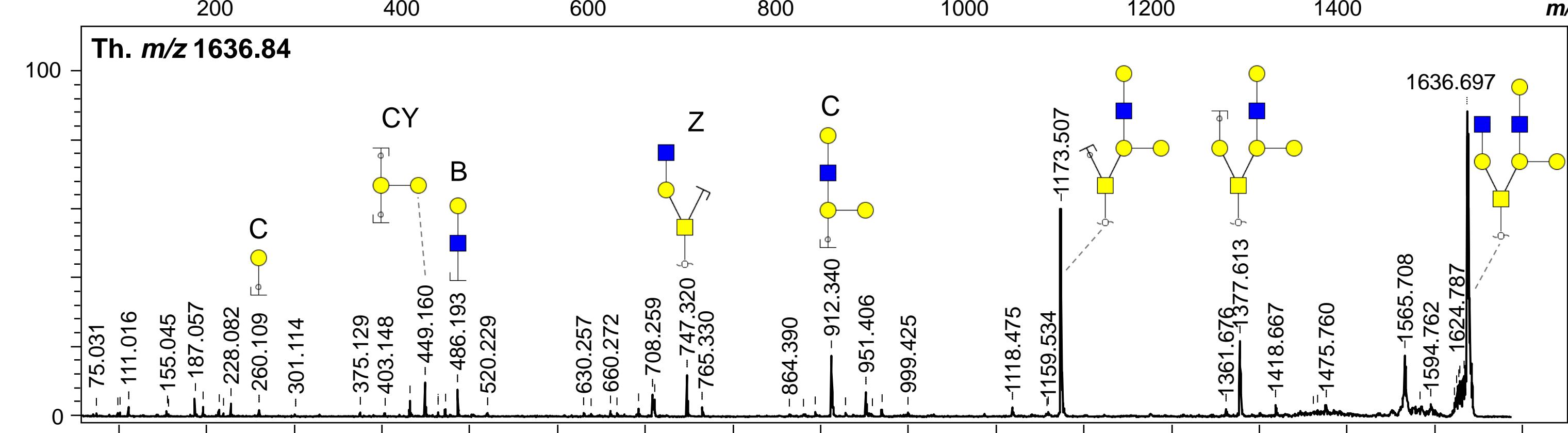
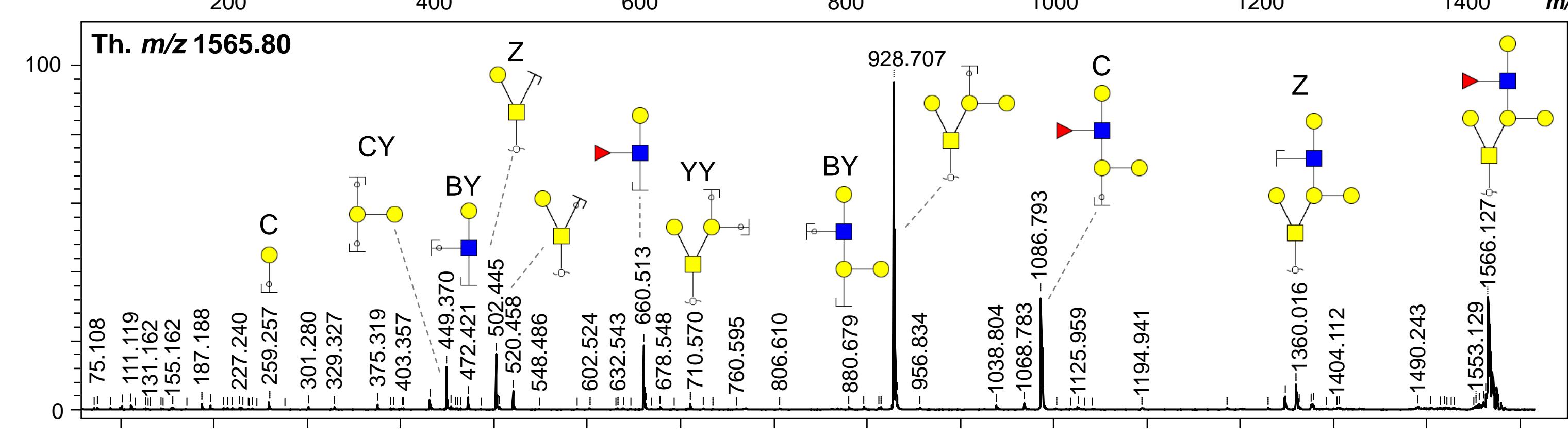
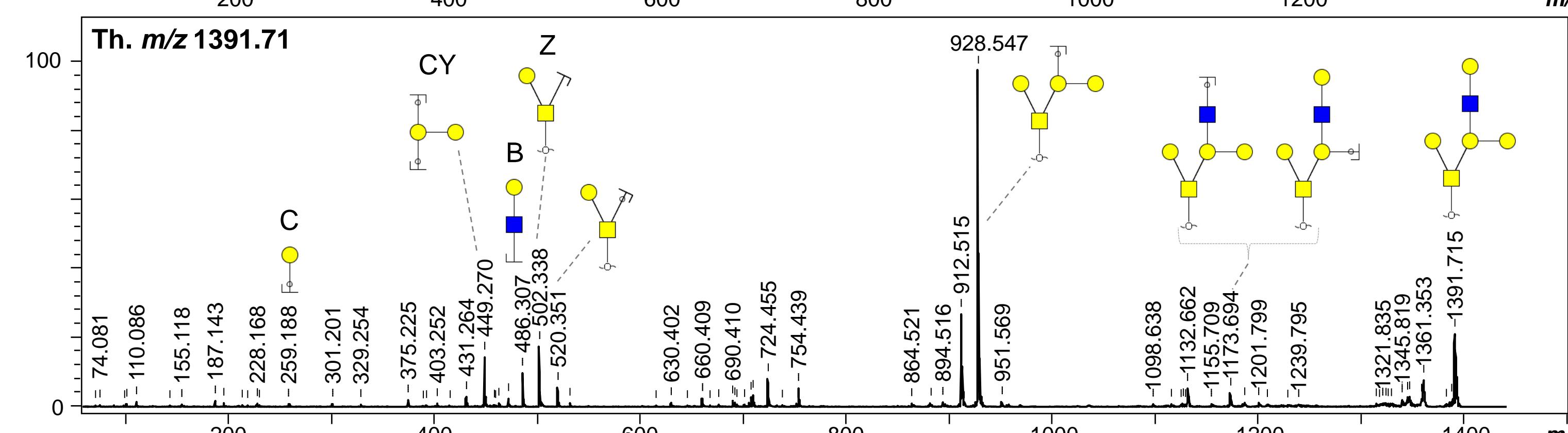
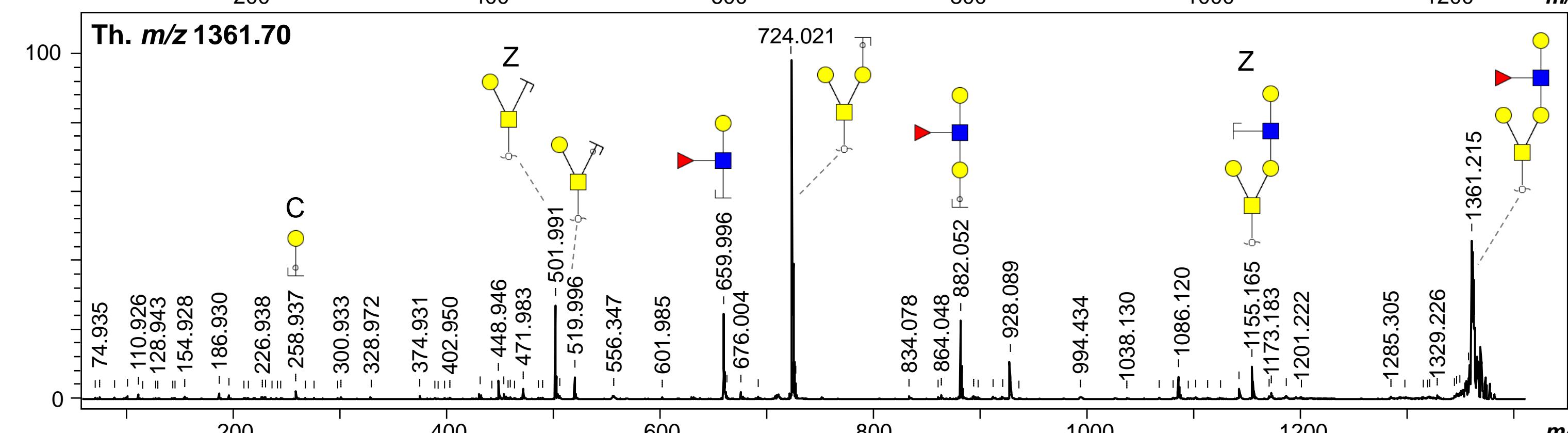
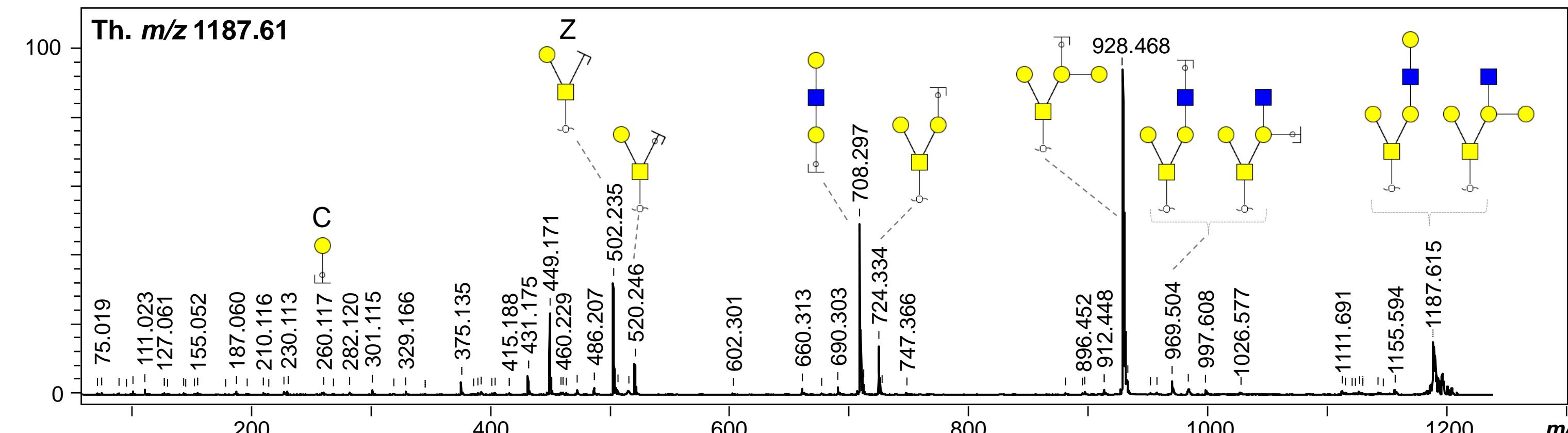
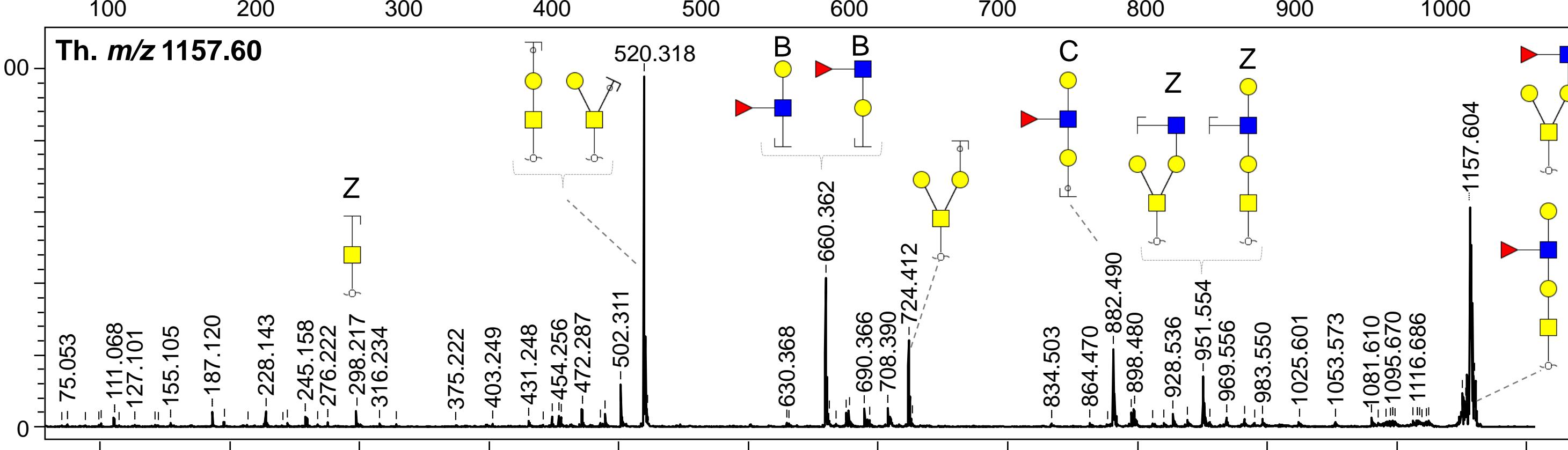
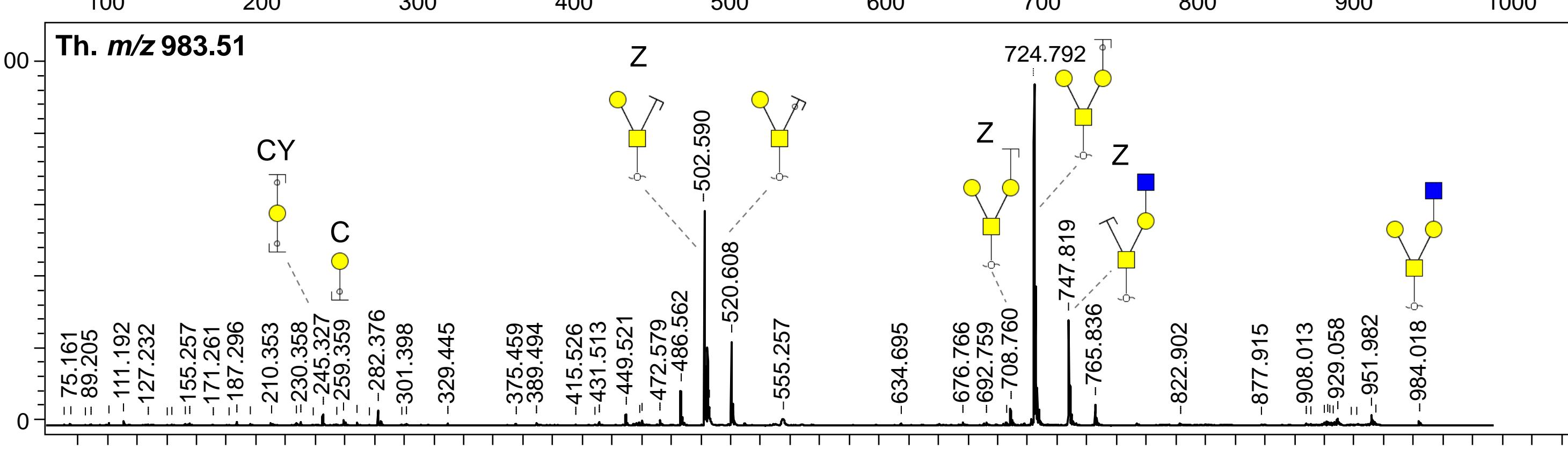
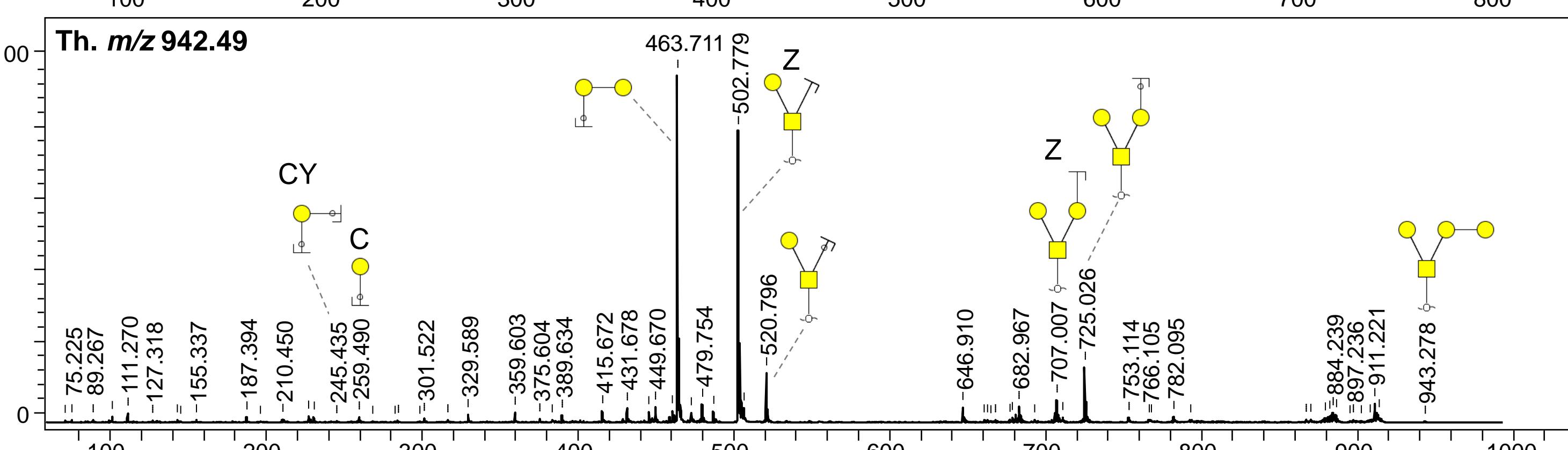
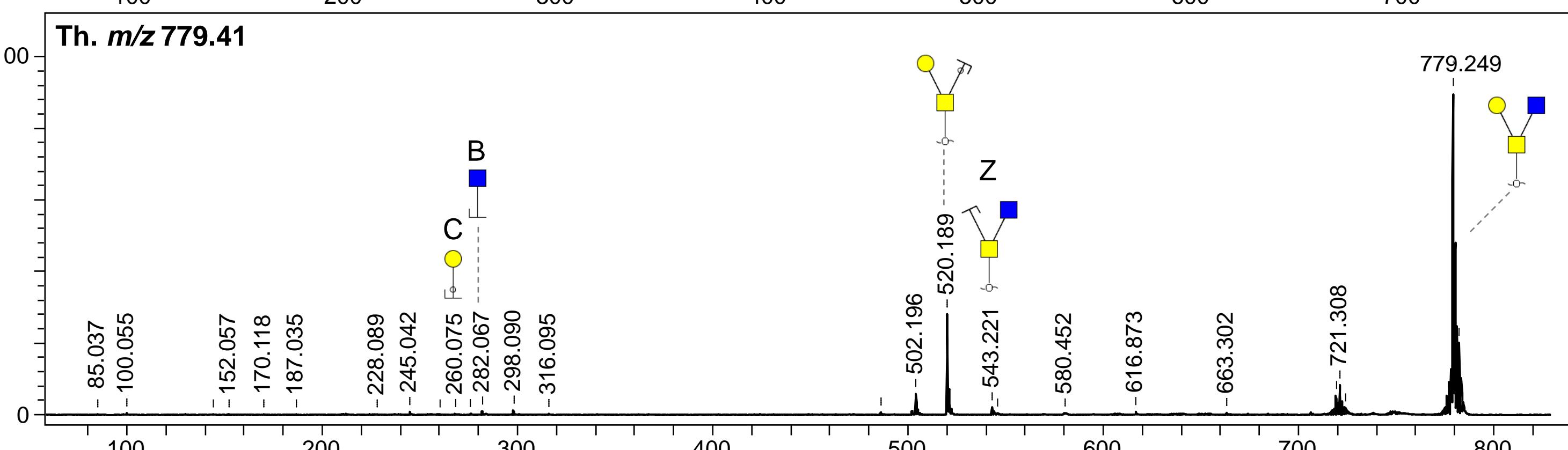
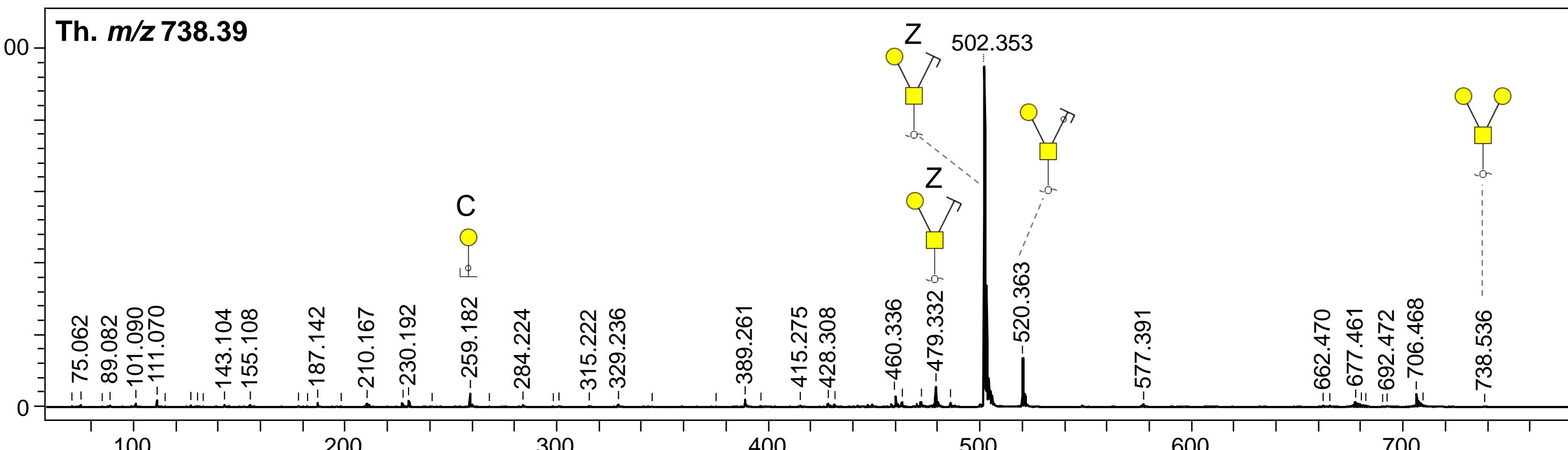
## Figure S7 – MALDI-TOF-MS/MS of *S. haematobium* O-glycans

O-glycans were chemically released using  $\beta$ -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.

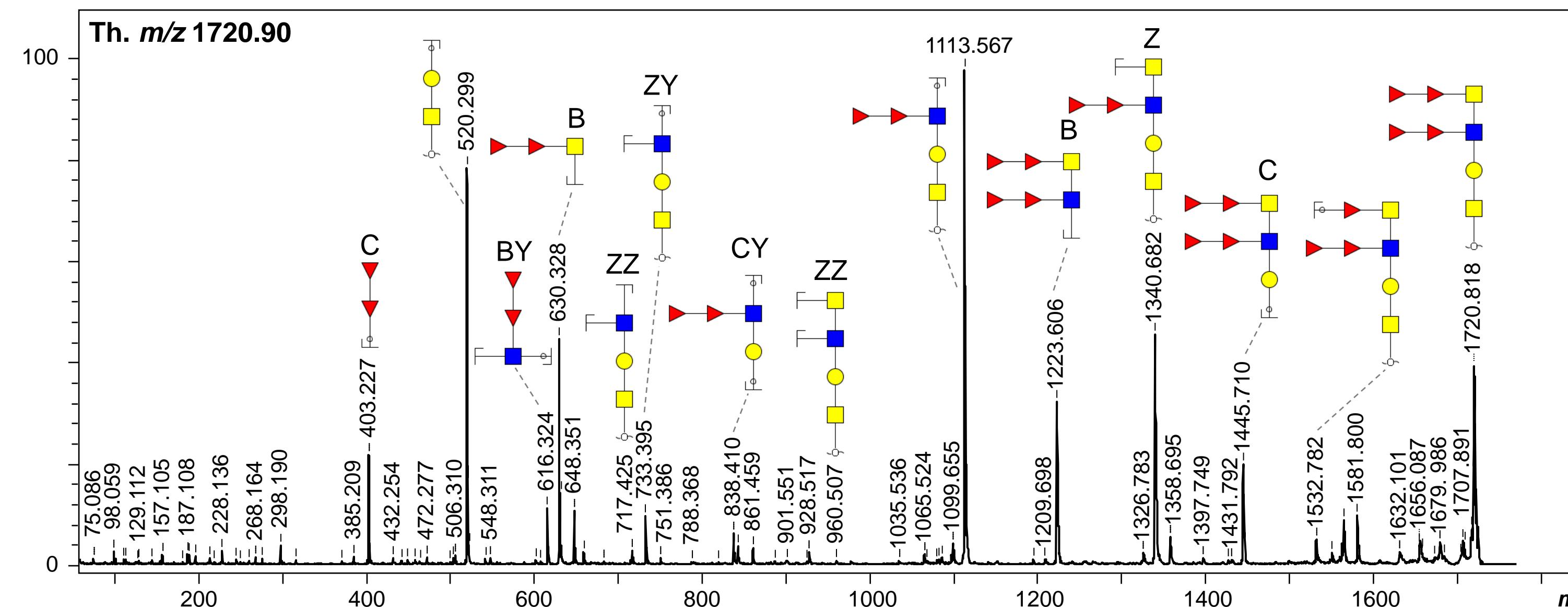
Fragments were registered as hydrogen or sodium adducts in positive-ion reflectron mode. Y-type ions, as defined by Domon and Costello (<https://doi.org/10.1007/BF01049915>) are represented, unless indicated otherwise (B = B type, C = C type, Z = Z type), using the CFG nomenclature (see symbol key insert below).



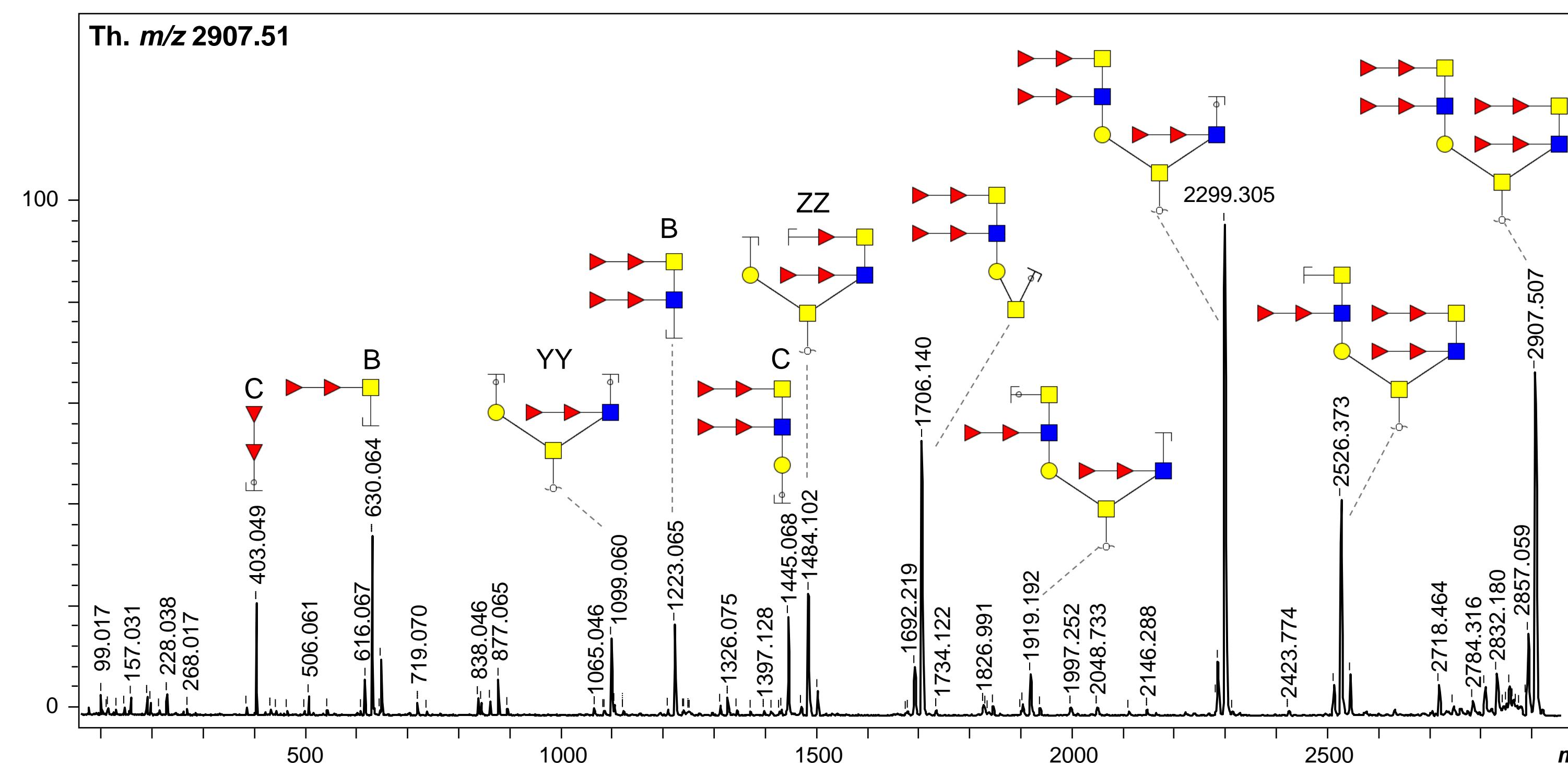
## (A) *S. haematobium* cercariae – MALDI-TOF-MS/M



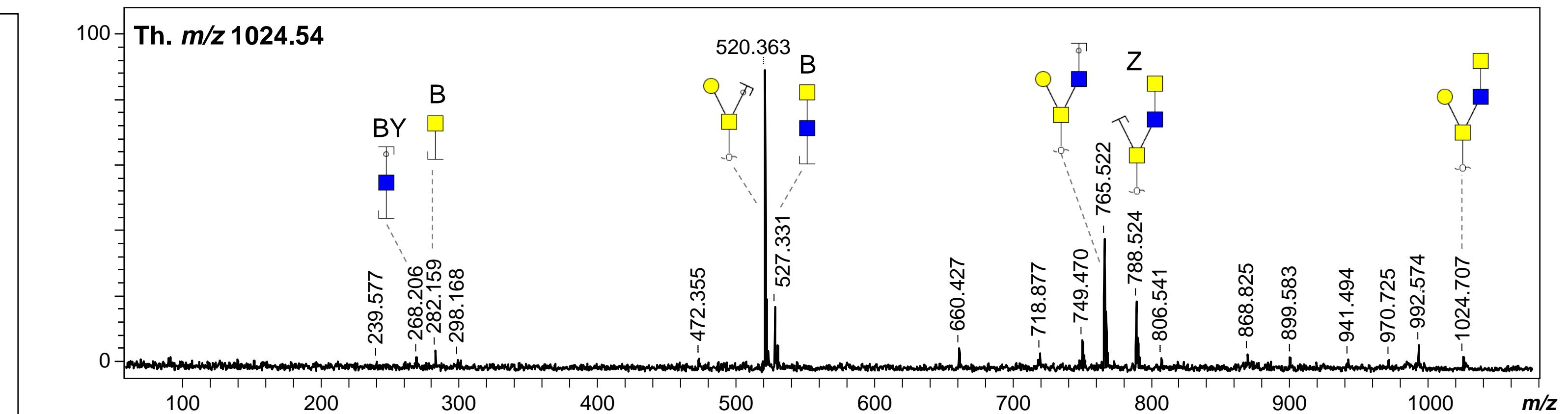
**(A) *S. haematobium* cercariae – MALDI-TOF-MS/MS, CONTINUED**



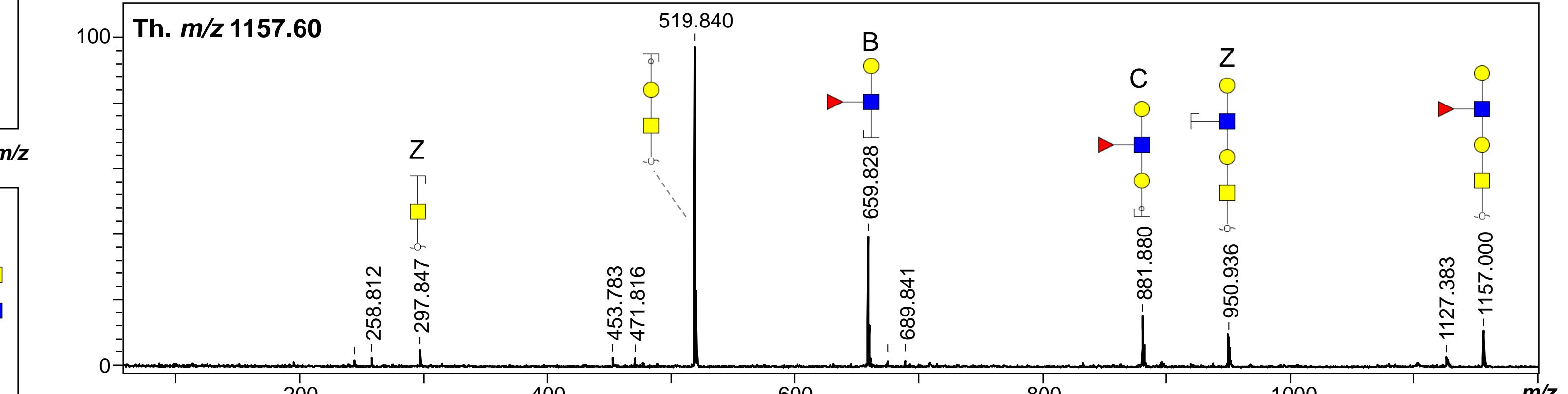
Th.  $m/z$  2907.51



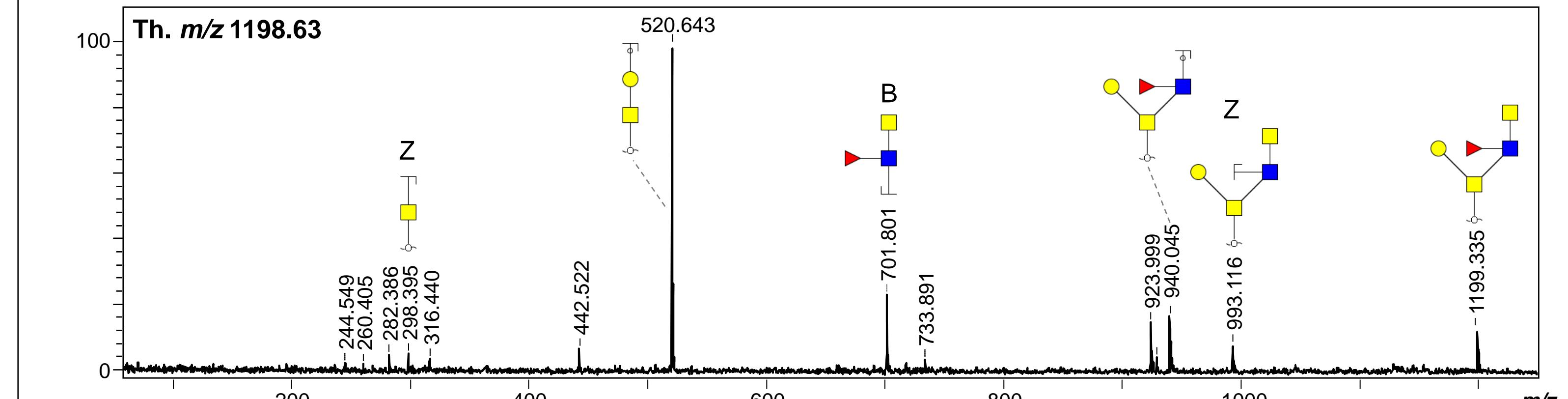
**(B) *S. haematobium* eggs – MALDI-TOF-MS/MS**



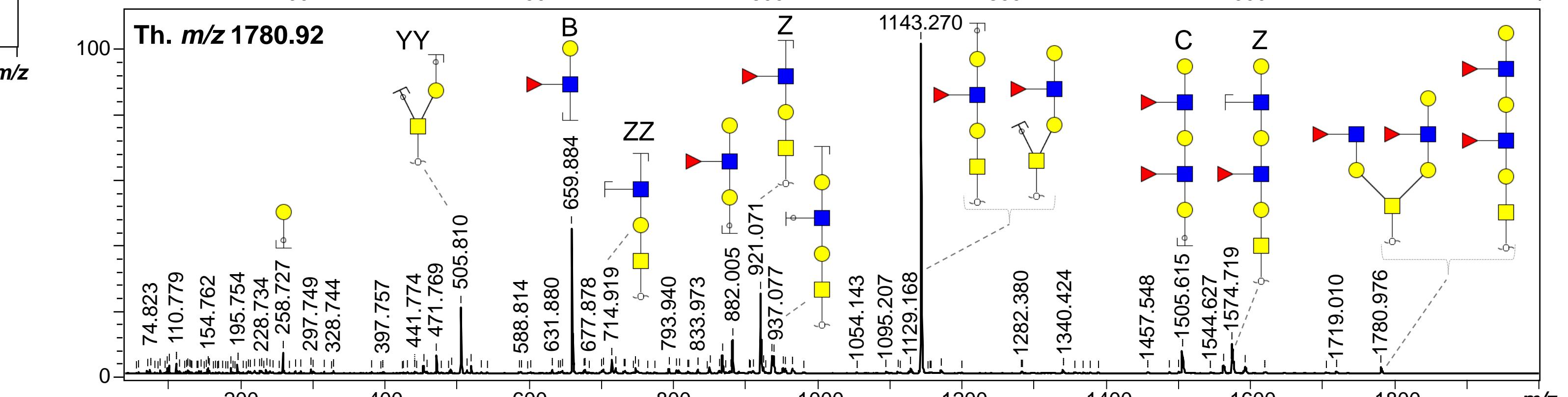
Th.  $m/z$  1157.60



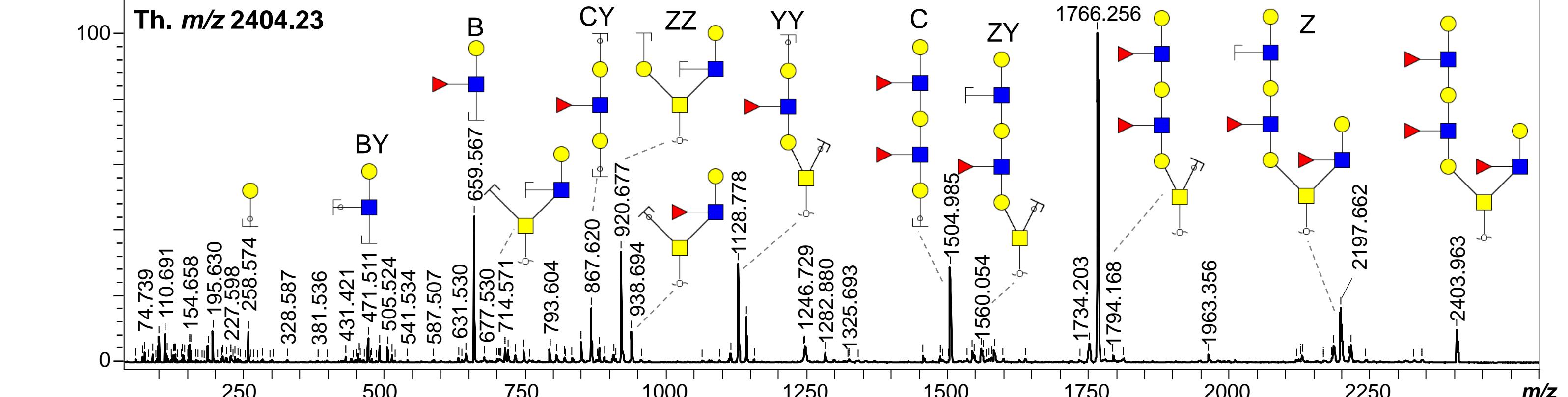
Th.  $m/z$  1198.63



Th.  $m/z$  1780.92



Th.  $m/z$  2404.23



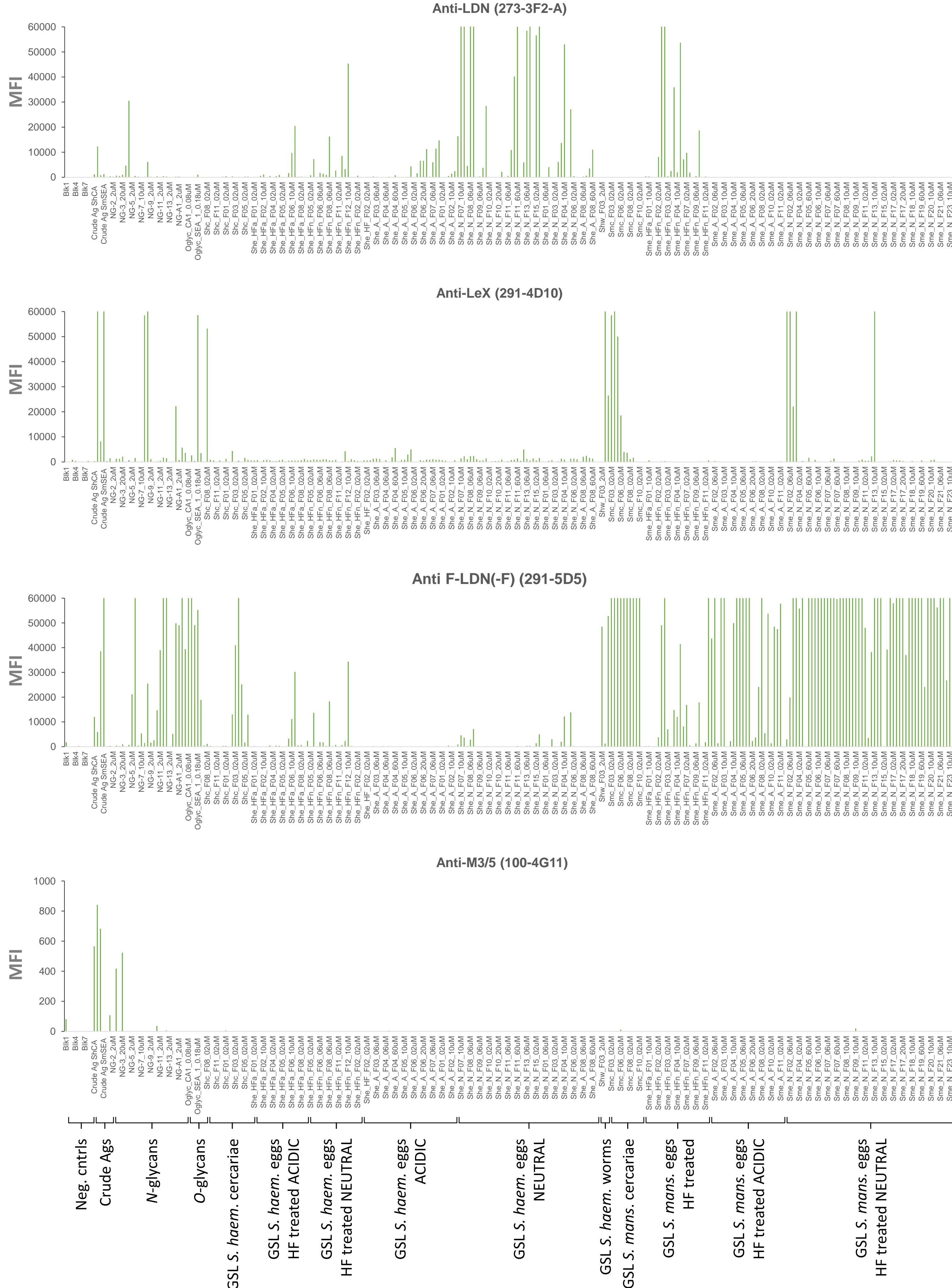
## Figure S8 – Glycan microarray validation using mAbs

Microarrays were incubated with mAbs 273-3F2-A, 291-2G3-A, 291-5D5-A and 100-4G11 respectively binding to LDN, LeX (<https://doi.org/10.1093/glycob/10.6.601>), fucosylated LDN (F-LDN(-F)) (<https://10.1017/s0031182004006390>; <https://10.1016/j.ab.2010.07.008>) and to the trimannosyl branching R- $\alpha$ 1-6Man( $\alpha$ 1-6Man) $\alpha$ 1-3Man (<https://10.1093/glycob/cwg025>) glycan motifs (**A**). Background-corrected median fluorescence intensities (MFIs, y-axis) are shown for each glycan fraction printed on the array (x-axis). 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 **Table S6** and fraction content detail in **Table S5**.

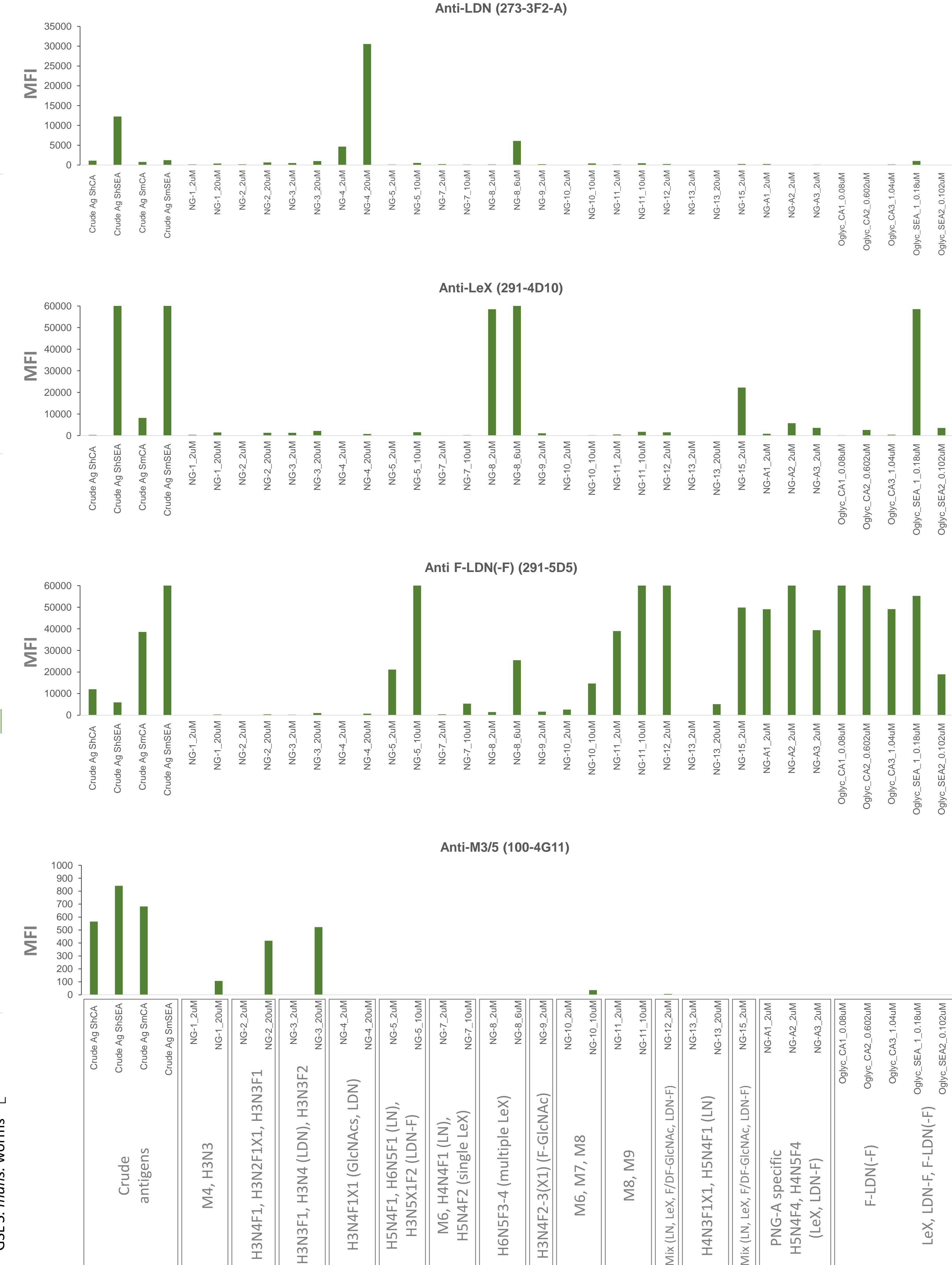
Panel (**B**) shows a closer view of mAb binding to crude antigens, N-linked and O-linked glycan-containing fractions. Details of major glycan structures present in fractions are provided below the x-axis: H = hexose, N = N-acetylhexosamine, F = fucose, X = core-xylose, M3-9 = mannosylated N-glycans with 3 to 9 mannose residues.

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 chemically 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. Expectedly, the mAb 291-5D5, showed an opposite pattern of binding, with high median fluorescence intensity (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, all of them containing fucosylated LDN motifs. 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 structural content. Finally, very low fluorescence levels were obtained from incubation with 100-4G11 a mAb recognizing the trimannosyl branching Man $\alpha$ 1-3(Man $\alpha$ 1-6)Man $\alpha$ 1-6-R (Van Remoortere *et al.*, 2003) present in Man3/5, consistent with this epitope only being present in low amounts in a few N-glycan-containing fractions.

## (A) mAb binding to glycan array – all fractions

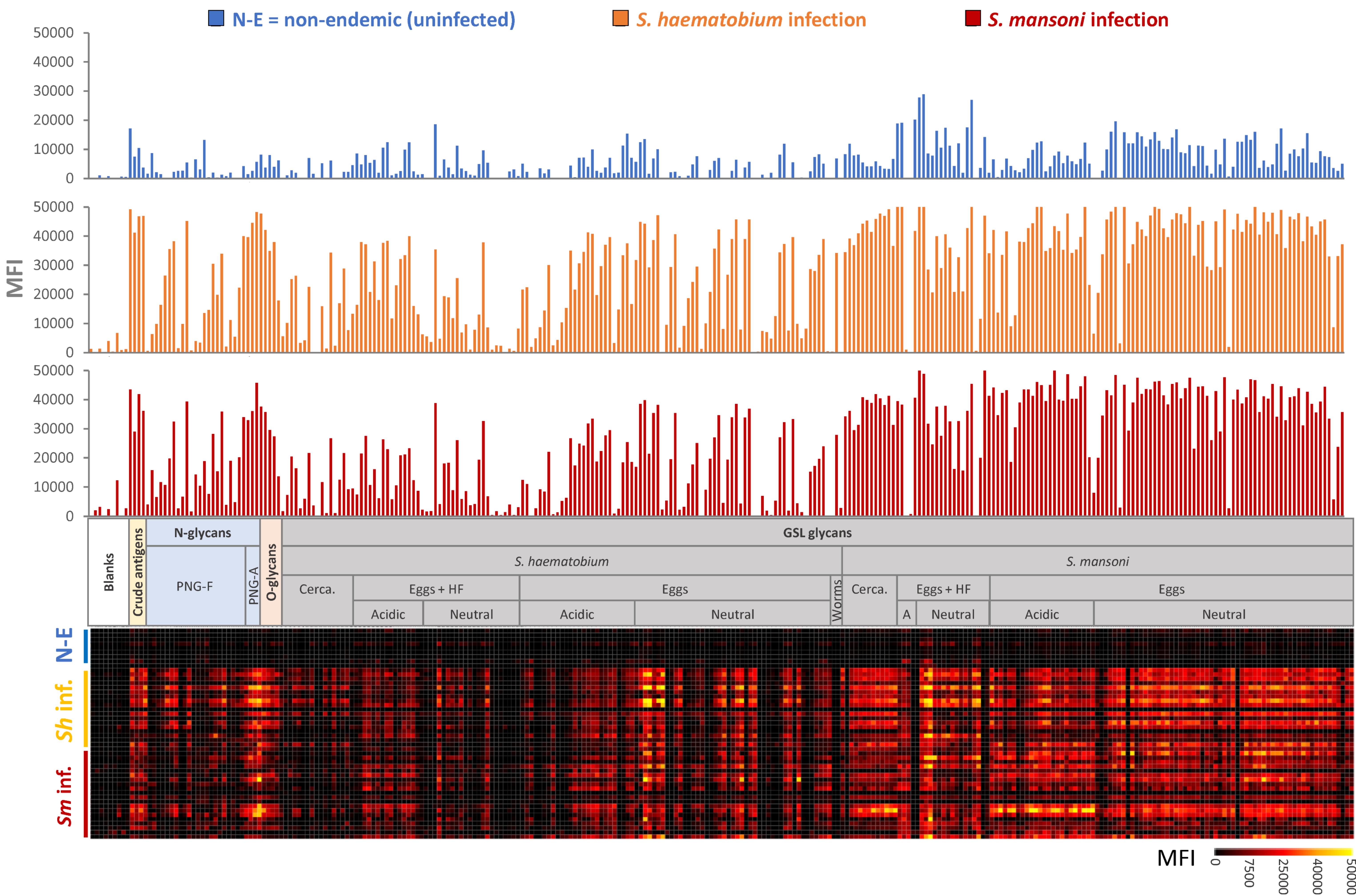


## (B) mAb binding to crude antigens and *N/O*-glycan-containing fractions



## Figure S9 – 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 [Table S8](#). 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).



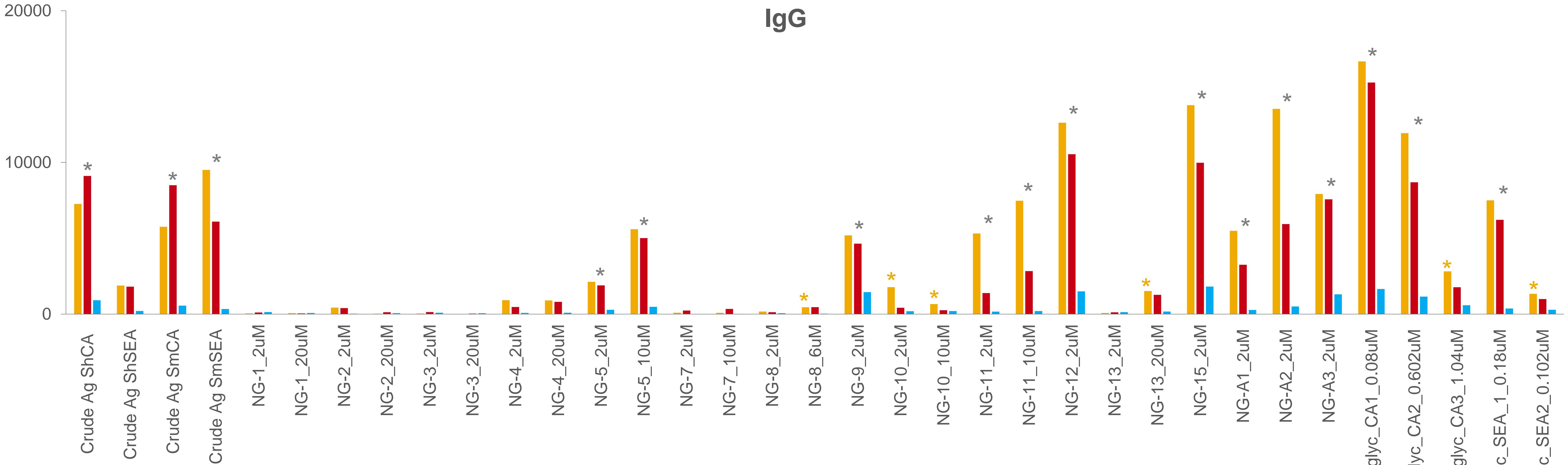
## Figure S10 – IgG and IgM responses to crude antigens, N-linked and O-linked glycans

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 (Table S7 (IgG) and Table S8 (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. P-values  $< 0.05$  that indicate a significant difference in MFI values between groups are indicated 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.

■ *S. haematobium* infection■ *S. mansoni* infection

■ N-E = non-endemic (uninfected)

**IgG****IgM**