

SUPPORTING INFORMATION for the article

“Novel O-linked Sialoglycan Structures in Human Urinary Glycoproteins”

A. Pap, E. Tasnadi, K.F. Medzihradzsky and Z. Darula

The most likely structure of new glycoforms was deciphered manually, using both HCD and EThcD data. Obviously, the identity of the sugar units and their linkage positions cannot be assigned from these data. However, this supplement demonstrates that **EThcD spectra acquired with 15% NCE may deliver structural information on the glycan part of a glycopeptide**. All AVAVTLQSH spectra presented here have proof for the site assignment (modified c_7 and/or unmodified z_4 peptide fragments).

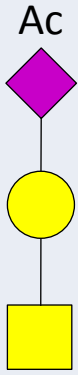
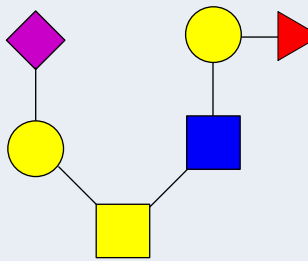
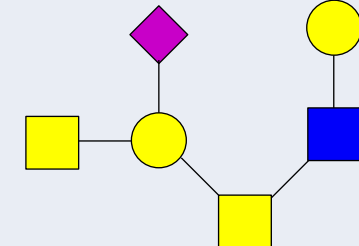
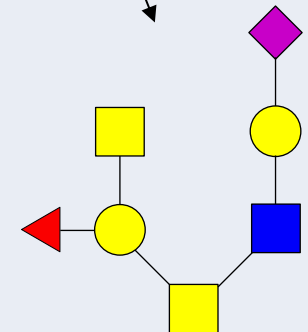
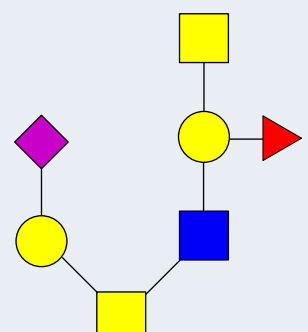
Table of Contents

- Cover page
- **Table S-1** - Novel glycans S-2
- Internal & non-reducing end glycan fragments (calculated masses) S-5
- Legends and Annotations S-6
- Manually annotated HCD and EThcD spectra of different glycoforms of AVAVTLQSH and DFTAAFPR peptides, and schemes with alternative structures, when there is any, pointing out the fragmentation differences S-7

Table S-1 Novel glycans

This isomer has been reported in [Darula, Z.; Pap A.; Medzihradzsky, K. F. *J Proteome Res.* 2019, 18, 280-291.]

N: HexNAc, H: Hex, F: Fuc, A: NeuAc, AAC: NeuAcAc

Monoisotopic additive mass	698.2382	1167.4177	1224.4392	1370.4971	
<i>Composition</i>	<i>NHAAc</i>	<i>N₂H₂FA</i>	<i>N₃H₂A</i>	<i>N₃H₂FA</i>	
<i>Structure</i>					

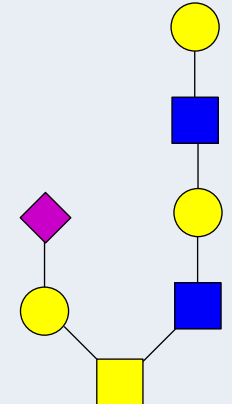
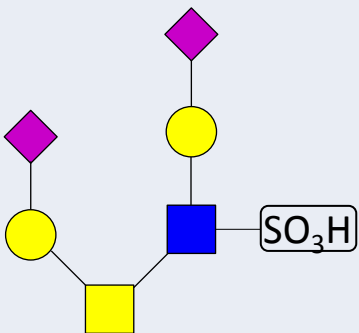
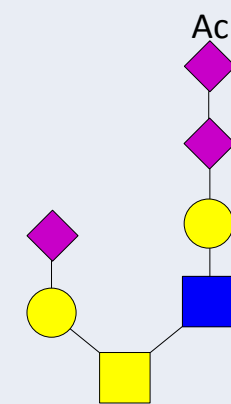
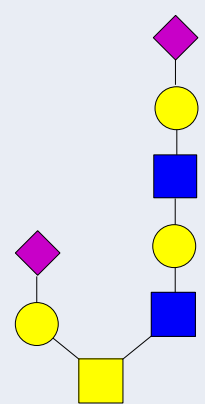
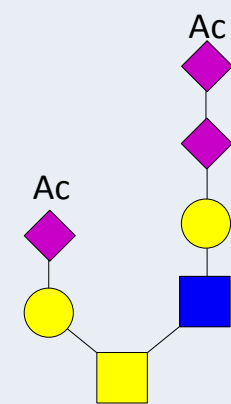
Monoisotopic additive mass	1386.4920	1392.4120	1645.5612	1677.5874	1687.5718
<i>Composition</i>	<i>N₃H₃A</i>	<i>N₂H₂A₂Sulfo</i>	<i>N₂H₂A₂AAC</i>	<i>N₃H₃A₂</i>	<i>N₂H₂AAAc₂</i>
<i>Structure</i>					

Table S-1 Novel glycans (contd.)

N: HexNAc, H: Hex, F: Fuc, A: NeuAc

Monoisotopic additive mass	1735.6293	1751.6242	1969.7032	2042.7196	2116.7564
Composition	N_4H_3FA	N_4H_4A	$N_3H_3F_2A_2$	$N_4H_4A_2$	N_5H_5A
Structure					

Table S-1 Novel glycans (contd.)

N: HexNAc, H: Hex, F: Fuc, A: NeuAc

Monoisotopic additive mass	2333.8150	2391.8569	2407.8518
Composition	$N_4H_4A_3$	$N_5H_4FA_2$	$N_5H_5A_2$
Structure			

m/z	Glycan composition	m/z	Glycan composition
186.0761	HexNAc-H ₂ O	569.2189	HexNAc ₂ Hex
204.0867	HexNAc	657.2349	HexNAcHexNeuAc
274.0921	NeuAc-H ₂ O	698.2615	HexNAc ₂ NeuAc
292.1027	NeuAc	699.2455	HexNAcHexNeuAcAc
314.0846	NeuAc+Na	715.2768	HexNAc ₂ HexFuc
316.1027	NeuAcAc-H ₂ O	731.2717	HexNAc ₂ Hex ₂
334.1133	NeuAcAc	803.2928	HexNAcHexFucNeuAc
366.1395	HexNAcHex	1022.3671	HexNAc ₂ Hex ₂ NeuAc
388.1214	HexNAcHex+Na	1096.4039	HexNAc ₃ Hex ₃
407.1661	HexNAc ₂	1168.4250	HexNAc ₂ Hex ₂ FucNeuAc
454.1555	HexNeuAc	1387.4993	HexNAc ₃ Hex ₃ NeuAc
476.1374	HexNeuAc+Na	1371.5044	HexNAc ₃ Hex ₂ FucNeuAc
495.1821	HexNAcNeuAc	1678.5947	HexNAc ₃ Hex ₃ NeuAc ₂
512.1974	HexNAcHexFuc		
528.1923	HexNAcHex ₂		

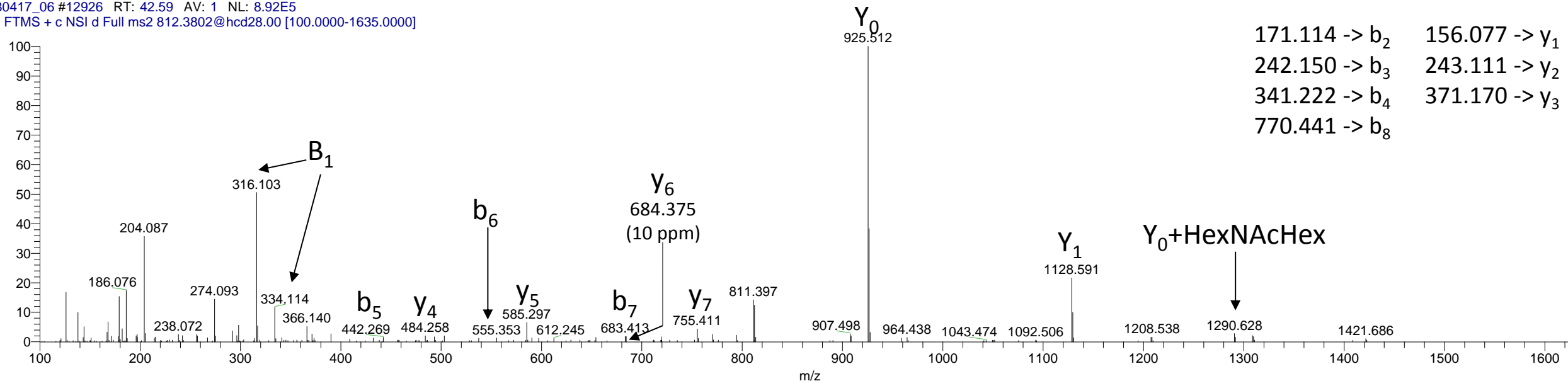
In **Figures S-1-14** HCD spectra are presented in the upper panels, EThcD data are shown in the lower panels. Peptide fragments are labeled according to the Biemann nomenclature [Biemann K. Nomenclature for peptide fragment ions (positive ions). *Methods Enzymol.* **1990**, *193*, 886-887.]. In most cases the peptide fragments are rather weak and the HCD data may reveal the presence of coeluting/cofragmenting components. Glycan fragments are labeled according to the Domon-Costello nomenclature [Domon, B.; Costello, C. E. *Glycoconjugate J.* **1988**, *5*, 397-409.], and the cleavages are indicated in the schemes included. Internal glycan fragments are labeled with cartoons, according to the CFG recommendations. In the EThcD spectra the precursor and its charge reduced forms are indicated as „pr(charge)”. The charge reduced form of coeluting molecules are labeled with „Coel.(charge)”. When a fragment is formed via multiple glycosidic bond cleavages than the following label format is used: $Y_{2\alpha,3\beta}$. When a particular mass may represent different structures, the following format is used: $Y_{2\alpha} | Y_{2\beta}$.

Figure S-15 – displays the EThcD data of isomeric blood-type A antigens, Type 3 (upper panel), Type 1 or 2 (lower panel).

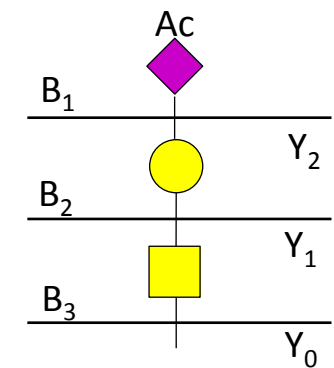
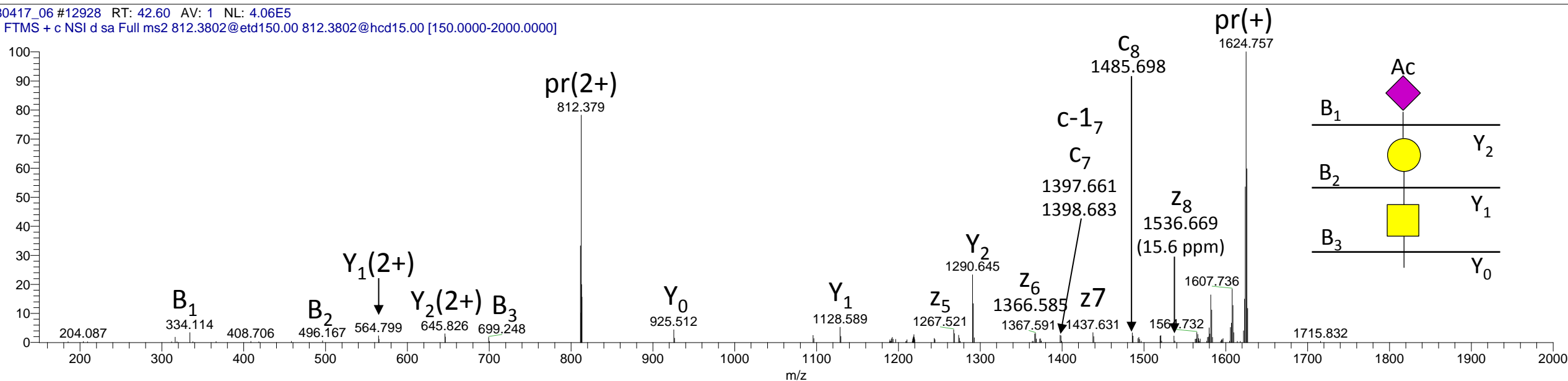
As we reported earlier [Darula, Z.; Pap, Á.; Medzihradszky, K. F. *J Proteome Res.* **2019**, *18*, 280–291.], in some of the spectra we observed fragments that indicated sialic acid or fucose migration, a known phenomenon in glycan analysis [Wuhrer, M.; Deelder, A. M.; van der Burgt, Y. E. *Mass Spectrom Rev.* **2011**, *30*, 664-680.]. Such fragments are clearly annotated.

180417_06 #12926 RT: 42.59 AV: 1 NL: 8.92E5
 T: FTMS + c NSI d Full ms2 812.3802@hcd28.00 [100.0000-1635.0000]

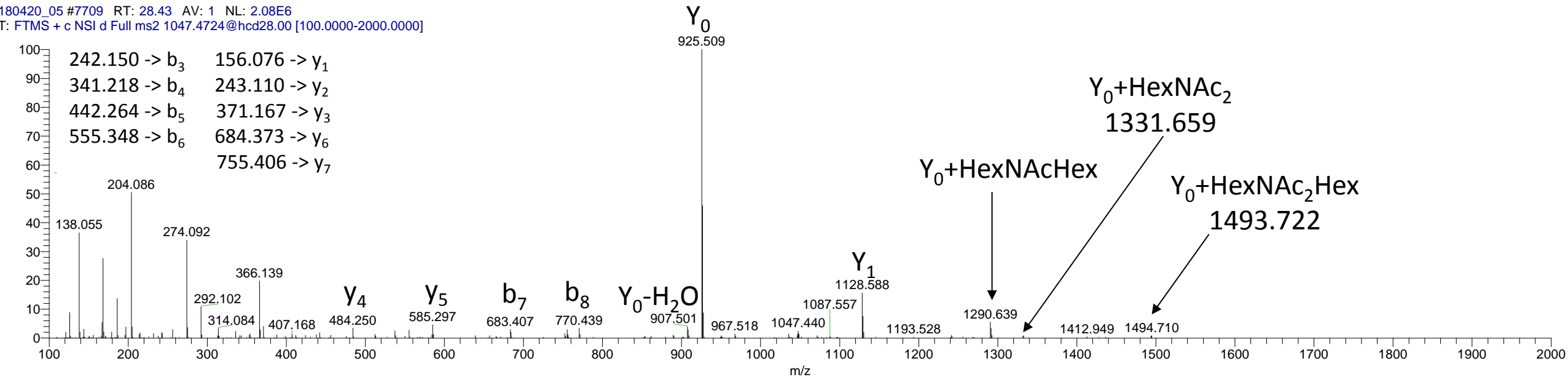
171.114 → b₂ 156.077 → y₁
 242.150 → b₃ 243.111 → y₂
 341.222 → b₄ 371.170 → y₃
 770.441 → b₈



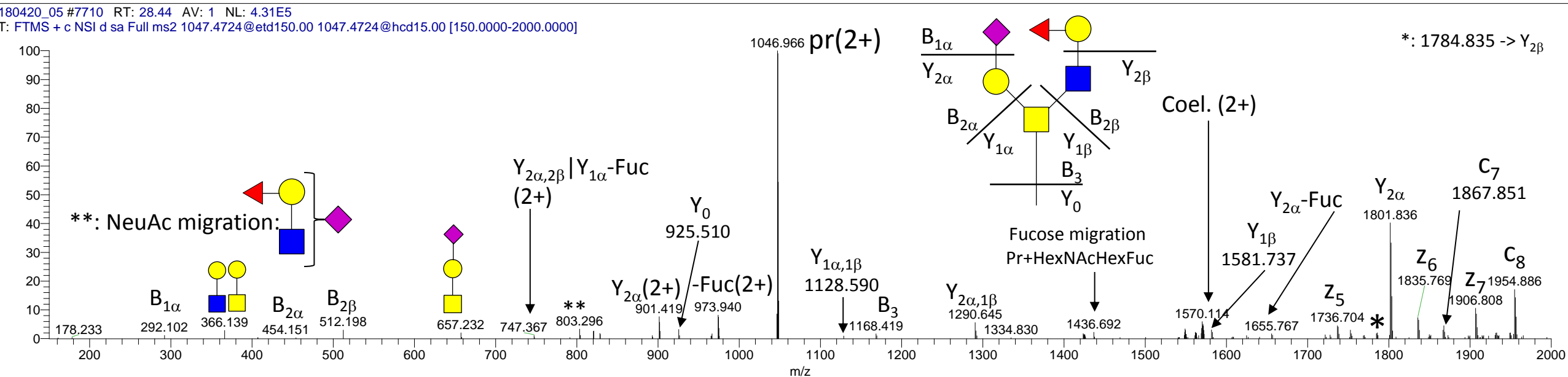
180417_06 #12928 RT: 42.60 AV: 1 NL: 4.06E5
 T: FTMS + c NSI d sa Full ms2 812.3802@etd150.00 812.3802@hcd15.00 [150.0000-2000.0000]



180420_05 #7709 RT: 28.43 AV: 1 NL: 2.08E6
 T: FTMS + c NSI d Full ms2 1047.4724@hcd28.00 [100.0000-2000.0000]

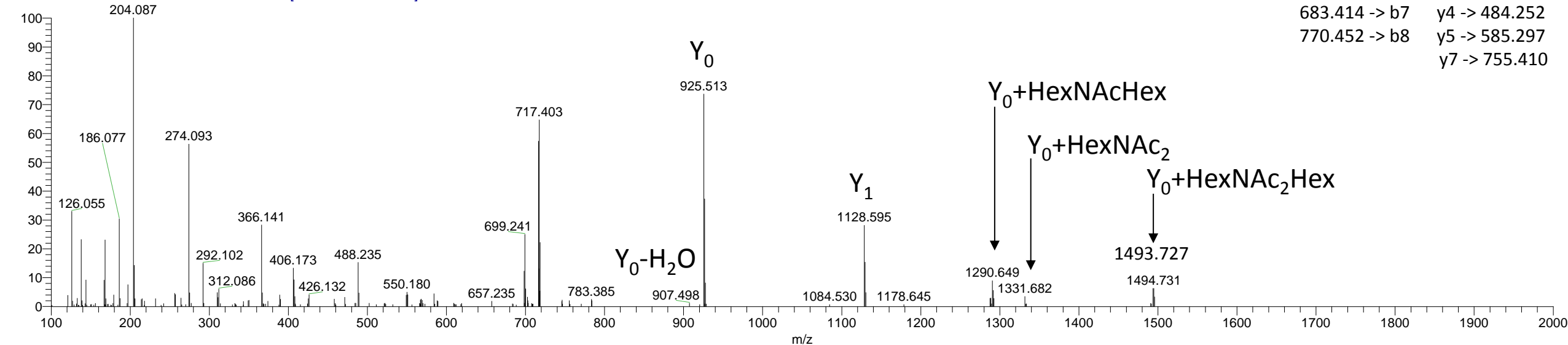


180420_05 #7710 RT: 28.44 AV: 1 NL: 4.31E5
 T: FTMS + c NSI d sa Full ms2 1047.4724@etd150.00 1047.4724@hcd15.00 [150.0000-2000.0000]

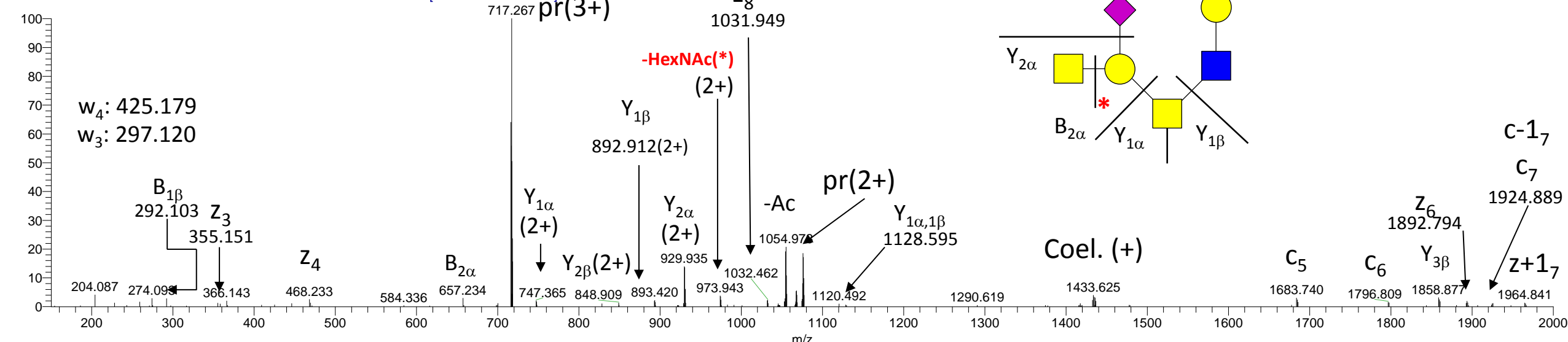


180417_13 #11762 RT: 31.89 AV: 1 NL: 7.39E5
 T: FTMS + c NSI d Full ms2 717.3212@hcd28.00 [100.0000-2000.0000]

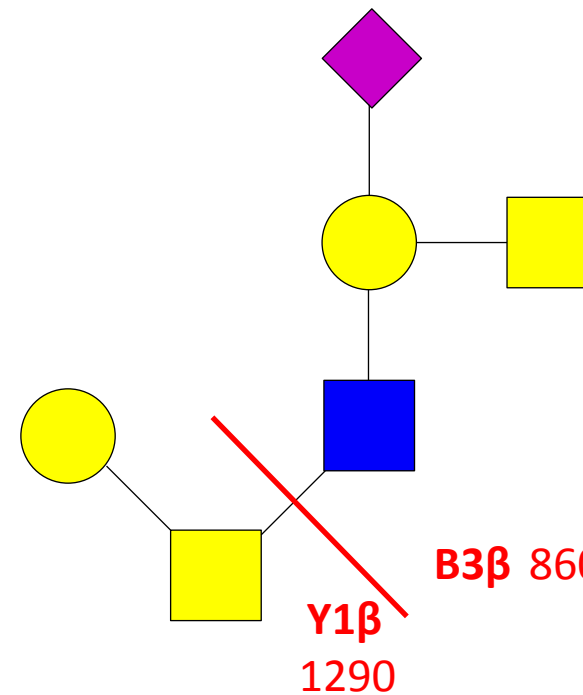
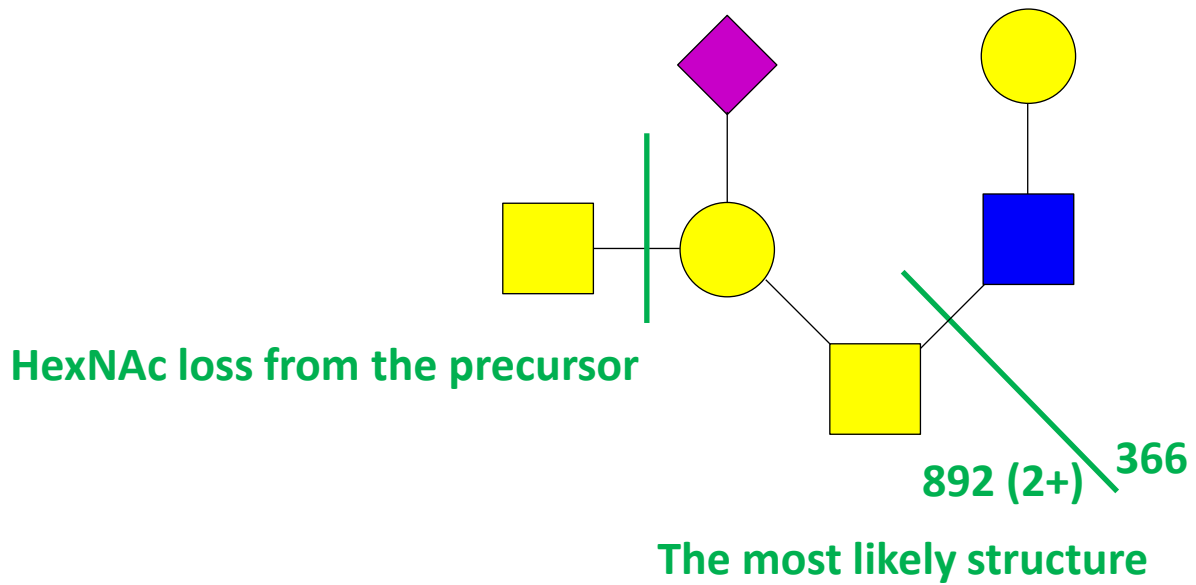
242.151 → b2 y1 → 156.077
 683.414 → b7 y4 → 484.252
 770.452 → b8 y5 → 585.297
 y7 → 755.410



180417_13 #11764 RT: 31.89 AV: 1 NL: 7.39E5
 T: FTMS + c NSI d sa Full ms2 717.3212@etd70.13 717.3212@hcd15.00 [150.0000-2000.0000]

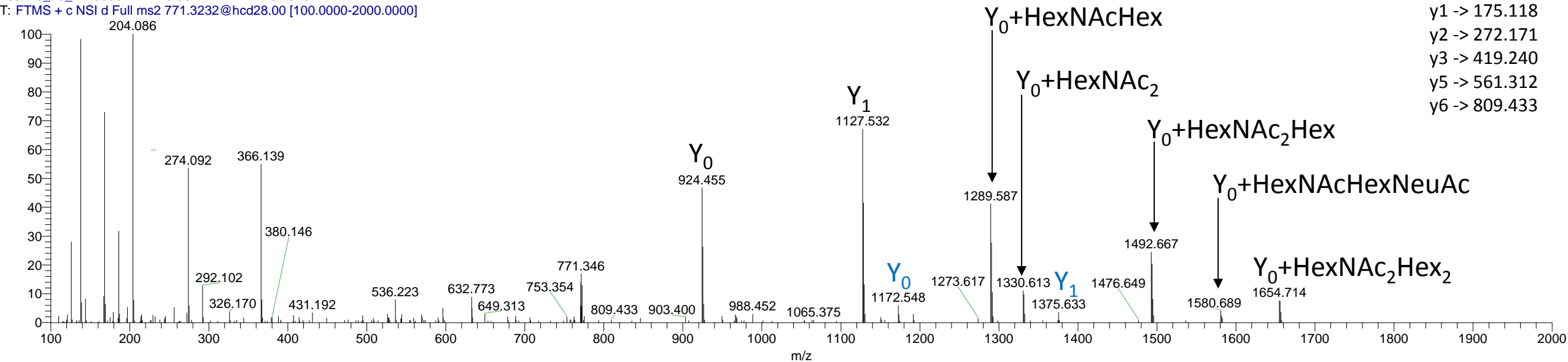


The other option



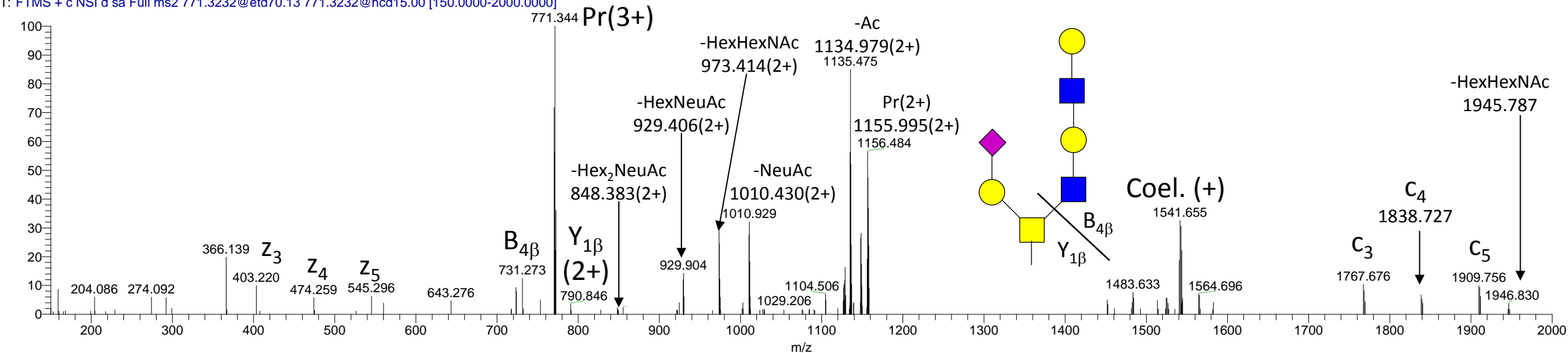
These fragment ions were not detected

180420_19_i #15300 RT: 43.98 AV: 1 NL: 8.43E5
 T: FTMS + c NSI d Full ms2 771.3232@hcd28.00 [100.0000-2000.0000]

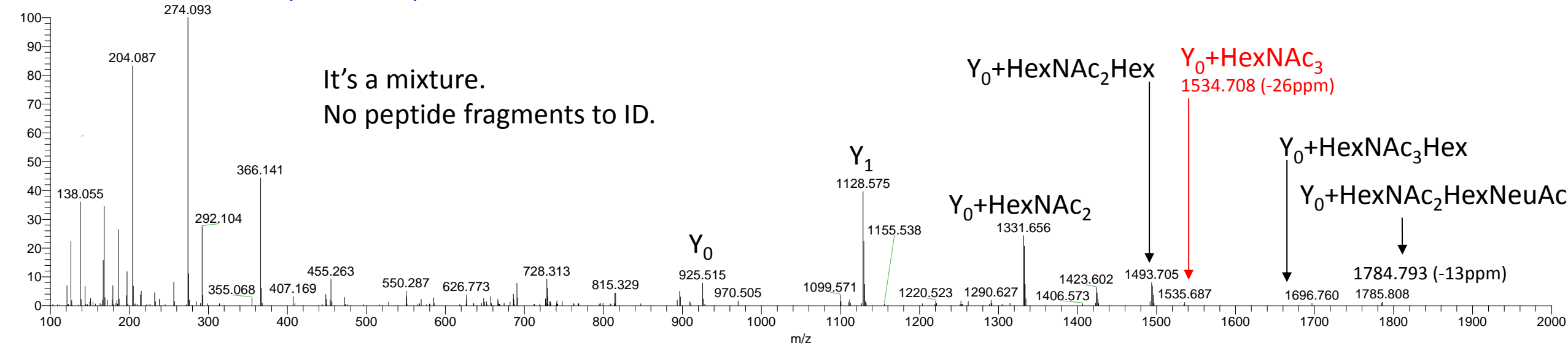


y1 -> 175.118
 y2 -> 272.171
 y3 -> 419.240
 y5 -> 561.312
 y6 -> 809.433

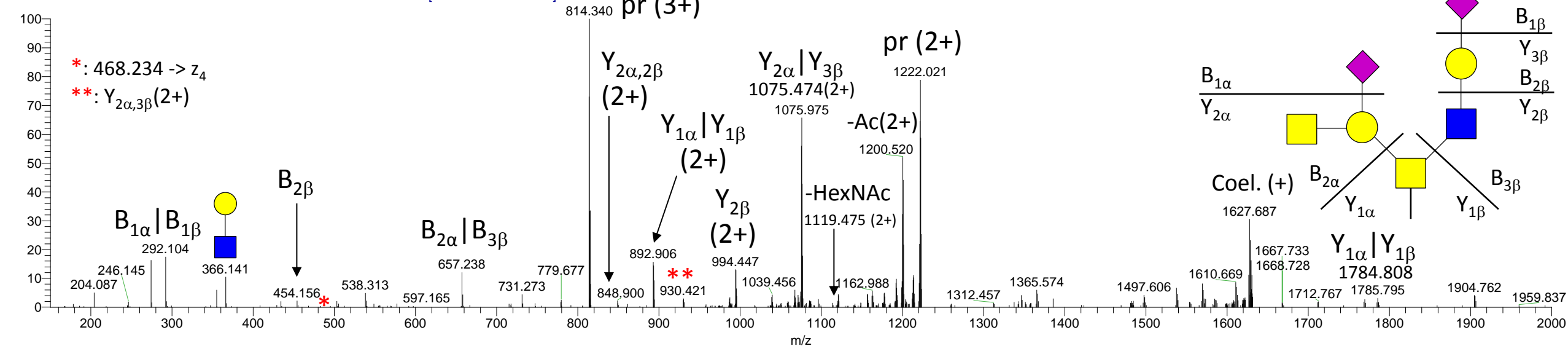
180420_19_i #15302 RT: 43.99 AV: 1 NL: 3.92E5
 T: FTMS + c NSI d sa Full ms2 771.3232@etd70.13 771.3232@hcd15.00 [150.0000-2000.0000]



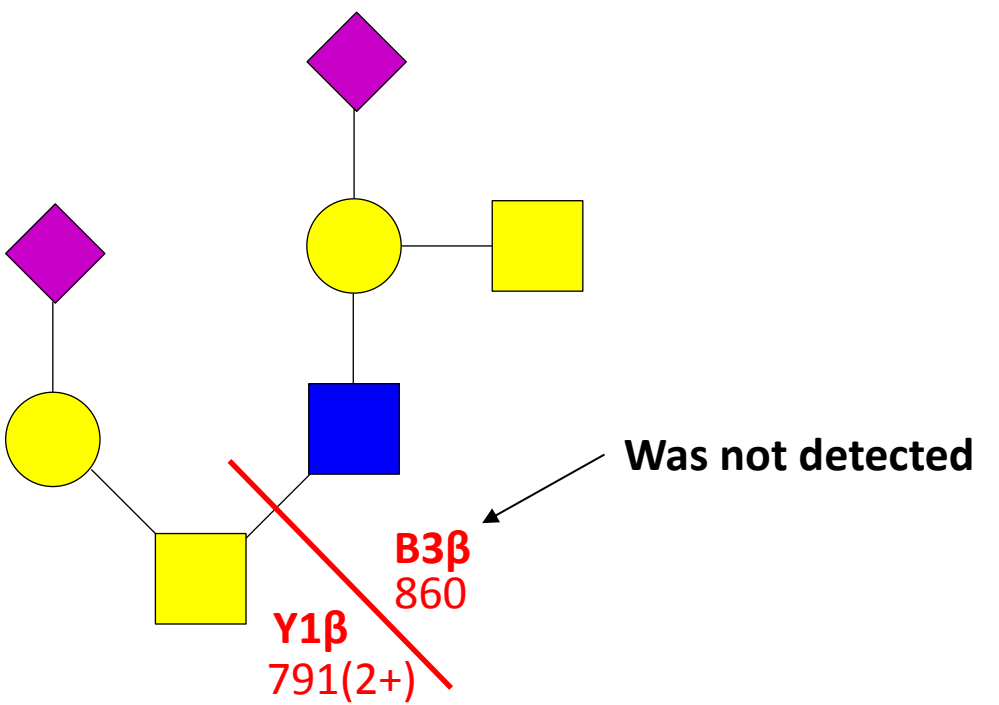
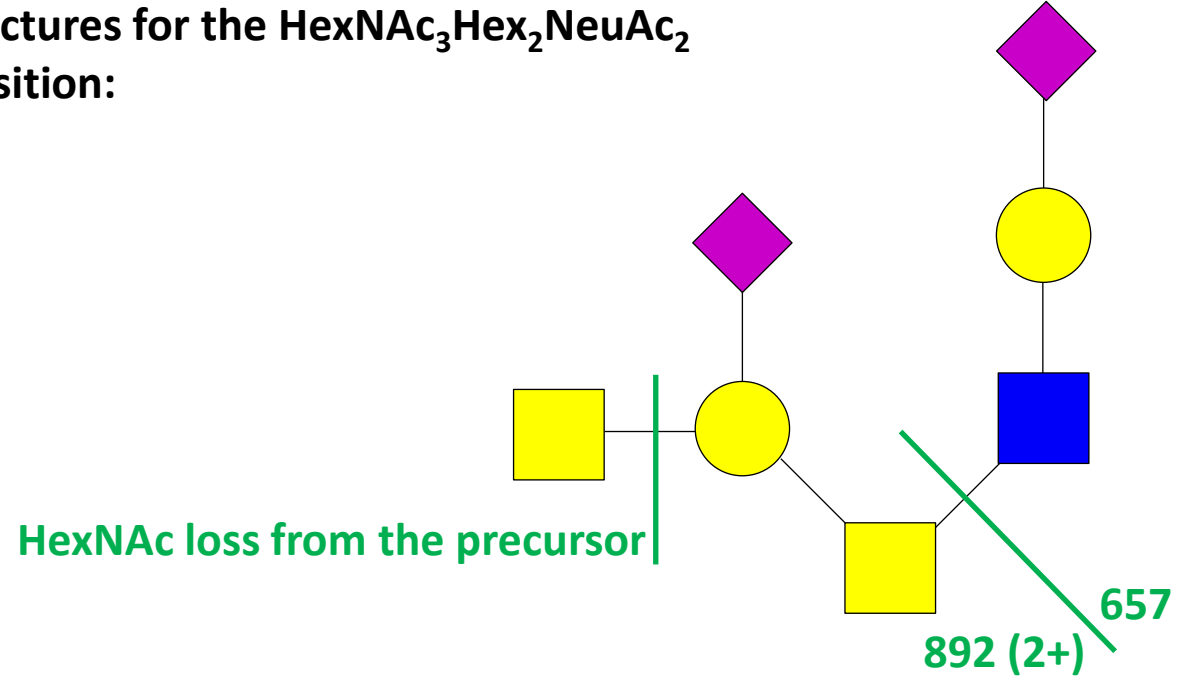
180417_05 #10057 RT: 28.93 AV: 1 NL: 6.97E5
 T: FTMS + c NSI d Full ms2 814.6827@hcd28.00 [100.0000-2000.0000]



180417_05 #10059 RT: 28.94 AV: 1 NL: 2.08E5
 T: FTMS + c NSI d sa Full ms2 814.6827@etd70.13 814.6827@hcd15.00 [150.0000-2000.0000]

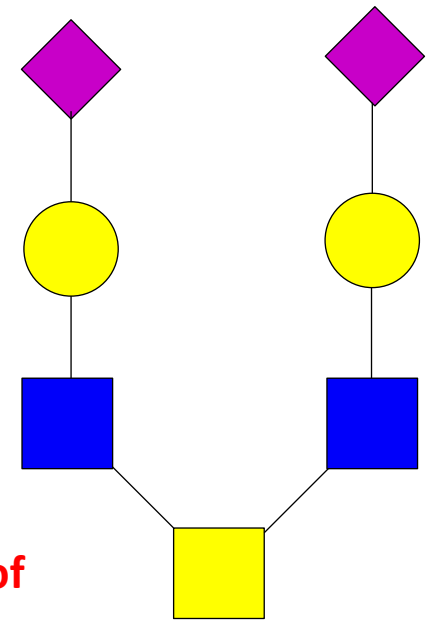


Potential other structures for the HexNAc₃Hex₂NeuAc₂ (1515.5346) composition:



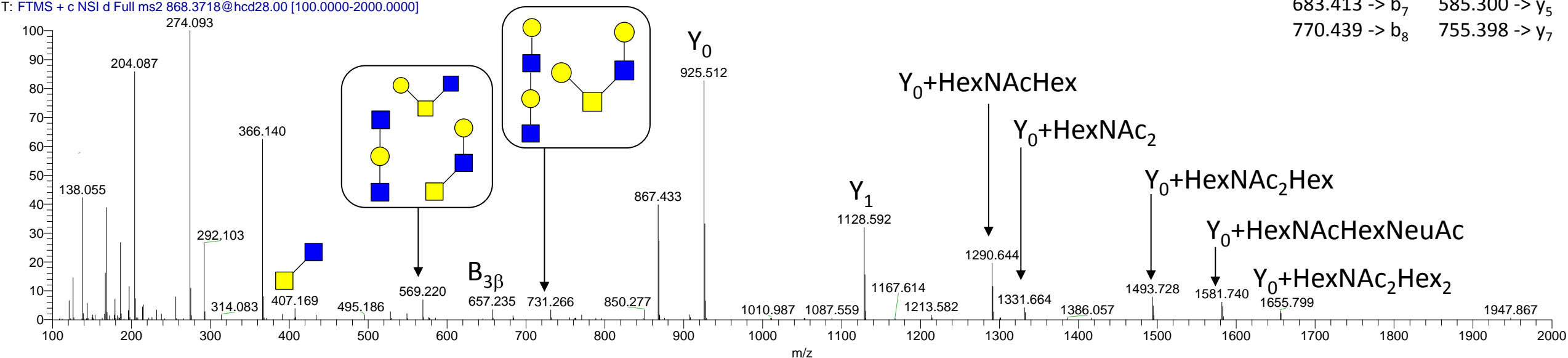
Other options

The HCD fragment that could prove this structure (peptide+HexNAc₃) is outside of the acceptable mass accuracy range.



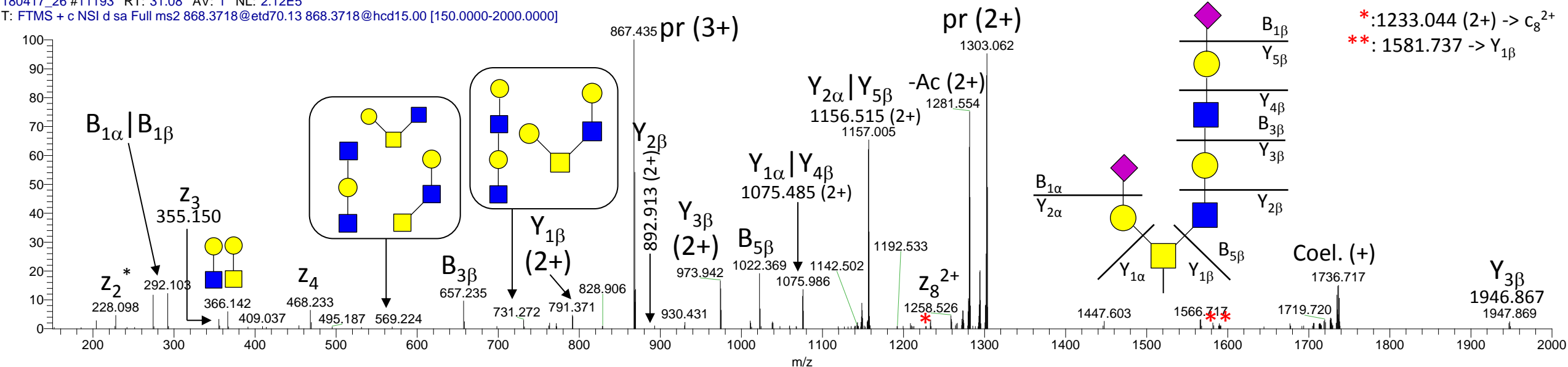
180417_26 #11192 RT: 31.07 AV: 1 NL: 6.42E5
 T: FTMS + c NSI d Full ms2 868.3718@hcd28.00 [100.0000-2000.0000]

683.413 → b₇ 585.300 → y₅
 770.439 → b₈ 755.398 → y₇



180417_26 #11193 RT: 31.08 AV: 1 NL: 2.12E5
 T: FTMS + c NSI d sa Full ms2 868.3718@etd70.13 868.3718@hcd15.00 [150.0000-2000.0000]

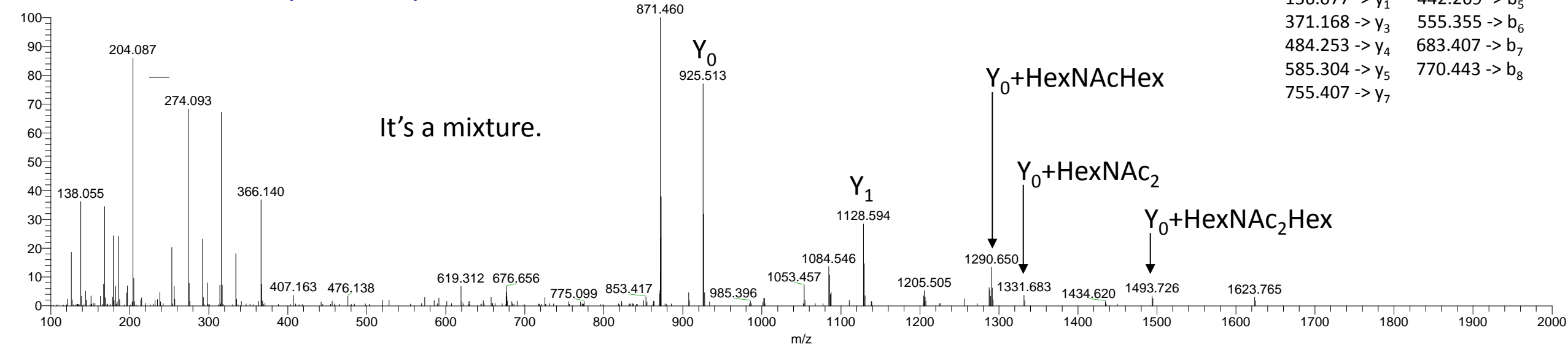
*:1233.044 (2+) → c₈²⁺
 **:1581.737 → Y_{1β}



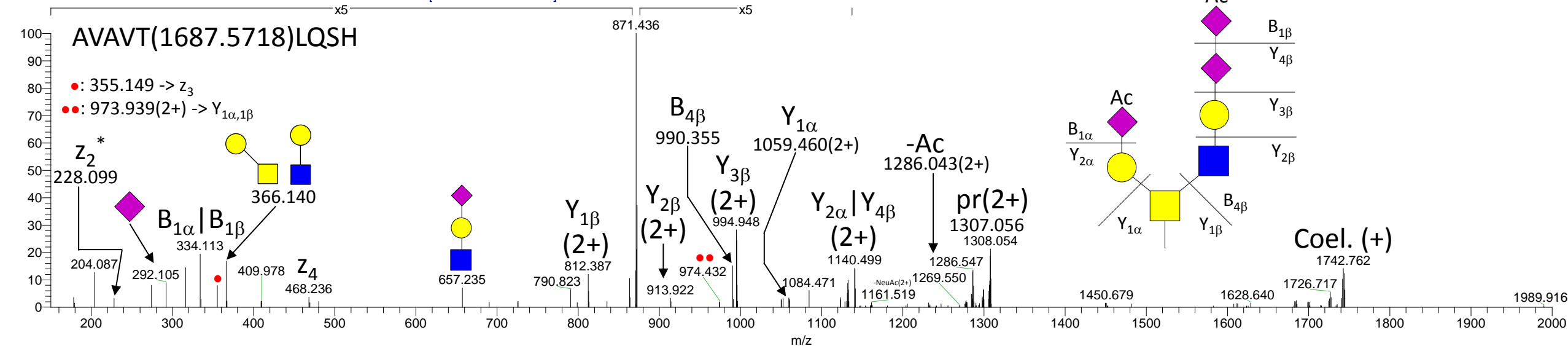
180417_07 #16150 RT: 41.52 AV: 1 NL: 6.35E5
 T: FTMS + c NSI d Full ms2 871.7020@hcd28.00 [100.0000-2000.0000]

156.077 → y₁ 442.269 → b₅
 371.168 → y₃ 555.355 → b₆
 484.253 → y₄ 683.407 → b₇
 585.304 → y₅ 770.443 → b₈
 755.407 → y₇

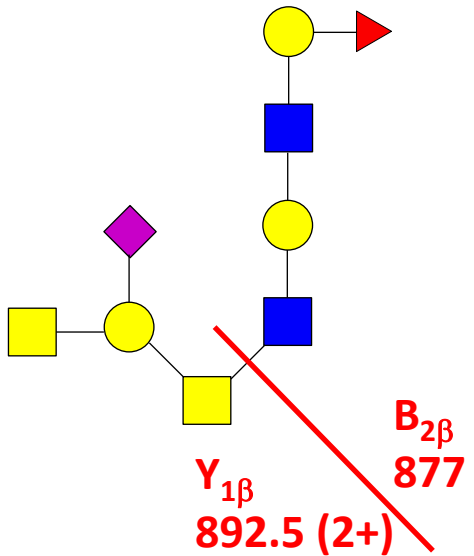
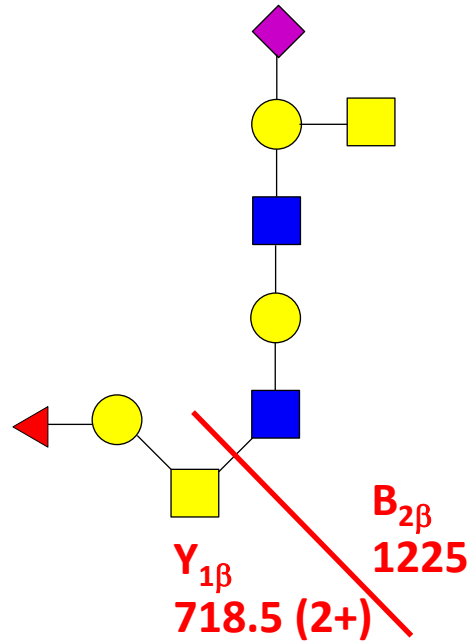
It's a mixture.



180417_07 #16152 RT: 41.52 AV: 1 NL: 9.44E5
 T: FTMS + c NSI d sa Full ms2 871.7020@etd70.13 871.7020@hcd15.00 [150.0000-2000.0000]

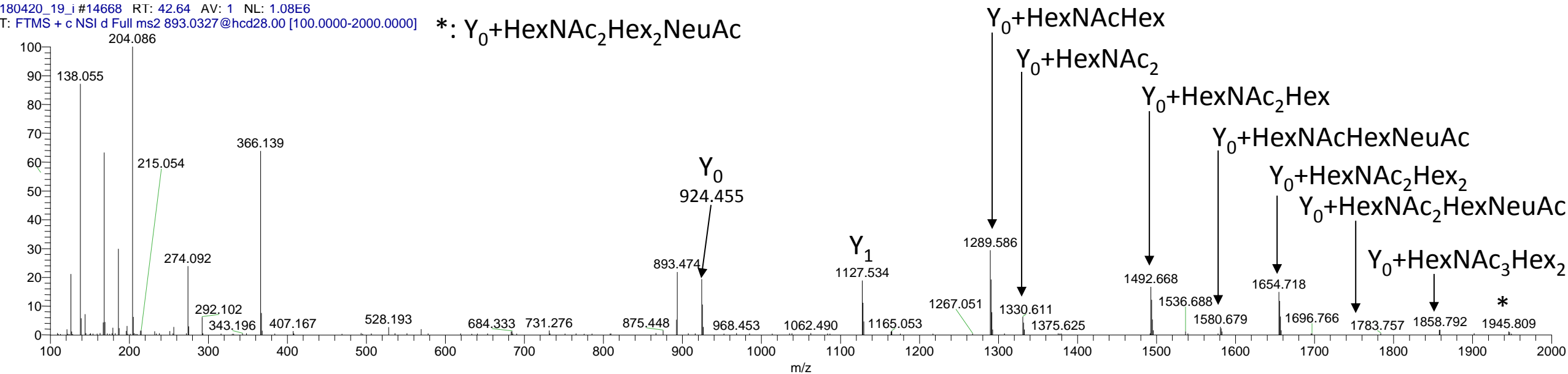


Another potential structure for the HexNAc₄Hex₃FucNeuAc (1735.6293) composition:
(The most likely structure and the supporting EThcD spectrum can be found in the article as Figure 2.)



180420_19_i#14668 RT: 42.64 AV: 1 NL: 1.08E6
 T: FTMS + c NSI d Full ms2 893.0327@hcd28.00 [100.0000-2000.0000]

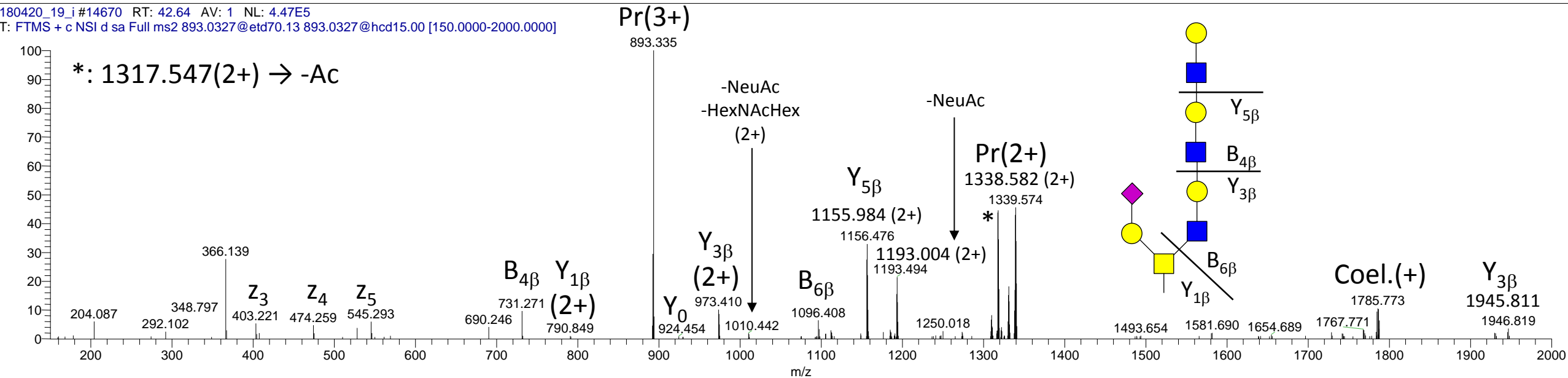
*: Y₀+HexNAc₂Hex₂NeuAc



180420_19_i#14670 RT: 42.64 AV: 1 NL: 4.47E5
 T: FTMS + c NSI d sa Full ms2 893.0327@etd70.13 893.0327@hcd15.00 [150.0000-2000.0000]

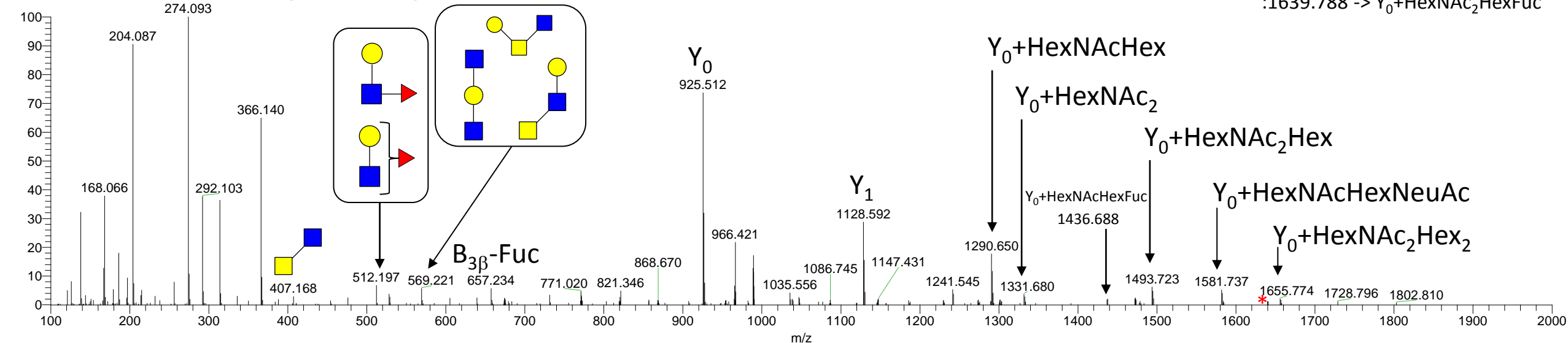
Pr(3+)

*: 1317.547(2+) → -Ac



180417_07 #9478 RT: 28.95 AV: 1 NL: 1.31E6
T: FTMS + c NSI d Full ms2 966.0764@hcd28.00 [100.0000-2000.0000]

*:1639.788 → Y₀+HexNAc₂HexFuc

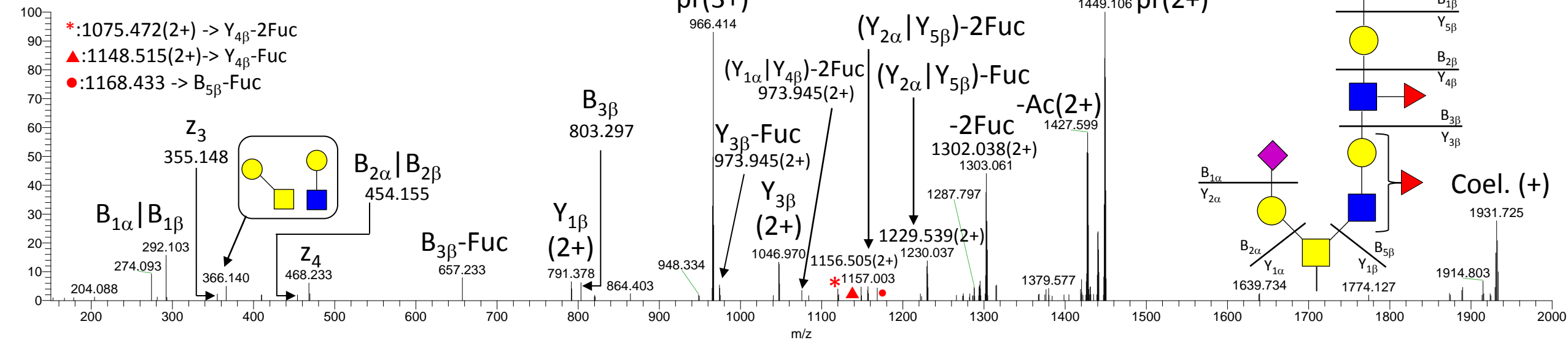


180417_07 #9480 RT: 28.96 AV: 1 NL: 6.10E5
T: FTMS + c NSI d sa Full ms2 966.0764@etd70.13 966.0764@hcd15.00 [150.0000-2000.0000]

*:1075.472(2+) → Y_{4β}-2Fuc

▲:1148.515(2+) → Y_{4β}-Fuc

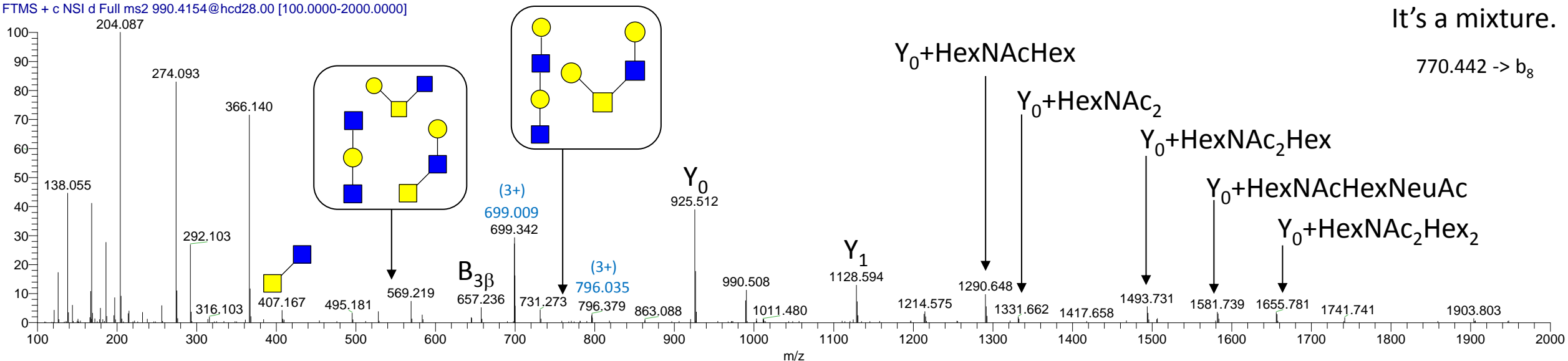
●:1168.433 → B_{5β}-Fuc



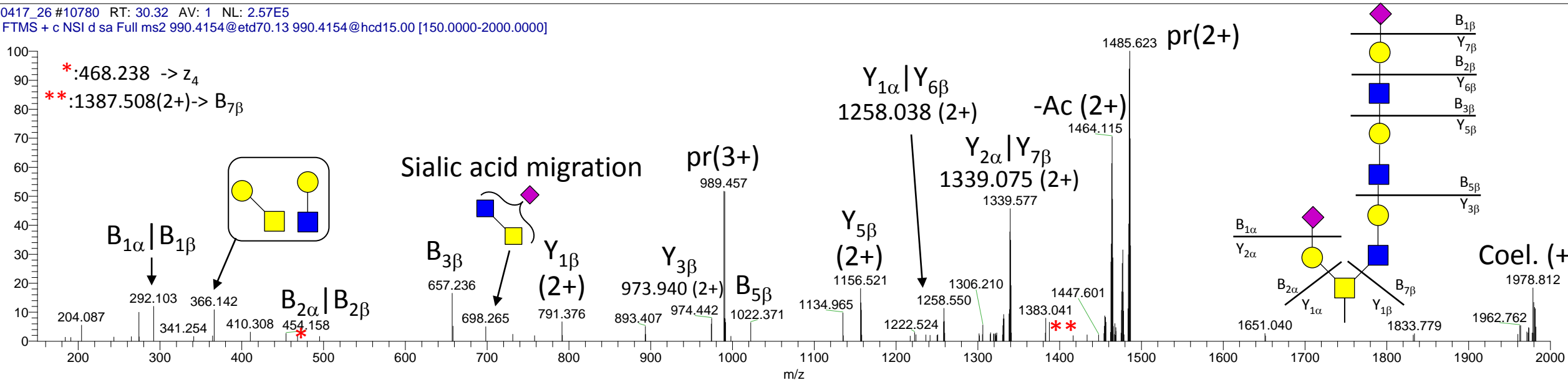
180417_26 #10778 RT: 30.31 AV: 1 NL: 8.06E5
 T: FTMS + c NSI d Full ms2 990.4154@hcd28.00 [100.0000-2000.0000]

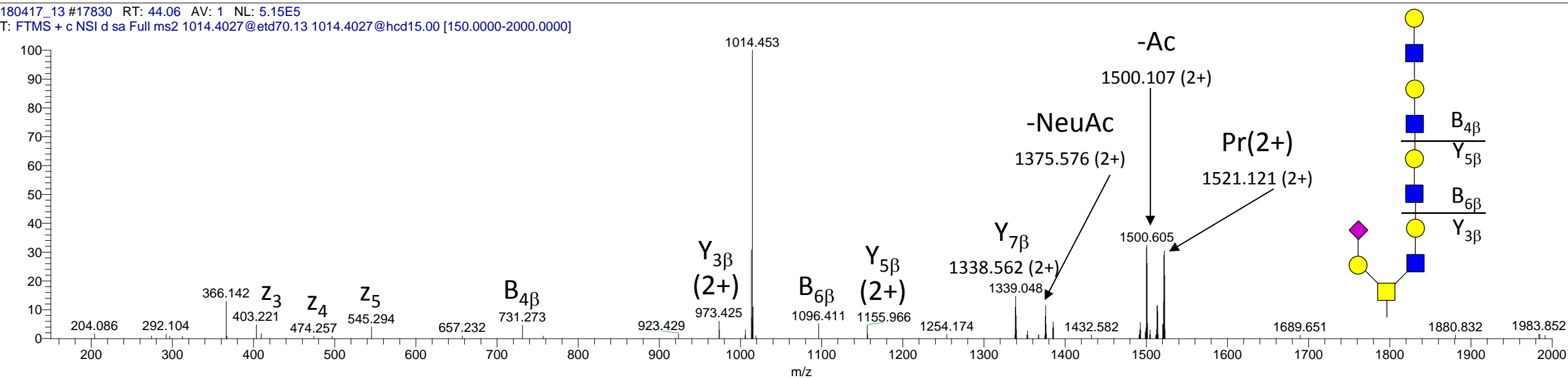
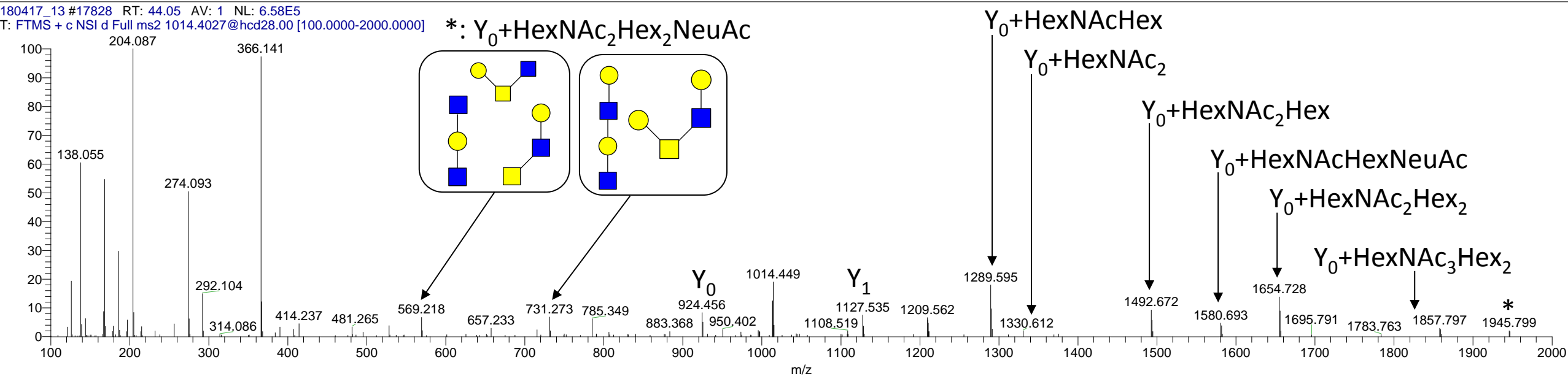
It's a mixture.

770.442 → b₈

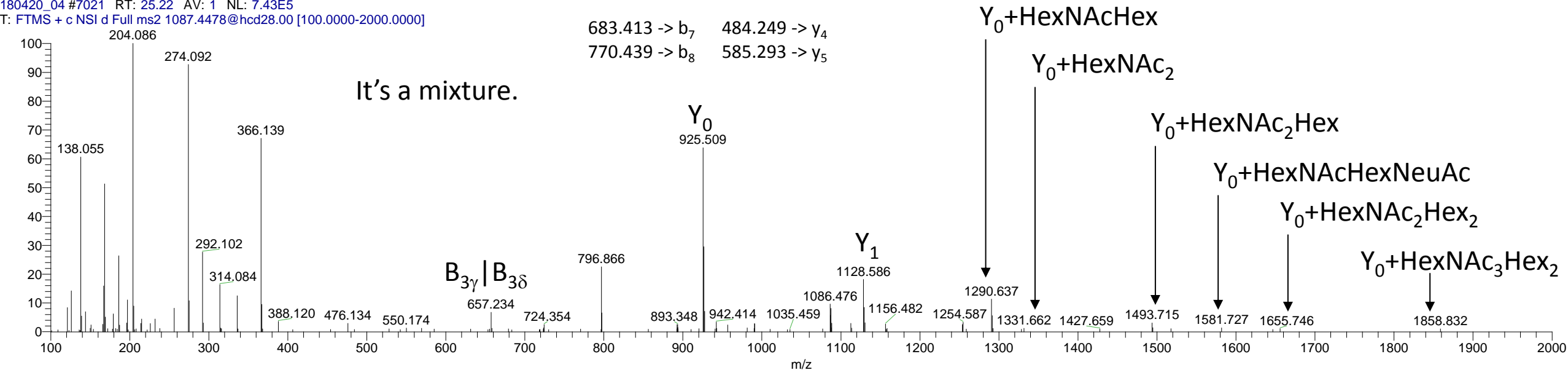


180417_26 #10780 RT: 30.32 AV: 1 NL: 2.57E5
 T: FTMS + c NSI d sa Full ms2 990.4154@etd70.13 990.4154@hcd15.00 [150.0000-2000.0000]

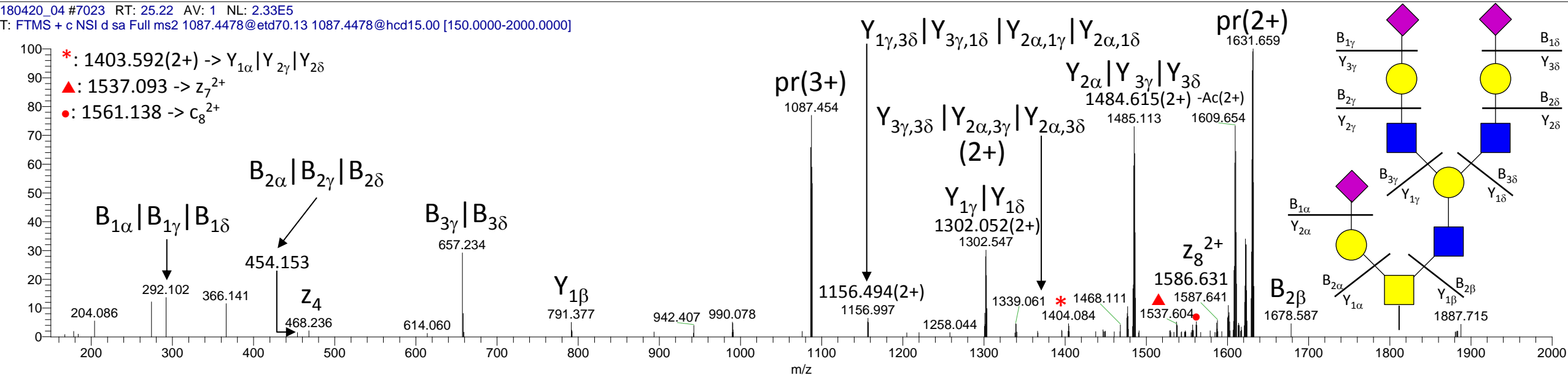




180420_04 #7021 RT: 25.22 AV: 1 NL: 7.43E5
 T: FTMS + c NSI d Full ms2 1087.4478@hcd28.00 [100.0000-2000.0000]

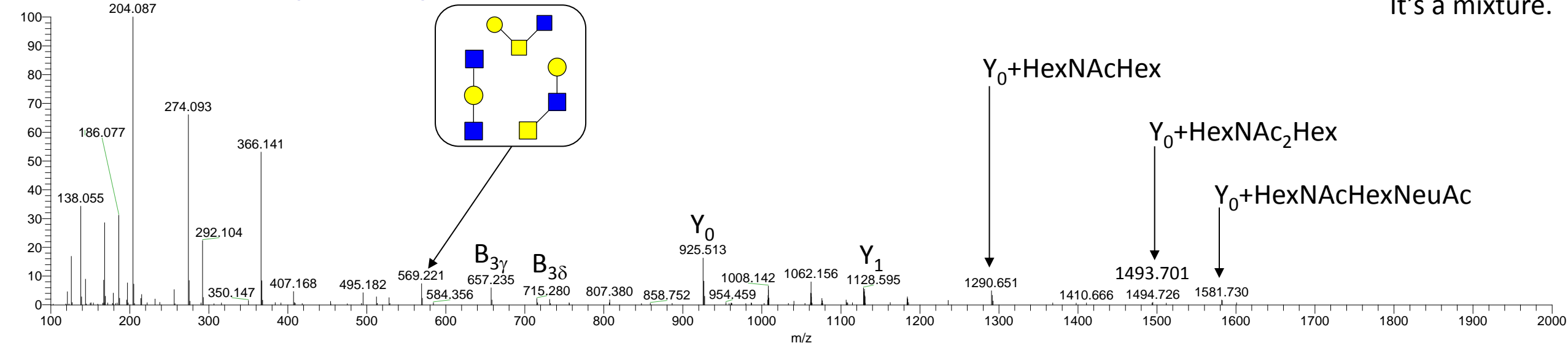


180420_04 #7023 RT: 25.22 AV: 1 NL: 2.33E5
 T: FTMS + c NSI d sa Full ms2 1087.4478@etd70.13 1087.4478@hcd15.00 [150.0000-2000.0000]

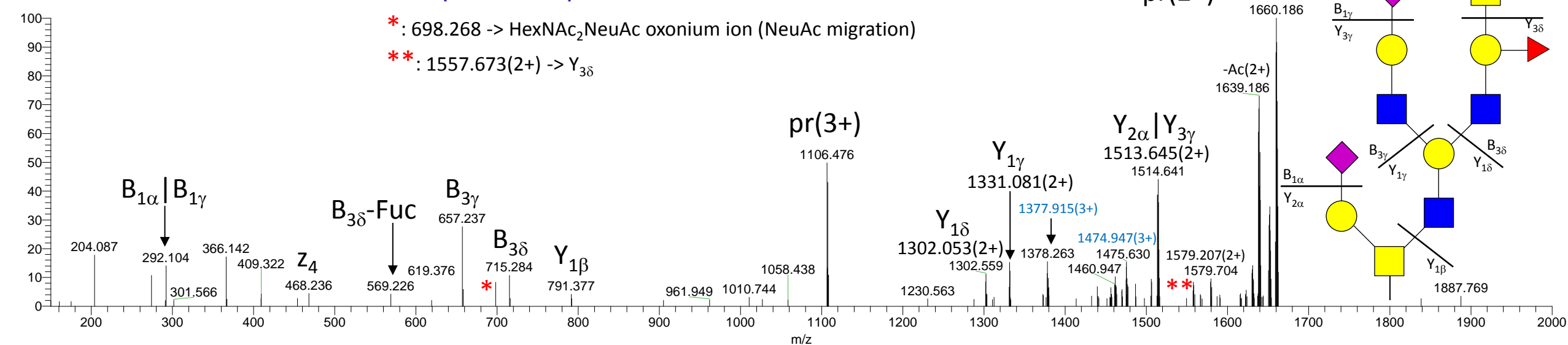


180417_13 #10026 RT: 28.53 AV: 1 NL: 1.24E6
 T: FTMS + c NSI d Full ms2 1107.4537@hcd28.00 [100.0000-2000.0000]

It's a mixture.

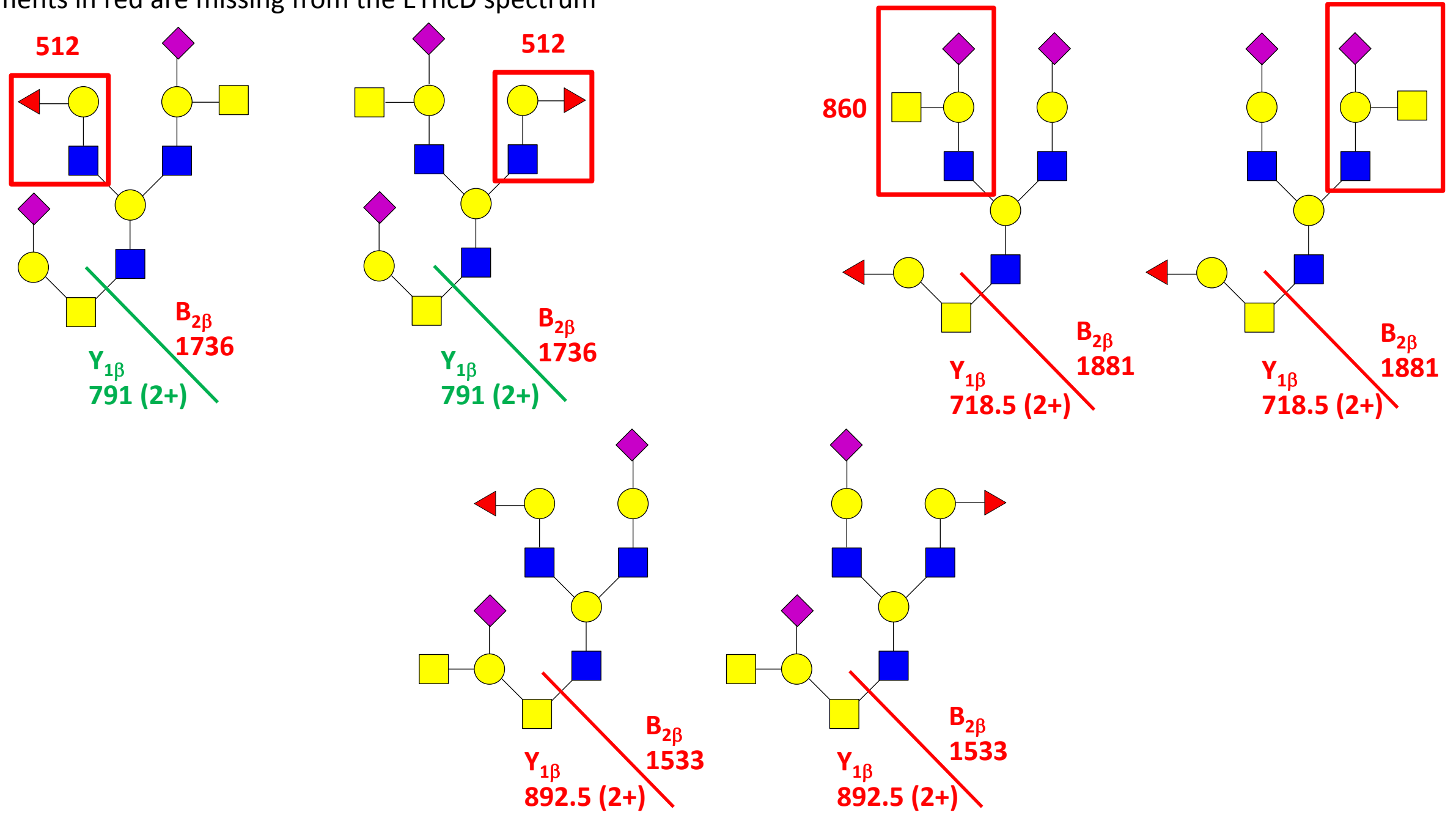


180417_13 #10029 RT: 28.53 AV: 1 NL: 1.68E5
 T: FTMS + c NSI d sa Full ms2 1107.4537@etd70.13 1107.4537@hcd15.00 [150.0000-2000.0000]



Additional potential structures for the HexNAc₅Hex₄FucNeuAc₂ (2391.8569) composition:

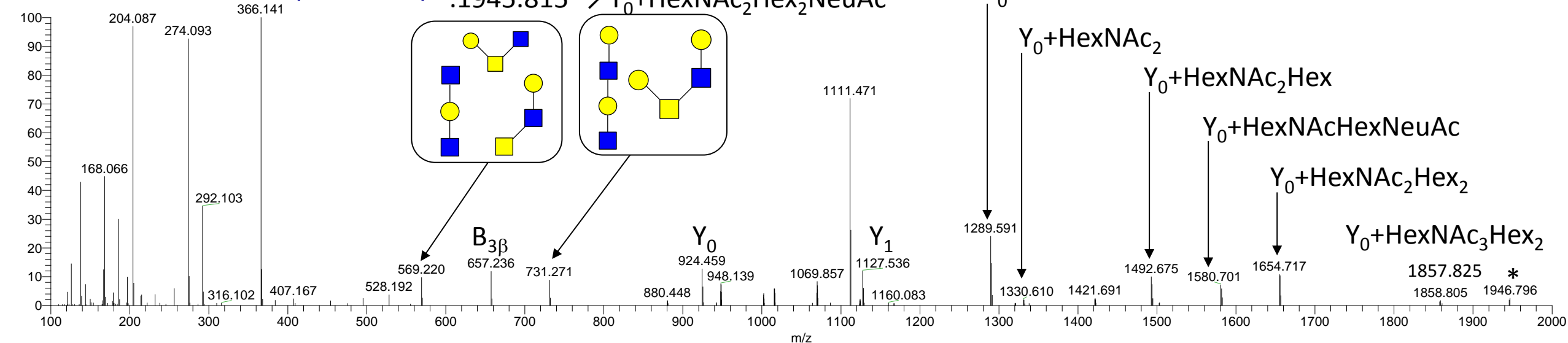
Fragments in red are missing from the EThcD spectrum



180417_13 #17623 RT: 43.64 AV: 1 NL: 1.30E6

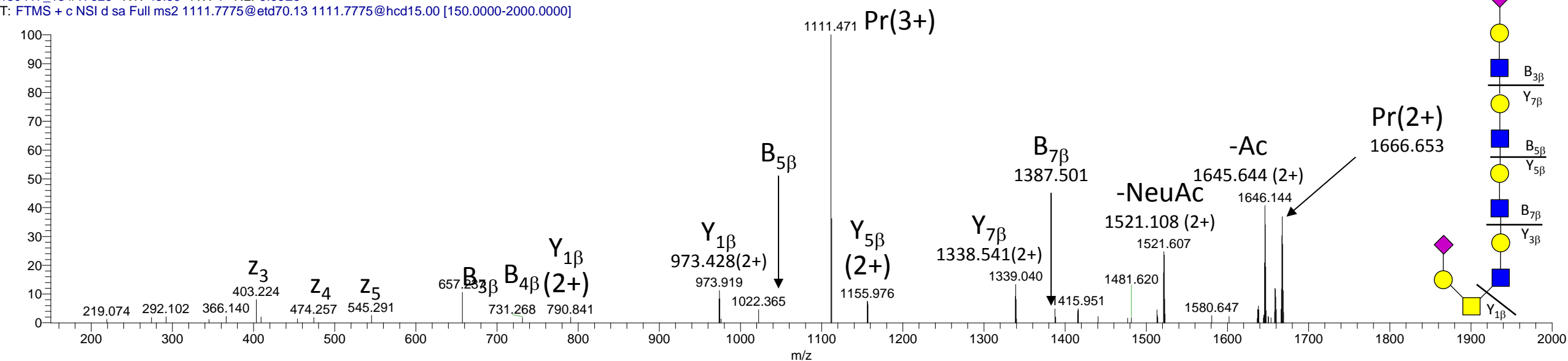
T: FTMS + c NSI d Full ms2 1111.7775@hcd28.00 [100.0000-2000.0000]

*:1945.815 → Y₀+HexNAC₂Hex₂NeuAc

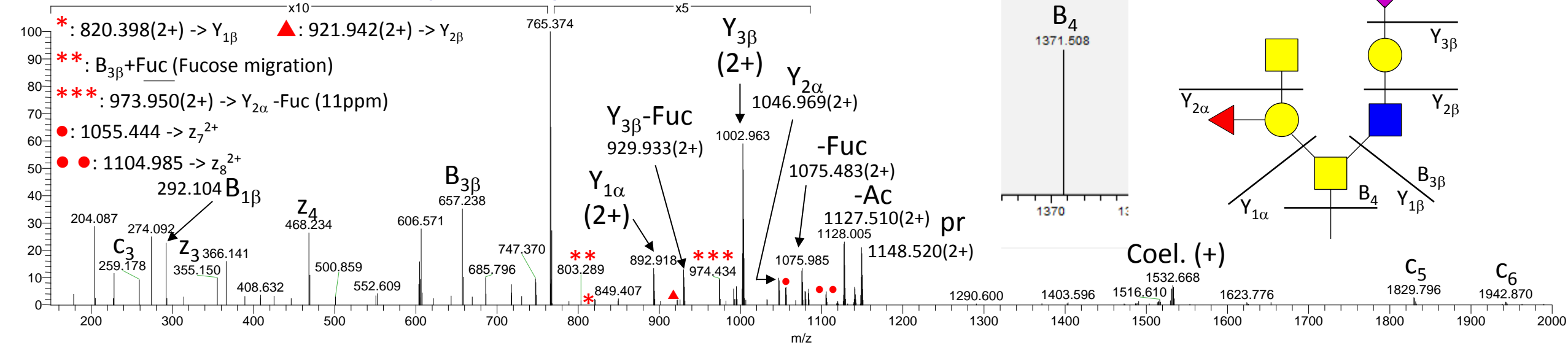


180417_13 #17626 RT: 43.65 AV: 1 NL: 9.66E5

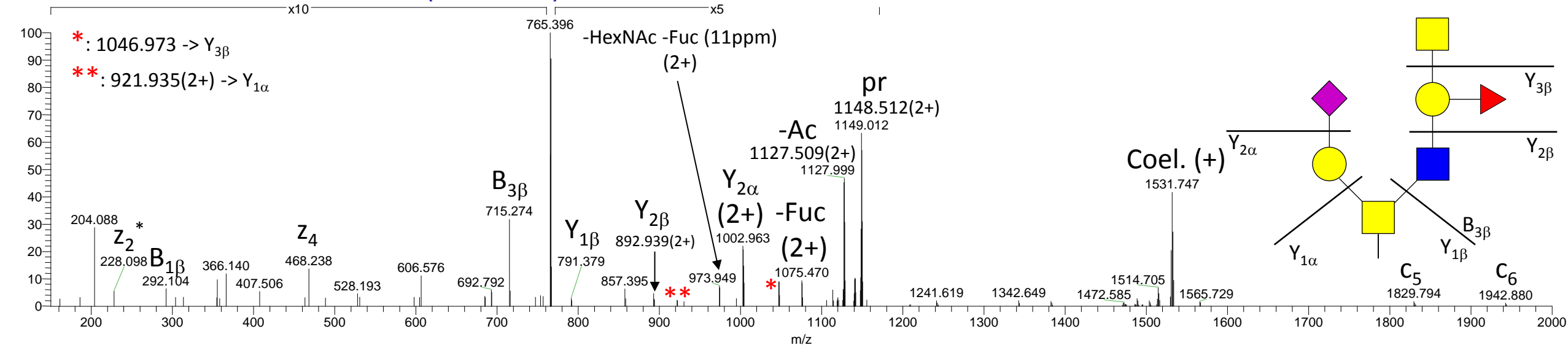
T: FTMS + c NSI d sa Full ms2 1111.7775@etd70.13 1111.7775@hcd15.00 [150.0000-2000.0000]



180417_15 #9852 RT: 29.30 AV: 1 NL: 1.16E6
T: FTMS + c NSI d sa Full ms2 766.3433@etd70.13 766.3433@hcd15.00 [150.0000-2000.0000]



180417_15 #11112 RT: 31.65 AV: 1 NL: 1.14E6
T: FTMS + c NSI d sa Full ms2 766.0082@etd70.13 766.0082@hcd15.00 [150.0000-2000.0000]



Additional potential structures for the HexNAc₃Hex₂FucNeuAc (1370.4971) composition:

Fragments in red are missing from the EThcD spectrum

Neutral loss of HexNeuAc (921 (2+)) from the precursor is present which cannot be explained with this structure.

