

Improved identification of O-linked glycopeptides from ETD data with optimized scoring for different charge states and cleavage specificities

Amino Acids

Zs. Darula¹, R.J. Chalkley², A. Lynn², P. Baker² & K.F. Medzihradszky^{1,2} *

¹ Proteomics Research Group, Biological Research Center, Szeged, Hungary

² Mass Spectrometry Facility, Department of Pharmaceutical Chemistry,
University of California San Francisco, USA

* Corresponding author e-mail: folkl@cgl.ucsf.edu

Legends

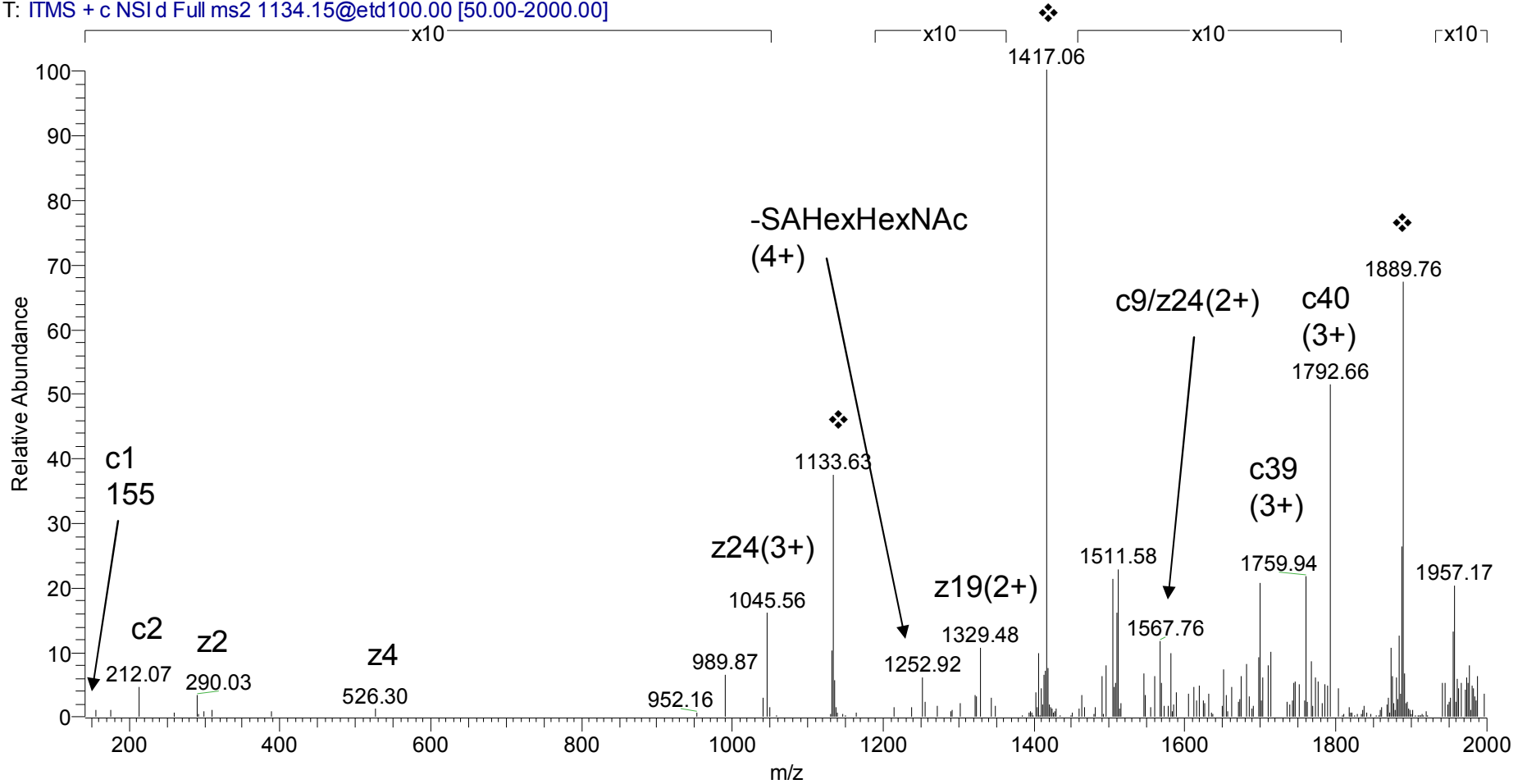
For ETD data

- $z^* = z+1$;
- $c^* = c-1$;
- ❖ = charge reduced ions
- * = 2x precursor ion;
- **bold** = z ion detected;
- underlined = c ion detected;
- *italic* = b or y ion detected

A5D7R6, ITIH2 protein

1133.5541(5+)

T8110317 #3466 RT: 39.40 AV: 1 NL: 4.12E4
T: ITMS + c NSI d Full ms2 1134.15@etd100.00 [50.00-2000.00]



⁶⁵⁷HGSKVS(GaINAcGalSA)PNSVPSWVNPSPAPVLPMPAVGAQVLES(GaINAcGalSA)-
TPPPHVMR⁶⁹⁸

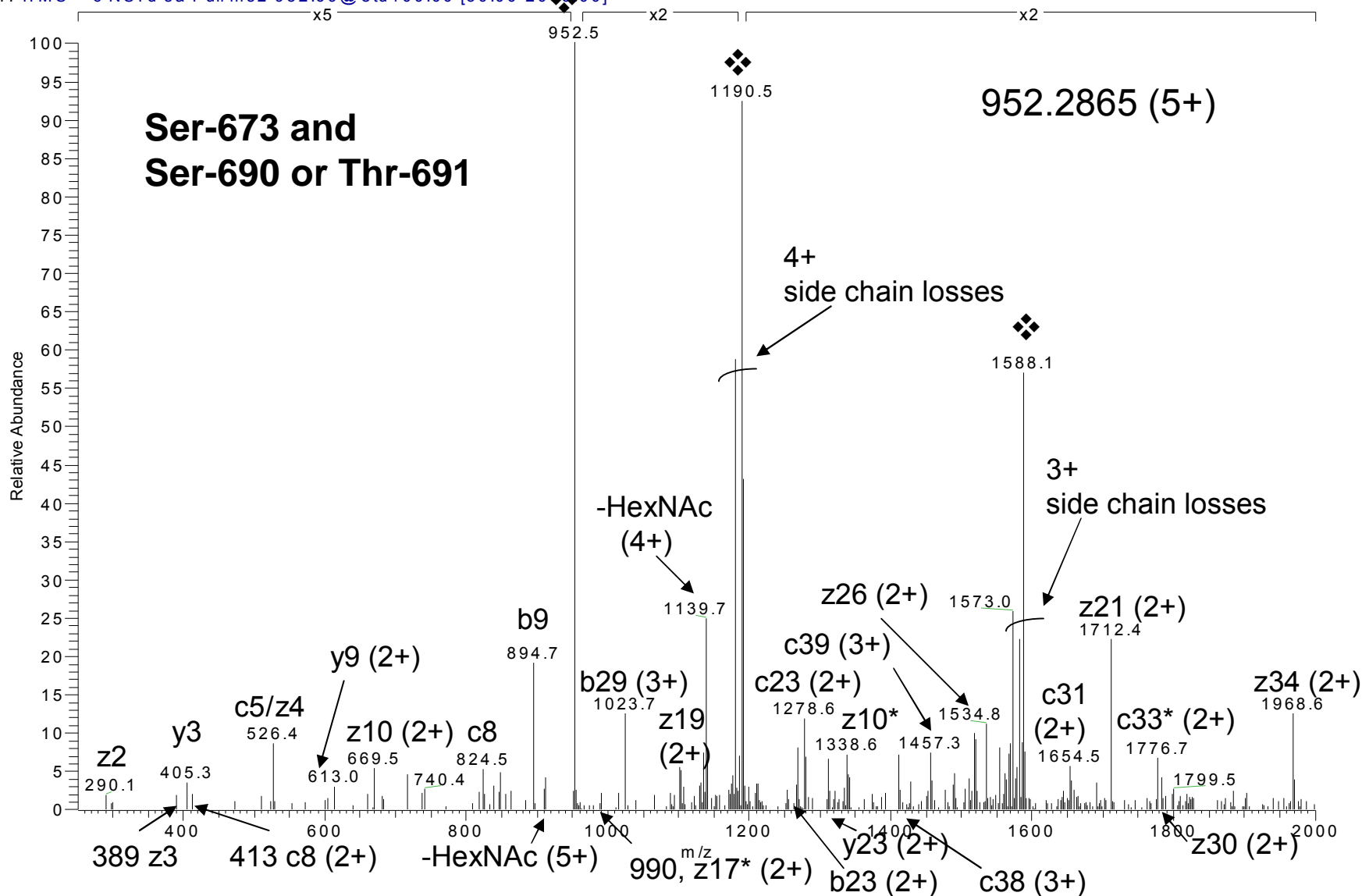
**Ser-659 or 662 or 665 or 668 or 673
and Ser-690 or Thr-691**

Supplemental Figure 1

A5D7R6, ITIH2 protein

Supplemental Figure 2

T8102002 #3530 RT: 37.10 AV: 1 NL: 1.10E5
T: ITMS + c NSI d sa Full ms2 952.69@etd100.00 [50.00-2000.00]



HGSKVSPNSVPSWVNPS(HexNAc)PAPVLPMPAVGAQVLEST(HexNAc)PPPHVMR

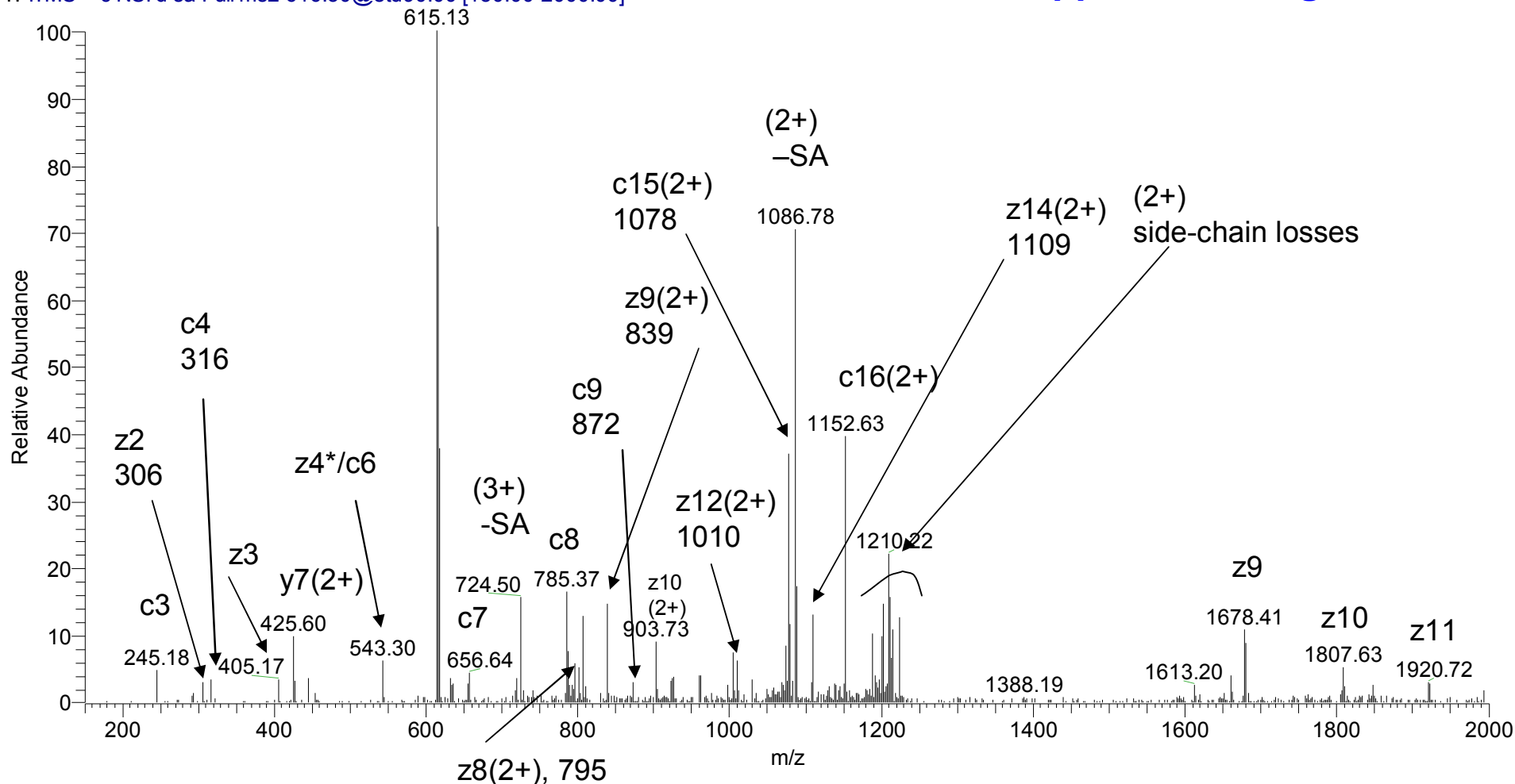
A5D7R6, ITIH2 protein

Thr-691

616.0485 (4+)

Katalin_Szuszanna_53 #2537 RT: 28.86 AV: 1 NL: 1.54E3
T: ITMS + c NSI d sa Full ms2 616.30@etd90.00 [150.00-2000.00]

Supplemental Figure 3



AVGAQVLEST(HexNAcHexSA)PPPHVM(Oxidation)R

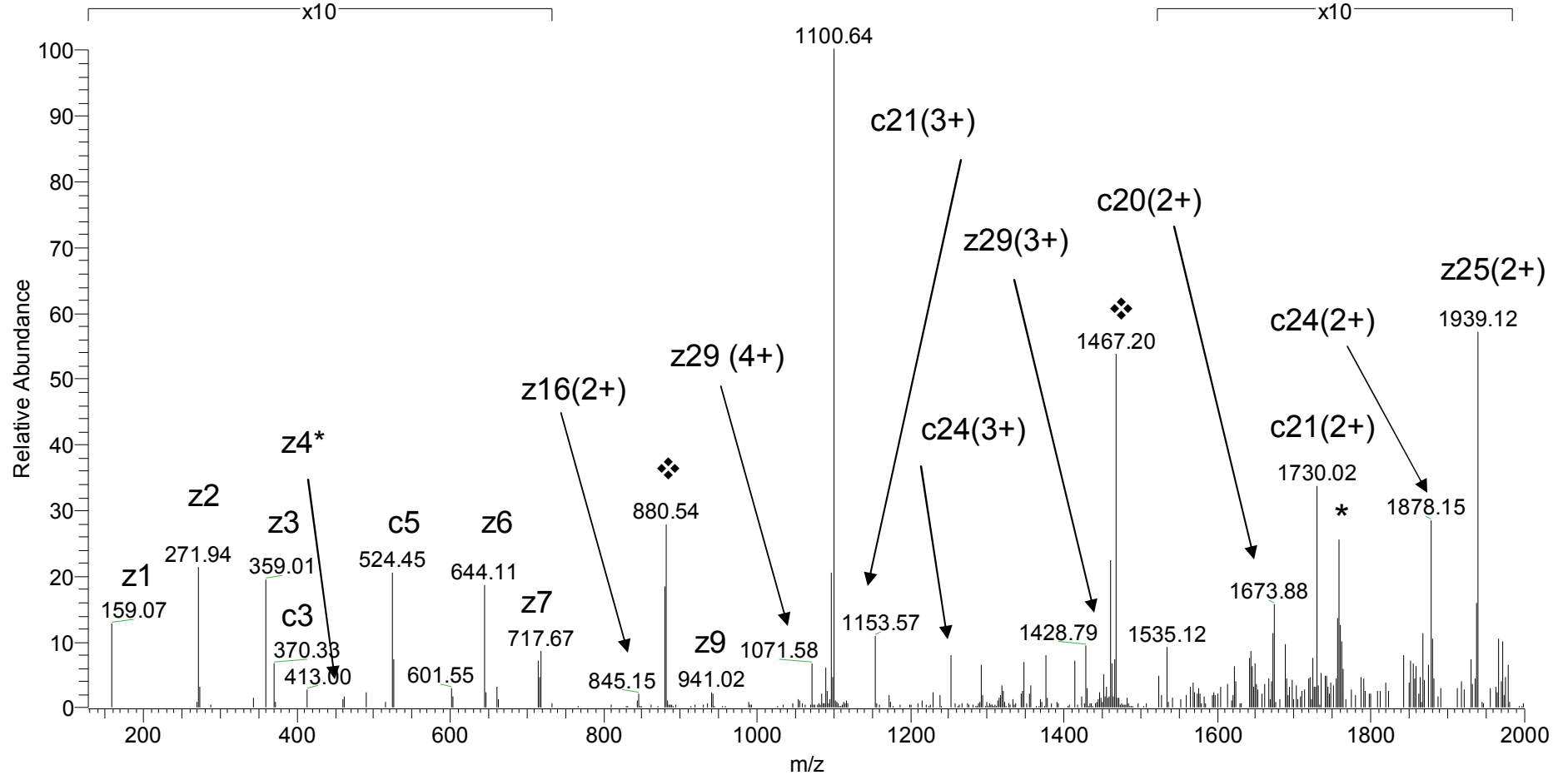
A4IFA5, VASN protein

Ser-455 & Thr-460

Supplemental Figure 4

880.2454 (5+)

T8110317 #2727 RT: 33.97 AV: 1 NL: 7.20E3
T: ITMS + c NSI d Full ms2 880.85@etd100.00 [50.00-2000.00]



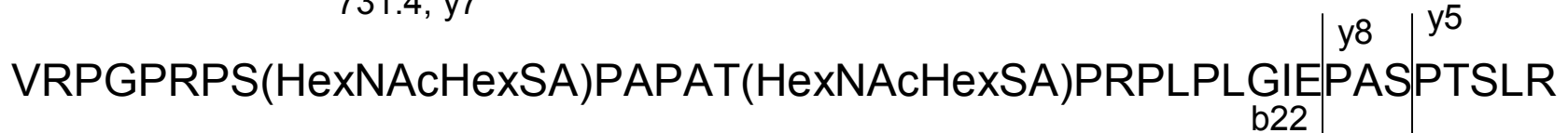
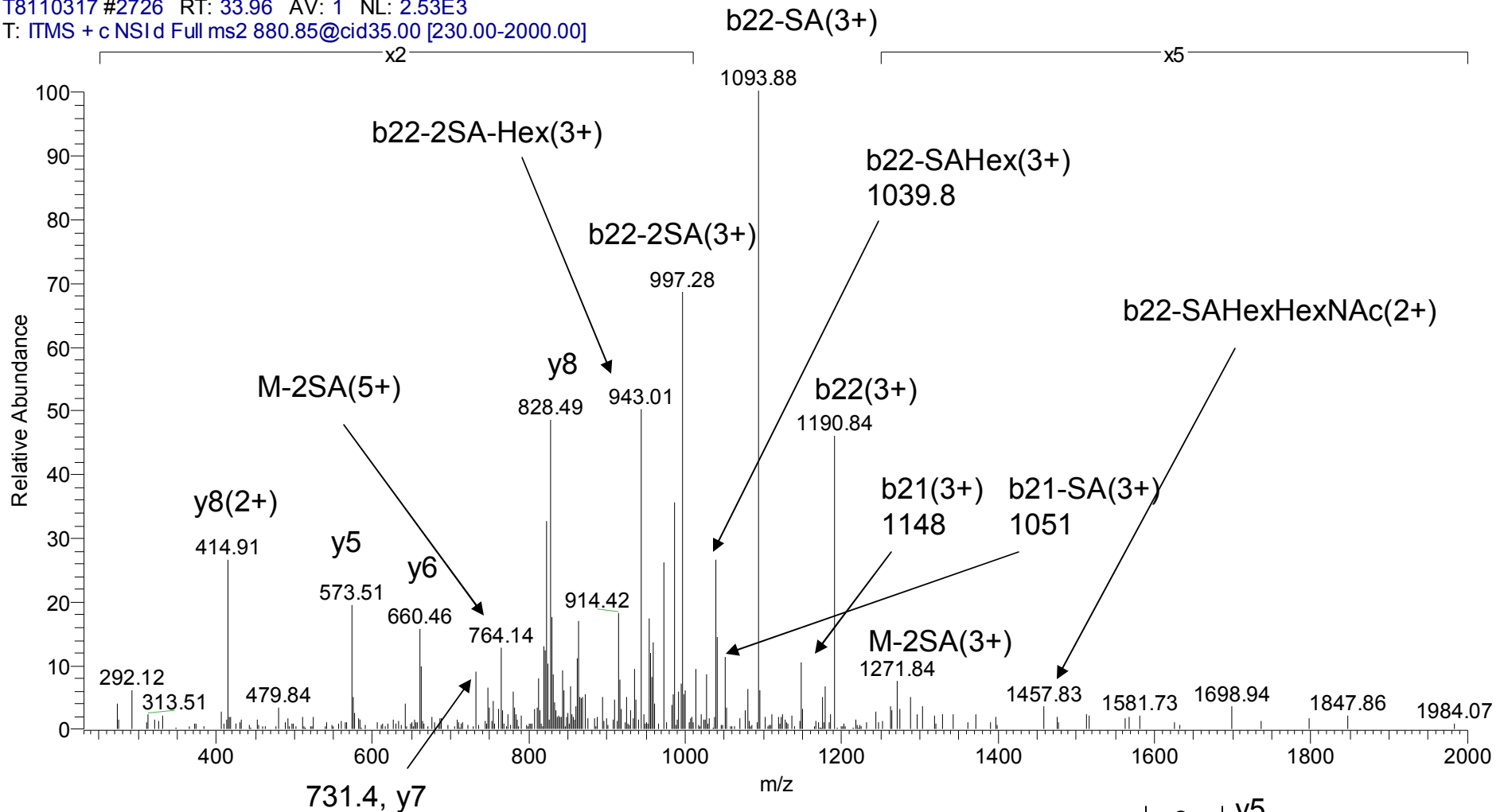
VRPGPRPS(HexNAcHexSA)PAPAT(HexNAcHexSA)PRPLPLGIEPASPTSLR

A4IFA5, VASN protein

880.2454 (5+)

Supplemental Figure 5A

T8110317 #2726 RT: 33.96 AV: 1 NL: 2.53E3
T: ITMS + c NSI d Full ms2 880.85@cid35.00 [230.00-2000.00]

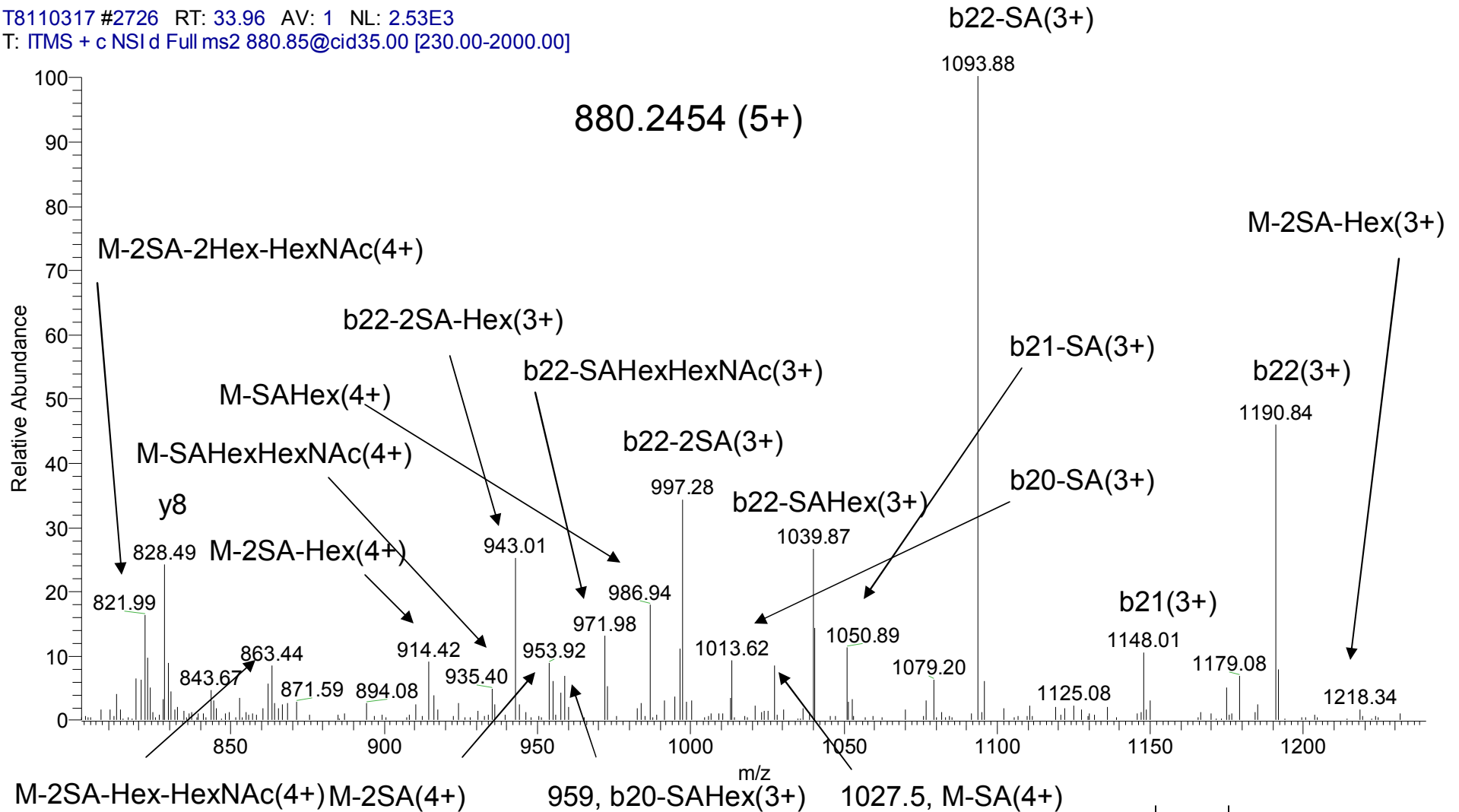


CID !

A4IFA5, VASN protein

Supplemental Figure 5B

T8110317 #2726 RT: 33.96 AV: 1 NL: 2.53E3
 T: ITMS + c NSI d Full ms2 880.85@cid35.00 [230.00-2000.00]



VRPGPRPS(HexNAcHexSA)PAPAT(HexNAcHexSA)PRPLPLGIEPASPTSLR
 b22

CID!

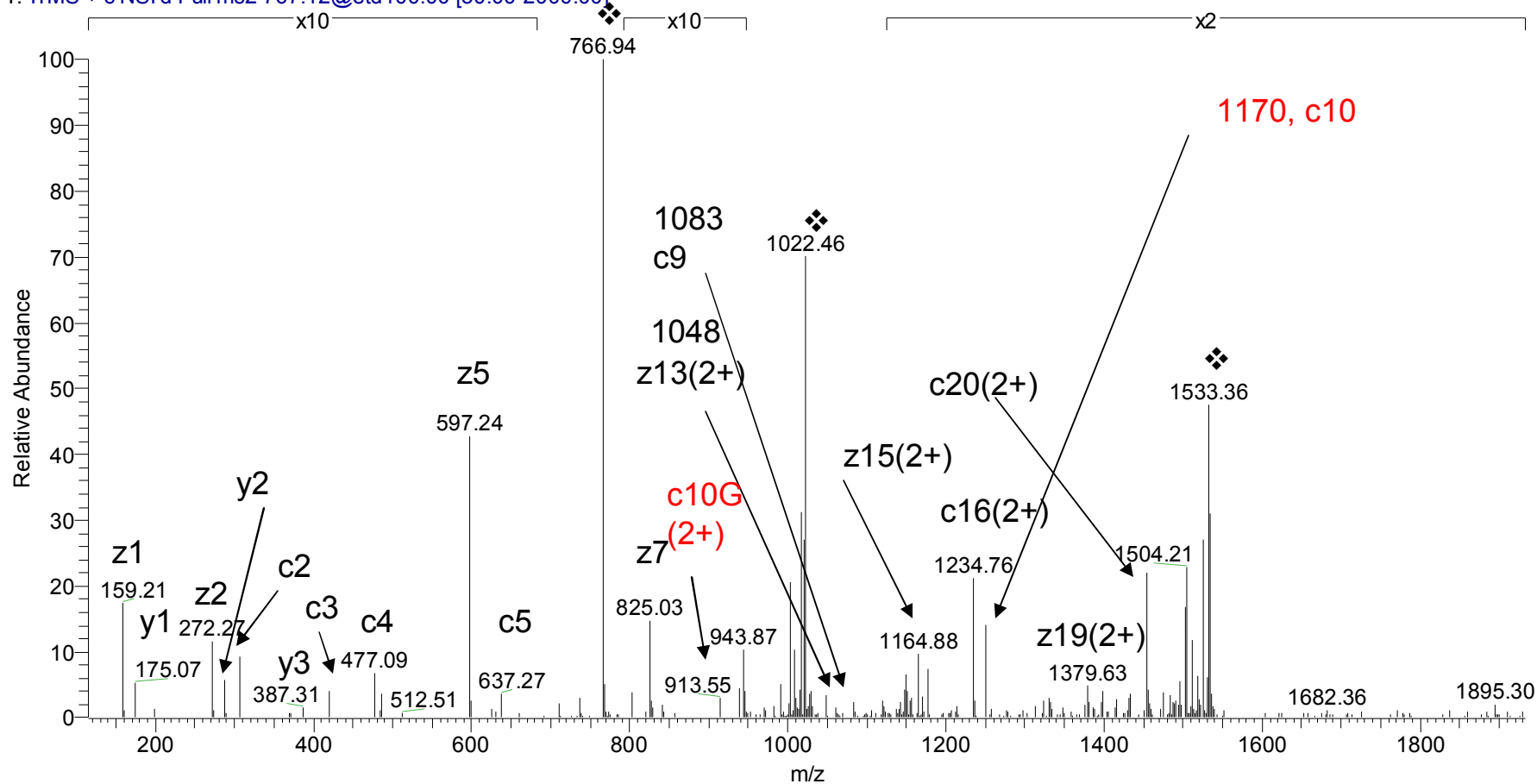
P01044, Kininogen-1

766.6159 (4+)

Supplemental Figure 6

T8110313 #2371 RT: 32.79 AV: 1 NL: 5.91E3
T: ITMS + c NSI d Full ms2 767.12@etd100.00 [50.00-2000.00]

Ser-149 or Thr-150/Ser-152



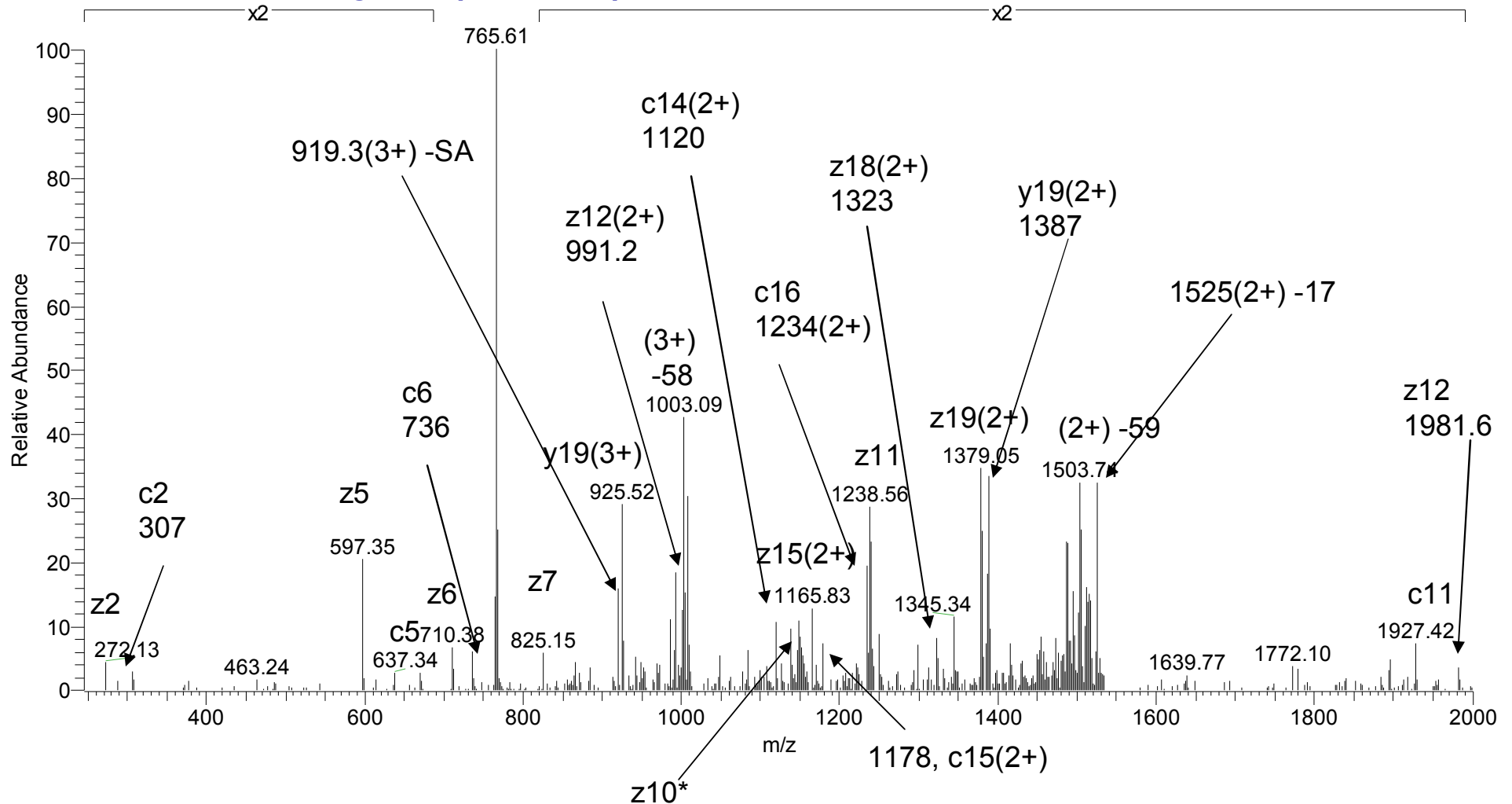
EC(Carbamidomethyl)LGC(Carbamidomethyl)VHPIS(HexNAcHexSA)TKSPDLEPVLR⁺⁴

Since c10 is present with and without the sugar, this is most likely a mixture

P01044, Kininogen-1 Ser-149

766.6168 (4+)

Katalin_Szuszanna_53 #3785 RT: 41.46 AV: 1 NL: 1.09E3
 T: ITMS + c NSI d sa Full ms2 767.12@etd90.00 [150.00-2000.00]



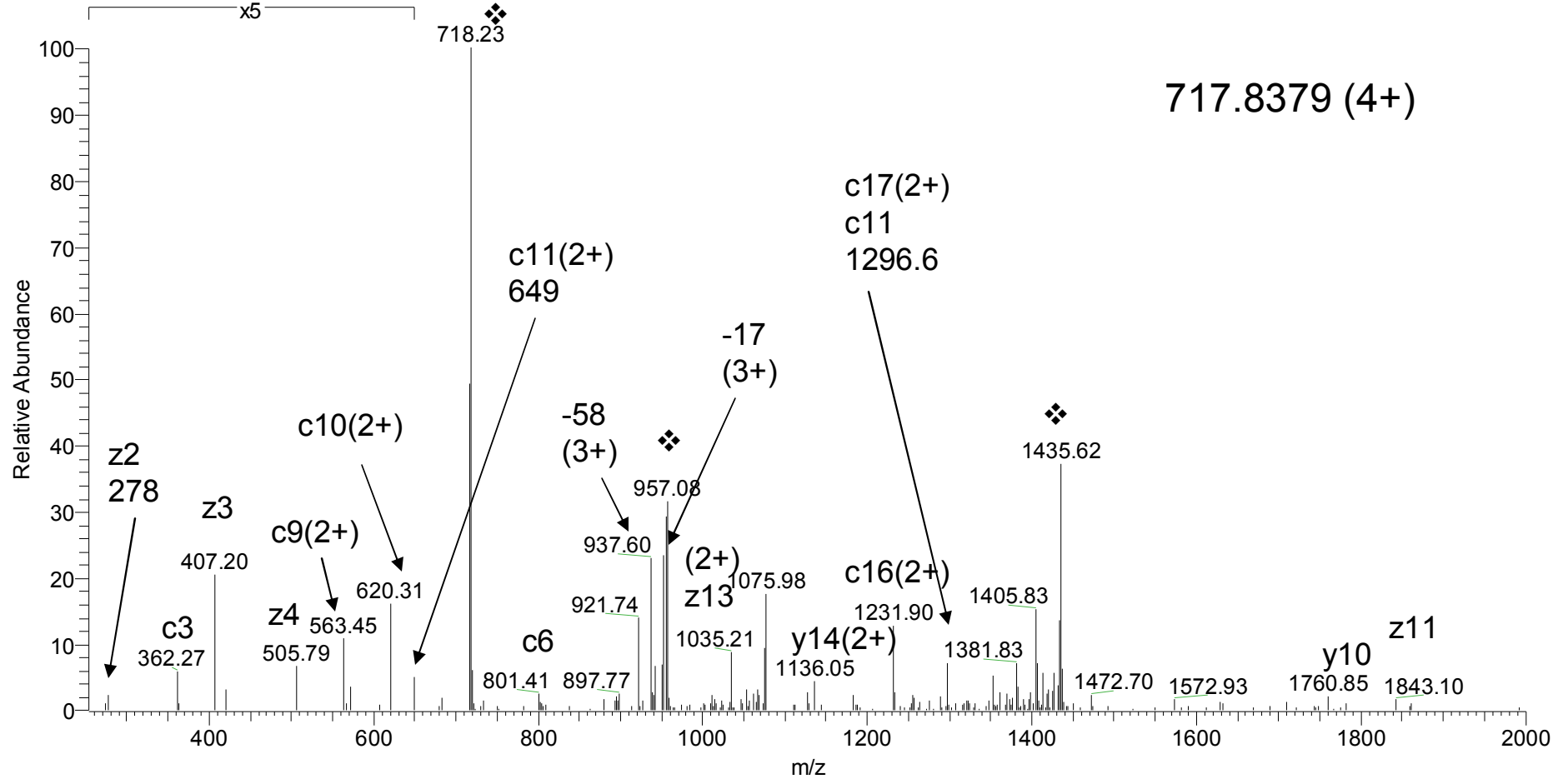
EC(Carbamidomethyl)LGC(Carbamidomethyl)VHPIS(HexNAcHexSA)TKSPDLEPVLR⁺⁴

There is a weak m/z 1170, that may correspond to c10 without the sugar! + a weak m/z 1896, z11 if Thr-150/Ser-152 is glycosylated)

P01044, Kininogen-1

Thr-605

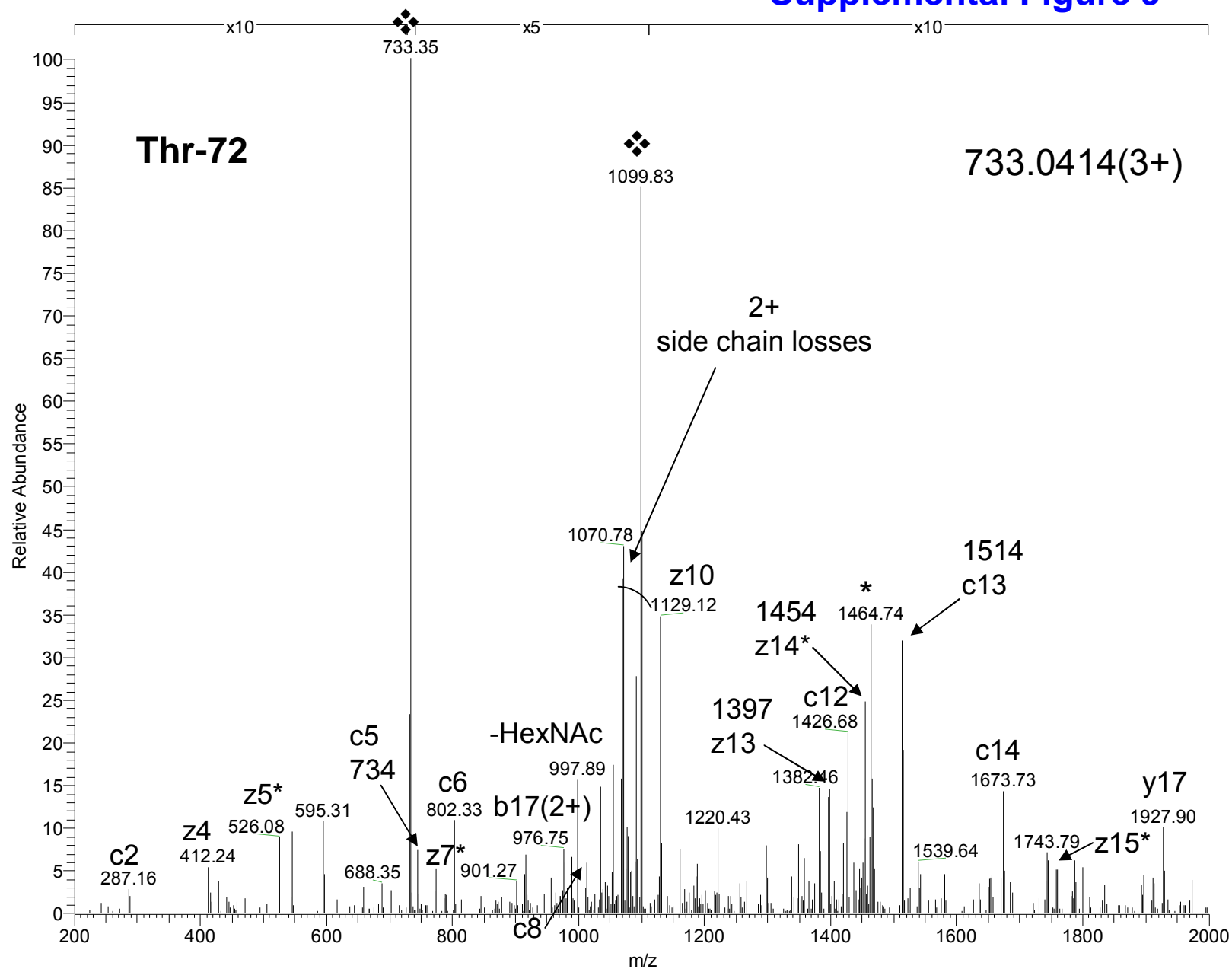
T8110314 #1483 RT: 25.77 AV: 1 NL: 1.98E3
 T: ITMS + c NSI d Full ms2 718.09@etd100.00 [50.00-2000.00]



C(Carbamidomethyl)PSRPWKPVNGVNPT(HexNAcHexSA)VEM(Oxidation)K

Q2KIU3, Putative uncharacterized protein

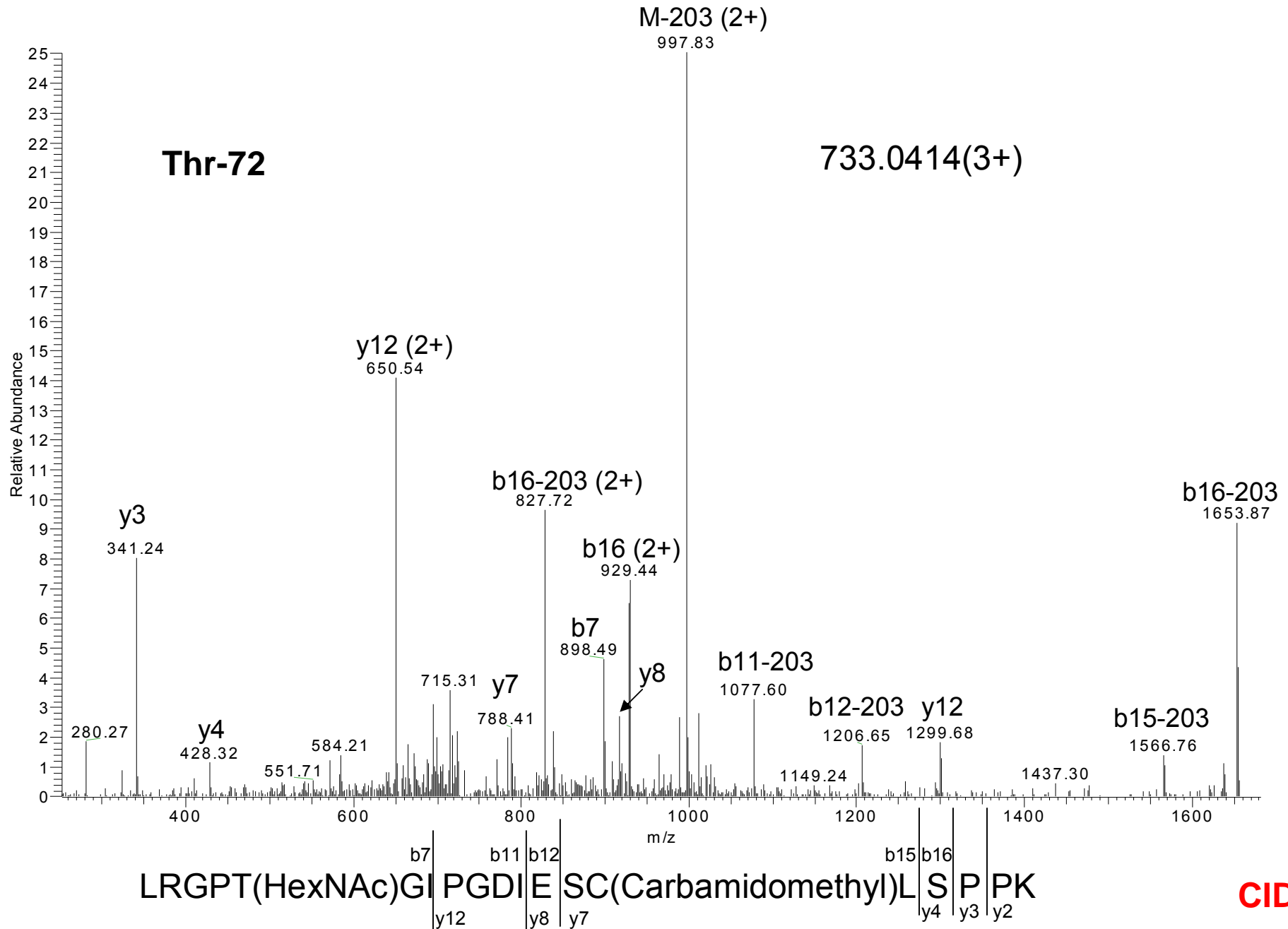
Supplemental Figure 9



LRGPT(HexNAc)GIPGDIESC(Carbamidomethyl)LSPPK

Q2KIU3, Putative uncharacterized protein

Supplemental Figure 10

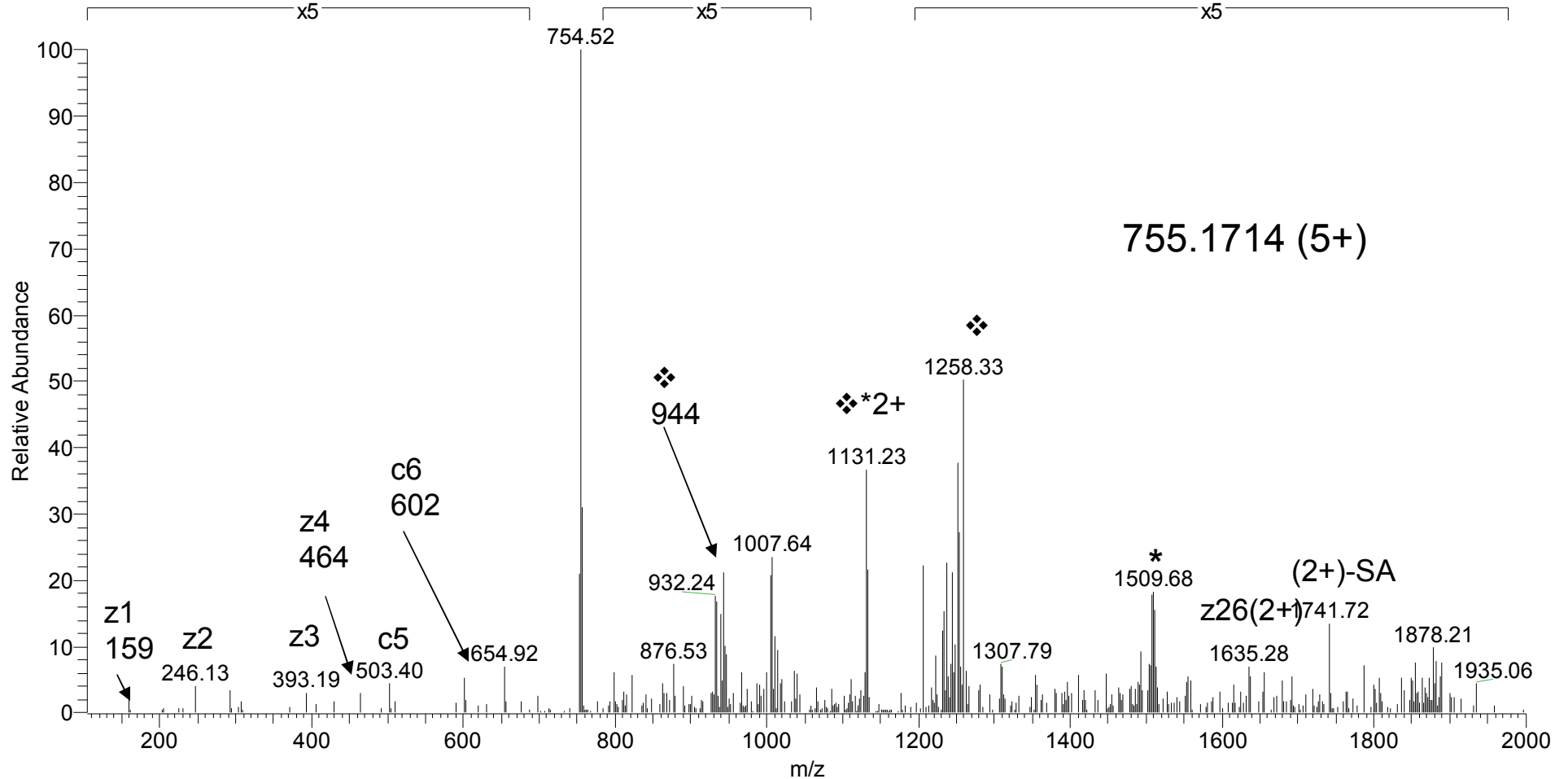


CID !

Supplemental Figure 11

❖ *also overlapping (3+)

T8110316 #2794 RT: 34.84 AV: 1 NL: 5.92E3
T: ITMS + c NSI d Full ms2 755.57@etd100.00 [50.00-2000.00]



LVVSSVQHQSALELSEAGVQAAAATST(HexNAcHexSA)AM(Oxidation)SR

If the ID is accepted, still seven potential glycosylation sites

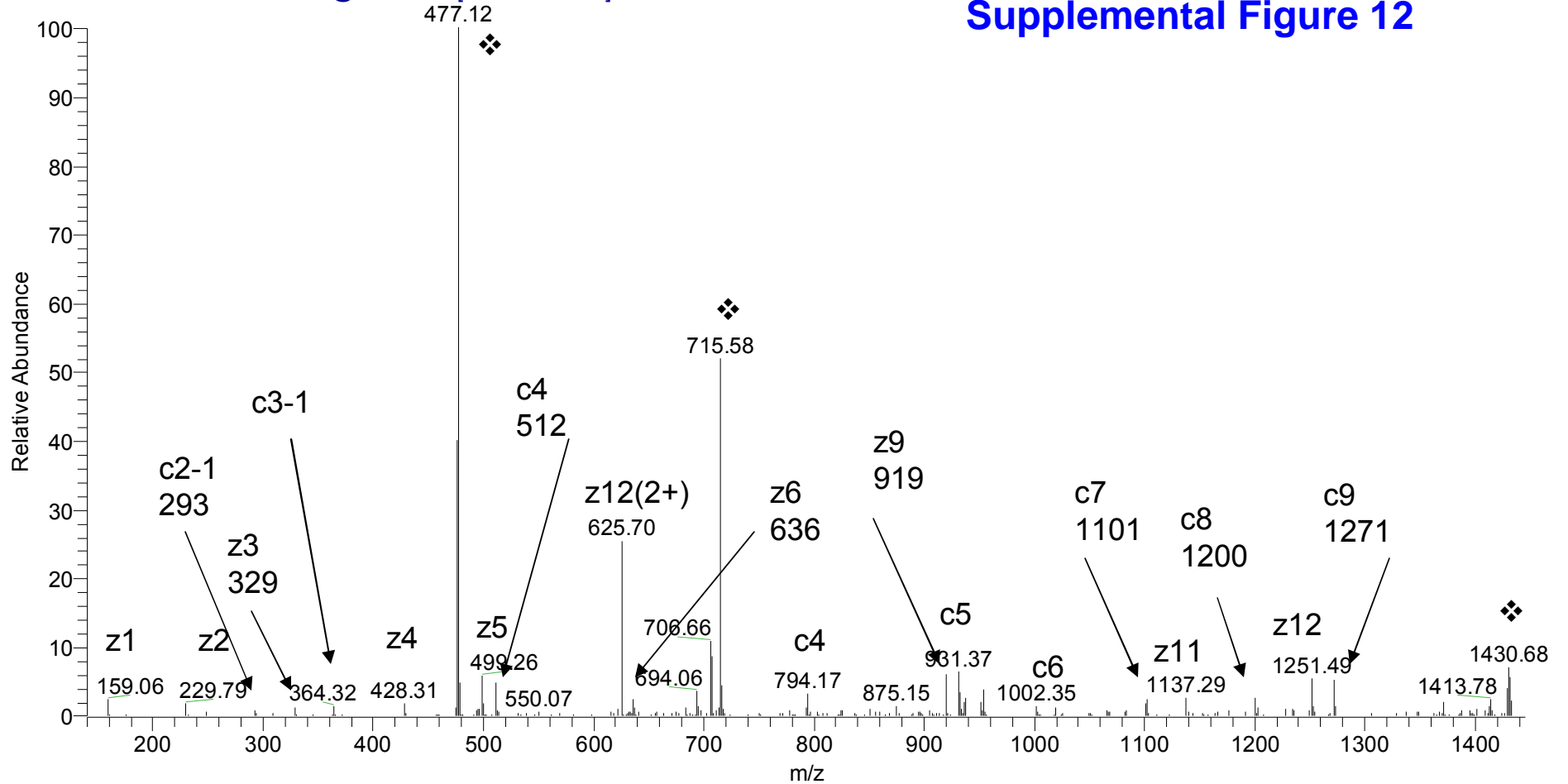
~~B3VTM3, Lactoferrin~~

Q29443, serotransferrin

476.9216 (3+)

T8110314 #1227 RT: 23.82 AV: 1 NL: 4.82E3
T: ITMS + c NSI d Full ms2 476.92@etd100.00 [50.00-1445.00]

Supplemental Figure 12



QVPS(HexNAcHex)HAVVAR

This is NOT a glycopeptide!

(C) YLAM(O)VPSHAVVAR