Electronic Supplementary Material (ESI) for Molecular Omics.

Supplementary Information

β-catenin/CBP Inhibition Alters Epidermal Growth Factor Receptor Fucosylation Status in Oral Squamous Cell Carcinoma

Kevin Brown Chandler[1], Khalid Alamoud[2], Vanessa L Stahl[1], Bach-Cuc Nguyen[2], Vinay Kartha[3], Manish V. Bais[2], Kenichi Nomoto[4], Takashi Owa[4], Stefano Monti[3], Maria A Kukuruzinska[2*], and Catherine E Costello[1*]

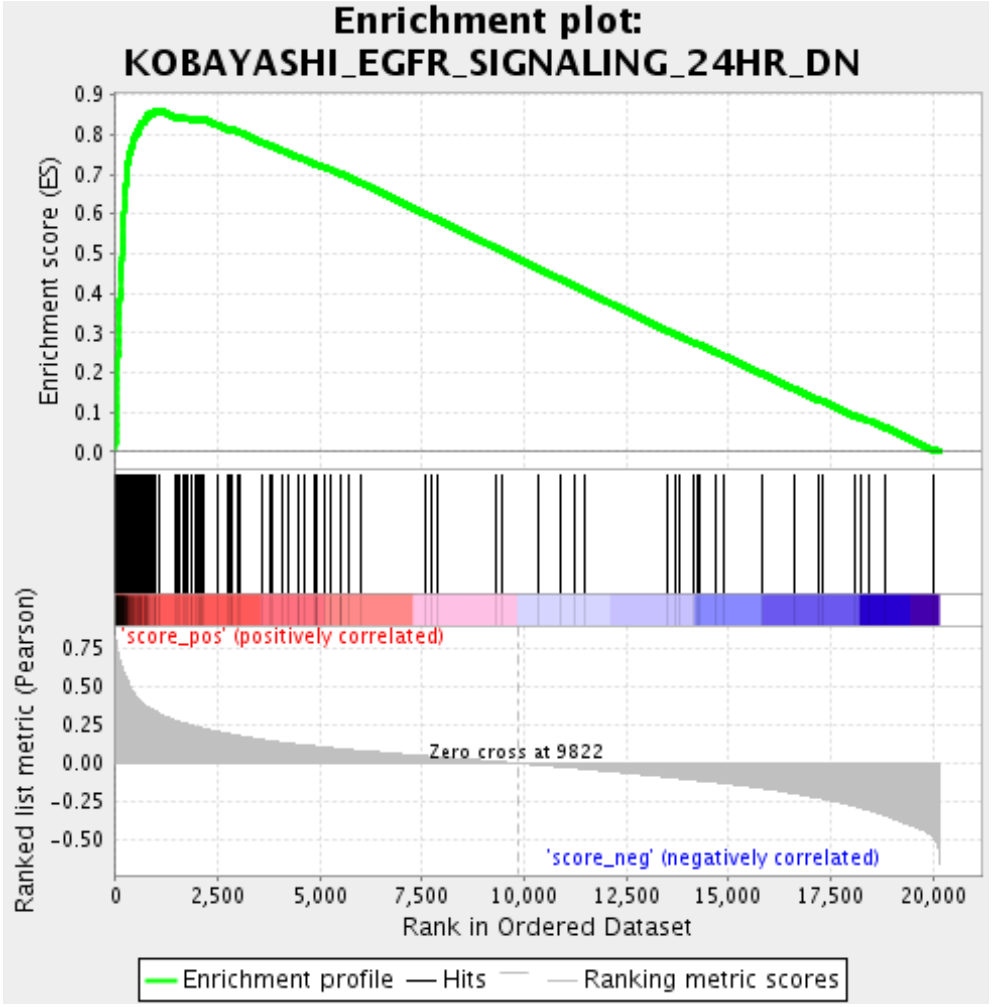
[1]Center for Biomedical Mass Spectrometry, Department of Biochemistry, Boston University School of Medicine, [2]Department of Translational Dental Medicine, Boston University Henry M. Goldman School of Dental Medicine, [3]Division of Computational Biomedicine, Boston University School of Medicine, Boston, MA, 02118, and [4]Eisai Inc., Woodcliff Lake, NJ 07677

***Correspondence to:**
Catherine E. Costello
Center for Biomedical Mass Spectrometry
Boston University School of Medicine
670 Albany St, Rm 511
Boston, MA 02118-2646 USA
Ph.: (+1) 617-358-2433
Fax: (+1) 617-358-2416
Email: cecmsms@bu.edu

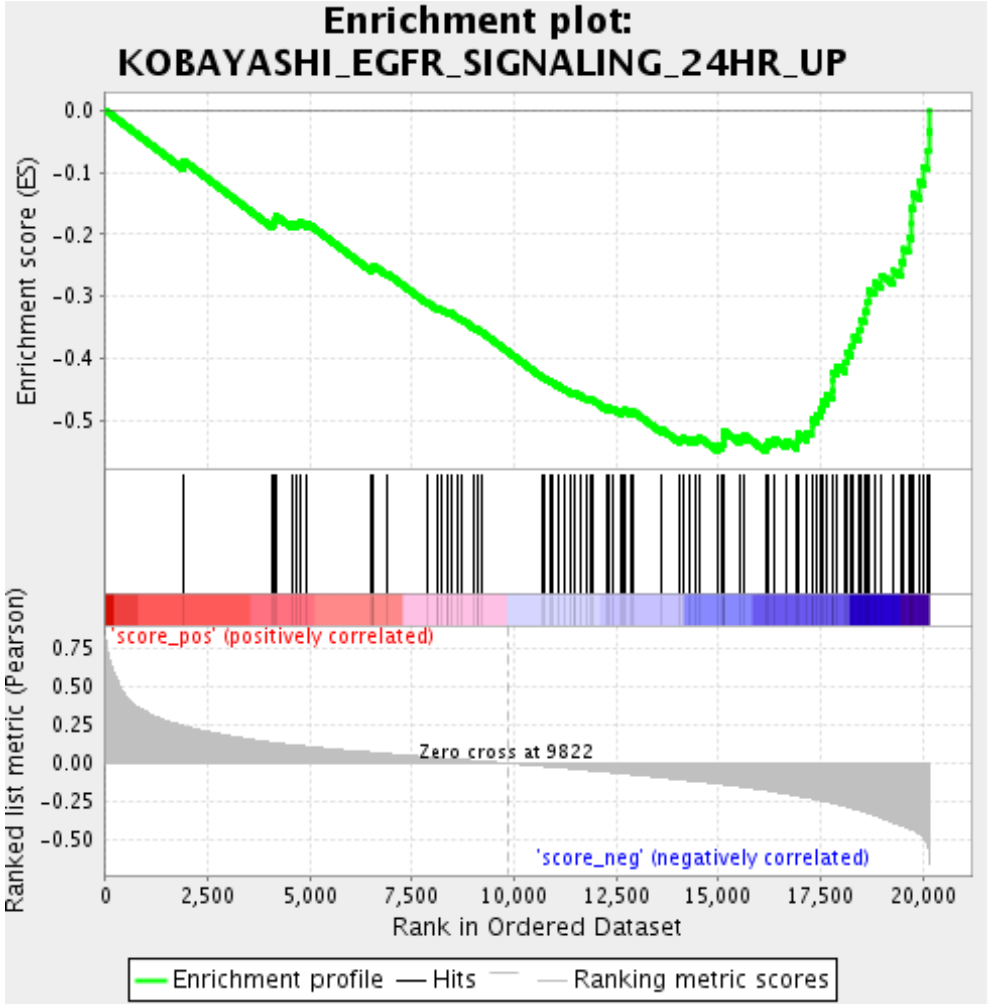
Maria A. Kukuruzinska
Department of Translational Dental Medicine
Boston University School of Dental Medicine
700 Albany Street, W201
Boston, MA 02118-2646 USA
Ph: (+1) 617-358-9690
Email: mkukuruz@bu.edu

Table of Contents

A.

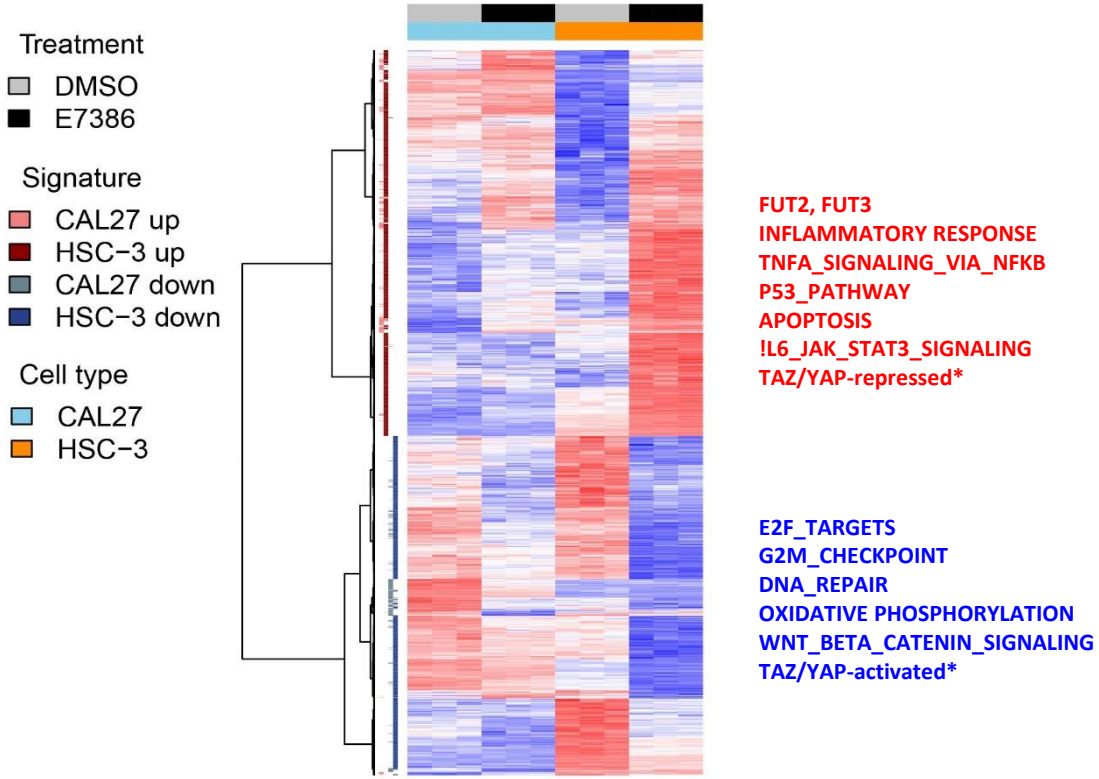


B.

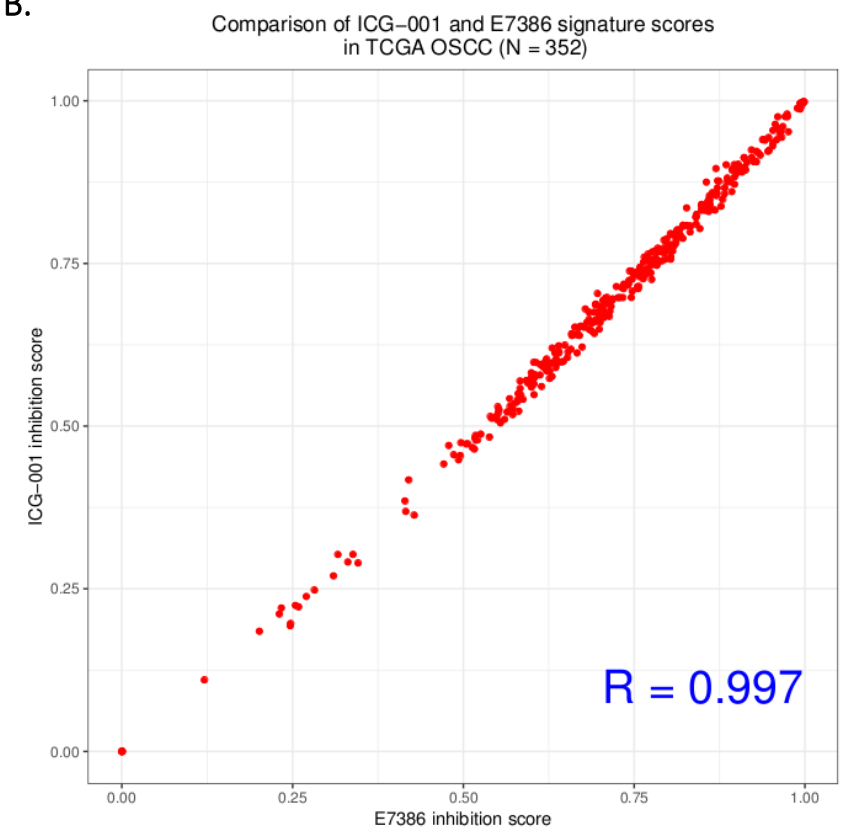


Supplementary Figure 1: ICG-001 Inhibition Positively Correlates with EGFR Inhibition. GSEA enrichment profiles corresponding to ICG-001 inhibition scores in TCGA OSCC data for gene sets pertaining to small-molecule-based EGFR inhibition. Genes that are down-regulated (**A**, left) are enriched towards higher ICG-001 inhibition levels (left; FDR q-value=1.31E-4), while genes that are up-regulated (**B**, right) are enriched towards lower ICG-001 inhibition levels (right; FDR q-value = 0.041), suggesting ICG-001 inhibition is positively correlates with EGFR inhibition in OSCC. A strong positive correlation between ICG-001 inhibition (in TCGA OSCCs) and EGFR inhibition is observed, suggesting that ICG-001 inhibits genes involved in EGFR signaling in primary oral cancer tumor tissues.

A.

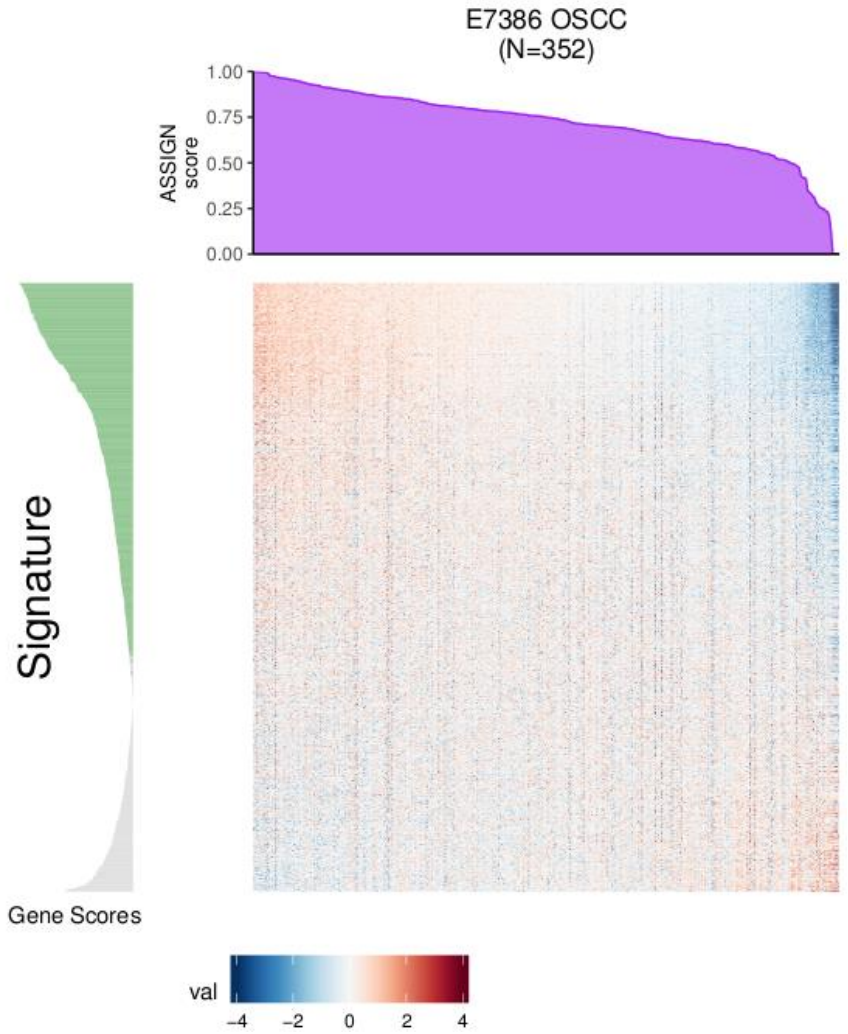


B.

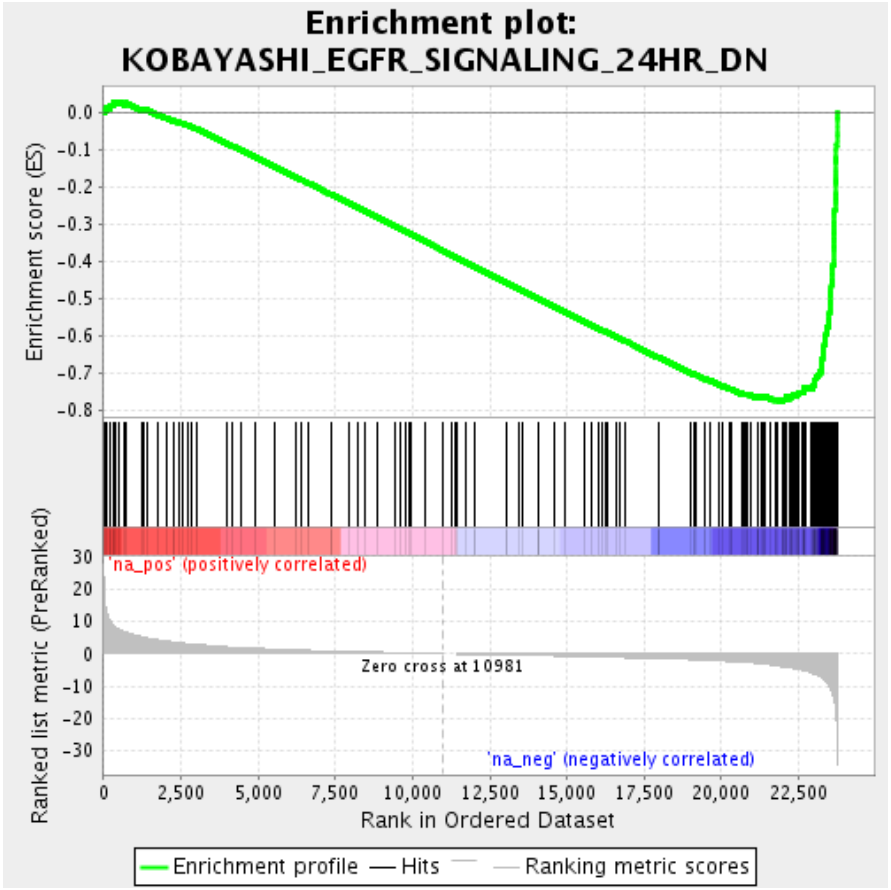


Supplementary Figure 2: E7386 Inhibition Signature. A. Heatmap of microarray gene expression treatment signatures derived from metastatic HSC-3 and non-metastatic CAL27 cells treated with E7386. RNA was extracted and profiled using microarrays from each cell line and gene expression profiles were compared by differential expression testing (E7386 treatment versus DMSO vehicle, n=3 per treatment group per cell line). Blue indicates an decrease with β-catenin/CBP inhibition, while red indicates an increase with β-catenin/CBP inhibition. Note the presence of FUT2, FUT3 in the latter group. **B.** Comparison of ICG-001 and E7386 anti-HNSCC activities. HSC-3 cell line–derived E7386 inhibition scores are highly correlated with ICG-001 in TCGA.

A.



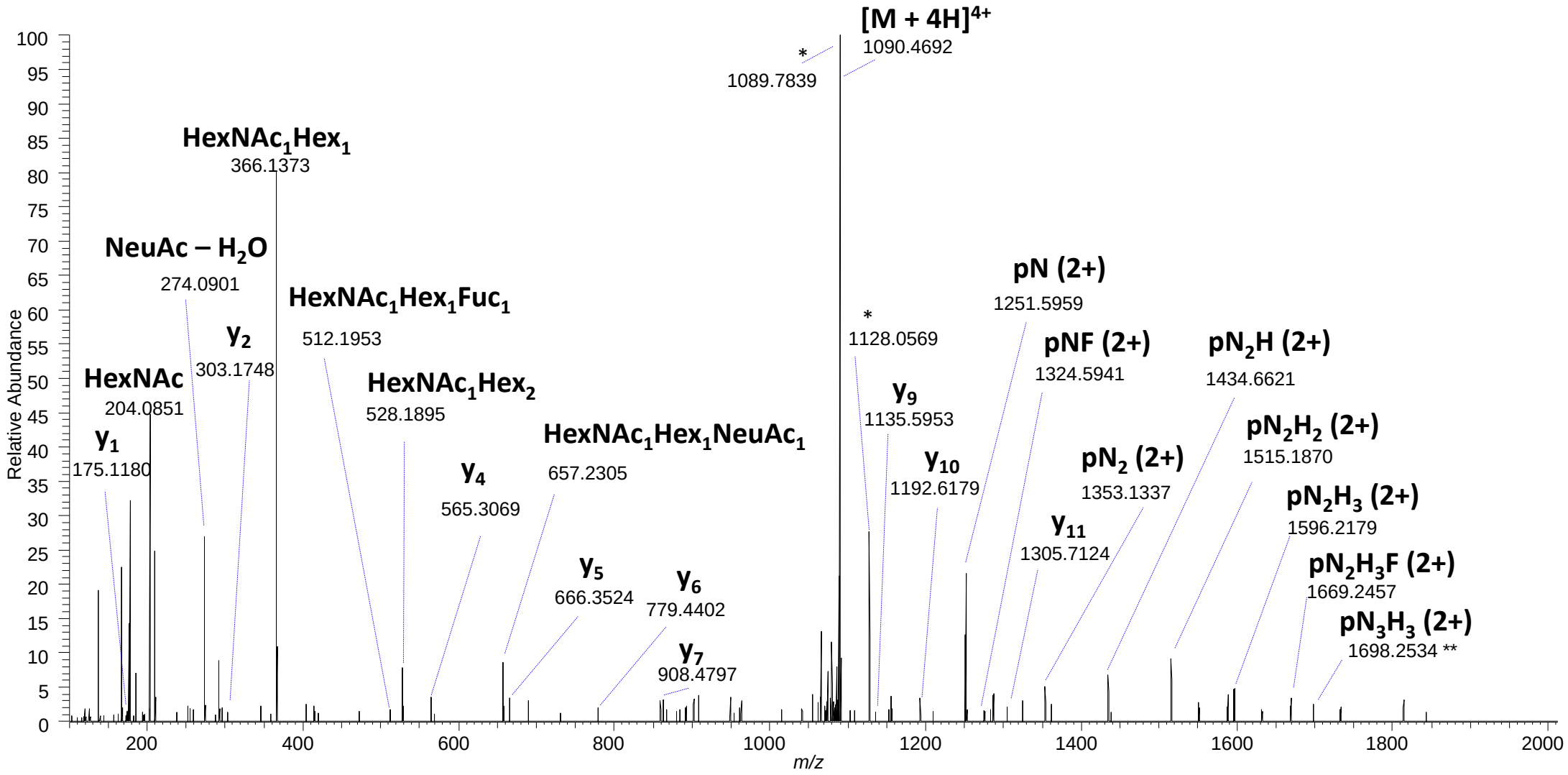
B.



Supplementary Figure 3: A. Transcriptional E7386 Inhibition Signature is Associated with Clinical Outcomes in Primary Human OSCC. Stratification of Human OSCCs (TCGA) by E7386 Inhibition Scores. Using ASSIGN, along with the TCGA OSCC data (*n*=352 samples) and E7386-repressed gene signature (*n*=976 genes) as input, we scored each sample based on the level of 'inhibition' exhibited for E7386-altered genes. **B. E7386 inhibition positively correlates with EGFR inhibition.** GSEA plots corresponding to E7386 inhibition scores in TCGA OSCC data for gene sets pertaining to small-molecule-based EGFR inhibition. Genes that are down-regulated are enriched towards higher E7386 inhibition levels, suggesting E7386 inhibition is positively correlated with EGFR inhibition in OSCC. A strong positive correlation between E7386 inhibition (in TCGA OSCCs) and EGFR inhibition suggests that E7386 inhibits genes involved in EGFR signaling in primary oral cancer tumor tissues.

EGFR Site N32 (Non-Canonical Sequon)

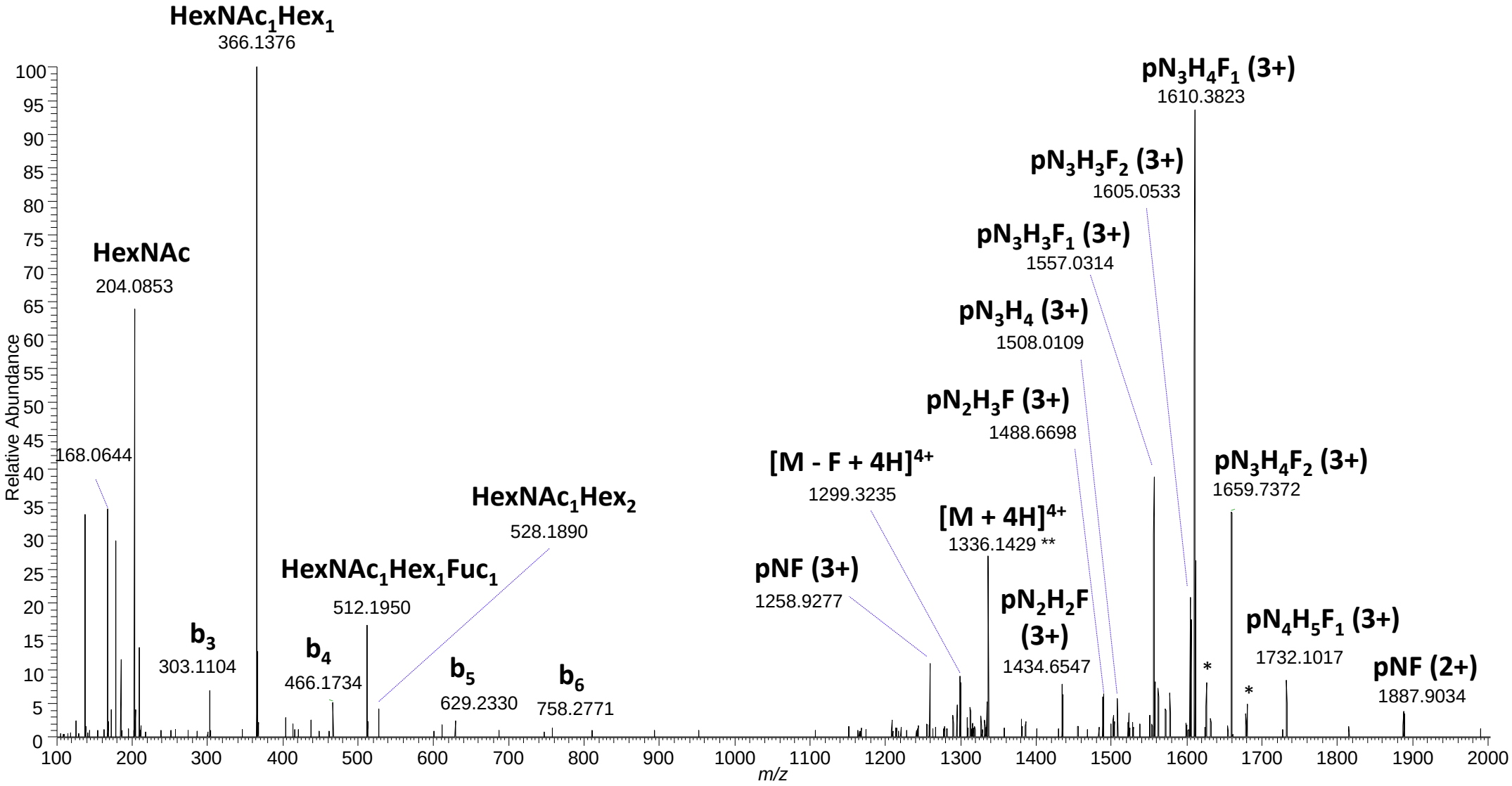
^{30}MFNNCEVVLGNLEITYVQR^{48} + $HexNAc_4Hex_5Fuc_1NeuAc_1$ $[M + 4H]^{4+}$ *m/z* 1090.4692



Supplementary Figure 4: HCD Spectrum of EGFR Site N32 Glycopeptide ^{30}MFNNCEVVLGNLEITYVQR^{48} + $HexNAc_4Hex_5Fuc_1NeuAc_1$. The glycopeptide HCD spectrum of the precursor with $[M + 4H]^{4+}$ *m/z* 1090.4692, is shown. *co-isolated precursor *m/z* 1089.7839 (z = 4+). ** second isotope is assigned, monoisotopic mass not detected.

EGFR Site N104

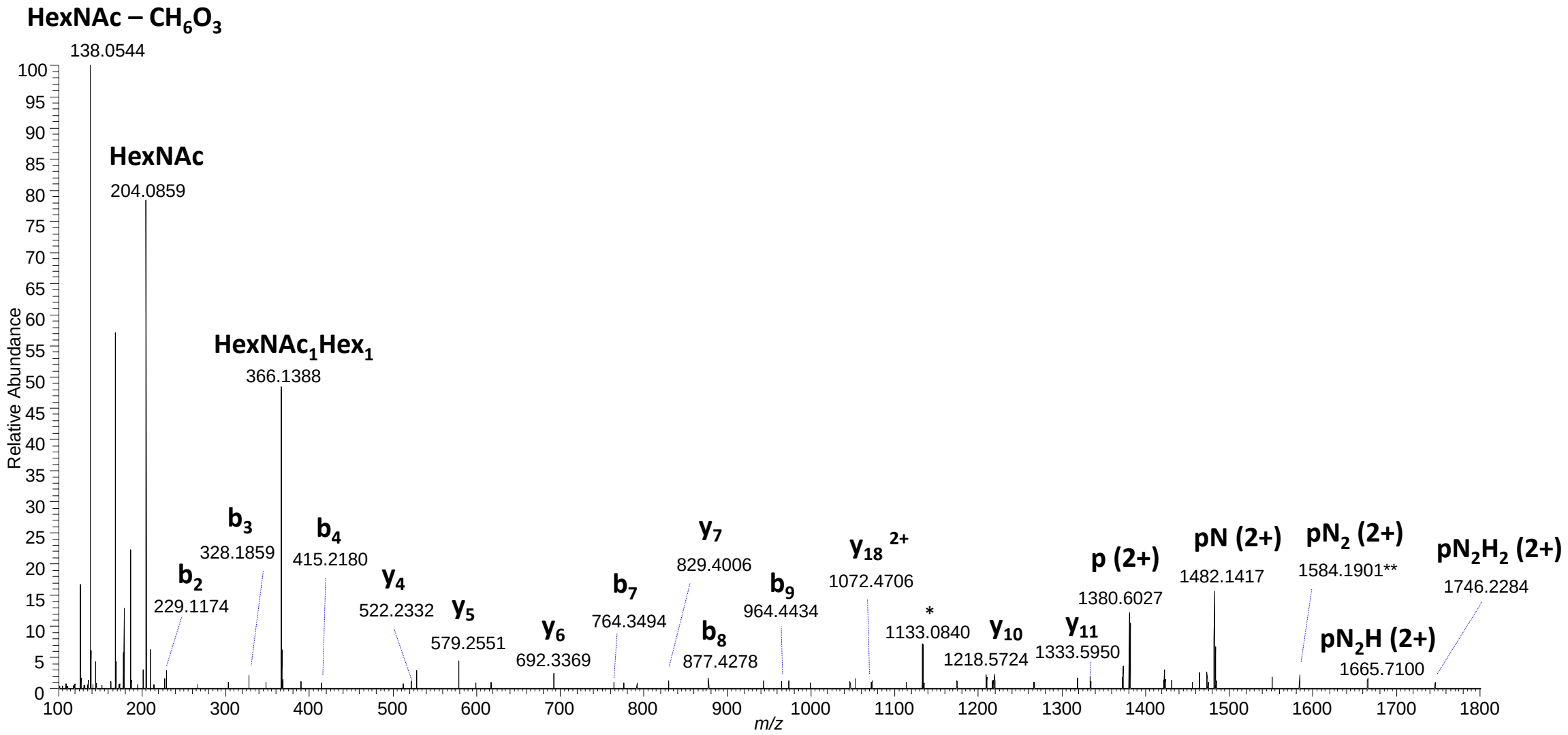
^{85}GNMYYENSYALAVLSNYDANKTGLKELPMR^{114} + $HexNAc_4Hex_5Fuc_2$ $[M + 4H]^{4+}$ *m/z* 1335.8433



Supplementary Figure 5: HCD Spectrum of EGFR Site N104 Glycopeptide ^{85}GNMYYENSYALAVLSNYDANKTGLKELPMR^{114} + $HexNAc_4Hex_5Fuc_2$. The glycopeptide HCD spectrum (20% collision energy) of the precursor with $[M + 4H]^{4+}$ *m/z* 1335.8433, is shown. *not assigned; may be from co-isolated ion. **co-isolated precursor with overlap of two isotope distributions.

EGFR Site N151

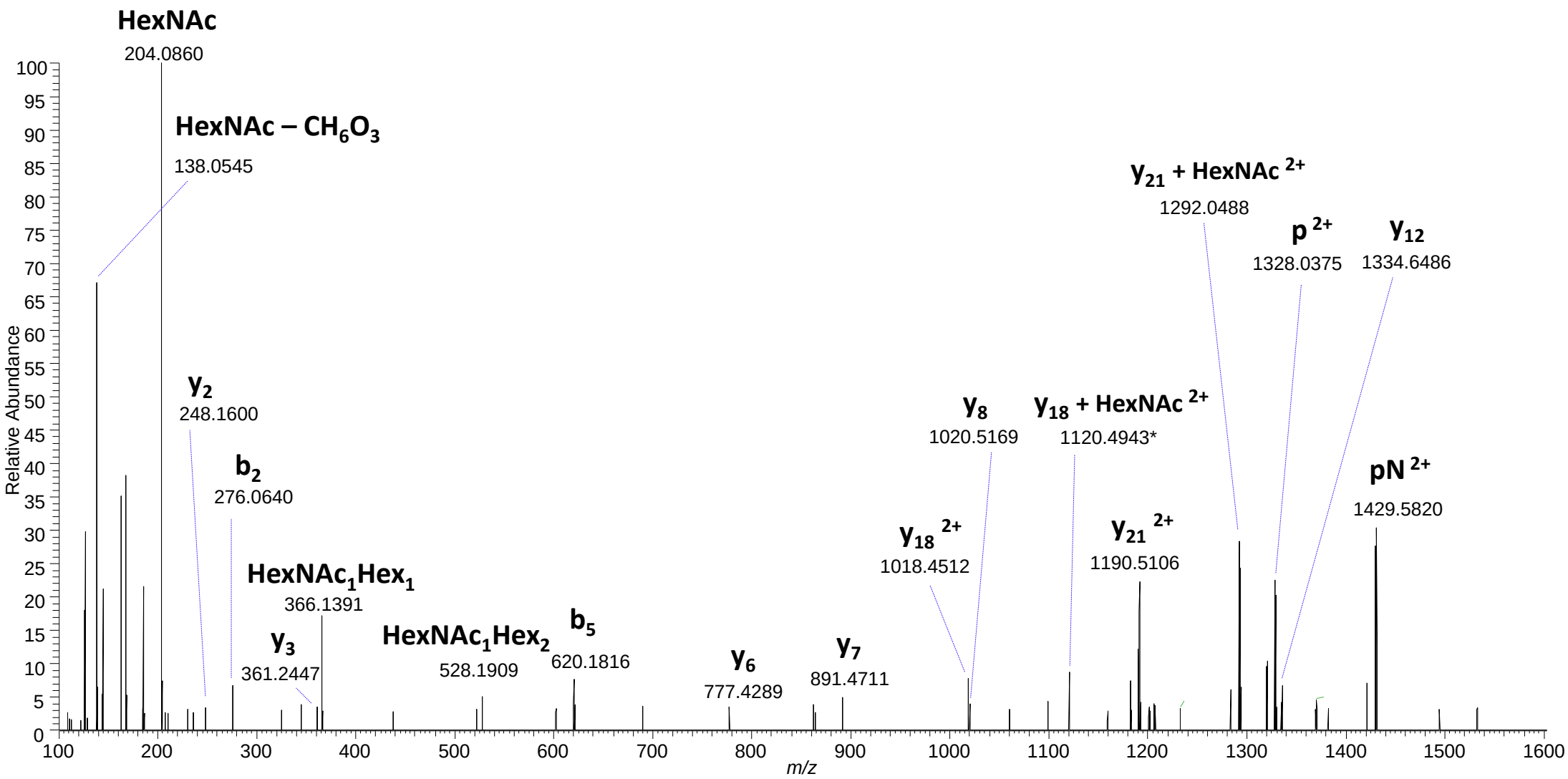
^{142}DIVSSDFLSNMSMDFQNHLGSCQK^{165} + $HexNAc_4Hex_5Fuc_1$ $[M + 4H]^{4+}$ *m/z* 1132.9656



Supplementary Figure 6: HCD Spectrum of EGFR Site N151 Glycopeptide ^{142}DIVSSDFLSNMSMDFQNHLGSCQK^{165} + $HexNAc_4Hex_5Fuc_1$. The glycopeptide HCD spectrum (30% collision energy) of the precursor with $[M + 4H]^{4+}$ *m/z* 1332.9656, is shown. *not assigned; may be from co-isolated ion. **second isotope is assigned, monoisotopic mass not detected.

EGFR Site N172

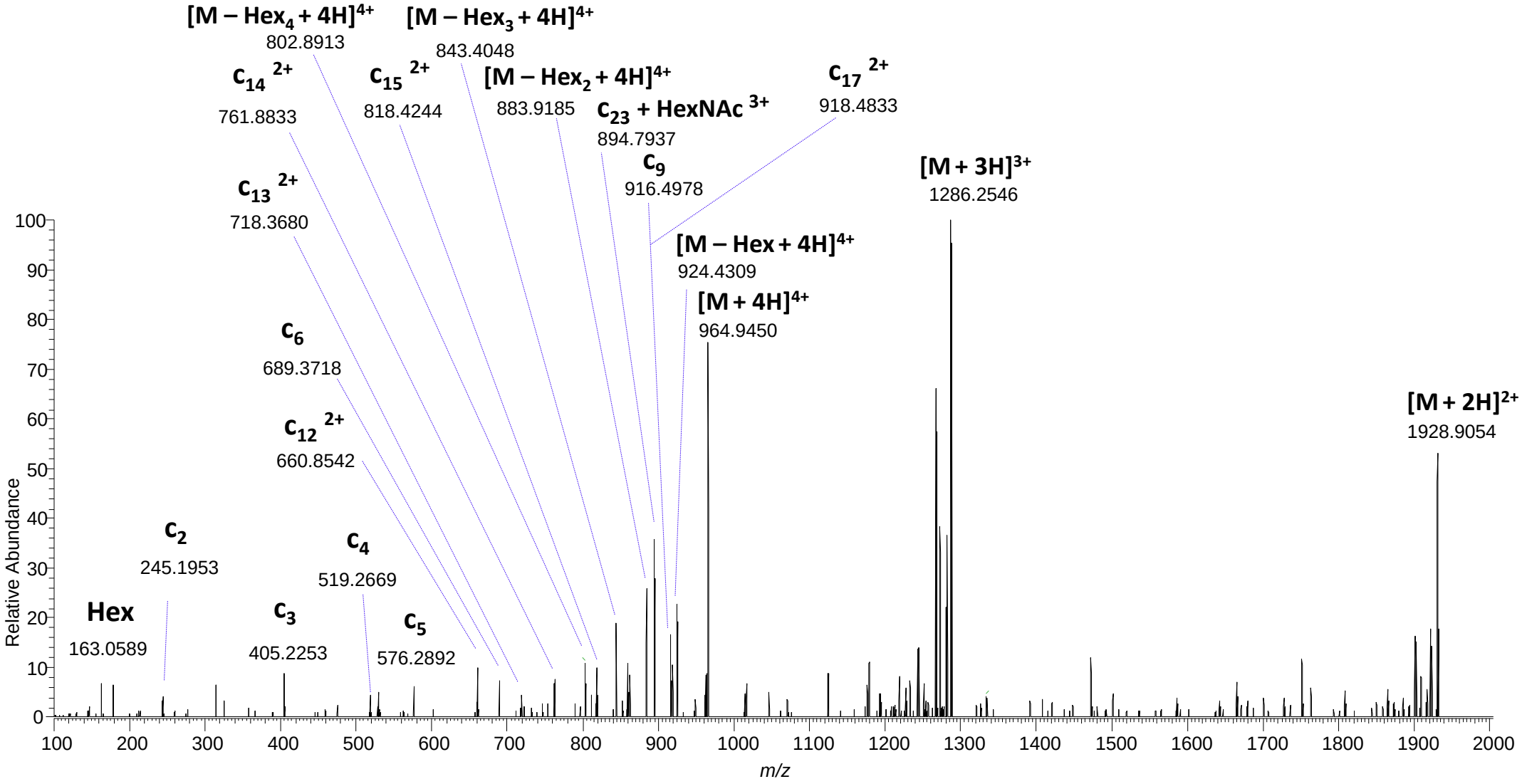
^{66}CDPSCPNGSCWGAGEENCQKLTK^{88} + $HexNAc_2Hex_6$ $[M + 3H]^{3+}$ *m/z* 1345.1873



Supplementary Figure 7: HCD Spectrum of EGFR Site N172 Glycopeptide ^{66}CDPSCPNGSCWGAGEENCQKLTK^{88} + $HexNAc_2Hex_6$. The glycopeptide HCD spectrum (30% collision energy) of the precursor with $[M + 3H]^{3+}$ *m/z* 1345.1873, is shown.

EGFR Site N328

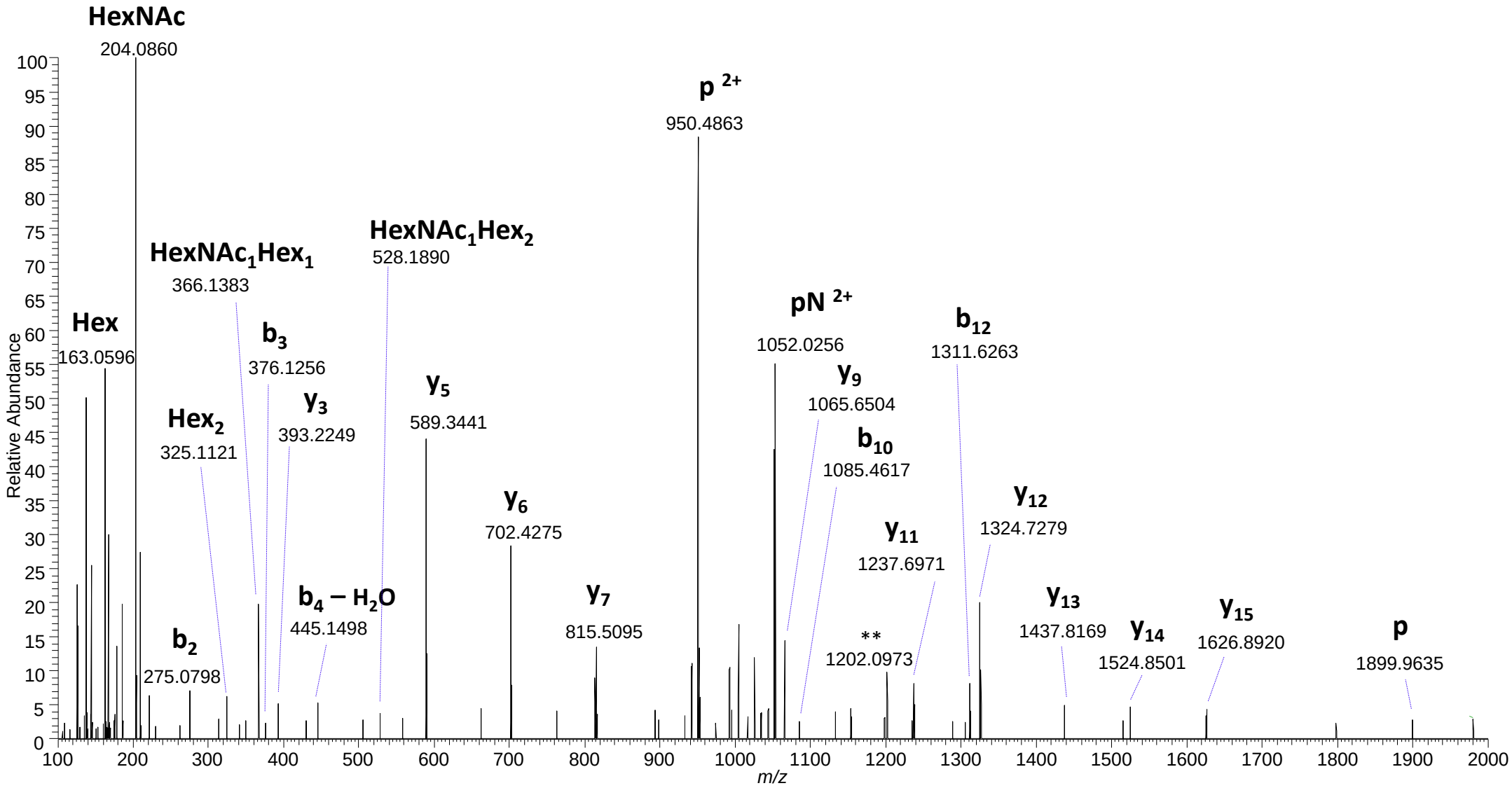
^{311}KVCNGIGIGEFKDSLSINATNIKHFK^{336} + $HexNAc_2Hex_6$ $[M + 4H]^{4+}$ *m/z* 964.9512



Supplementary Figure 8: EThcD Spectrum of EGFR Site N328 Glycopeptide ^{311}KVCNGIGIGEFKDSLSINATNIKHFK^{336} + $HexNAc_2Hex_6$. The glycopeptide EThcD spectrum (ETD 25%, HCD 15% supplemental energy) of the precursor with $[M + 4H]^{4+}$ *m/z* 964.9512, is shown.

EGFR Site N337

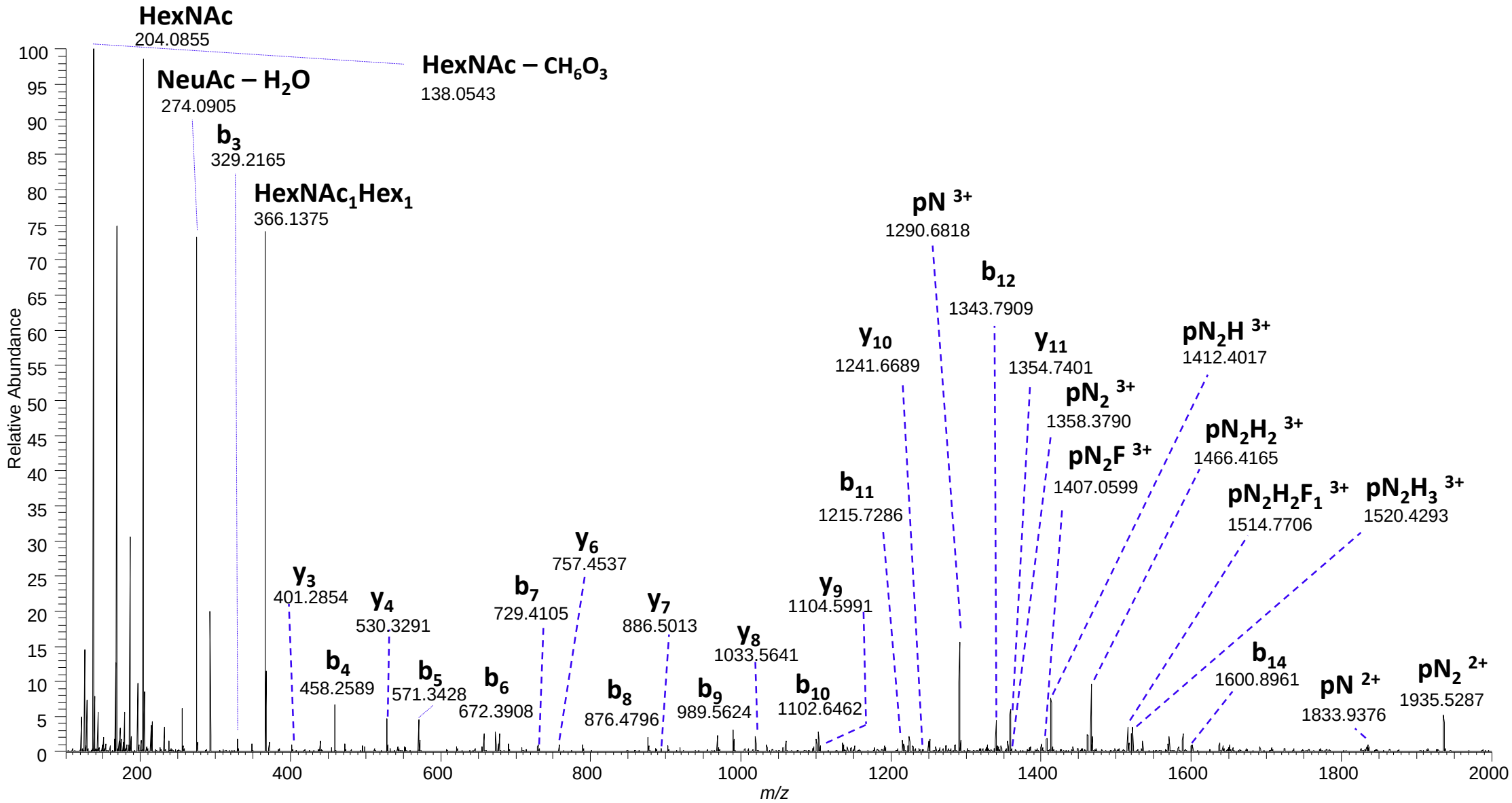
^{337}NCTSISGDLHILPVAFR^{353} + HexNAc$_2$Hex$_8$ [M + 3H]$^{3+}$ *m/z* 1201.5237



Supplementary Figure 9: HCD Spectrum of EGFR Site N337 Glycopeptide ^{337}NCTSISGDLHILPVAFR^{353} + HexNAc$_2$Hex$_8$. The glycopeptide HCD spectrum (30% collision energy) of the precursor with [M + 3H]$^{3+}$ *m/z* 1201.5237, is shown. **co-isolated precursor (overlap of two isotope distributions).

EGFR Site N389

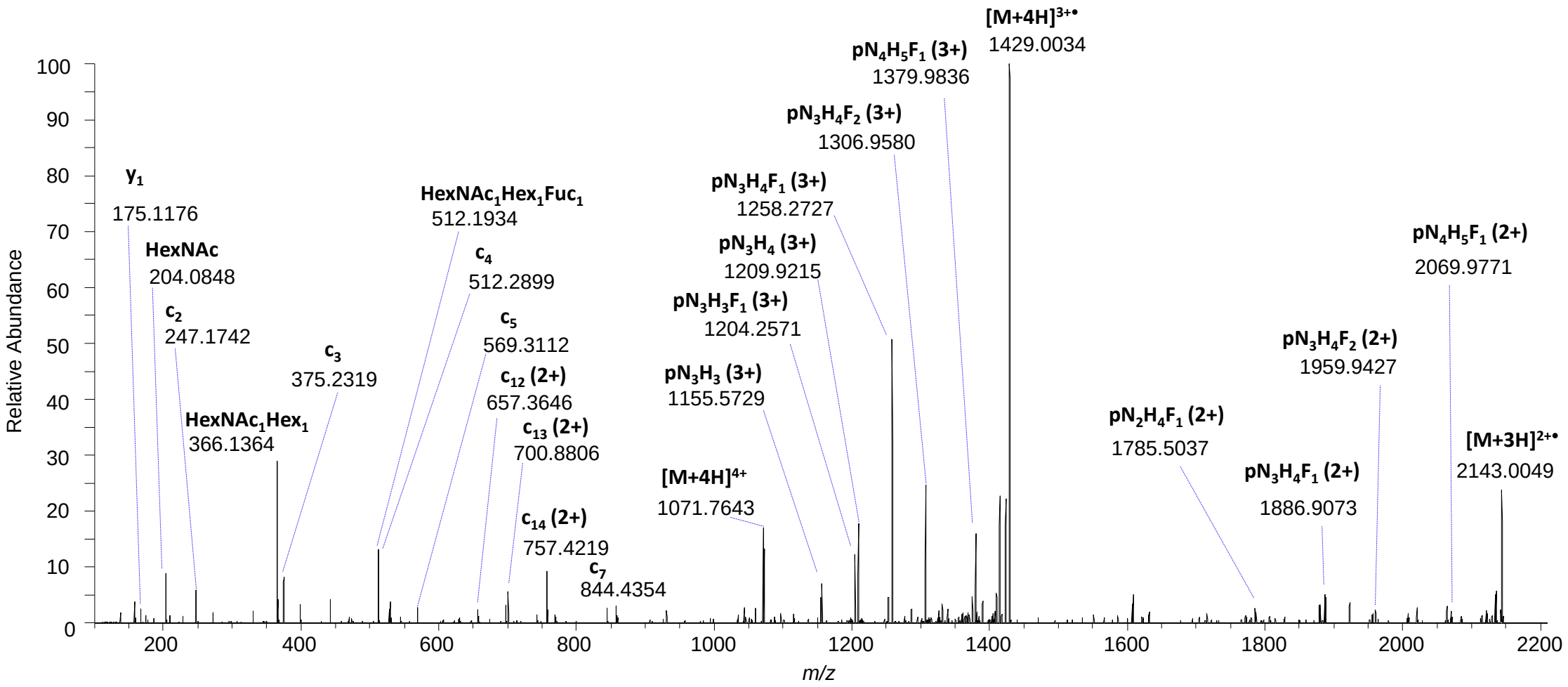
^{373}TVKEITGFLLIQAWPENRTDLHAFENLEIIR^{403} + $HexNAc_3Hex_4Fuc_1$ $[M + 5H]^{5+}$ *m/z* 1146.1460



Supplementary Figure 10: HCD Spectrum of EGFR Site N389 Glycopeptide ^{373}TVKEITGFLLIQAWPENRTDLHAFENLEIIR^{403} + $HexNAc_3Hex_4Fuc_1$. The glycopeptide HCD spectrum (28% collision energy) of the precursor with $[M + 5H]^{5+}$ *m/z* 1146.1460, is shown.

EGFR Site N420

^{406}TKQHGQFSLAVVSLNITSLGLR^{427} + $HexNAc_4Hex_5Fuc_2$ $[M + 4H]^{4+}$ *m/z* 1071.7526

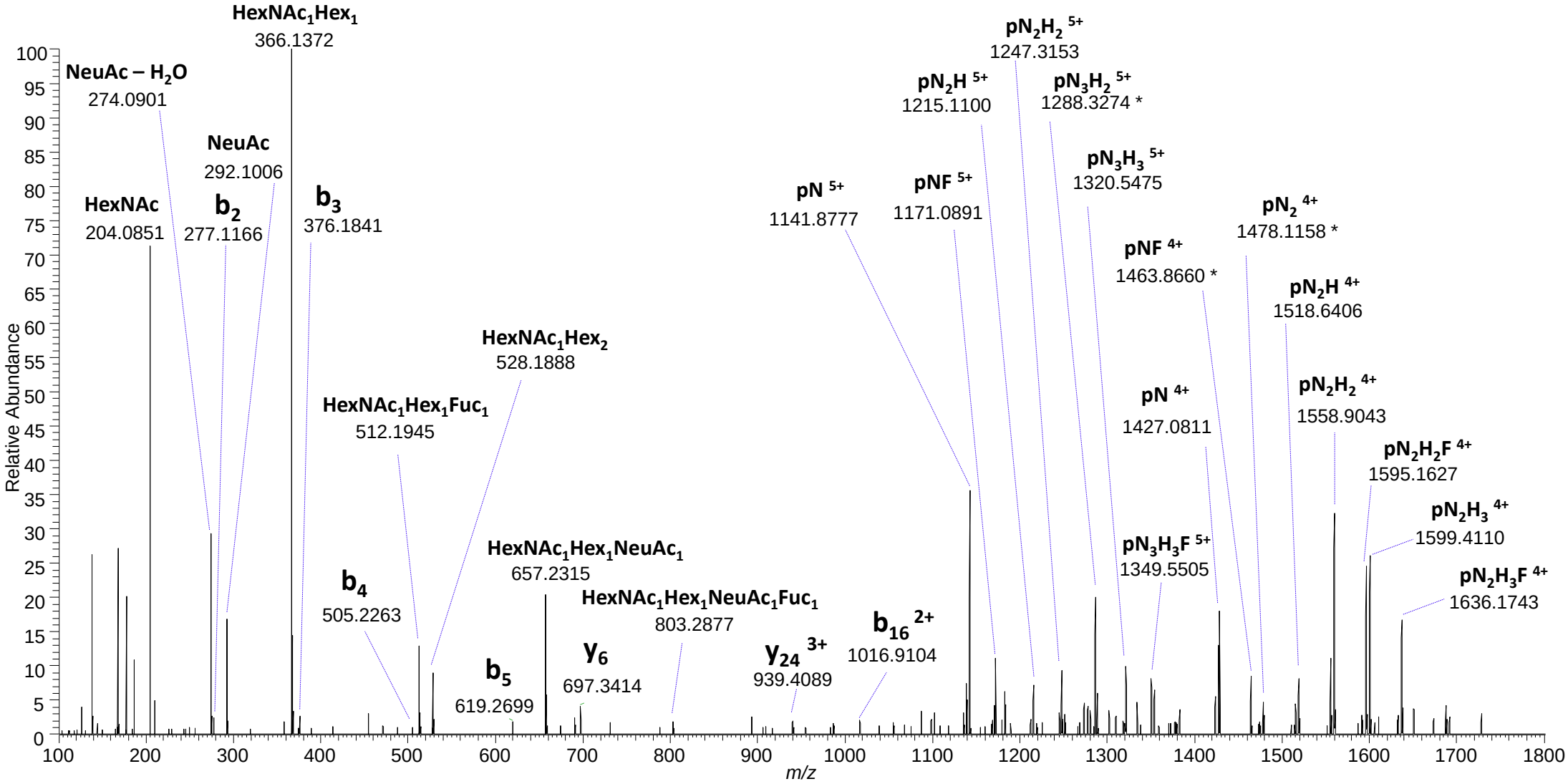


Supplementary Figure 11: EThcD Spectrum of EGFR Site N420 Glycopeptide ^{406}TKQHGQFSLAVVSLNITSLGLR^{427} + $HexNAc_4Hex_5Fuc_2$. The glycopeptide EThcD spectrum (ETD 26.6%, HCD 20% supplemental energy) of the precursor with $[M + 5H]^{5+}$ *m/z* 1146.1460, is shown.

EGFR Site N544

^{524}EFVENSECIQCHPECLPQAMNITCTGRGPDNCIQCAHYIDGPHCVK^{569} + $HexNAc_4Hex_5Fuc_2NeuAc_1$

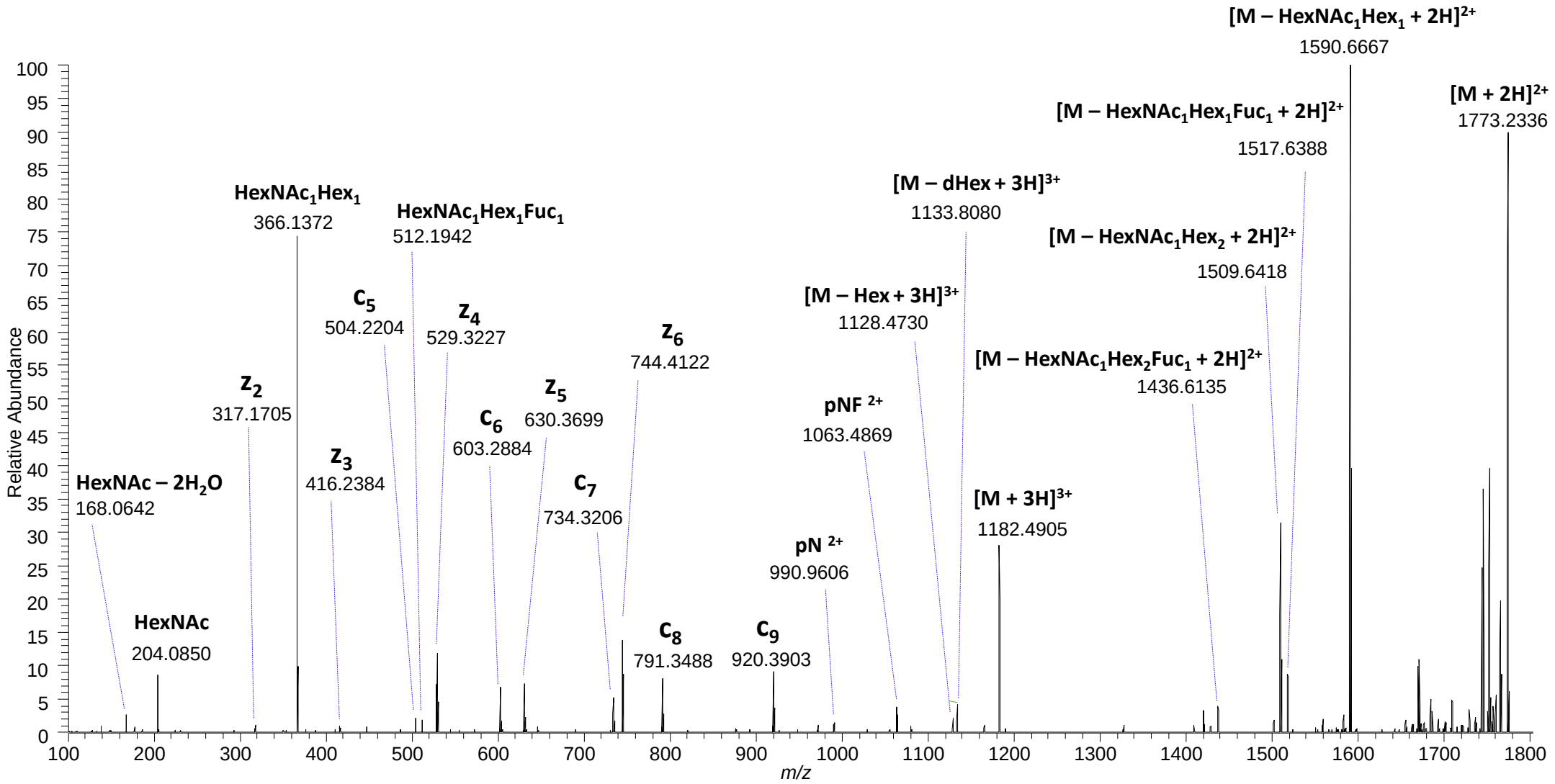
$[M + 6H]^{6+}$ *m/z* 1285.5280



Supplementary Figure 12: HCD Spectrum of EGFR Site N544 Glycopeptide ^{524}EFVENSECIQCHPECLPQAMNITCTGRGPDNCIQCAHYIDGPHCVK^{569} + $HexNAc_4Hex_5Fuc_2NeuAc_1$. The glycopeptide HCD spectrum (20% collision energy) of the precursor with $[M + 6H]^{6+}$ *m/z* 1285.5280, is shown. * second isotope is assigned, monoisotopic mass was not detected.

EGFR Site N579

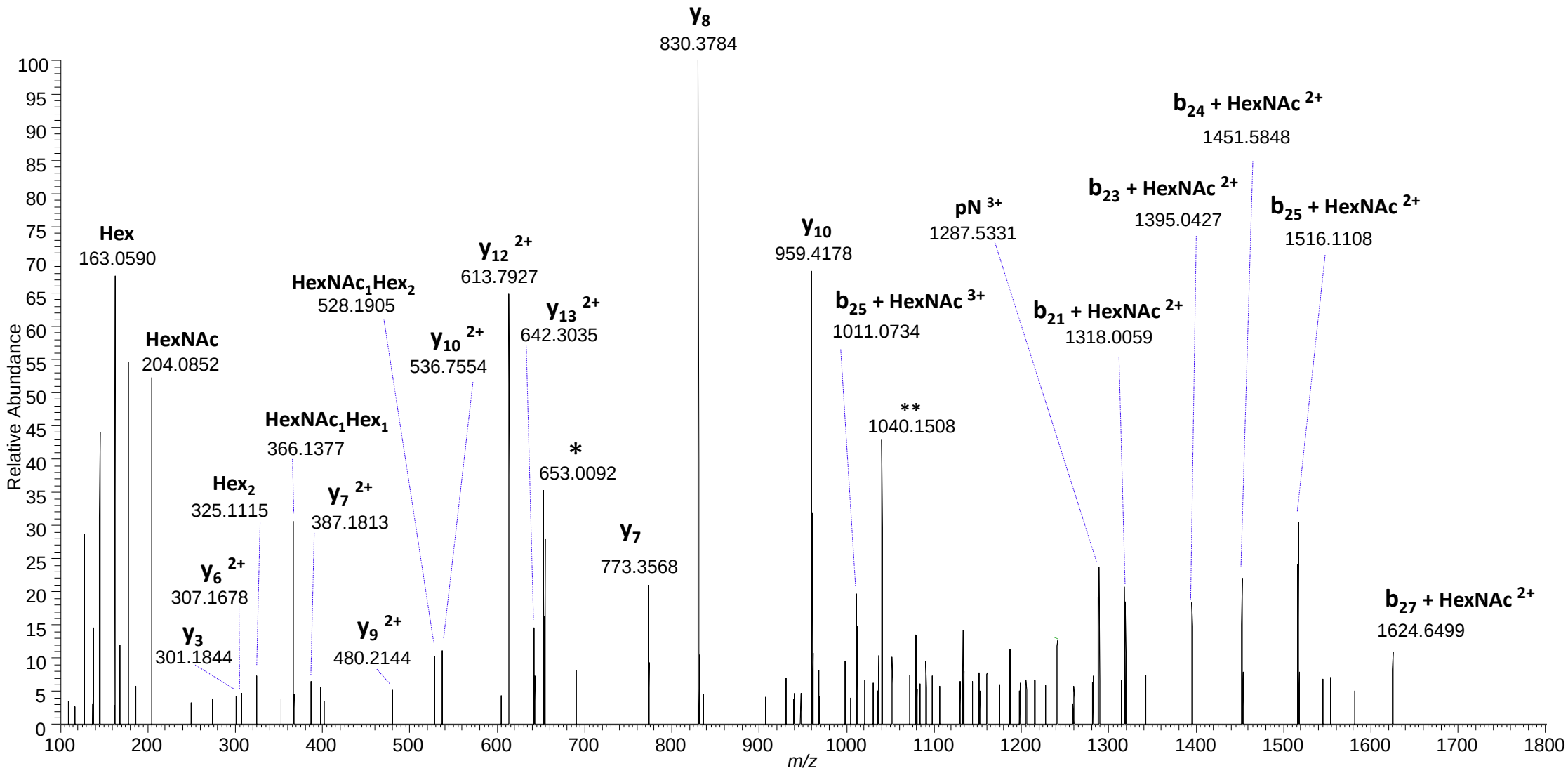
^{570}TCPAGVMGENNTLVWK^{585} + $HexNAc_4Hex_5Fuc_1$ $[M + 3H]^{3+}$ *m/z* 1182.4976



Supplementary Figure 13: EThcD Spectrum of EGFR Site N579 Glycopeptide ^{570}TCPAGVMGENNTLVWK^{585} + $HexNAc_4Hex_5Fuc_1$. The glycopeptide EThcD spectrum (ETD 45%, HCD 15% supplemental energy) of the precursor with $[M + 3H]^{3+}$ *m/z* 1182.4976, is shown.

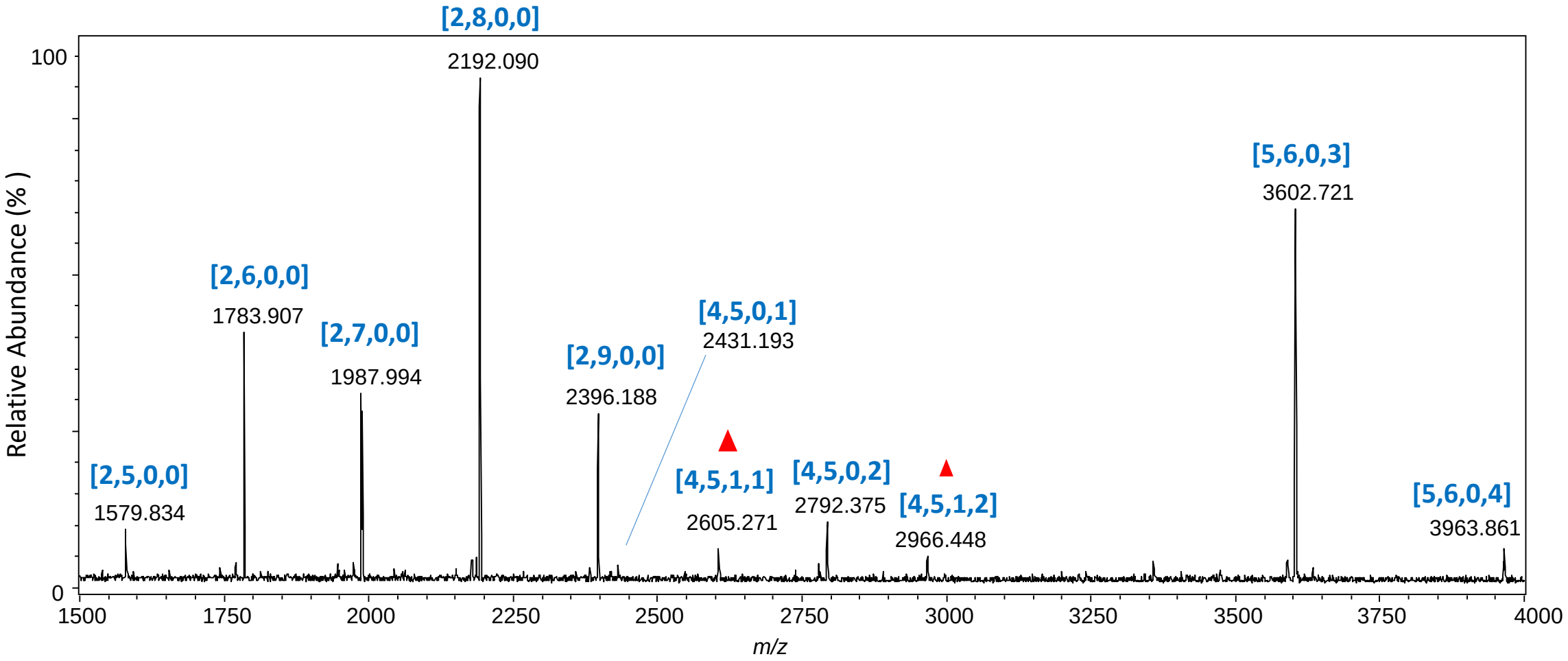
EGFR Site N599

^{586}YADAGHVCHLCHPNCTYGCTGPGLEGCPTNGPK^{618} + $HexNAc_2Hex_7$ $[M + 5H]^{5+}$ *m/z* 1040.4169

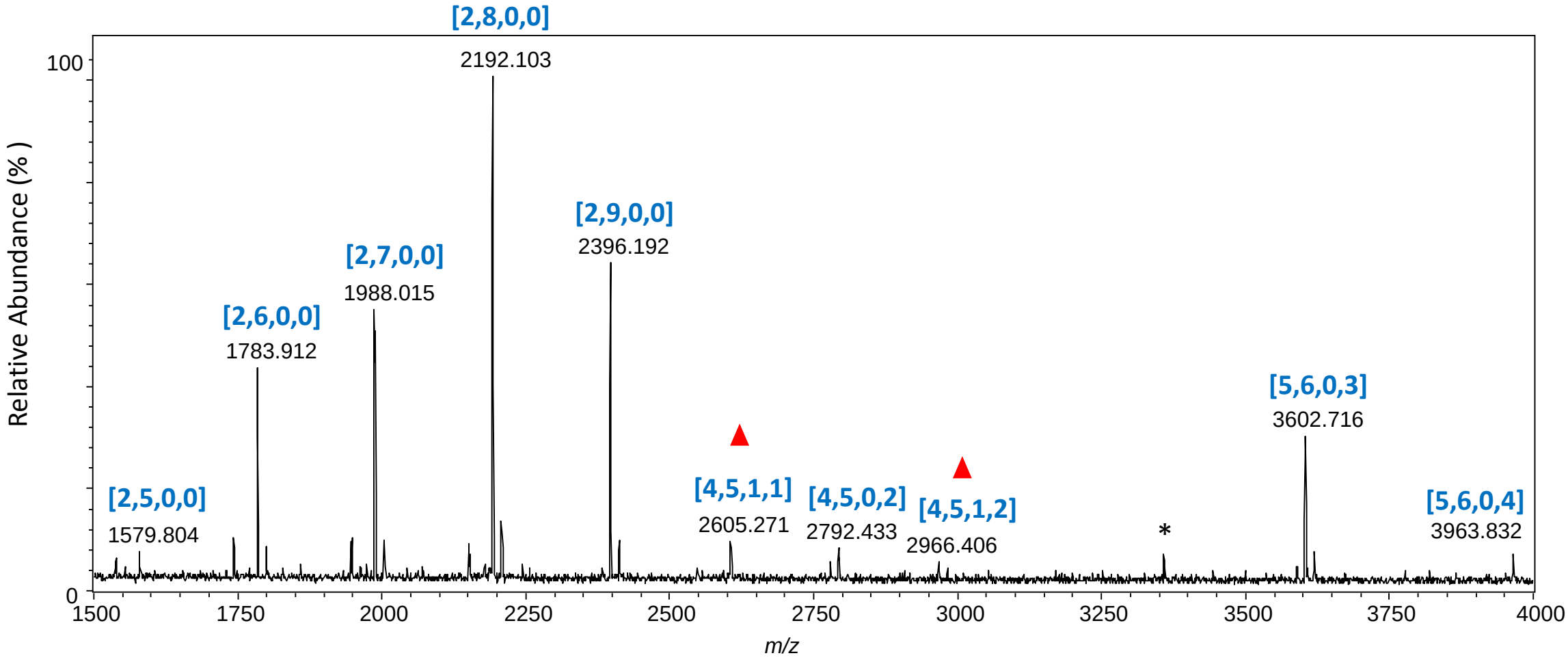


Supplementary Figure 14: HCD Spectrum of EGFR Site N599 Glycopeptide ^{586}YADAGHVCHLCHPNCTYGCTGPGLEGCPTNGPK^{618} + $HexNAc_2Hex_7$. The glycopeptide HCD spectrum (20% collision energy) of the precursor with $[M + 5H]^{5+}$ *m/z* 1040.4169, is shown. * unknown; **co-isolated precursor (overlap of two isotope distributions).

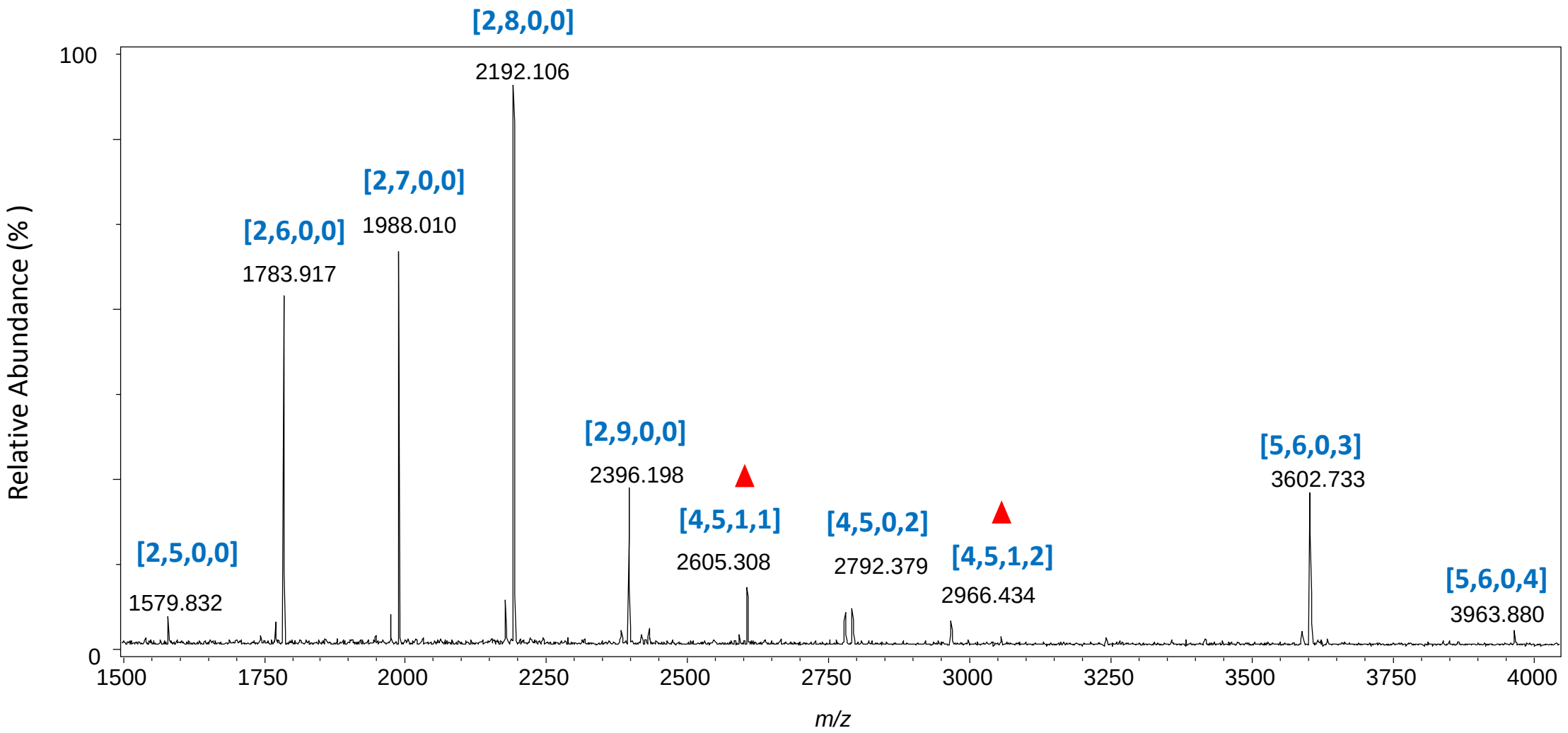
.



Supplemental Figure 15A: MALDI-TOF MS Spectrum of Released, Permethylated *N*-Glycans from Indolent CAL27 Cells. Released, sodiated, permethylated *N*-Glycans from untreated CAL27 whole cell lysate. The composition of each glycan is shown, in the format [*N*-acetylhexosamine (HexNAc), Hexose (Hex), deoxyhexose (Fuc), and *N*-acetylneuraminic acid (NeuAc)]. Red triangles indicate glycan compositions that correspond to fucosylated glycoforms.

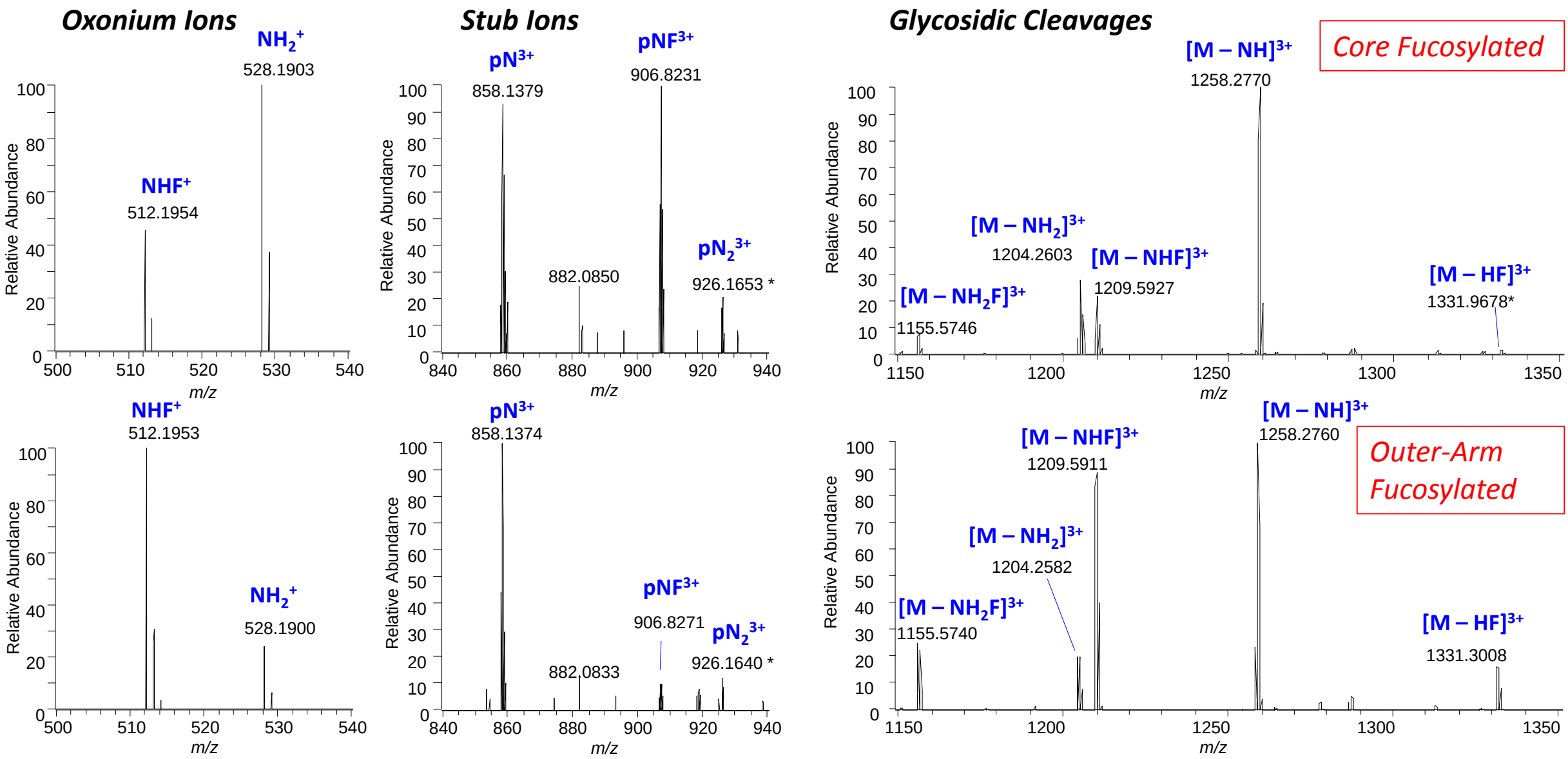


Supplemental Figure 15B: MALDI-TOF MS Spectrum of Released, Permethylated *N*-Glycans from Metastatic HSC-3 Cells. Released, sodiated, permethylated *N*-Glycans from untreated HSC-3 whole cell lysate. The compositions of each glycan are shown, in the format [*N*-acetylhexosamine (HexNAc), Hexose (Hex), deoxyhexose (Fuc), and *N*-acetylneuraminic acid (NeuAc)]. Red triangles indicate glycan compositions that correspond to fucosylated glycoforms; * unknown.



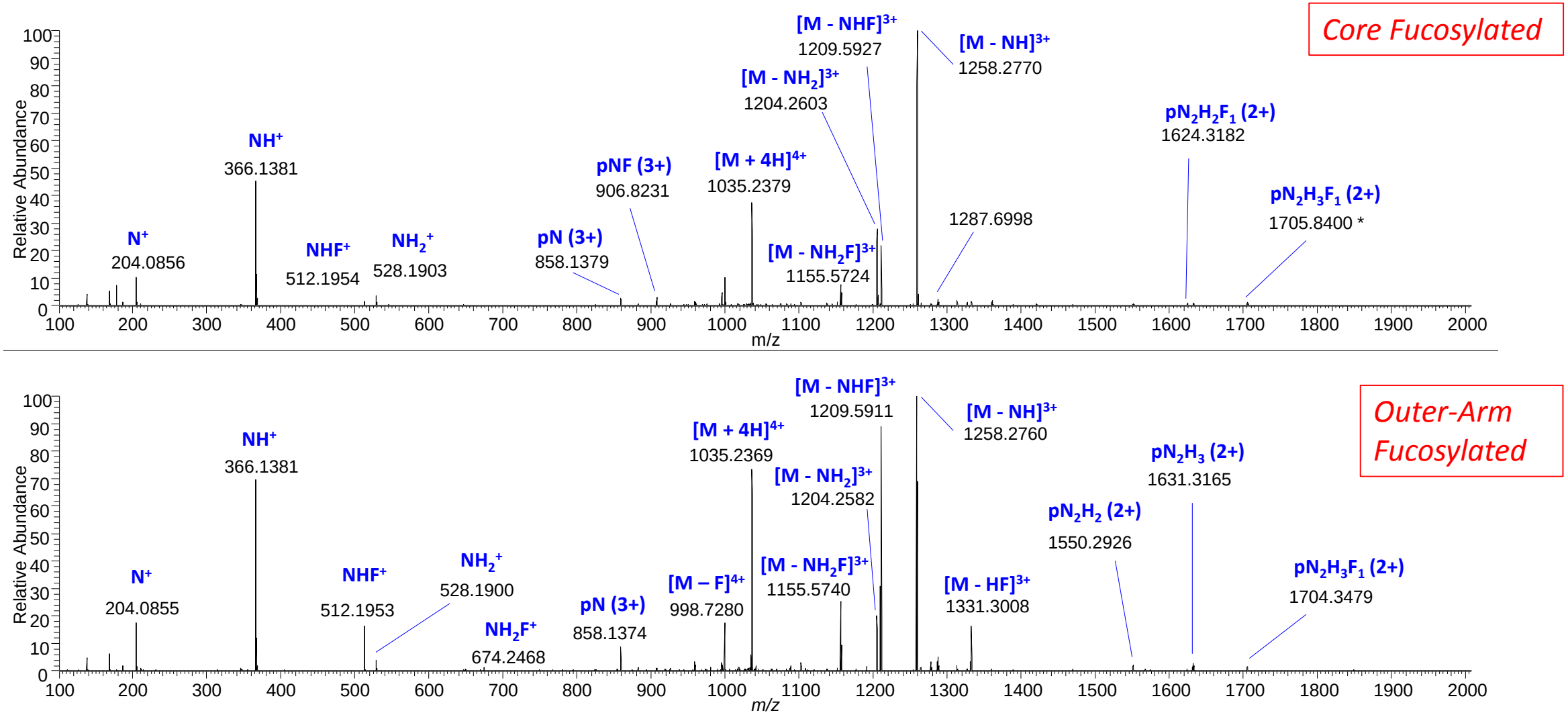
Supplemental Figure 15C: MALDI-TOF MS Spectrum of Released, Permethylated *N*-Glycans from Metastatic HSC-3 Cells Treated with ICG-001. Released, sodiated, permethylated *N*-Glycans from the whole cell lysate of HSC-3 cells treated with 10 µM ICG-001. The compositions of each glycan are shown, in the format [*N*-acetylhexosamine (HexNAc), Hexose (Hex), deoxyhexose (Fuc), and *N*-acetylneuraminic acid (NeuAc)]. Red triangles indicate glycan compositions that correspond to fucosylated glycoforms.

EGFR Site N420: ^{406}TKQHGQFSLAVVSLNITSLGLR^{427} + $HexNAc_4Hex_5Fuc_1$



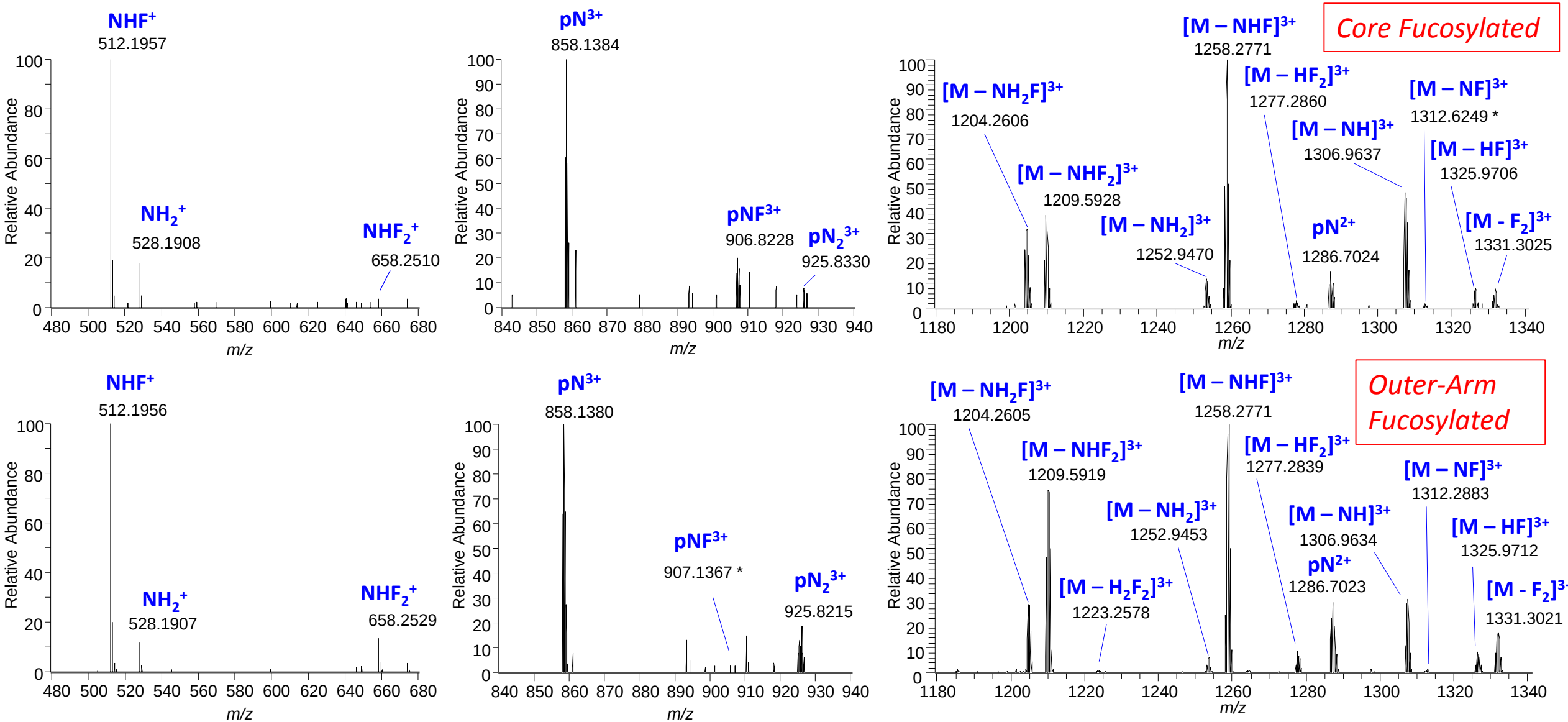
Supplementary Figure 16A: Comparison of Early and Late-Eluting EGFR Site N420 Glycopeptide ^{406}TKQHGQFSLAVVSLNITSLGLR^{427} + $HexNAc_4Hex_5Fuc_1$ HCD Spectra. Selected *m/z* regions of the glycopeptide HCD spectra (10% collision energy) are shown of the precursor with $[M + 4H]^{4+}$ *m/z* 1035.2483, from early eluting (top) and late eluting (bottom) peaks, are shown. N: *N*-acetylhexosamine; H: hexose; F: fucose.

EGFR Site N420: ^{406}TKQHGQFSLAVVSLNITSLGLR^{427} + $HexNAc_4Hex_5Fuc_1$



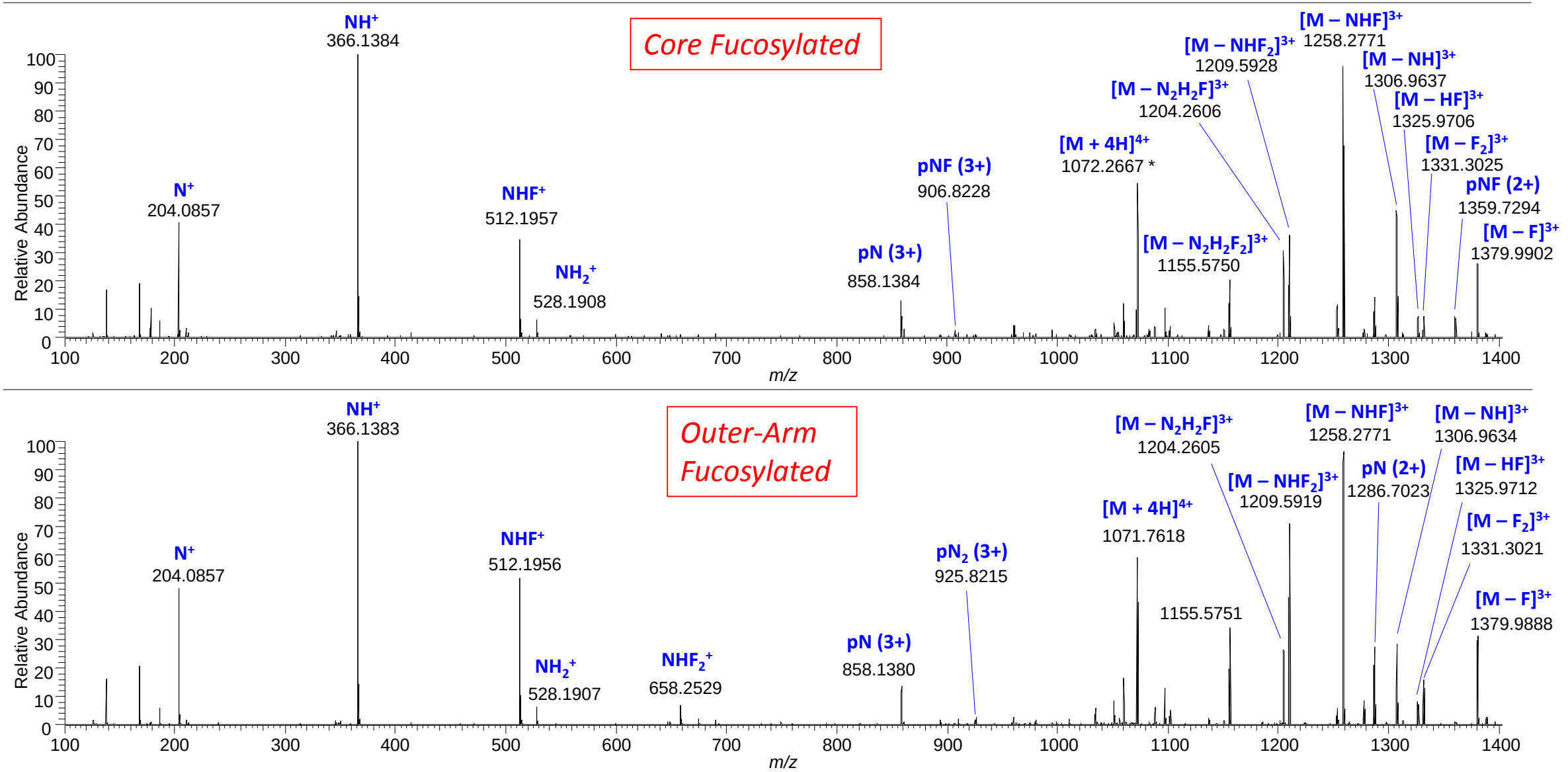
Supplementary Figure 16B: Comparison of Early and Late-Eluting EGFR Site N420 Glycopeptide ^{406}TKQHGQFSLAVVSLNITSLGLR^{427} + $HexNAc_4Hex_5Fuc_1$ HCD Spectra. The glycopeptide HCD spectra (10% collision energy) of the precursor with $[M + 4H]^{4+}$ *m/z* 1035.2483, from early eluting (top) and late eluting (bottom) peaks, are shown, *m/z* range 100 - 2000. N: *N*-acetylhexosamine; H: hexose; F: fucose; p: peptide; * monoisotopic *m/z* not detected.

EGFR Site N420: ^{406}TKQHGQFSLAVVSLNITSLGLR^{427} + $HexNAc_4Hex_5Fuc_2$



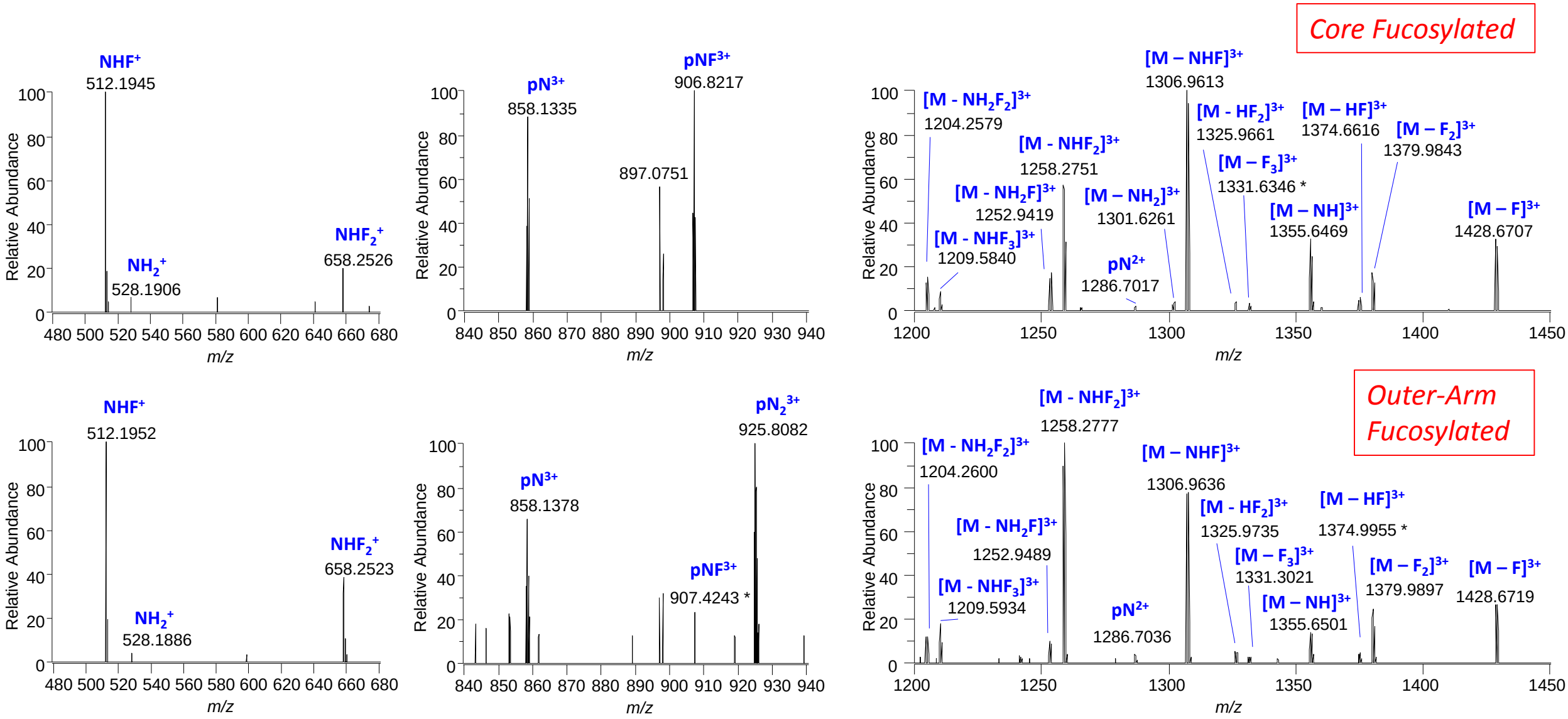
Supplementary Figure 17A: Comparison of Early and Late-Eluting EGFR Site N420 Glycopeptide ^{406}TKQHGQFSLAVVSLNITSLGLR^{427} + $HexNAc_4Hex_5Fuc_2$ HCD Spectra. Selected regions are shown (expanded) from the glycopeptide HCD spectra (15% collision energy) of the precursor with $[M + 4H]^{4+}$ *m/z* 1071.7624, from early eluting (top) and late eluting (bottom) peaks, are shown. N: *N*-acetylhexosamine; H: hexose; F: fucose; * monoisotopic *m/z* not detected.

EGFR Site N420: ^{406}TKQHGQFSLAVVSLNITSLGLR^{427} + $HexNAc_4Hex_5Fuc_2$



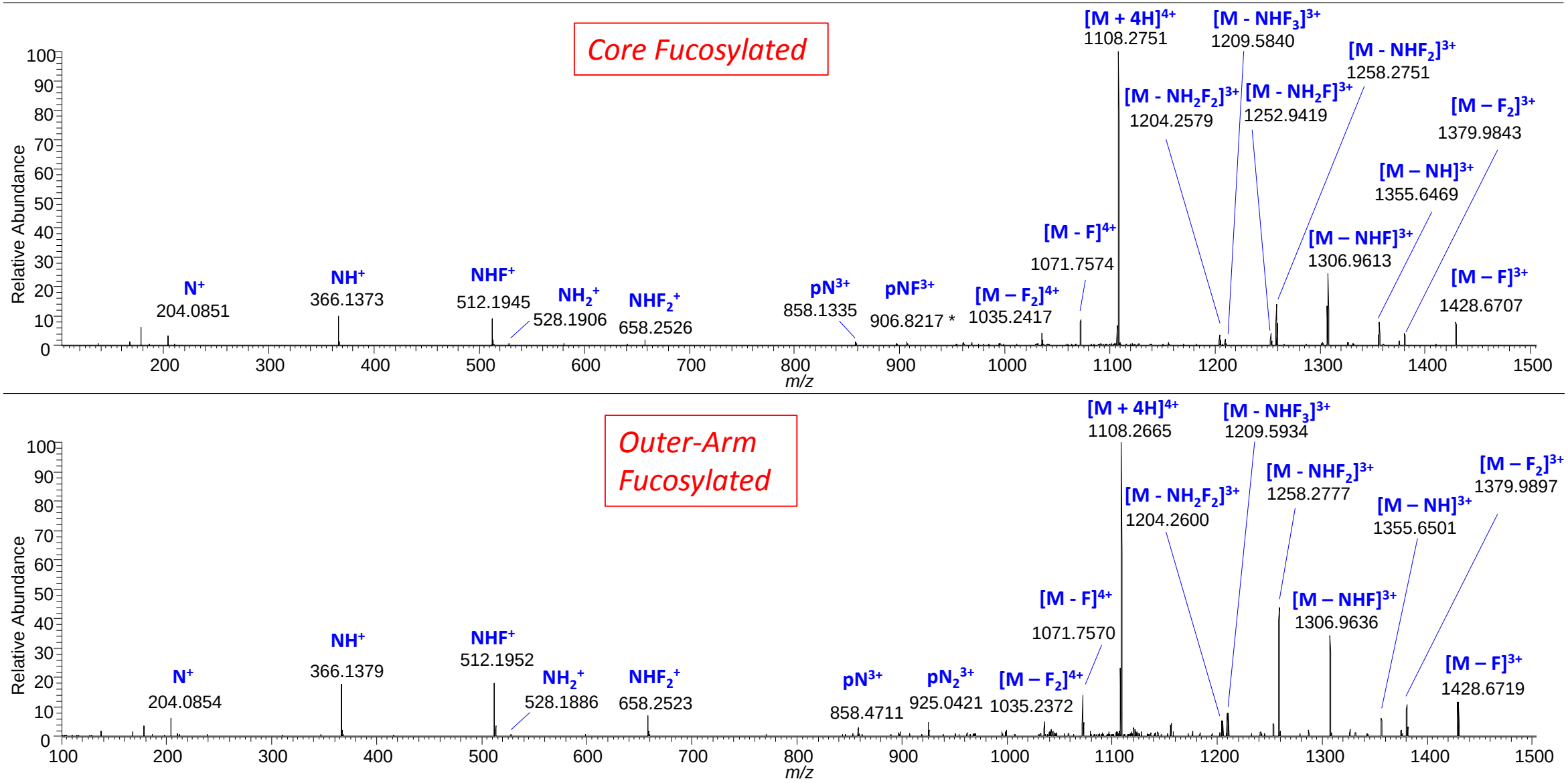
Supplementary Figure 17B: Comparison of Early and Late-Eluting EGFR Site N420 Glycopeptide ^{406}TKQHGQFSLAVVSLNITSLGLR^{427} + $HexNAc_4Hex_5Fuc_2$ HCD Spectra. The glycopeptide HCD spectra (15% collision energy) of the precursor with $[M + 4H]^{4+}$ *m/z* 1071.7624, from early eluting (top) and late eluting (bottom) peaks, *m/z* range 100-1400, are shown. N: *N*-acetylhexosamine; H: hexose; F: fucose ; * monoisotopic *m/z* not detected.

EGFR Site N420: ^{406}TKQHGQFSLAVVSLNITSLGLR^{427} + $HexNAc_4Hex_5Fuc_3$



Supplementary Figure 18A: Comparison of Early and Late-Eluting EGFR Site N420 Glycopeptide ^{406}TKQHGQFSLAVVSLNITSLGLR^{427} + $HexNAc_4Hex_5Fuc_3$ HCD Spectra. Selected regions from the glycopeptide HCD spectra (10% collision energy) of the precursor with $[M + 4H]^{4+}$ *m/z* 1108.2771, from early eluting (top) and late eluting (bottom) peaks, are shown. N: *N*-acetylhexosamine; H: hexose; F: fucose ; * monoisotopic *m/z* not detected.

EGFR Site N420: ^{406}TKQHGQFSLAVVSLNITSLGLR^{427} + $HexNAc_4Hex_5Fuc_3$



Supplementary Figure 18B: Comparison of Early and Late-Eluting EGFR Site N420 Glycopeptide ^{406}TKQHGQFSLAVVSLNITSLGLR^{427} + $HexNAc_4Hex_5Fuc_3$ HCD Spectra. The glycopeptide HCD spectra (10% collision energy) of the precursor with $[M + 4H]^{4+}$ *m/z* 1108.2771, from early eluting (top) and late eluting (bottom) peaks, *m/z* range 100-1500, are shown; * monoisotopic *m/z* not detected.