

# HLA Allele-specific Quantitative Profiling of Type 1 Diabetic B Lymphocyte Immunopeptidome

Putty-Reddy Sudhir, Tai-Du Lin, Qibin Zhang\*

## Figure captions

**Fig S1.** Scatter plots represent the LC-MS/MS data quality of HLA-I immunopeptidomics by pair-wise Pearson correlations ( $r$ ) of five replicate analyses of the HLA-I immunopeptides enriched from 2676 (T1D) and 2675 (HC) B cells.

**Fig S2.** Gibbscluster-2.0 was run separately using the HLA-I and HLA-II immunopeptides of two B cell lines (2676 (T1D) and 2675 (HC)) to identify their consensus binding motif sequences (this data is from Fig. 2E and 3E). The identified motifs were verified with the naturally presented ligand sequences of HLA alleles obtained from Motif viewer.

## Fig S3.

Scatter plots represent the LC-MS/MS data quality of HLA-II immunopeptidomics by pair-wise Pearson correlations ( $r$ ) of five replicate analyses of the HLA-II immunopeptides enriched from 2676 (T1D) and 2675 (HC) B cells.

**Fig. S4. Quantitative composition of cellular functions of the HLA-I and II antigens dysregulated in T1D.** *A*, Proteomap visualization of antigens presented on B cell HLA-I complexes. *B*, Proteomap visualization of antigens presented on B cell HLA-II complexes. The left-side panels show the cellular functions and right-side panels show the antigens associated with those cellular functions. The proteomaps were generated using the antigens, which possess the 2-fold differentially regulated immunopeptides (DIPs) in T1D B cells (2676 (T1D) vs. 2675 (HC), t-test FDR, 0.01). The cellular function regions dominate with the identified antigens whereas the antigen regions were weighted with the number of dysregulated immunopeptides in T1D.

**Fig S5.** List of HLA-I immunopeptides selected for the peptide sequence verification by parallel reaction monitoring (PRM) analysis of spiked stable isotope labelled synthetic peptides.

**Fig S6.** List of HLA-II immunopeptides selected for the peptide sequence verification by parallel reaction monitoring (PRM) analysis of spiked stable isotope labelled synthetic peptides and MixMHC2pred results.

**Fig S7.** The chromatograms represent the HLA-I immunopeptide sequences verified by parallel reaction monitoring (PRM) analysis of endogenous samples spiked stable isotope labelled synthetic peptides. The b or y ions coeluted or interfered with other ions were excluded from the chromatograms.

**Fig S8.** The chromatograms represent the HLA-II immunopeptide sequences verified by parallel reaction monitoring (PRM) analysis of endogenous samples spiked stable isotope labelled synthetic peptides. The b or y ions coeluted or interfered with other ions were excluded from the chromatograms. For some peptides, the HLA-II endogenous peptides identified in two cell lines experiment were compared with SIS peptide fragments by Mirror Plots, which were created using SigmaPlot.

**Fig S9.** Immunopeptides enriched from six additional B cell lines (2679 (HC), 2812 (HC), 2741 (HC), 2678 (T1D), 2765 (T1D), and 3027 (T1D)) were analyzed in two replicates (TR1 and TR2) by LC-MS/MS. Figure S8A shows the number of HLA-I (left panel) and HLA-II (right panel) immunopeptides detected in six B cell lines. Figure S8B represents the length of the HLA-I (left panel) and HLA-II (right panel) immunopeptides identified in six B cell lines listed in Figure 8AB.

**Fig S10.** HLA-I Immunopeptides identified in six B cell lines (2679 (HC), 2812 (HC), 2741 (HC), 2678 (T1D), 2765 (T1D), and 3027 (T1D)) were assigned to HLA allotypes based on their binding affinities to HLA-I alleles (Figure S9A). HLA-II Immunopeptides identified in six B cell lines were assigned to the different HLA allotypes based on their binding affinities to HLA-II alleles (Figure S9B). The binding affinities were predicted using NetMHCpan 4.1 and NetMHCIIpan 4.0 in case of HLA-I and HLA-II, respectively.

**Fig S11.** Gibbscluster-2.0 was run separately using the HLA-I immunopeptides of each of the six B cell lines (2679 (HC), 2812 (HC), 2741 (HC), 2678 (T1D), 2765 (T1D), and 3027 (T1D)) to identify their consensus binding motif sequences. The identified motifs were verified with the naturally presented ligand sequences of HLA alleles obtained from Motif viewer.

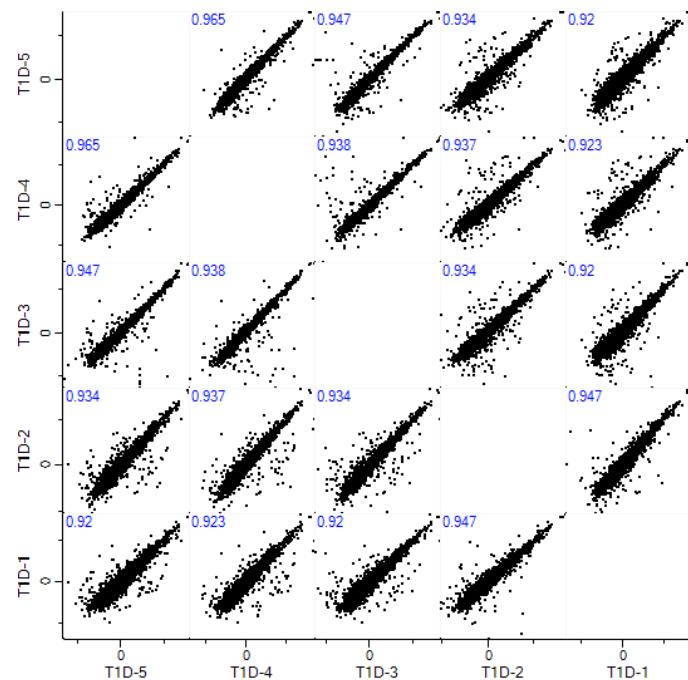
**Fig S12.** Gibbscluster-2.0 was run separately using the HLA-II immunopeptides list of each of the six B cell lines (2679 (HC), 2812 (HC), 2741 (HC), 2678 (T1D), 2765 (T1D), and 3027 (T1D)) to identify their consensus binding motif sequences. The identified motifs were verified with the naturally presented ligand sequences of HLA alleles obtained from Motif viewer.

**Fig S13.** LC-MS/MS analysis of HLA-I and HLA-II immunopeptides of T1D and healthy B cells resulted in identification of the peptides that are part of HLA molecules. These peptides were listed along with their information reported by previous studies in IEDB database.

<b>Figures</b>	<b><u>Page No.</u></b>
Fig. S1	04
Fig. S2	05
Fig. S3	06
Fig. S4	07
Fig. S5	08
Fig. S6	09
Fig. S7	10-37
Fig. S8	38-67
Fig. S9	68
Fig. S10	69
Fig. S11	70
Fig. S12	71
Fig. S13	72

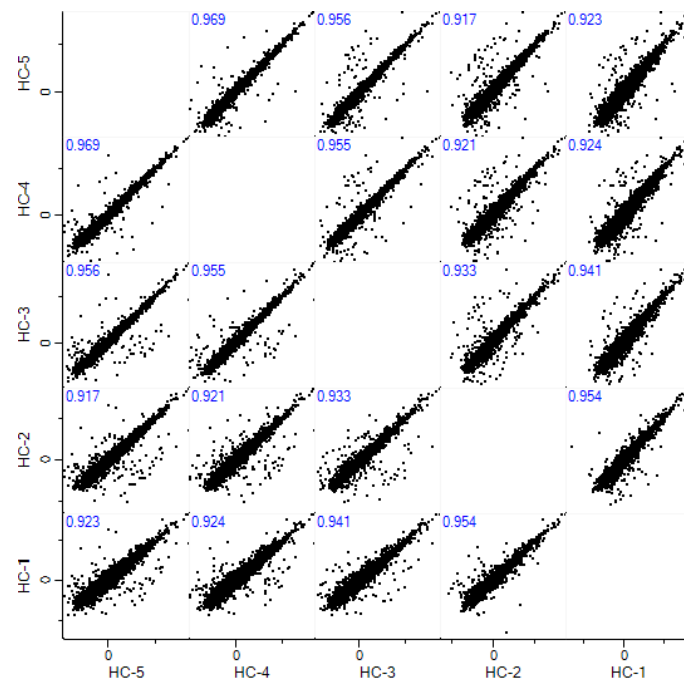
**Fig. S1**

**S1A**



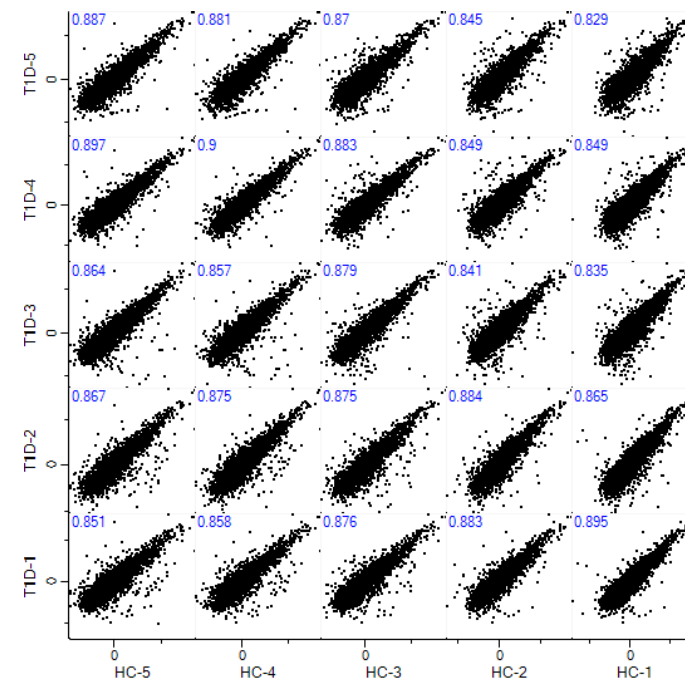
Correlation between five T1D replicates ( $r = 0.92 - 0.965$ )

**S1B**



Correlation between five HC replicates ( $r = 0.917 - 0.969$ )

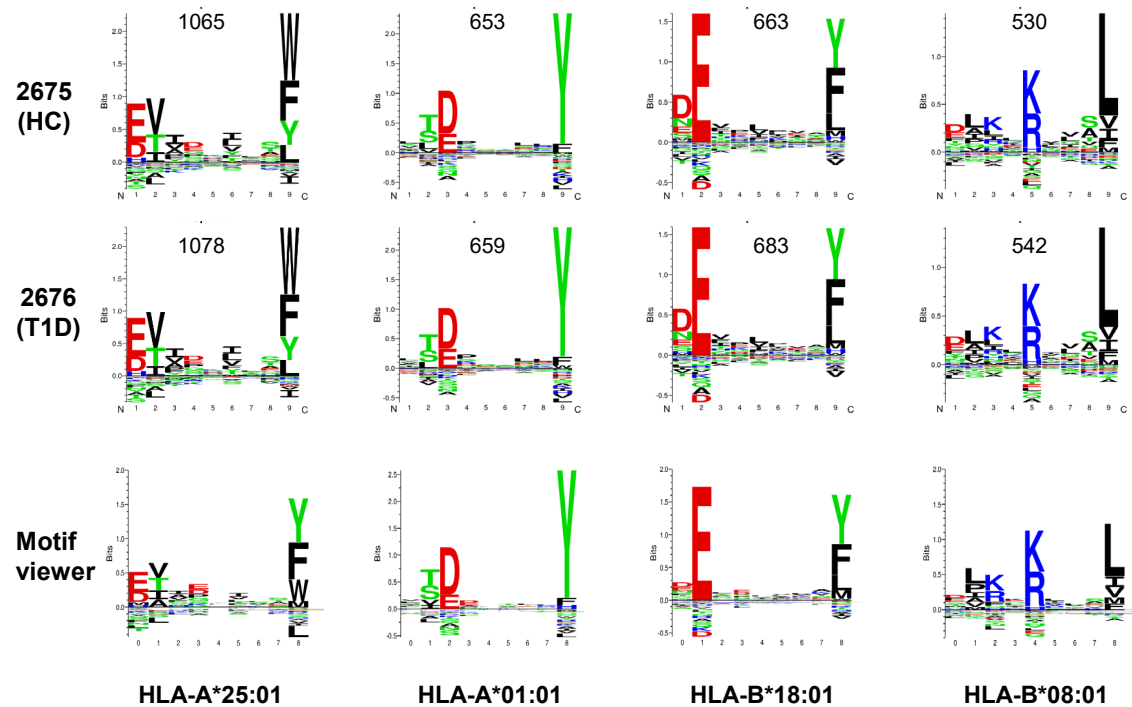
**S1C**



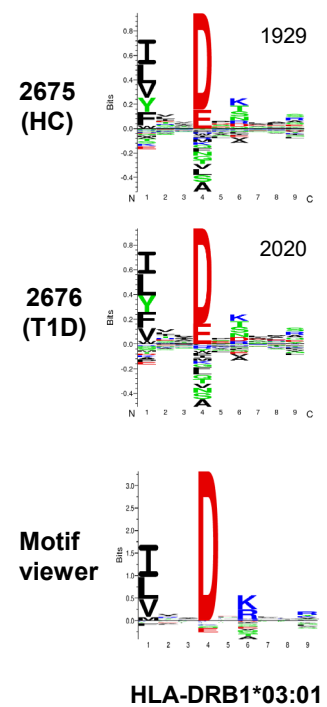
Correlation between five T1D and five HC replicates ( $r = 0.829 - 0.9$ )

Fig. S2

### HLA-I

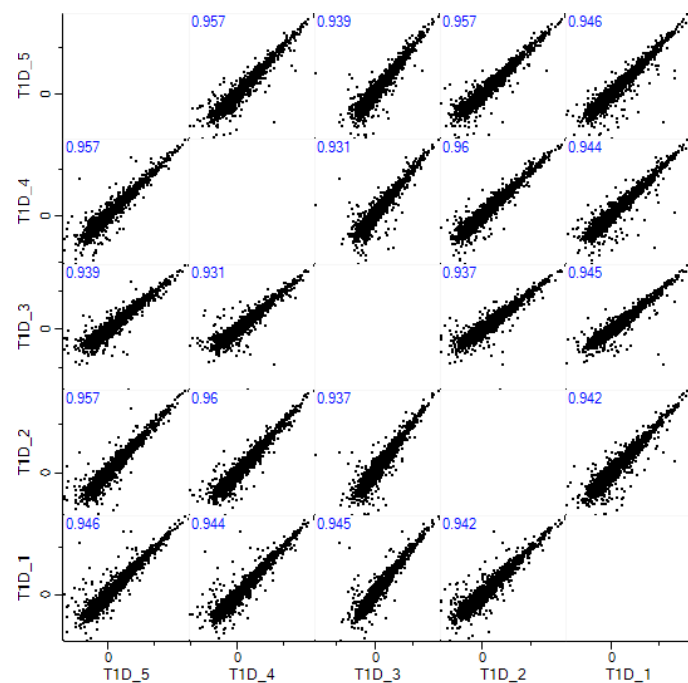


### HLA-II



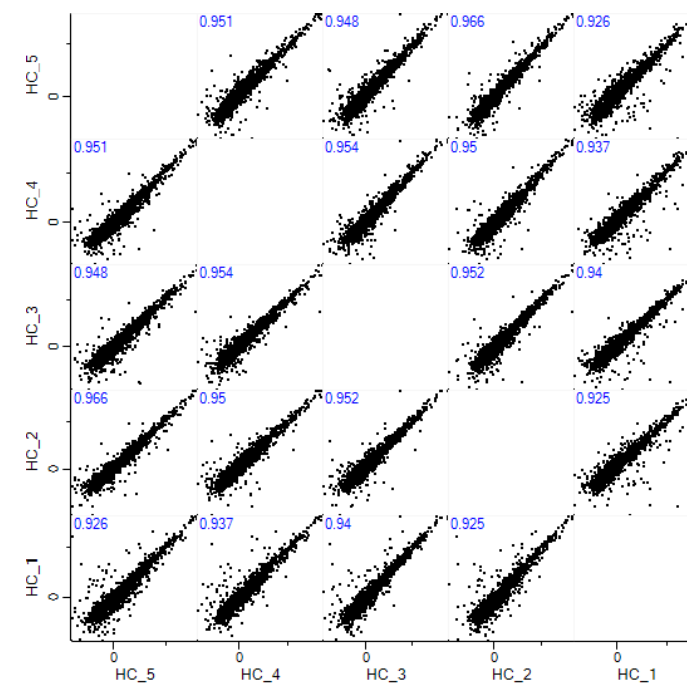
**Fig. S3**

**S3A**



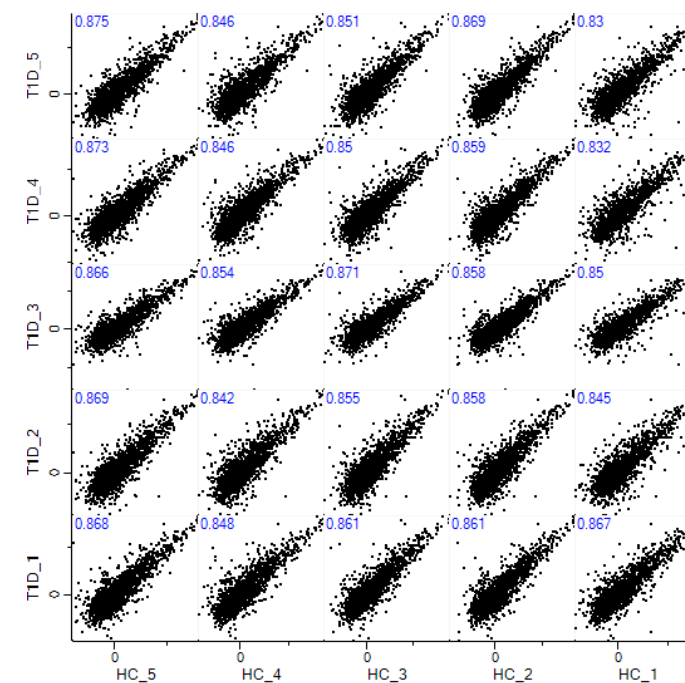
Correlation between five T1D replicates ( $r = 0.931 - 0.96$ )

**S3B**



Correlation between five HC replicates ( $r = 0.925 - 0.966$ )

**S3C**



Correlation between five T1D and five HC replicates ( $r = 0.83 - 0.875$ )



**Fig. S5**

No.	Uniprot ID	Protein symbol	Protein name	Sequence	Length	Log2 Fold change	HLA class	HLA alleles	Criteria		
									Significant difference	Biological significance *	ID in T1D only
1	Q8IV61	RASGRP3	Ras guanyl-releasing protein 3	EVITKFINV	9	2.11	I	HLA-B08:01	yes	&	
2	Q9H1A4	ANAPC1	Anaphase-promoting complex subunit 1	NTDPSIVMTY	10	1.91	I	HLA-A01:01	yes	& #	
3	Q02223	TNFRSF17	Tumor necrosis factor receptor superfamily member 17	DEIILPRGLEY	11	1.85	I	HLA-B18:01	yes	&	
4	Q96G74	OTUD5	OTU domain-containing protein 5	NEDEPIRVSY	10	1.66	I	HLA-B18:01	yes	&	
5	P01374	LTA	Lymphotoxin-alpha	HEVQLFSSQY	10	1.65	I	HLA-B18:01	yes	&	
6	P68104	EEF1A1	Elongation factor 1-alpha 1	ESFSDYPPLGRF	12	1.59	I	HLA-A01:01	yes	&	
7	Q9UJX3	ANAPC7	Anaphase-promoting complex subunit 7	LLGSLADLY	9	1.59	I	HLA-A01:01	yes	& #	
8	Q03518	TAP1	Antigen peptide transporter 1	TALPRIFSL	9	1.58	I	HLA-C12:03	yes	& #	
9	P42575	CASP2	Caspase-2	DMLVKVNAL	9	1.56	I	HLA-B08:01	yes	&	
10	P28062	PSMB8	Proteasome subunit beta type-8	VSDLLHQY	8	1.49	I	HLA-A01:01	yes	& #	
11	Q7L576	CYFIP1	Cytoplasmic FMR1-interacting protein 1	EVISGYEEL	9	1.45	I	HLA-A25:01	yes	&	
12	P42574	CASP3	Caspase-3	STAPGYYSW	9	1.28	I	HLA-C12:03	yes	&	
13	Q9NZ08	ERAP1	Endoplasmic reticulum aminopeptidase 1	ETIEENIGW	9	1.27	I	HLA-A25:01	yes	& #	
14	P40305	IFI27	Interferon alpha-inducible protein 27, mitochondrial	AAVPMVLSA	9	1.27	I	HLA-C12:03	yes	&	
15	Q01518	CAP1	Adenylyl cyclase-associated protein 1	NEFPVPEQF	9	1.27	I	HLA-B18:01	yes	&	
16	P42224	STAT1	Signal transducer and activator of transcription 1-alpha/beta	DSFPMERQY	10	1.21	I	HLA-A01:01	yes	&	
17	P41218	MNDA	Myeloid cell nuclear differentiation antigen	EVPNRIIEI	9	1.13	I	HLA-A25:01	yes	&	
18	P14625	HSP90B1	Endoplasmic	MMKLIINSL	9	1.12	I	HLA-B08:01	yes	&	
19	Q9UH77	KLHL3	Kelch-like protein 3	EVYDPGNTW	10	1.10	I	HLA-A25:01	yes	& #	
20	Q5SRE5	NUP188	Nucleoporin NUP188 homolog	EVAPSGTTL	9	1.06	I	HLA-A25:01	yes	&	
21	Q53GT1	KLHL22	Kelch-like protein 22	EEVLIHGVSY	10	1.05	I	HLA-B18:01	yes	& #	
22	Q96PU5	NEDD4L	E3 ubiquitin-protein ligase NEDD4-like	ASDPYVKLSLY	11	1.01	I	HLA-A01:01	yes	& #	
23	P12270	TPR	Nucleoprotein TPR	HLNTEKELL	8		I	HLA-B08:01		&	yes
24	O00463	TRAF5	TNF receptor-associated factor 5	EVIKSQEVF	9		I	HLA-A25:01			yes
25	O15439	ABCC4	Multidrug resistance-associated protein 4	EVITGIRII	9		I	HLA-A25:01			yes
26	Q9NX02	NLRP2	NACHT, LRR and PYD domains-containing protein 2	LLDEGAKLLY	10		I	HLA-A01:01			yes
27	P01850	TRBC1	T-cell receptor beta-1 chain C region	YEILLGKATLY	11		I	HLA-B18:01			yes

\* STRING database shown proteins associated with reactome pathways, immune system (&) and/or MHC-I antigen presenting and processing (#)

HLA-I alleles with relatively higher binding affinity for immunopeptides were listed.



**Fig. S6**

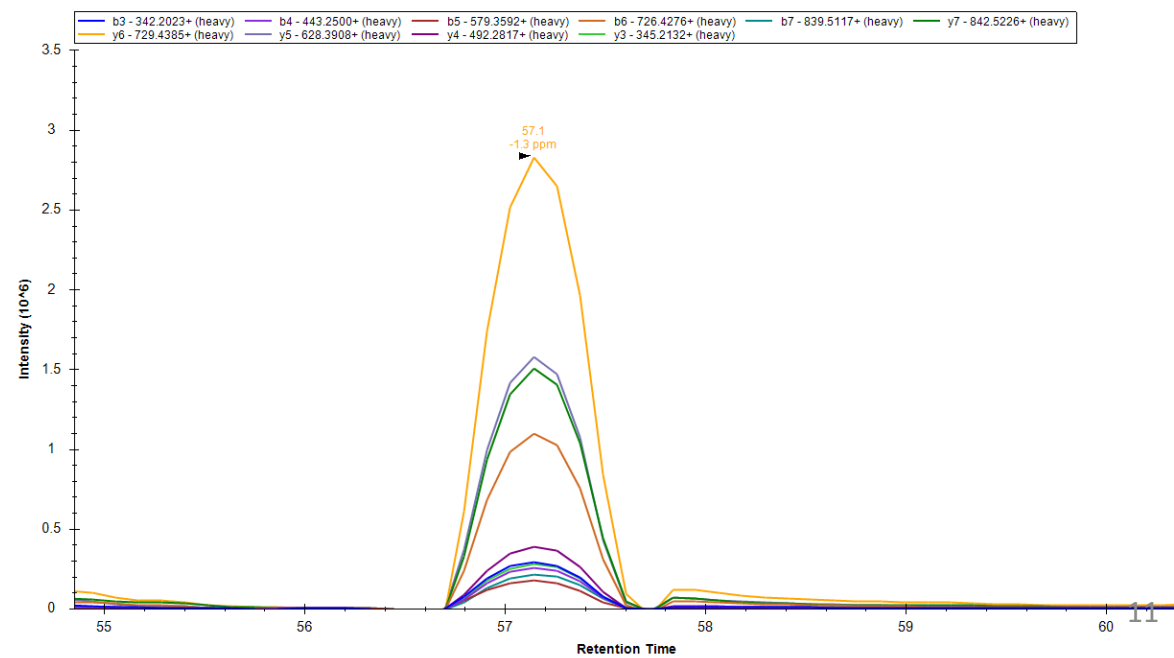
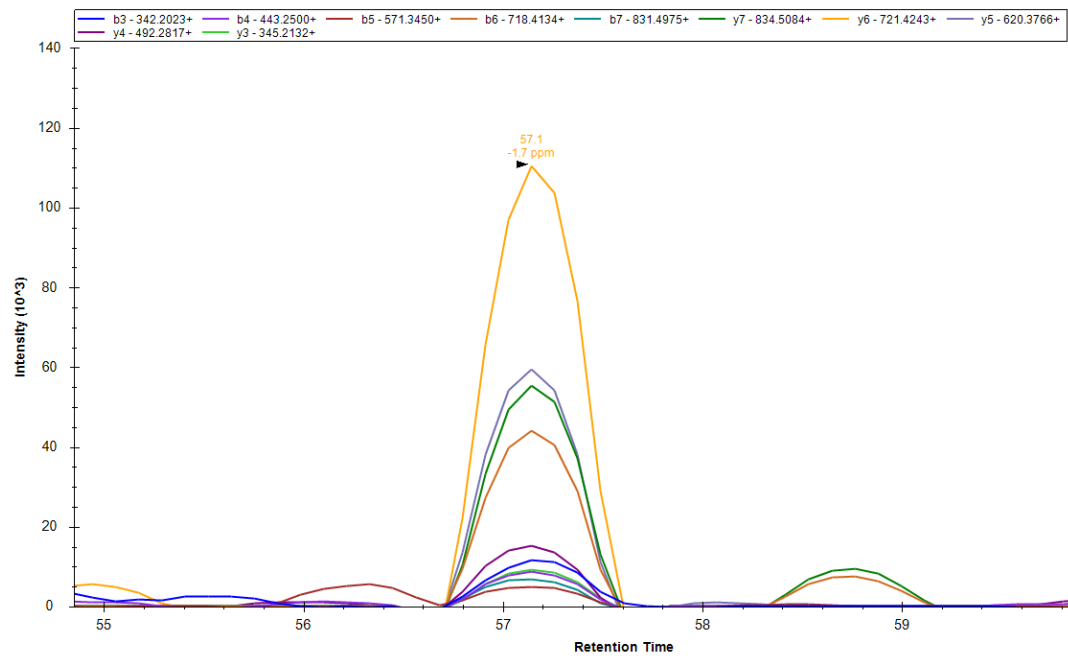
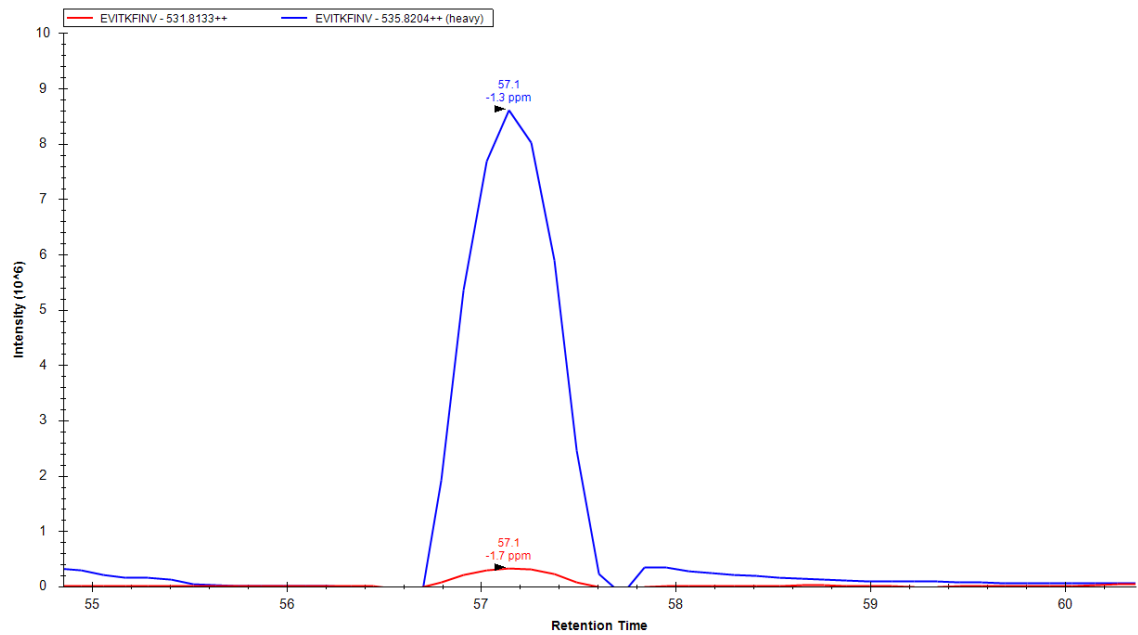
No.	Uniprot ID	Protein symbol	Protein name	Sequence	Length	Log2 Fold change	HLA class	HLA alleles	Criteria			MixMHC2pred %Rank_best score	PRM
									Significant difference	Biological significance*	ID in T1D only		
1	P09668	CTSH	Pro-cathepsin H	LPSQAFEYILYNGKIM	16	3.23	II	HLA-DPA1*01:03-DPB1*04:01	yes	& #		0.039	yes
2	P61769	B2M	Beta-2-microglobulin	YLLYYTEFTPTKDEY	16	2.52	II	HLA-DPA1*01:03-DPB1*04:01	yes	&		0.482	yes
3	O75787	ATP6AP2	Renin receptor	DASKILVDALQKFADD	16	2.42	II	HLA-DRB1*03:01	yes	&		0.129	yes
4	P26951	IL3RA	Interleukin-3 receptor subunit alpha	GPGAPADVQYDLYLNVANR	19	2.31	II	HLA-DPA1*01:03-DPB1*04:01	yes	&		0.286	yes
5	P01374	LTA	Lymphotoxin-alpha	HSMYHGAAFLTQGGDQ	16	2.24	II	HLA-DQA1*05:01-DQB1*03:02	yes	&		NA	
6	O14672	ADAM10	Disintegrin and metalloproteinase domain-containing protein 10	FPNIGVEKFLELNSE	15	2.11	II	HLA-DPA1*01:03-DPB1*04:01	yes	&		0.00988	yes
7	P07437	TUBB	Tubulin beta chain	GSDLQLDRISVYYNEA	17	1.93	II	HLA-DRB1*03:01	yes	&		0.737	yes
8	P11142	HSPA8	Heat shock cognate 71 kDa protein	VLRIINEPTAAAIAYG	16	1.78	II	HLA-DQA1*05:01-DQB1*03:02	yes	&		NA	yes
9	P09668	CTSH	Pro-cathepsin H	LPSQAFEYILYNGKI	15	1.77	II	HLA-DPA1*01:03-DPB1*04:01	yes	& #		0.47	yes
10	P13164	IFITM1	Interferon-induced transmembrane protein 1	VPDHVVWSLFNTL	13	1.51	II	HLA-DPA1*01:03-DPB1*02:01	yes	&		0.0548	yes
11	P09668	CTSH	Pro-cathepsin H	LPSQAFEYILYNGK	14	1.38	II	HLA-DPA1*01:03-DPB1*04:01	yes	& #		0.288	yes
12	Q12913	PTPRJ	Receptor-type tyrosine-protein phosphatase eta	DVYGIYDLRMHRP	14	1.37	II	HLA-DRB1*03:01	yes	&		0.00412	
13	O15143	ARPC1B	Actin-related protein 2/3 complex subunit 1B	ALTFITDNSLVAAGHD	16	1.28	II	HLA-DRB1*04:01	yes	&		0.0195	yes
14	Q13114	TRAF3	TNF receptor-associated factor 3	GDAFKPDPNSSSFKKPT	17	1.21	II	HLA-DRB1*04:01	yes	&		0.316	yes
15	P11142	HSPA8	Heat shock cognate 71 kDa protein	ERAMTKDNLLGKFEL	16	1.15	II	HLA-DRB1*03:01	yes	&		0.382	yes
16	Q9H3Z4	DNAJC5	DnaJ homolog subfamily C member 5	TTQLTADSHPSYHTDG	16	1.11	II	HLA-DRB1*03:01	yes	&		0.757	yes
17	P25774	CTSS	Cathepsin S	NGGFMTTAFQYIIDNK	16	1.10	II	HLA-DPA1*01:03-DPB1*04:01	yes	& #		0.0119	yes
18	P11142	HSPA8	Heat shock cognate 71 kDa protein	VLRIINEPTAAAIAY	15	1.04	II	HLA-DQA1*05:01-DQB1*03:02	yes	&		NA	yes
19	P14618	PKM	Pyruvate kinase PKM	DENILWLDYKNICKVVE	17		II	HLA-DRB1*03:01		&	yes	0.497	
20	P21580	TNFAIP3	Tumor necrosis factor alpha-induced protein 3	EINLVDDYFELVQHE	15		II	HLA-DQA1*03:01-DQB1*02:01		&	yes	NA	yes
21	P52907	CAPZA1	F-actin-capping protein subunit alpha-1	FNEVFNDVRLNNDN	16		II	HLA-DRB1*03:01		& #	yes	0.00543	
22	P10147	CCL3	C-C motif chemokine 3	IPQNFADYFETSSQ	15		II	HLA-DPA1*01:03-DPB1*04:01		&	yes	0.0053	yes
23	O00754	MAN2B1	Lysosomal alpha-mannosidase	IRATFDPDTGLLMEIM	16		II	HLA-DRB1*03:01		&	yes	0.117	
24	P26951	IL3RA	Interleukin-3 receptor subunit alpha	GPGAPADVQYDLYLNVA	17		II	HLA-DPA1*01:03-DPB1*04:01		&	yes	0.434	
25	P09668	CTSH	Pro-cathepsin H	LPSQAFEYILYNGKIMGED	19		II	HLA-DPA1*01:03-DPB1*04:01		& #	yes	0.754	yes

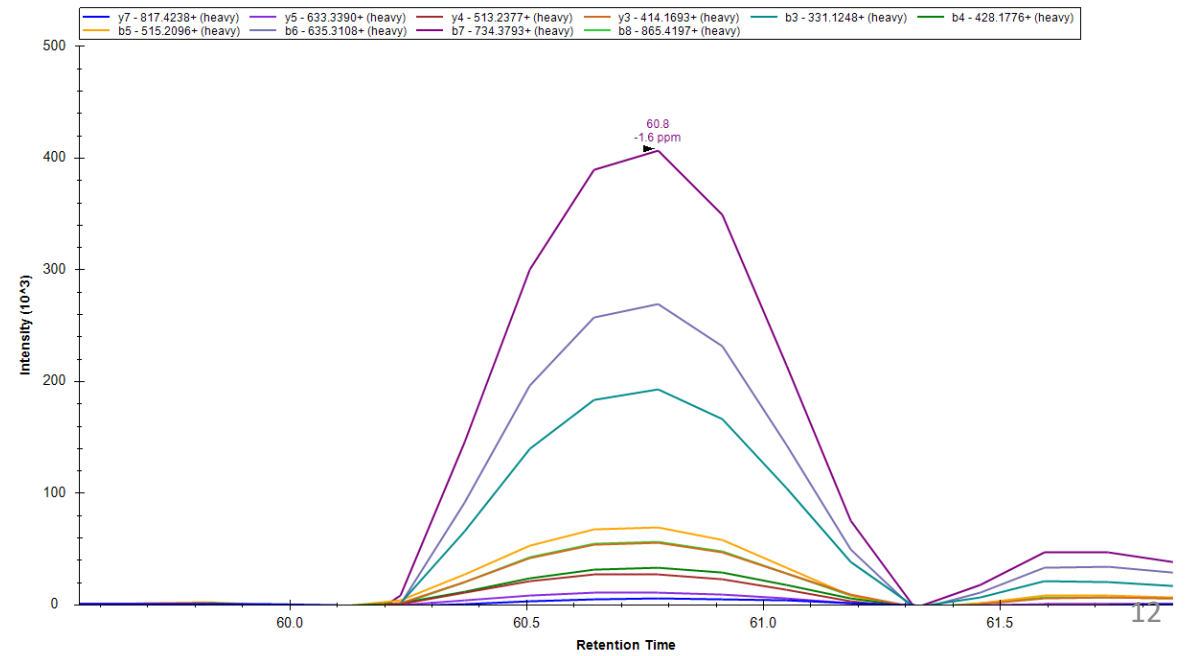
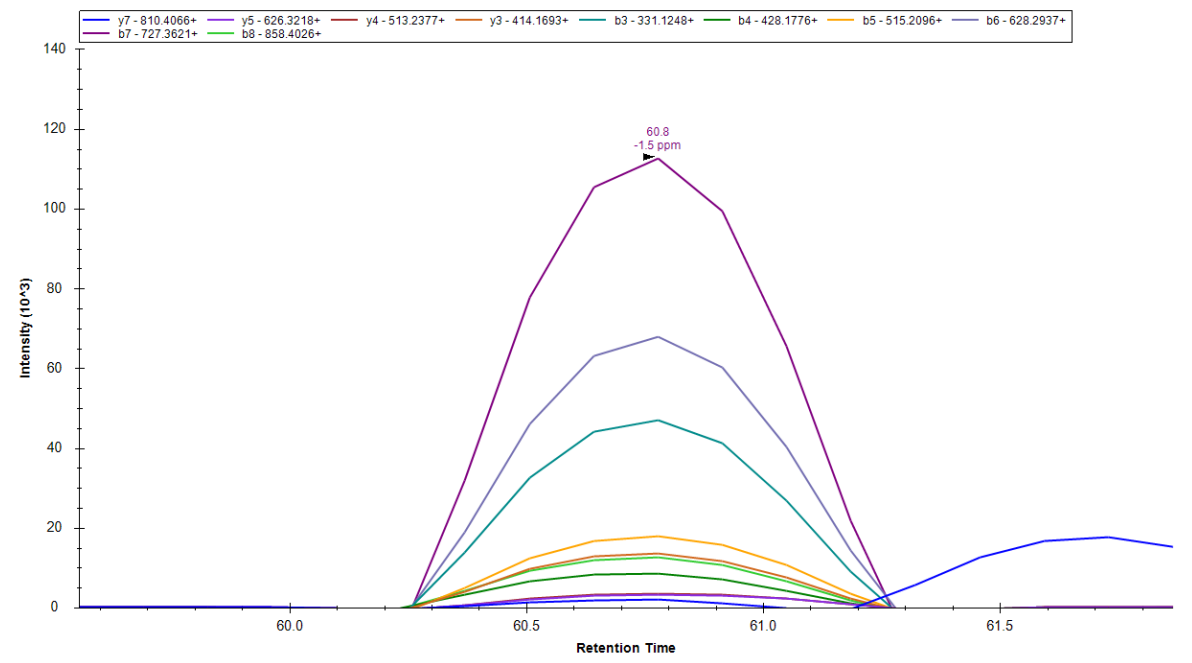
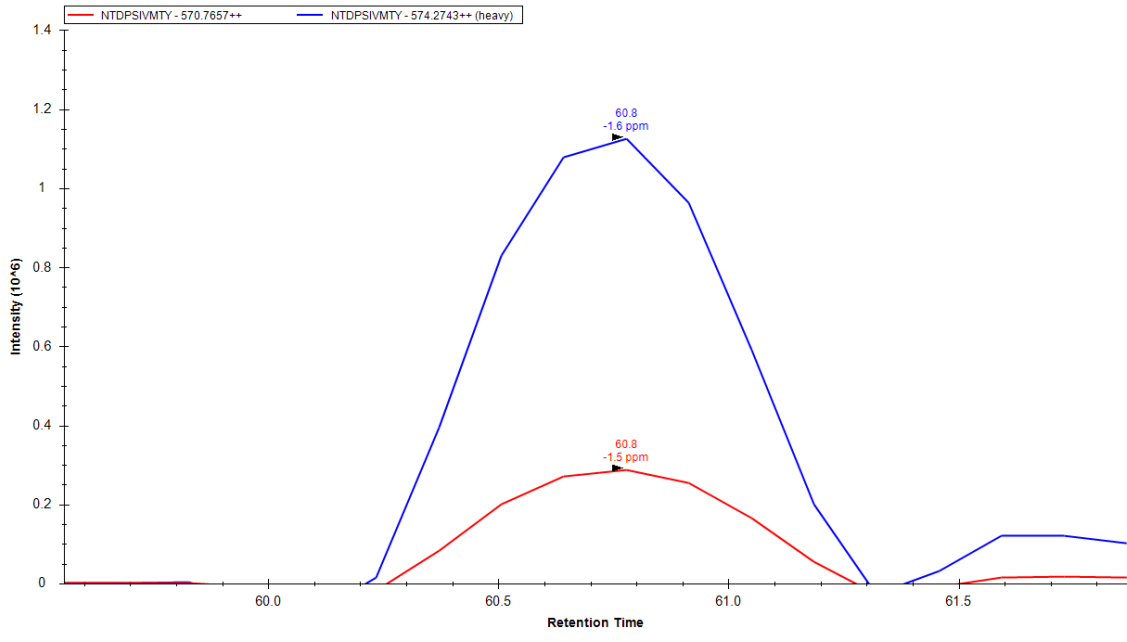
\* STRING database shown proteins associated with reactome pathways, immune system (&) and/or MHC-II antigen presentation (#); NA, allele not included in MixMHC2pred.

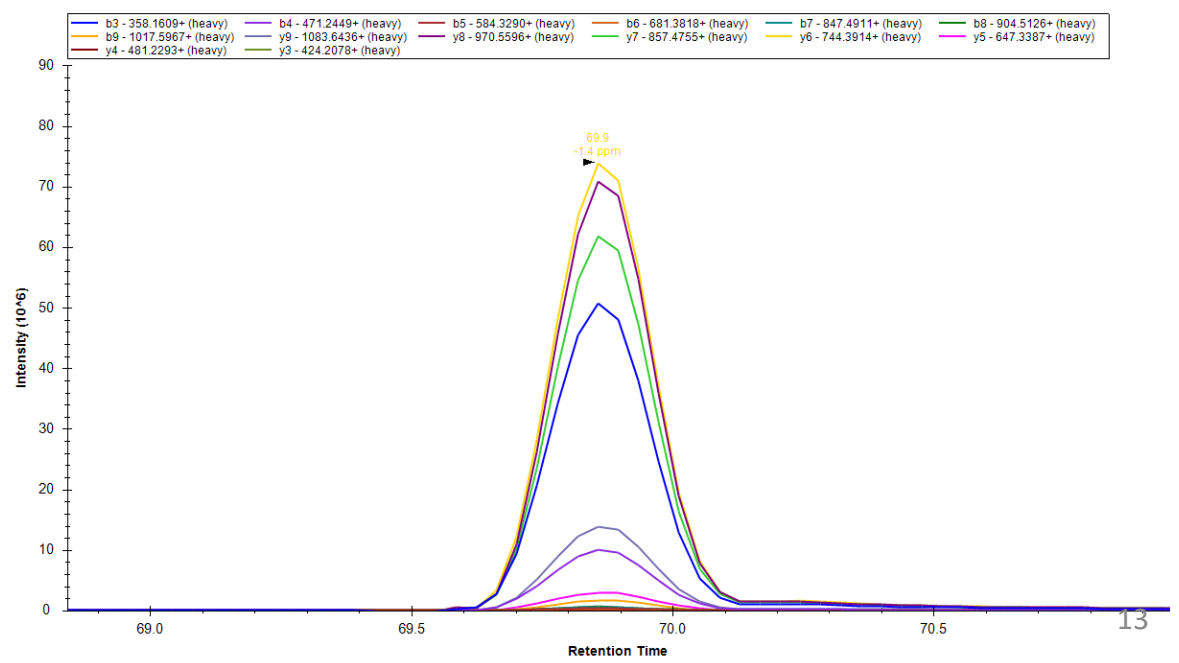
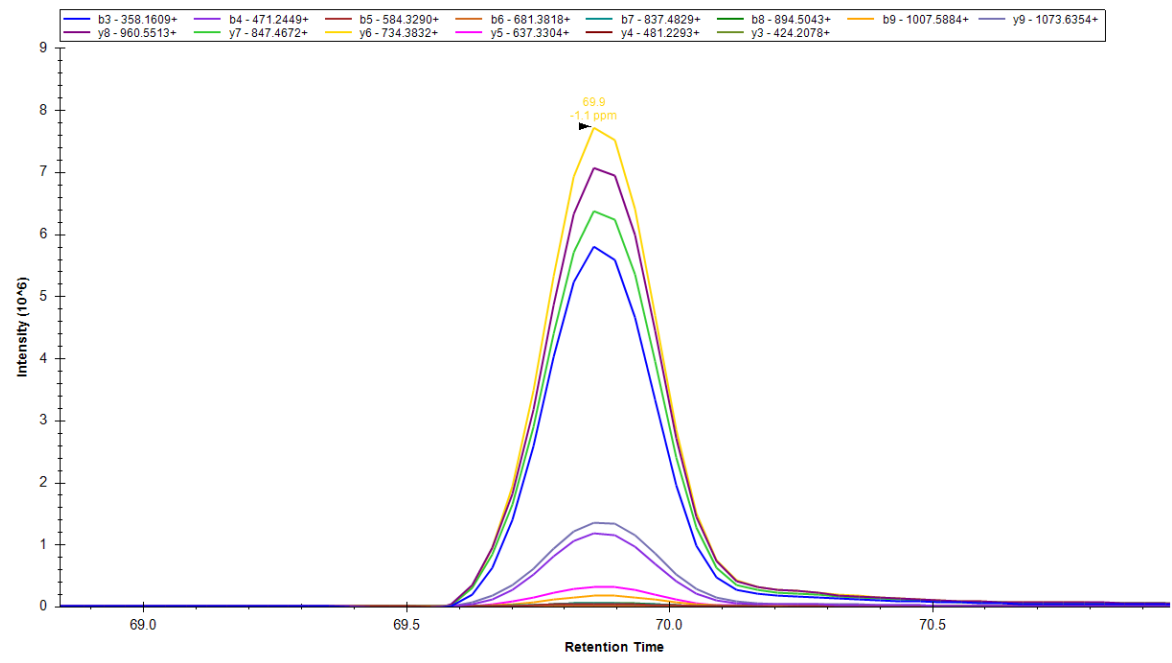
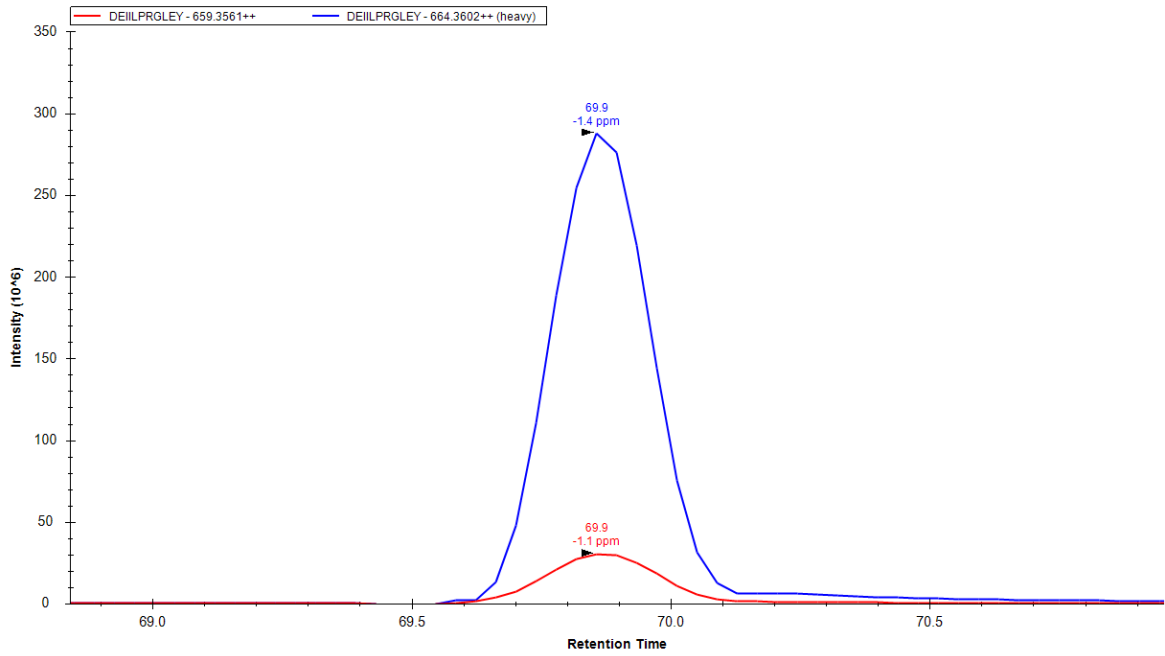
HLA-II alleles with relatively higher binding affinity for immunopeptides were listed.

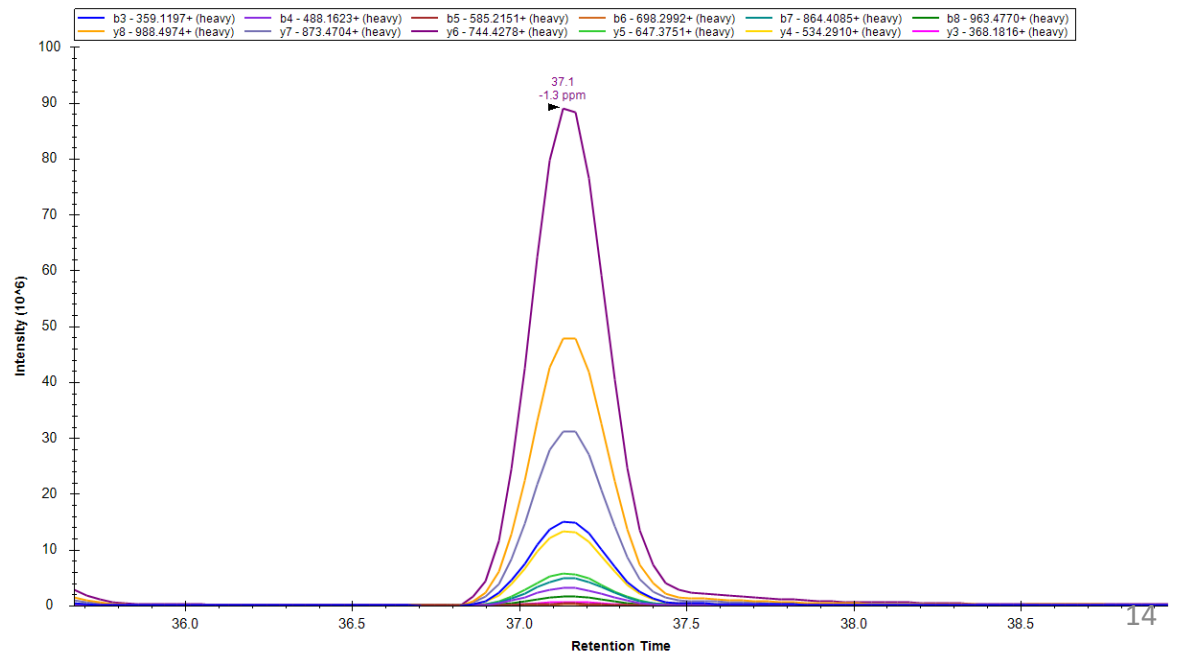
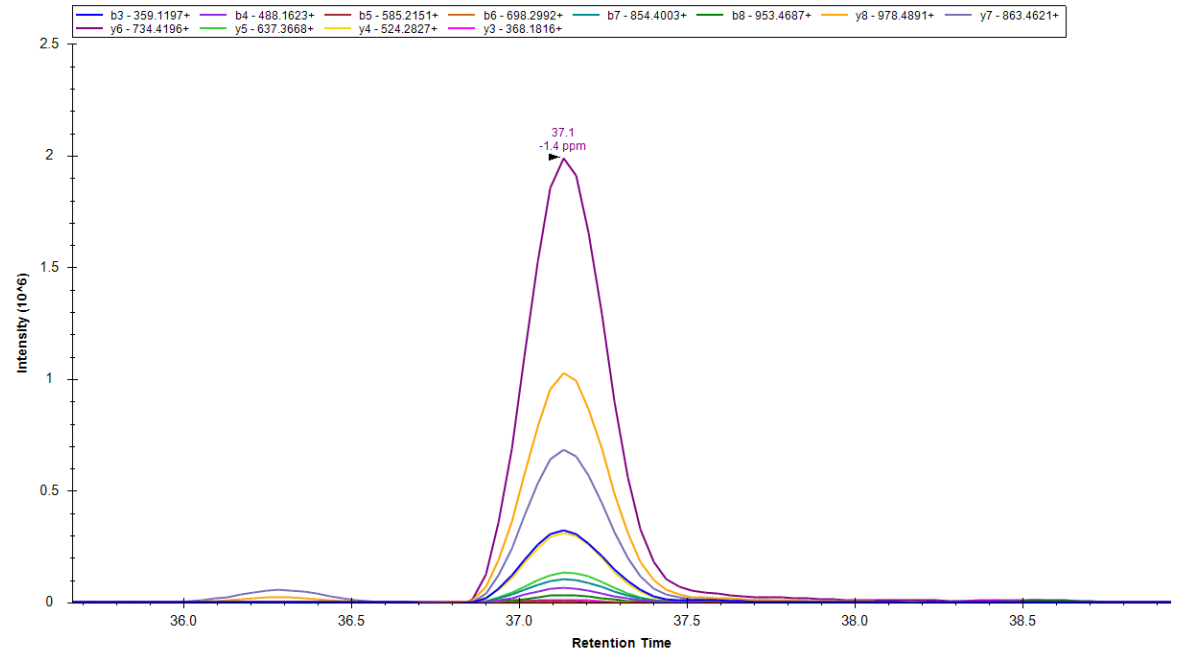
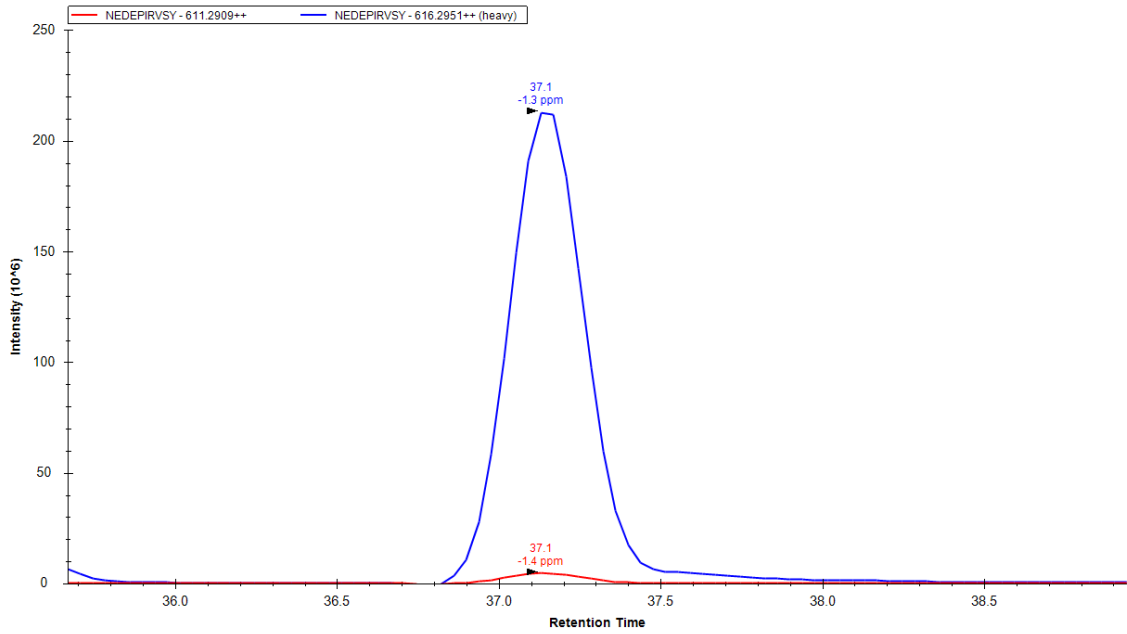
**Fig. S7**

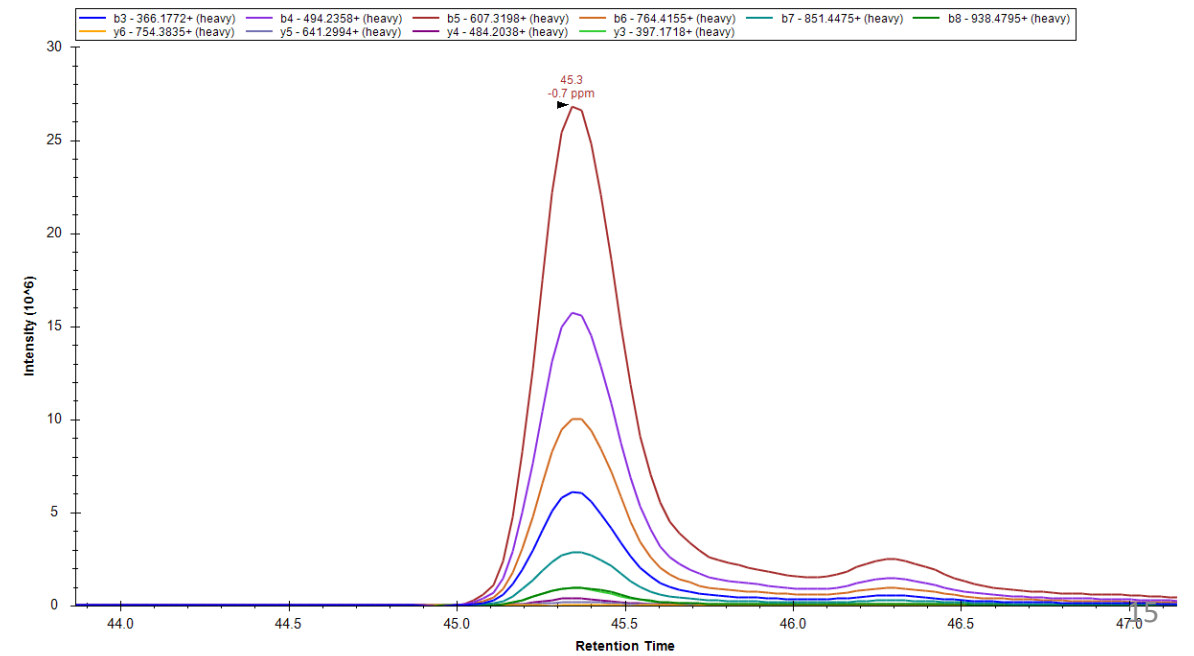
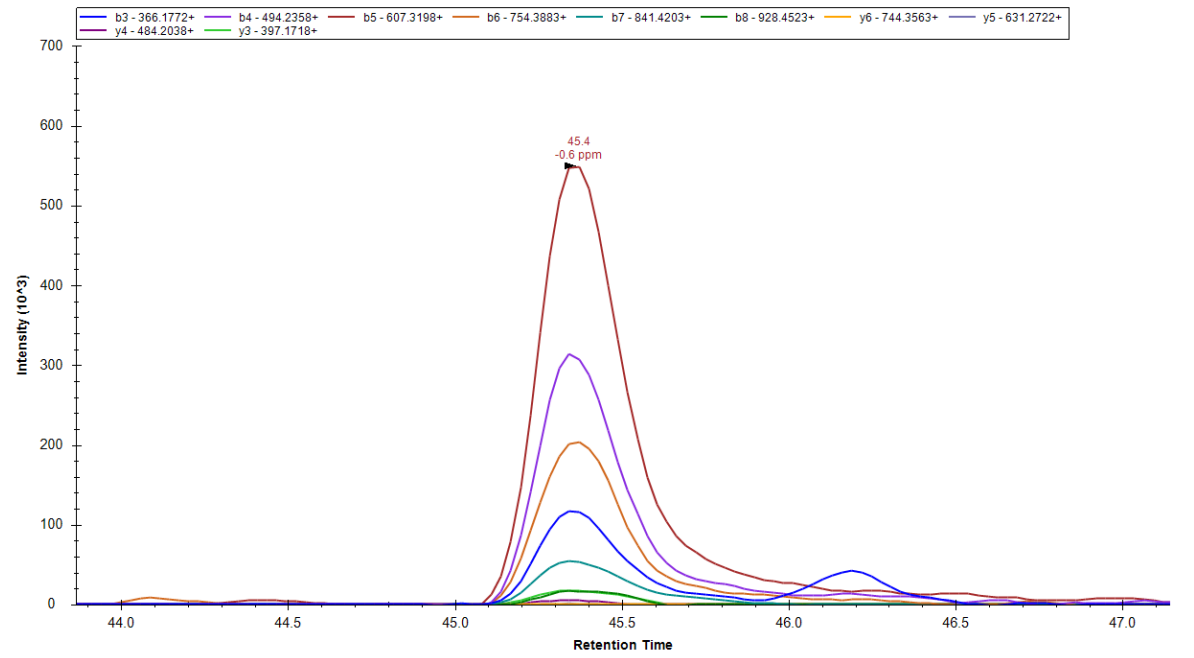
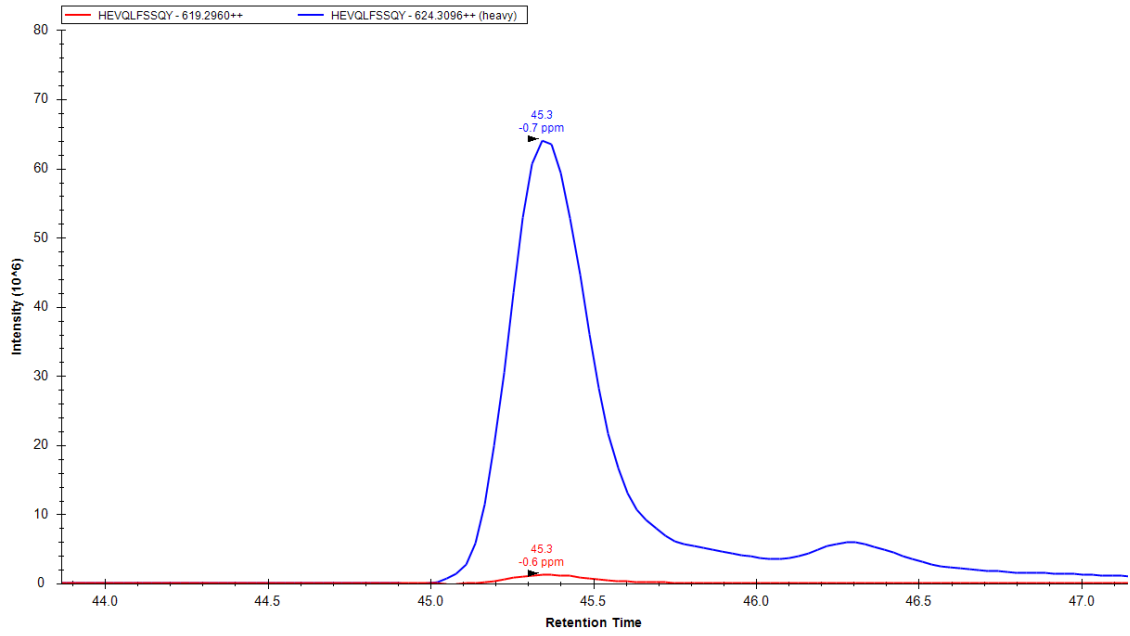
**Pages: 10 - 37**

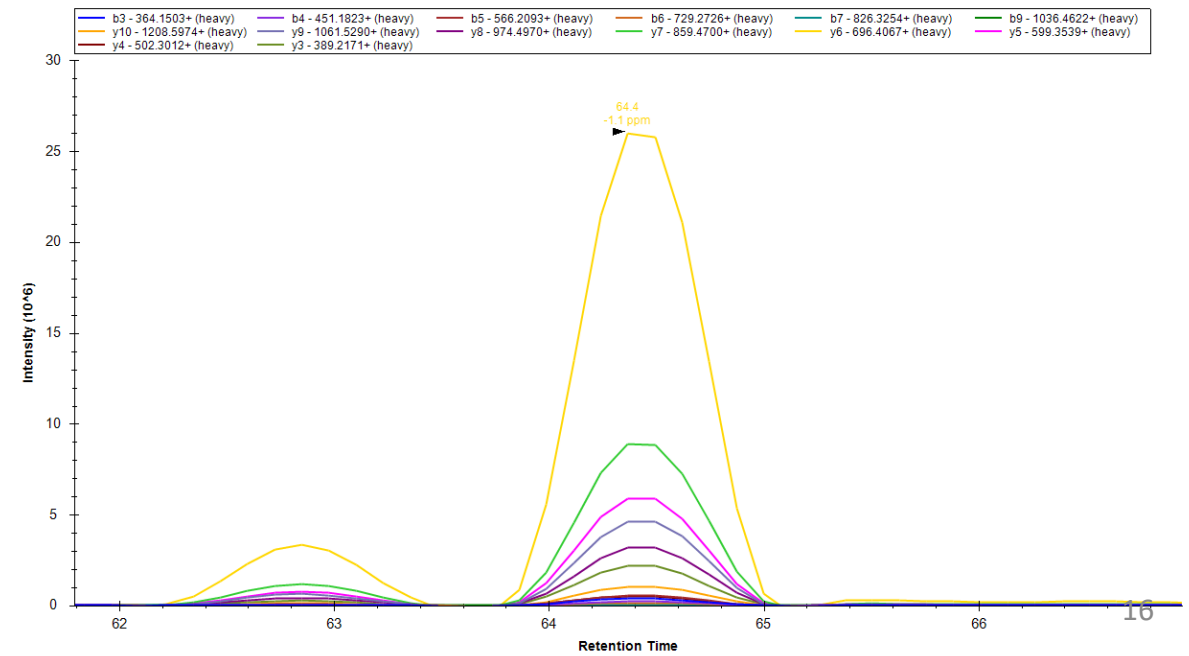
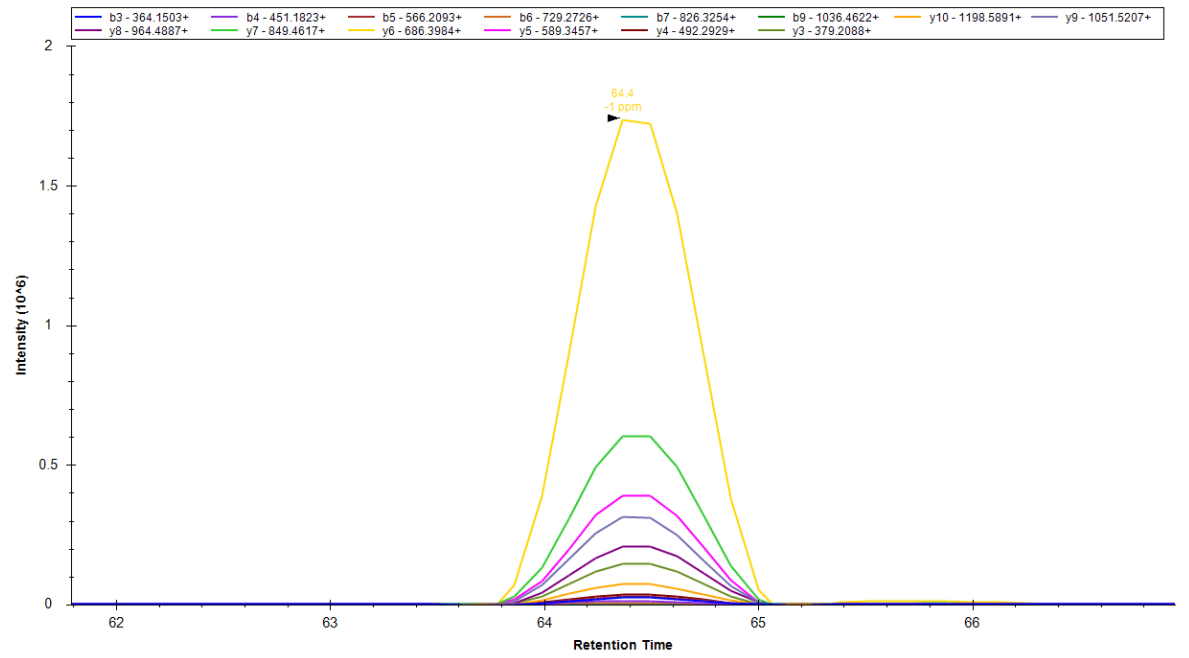
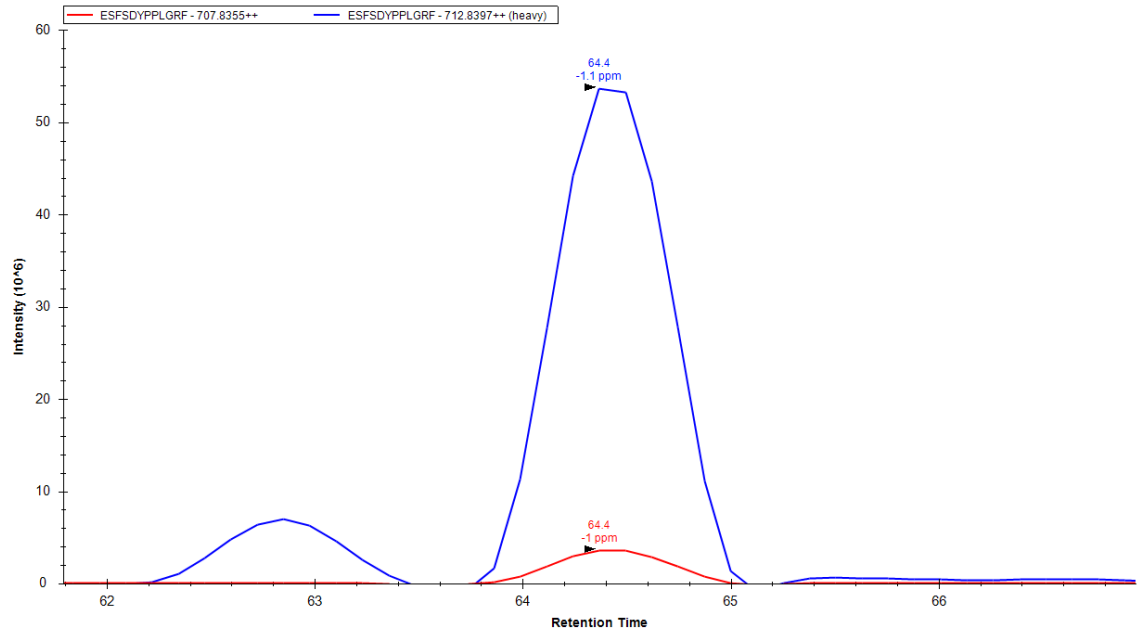




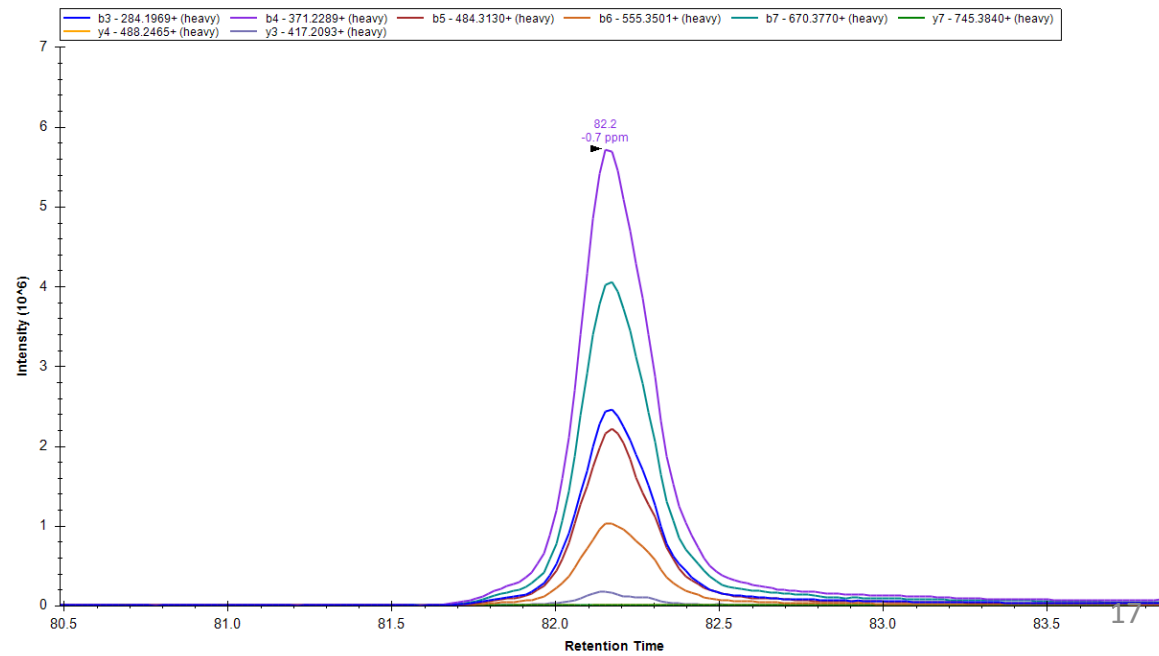
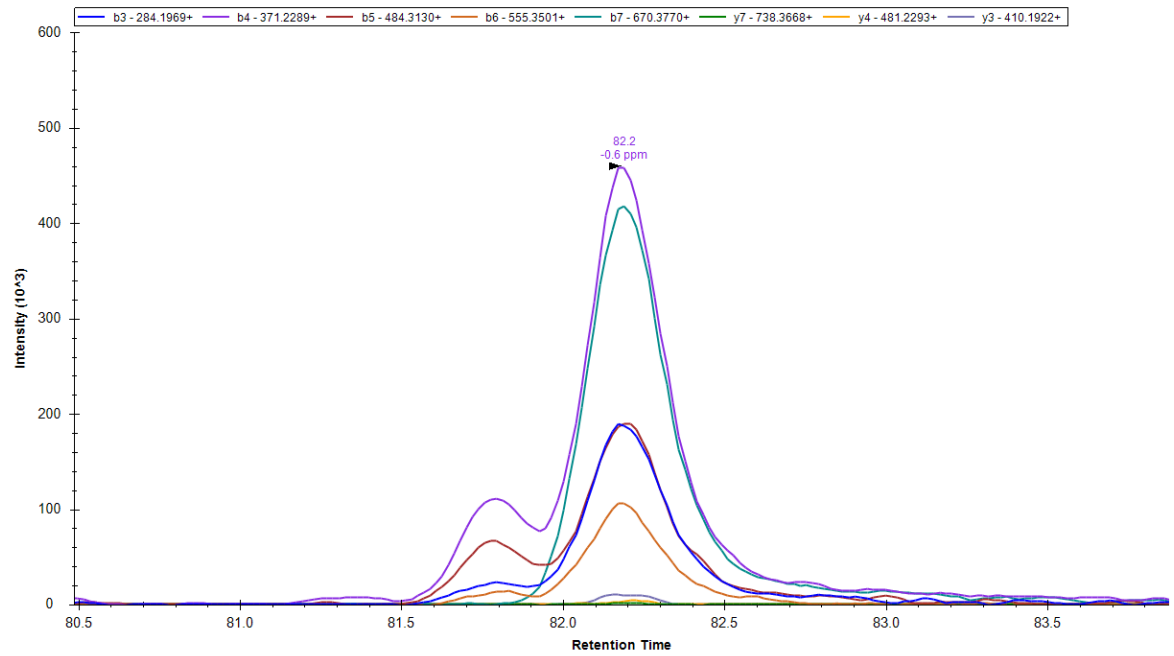
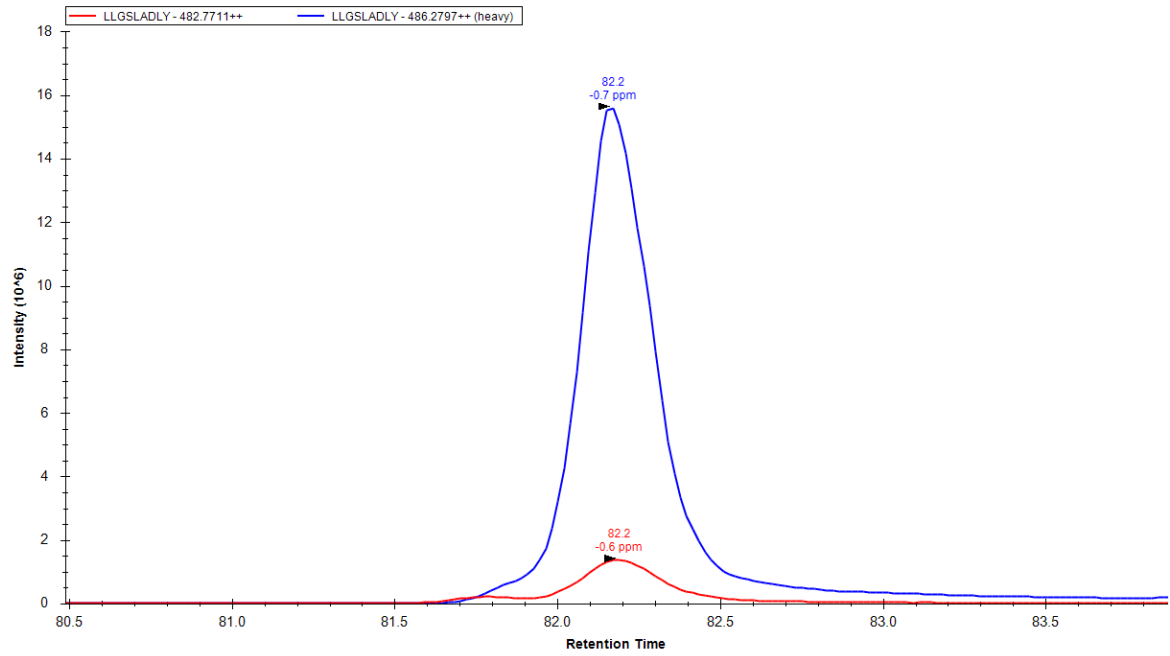


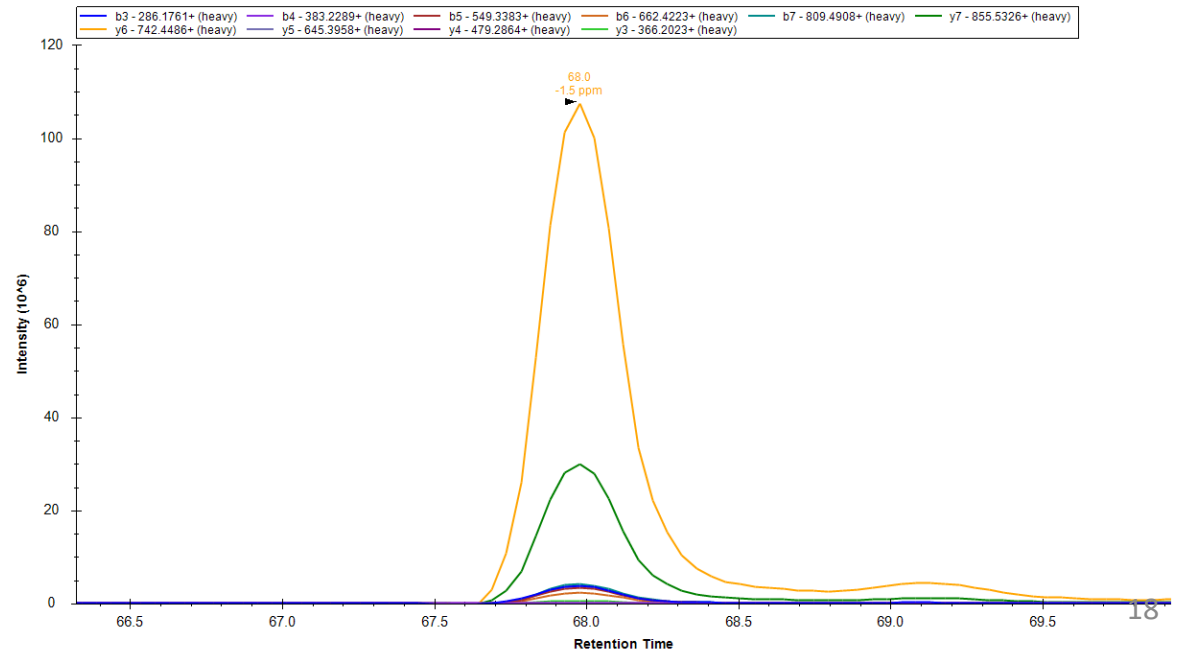
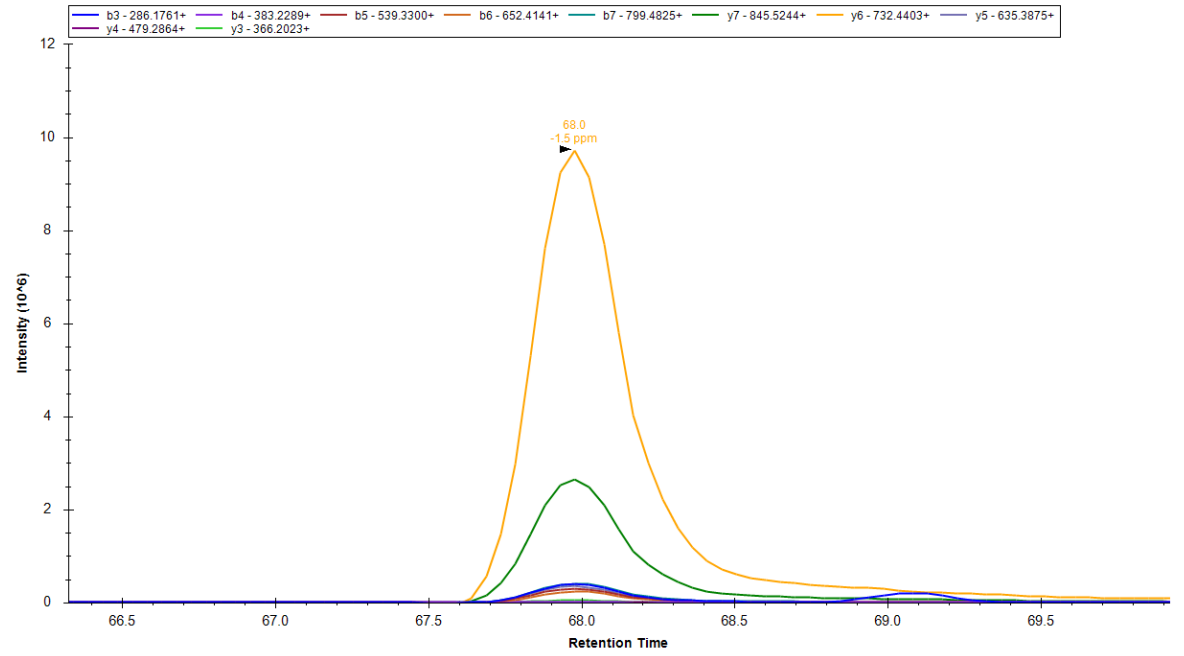
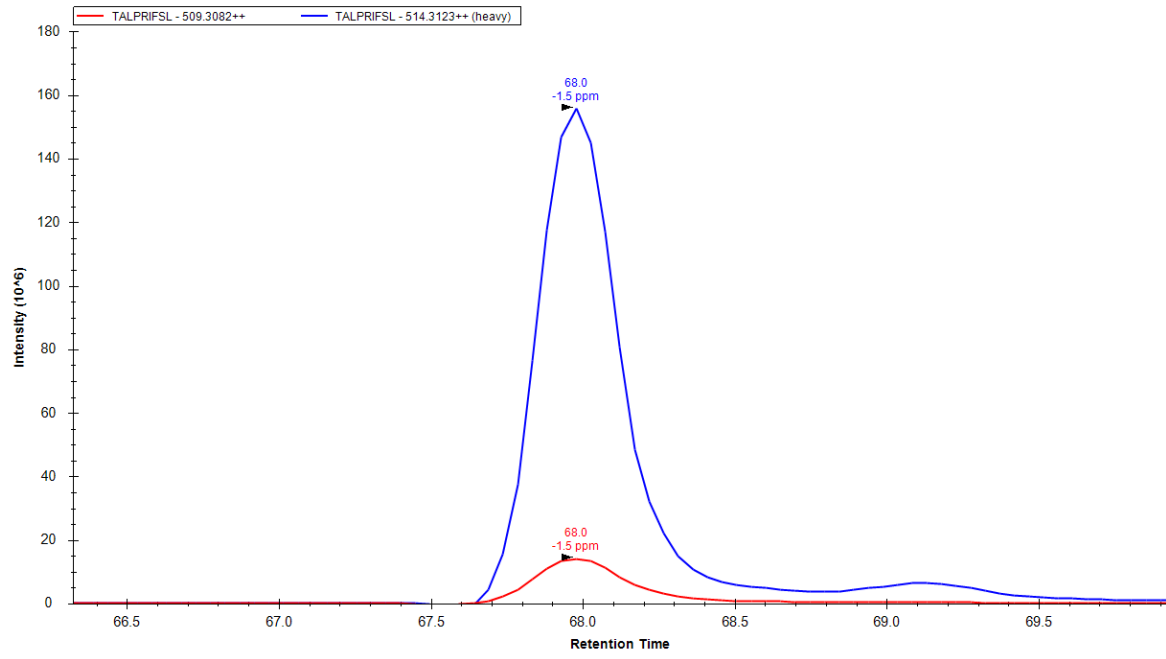


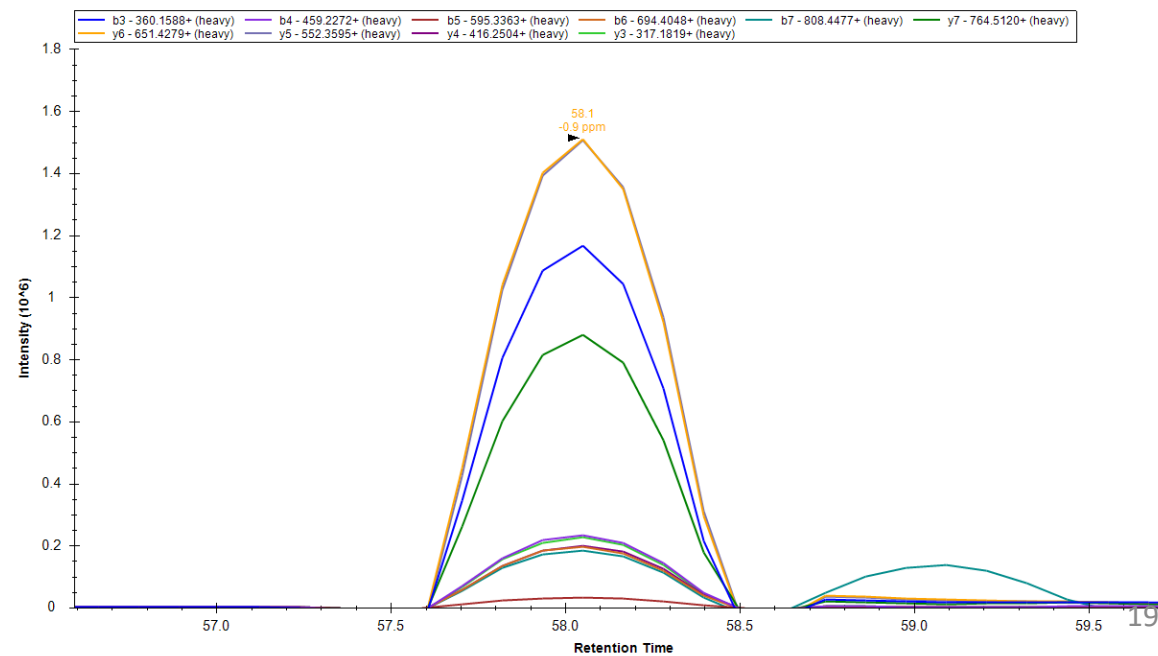
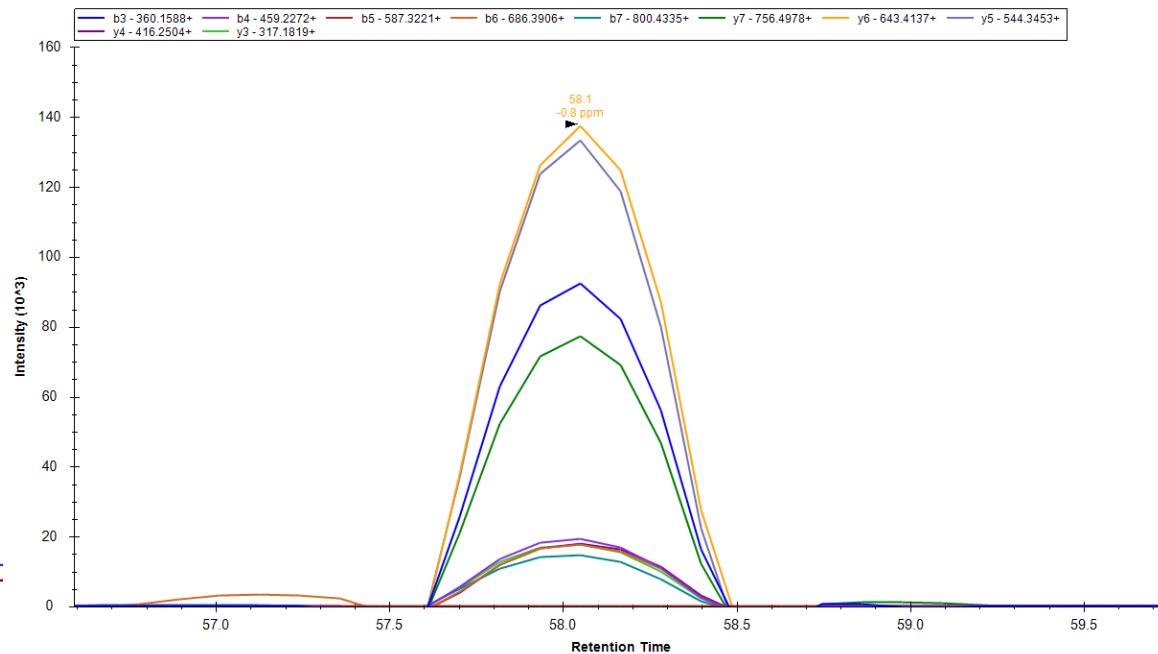
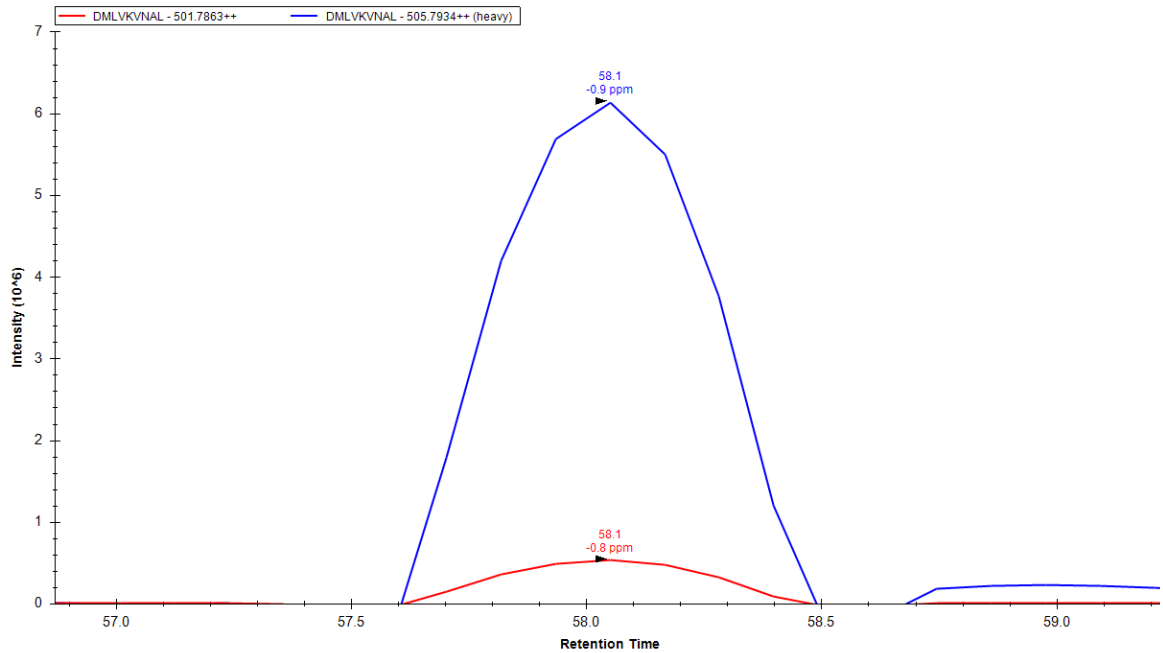


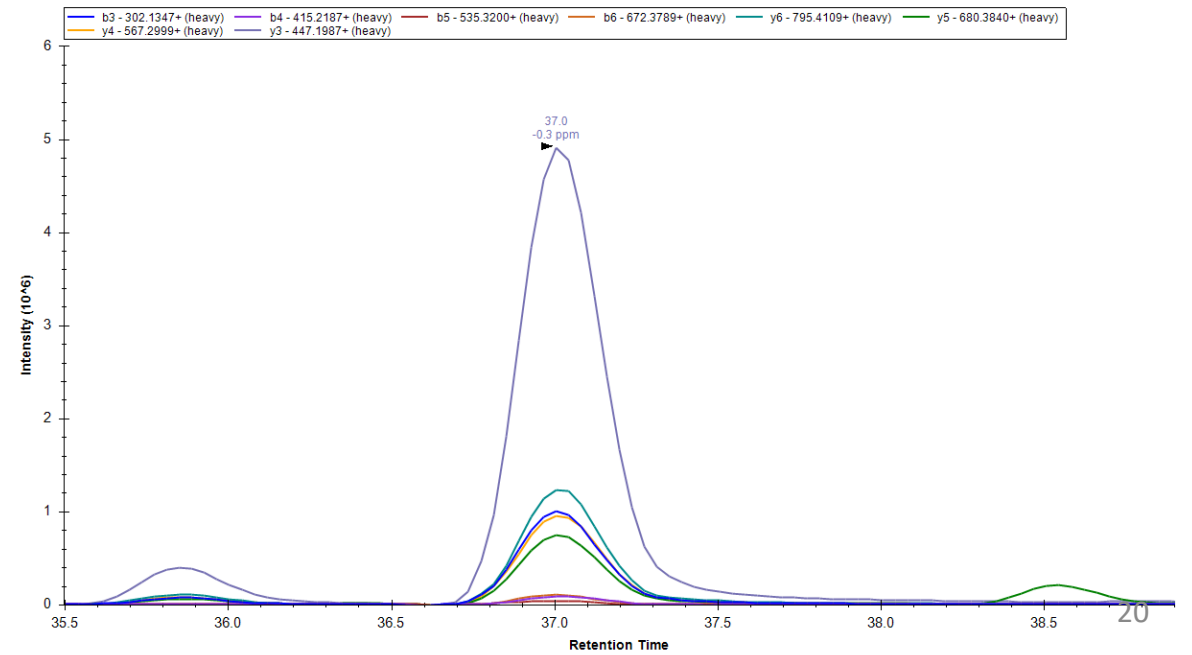
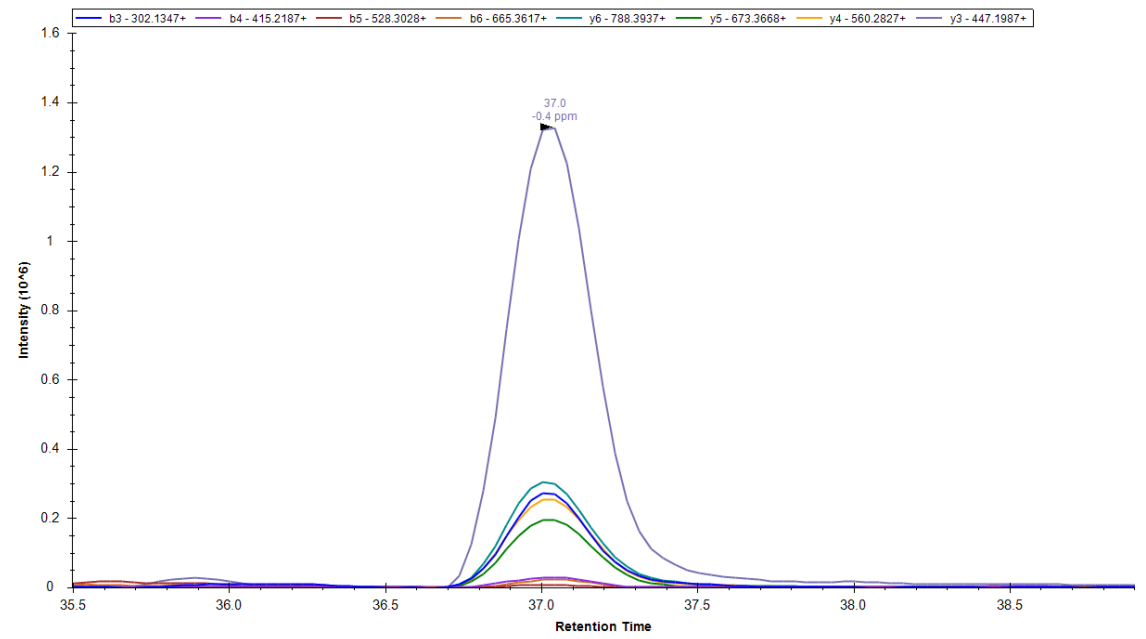
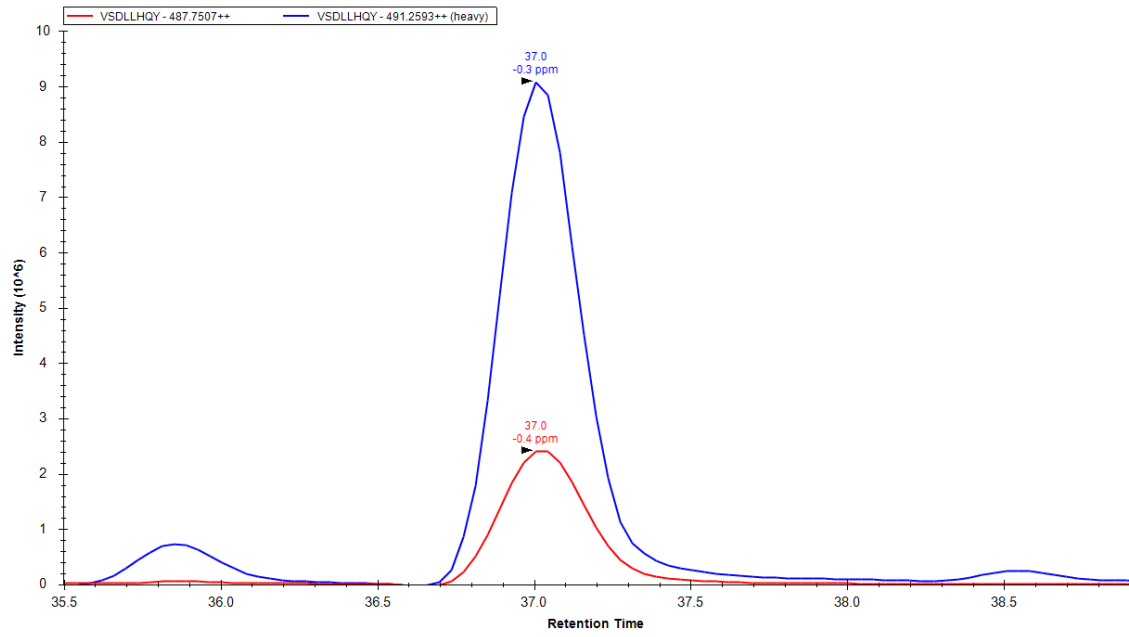


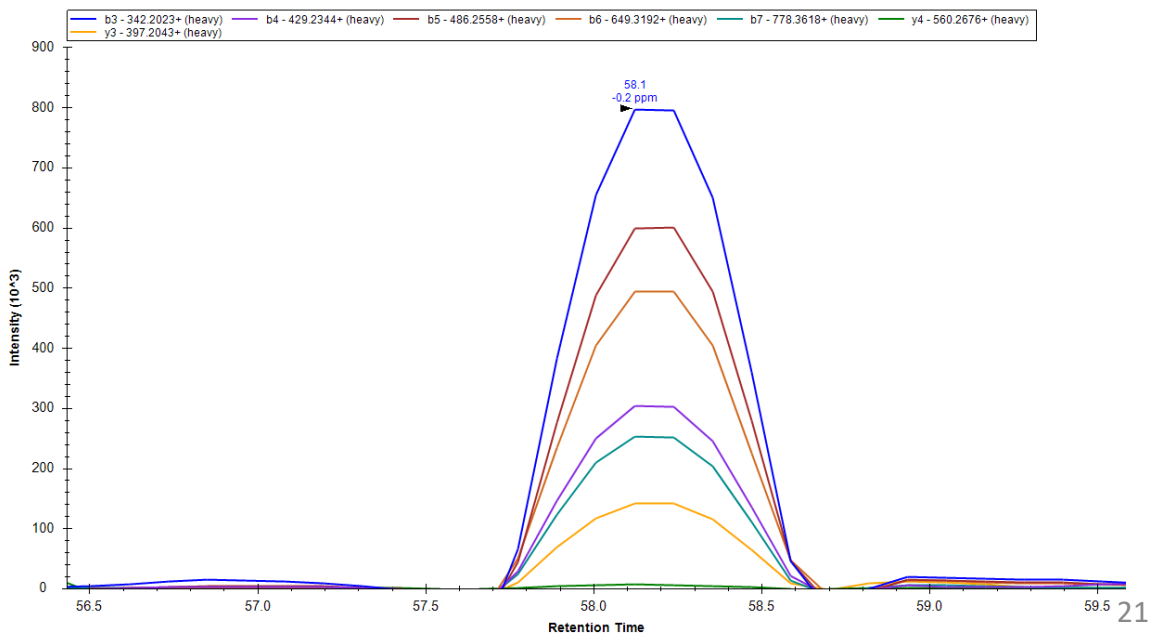
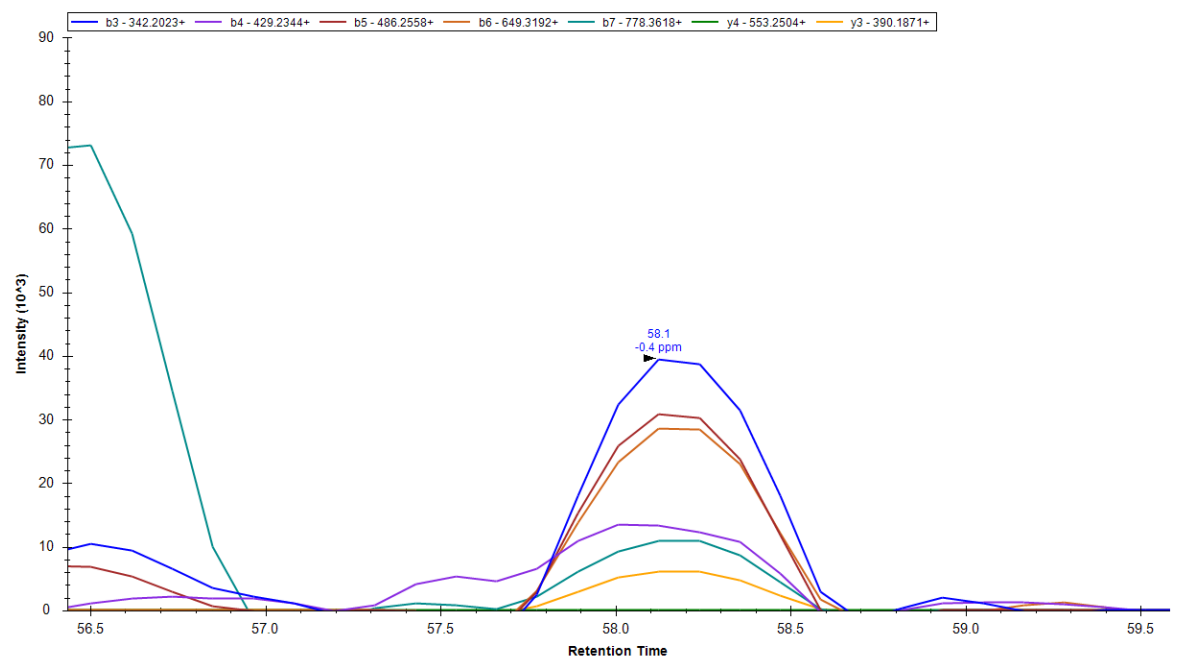
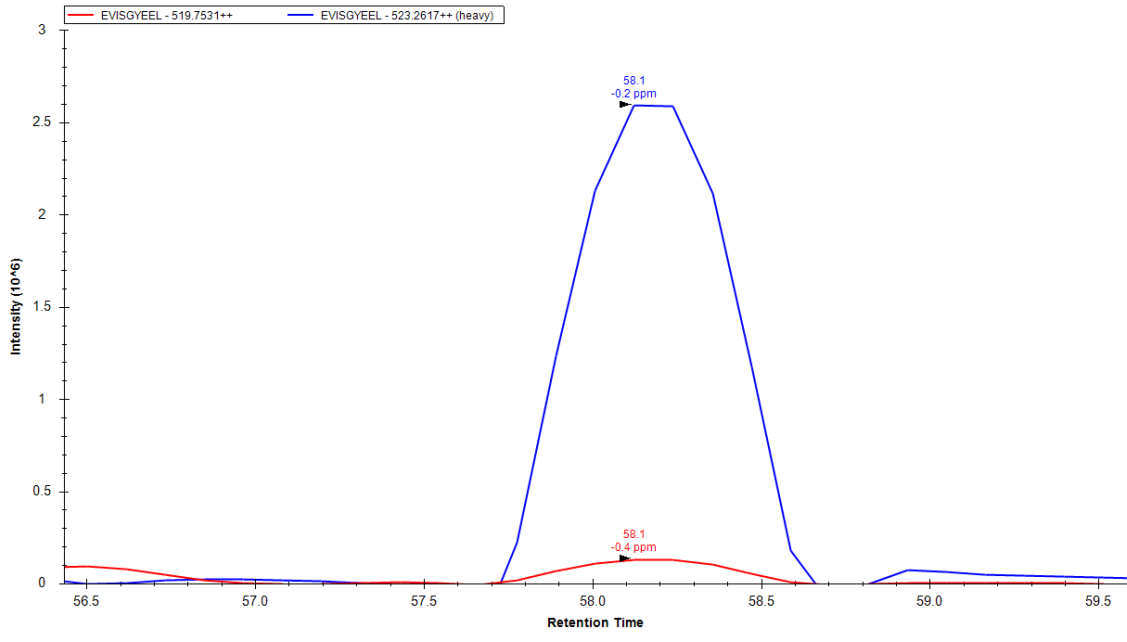


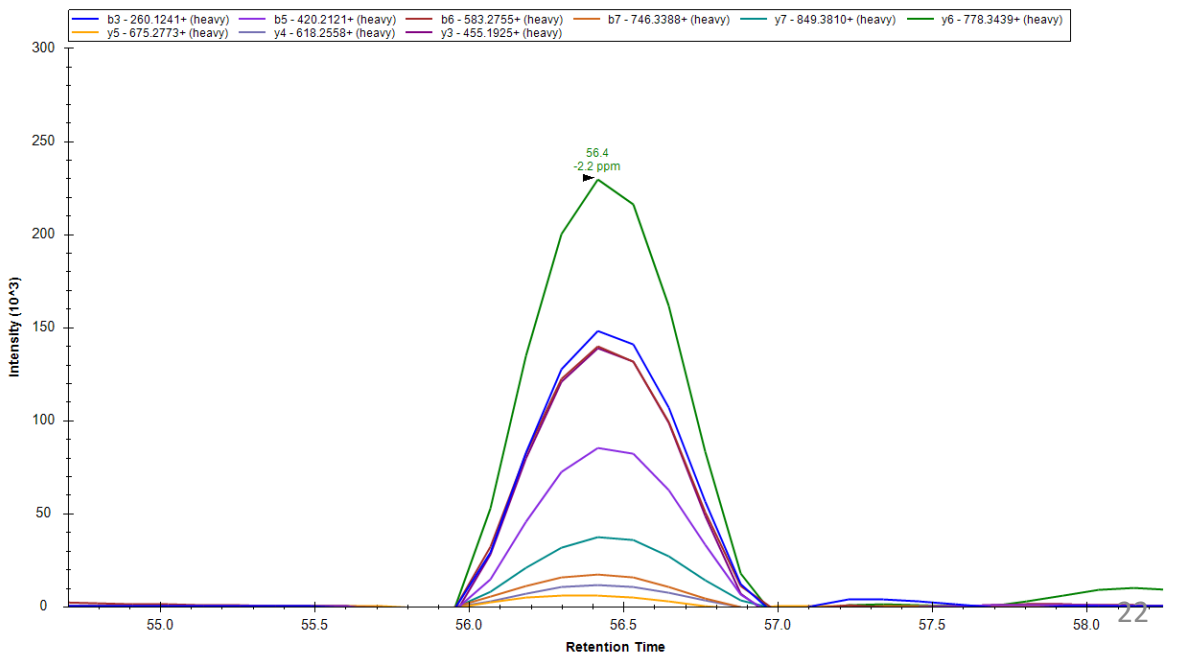
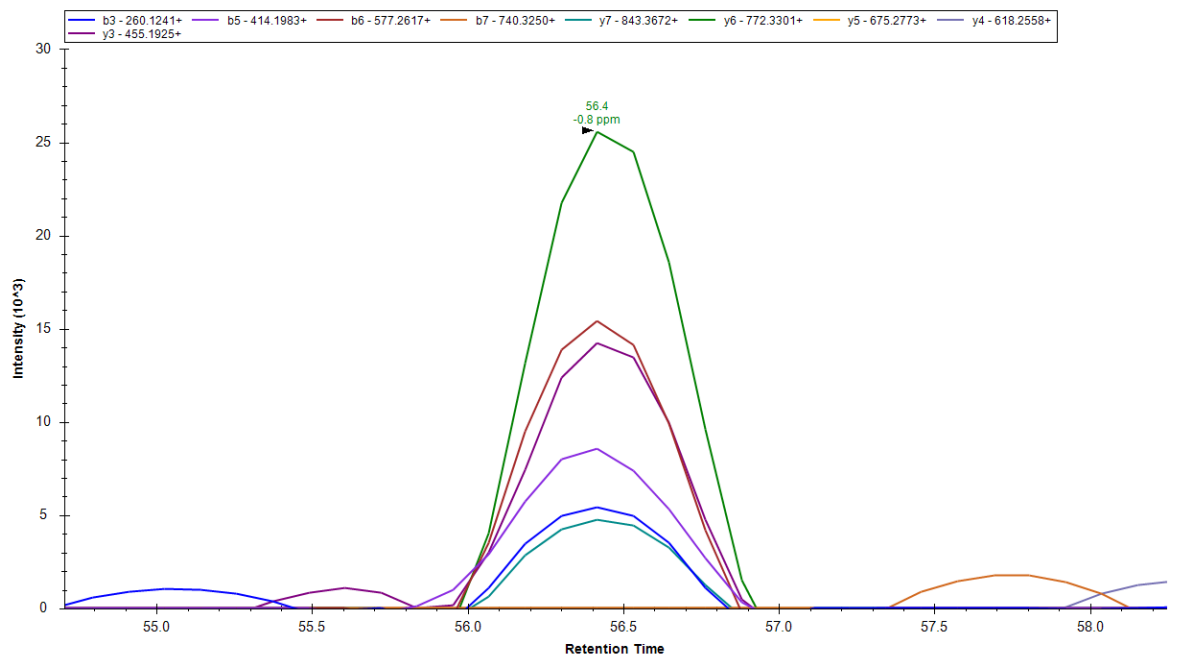
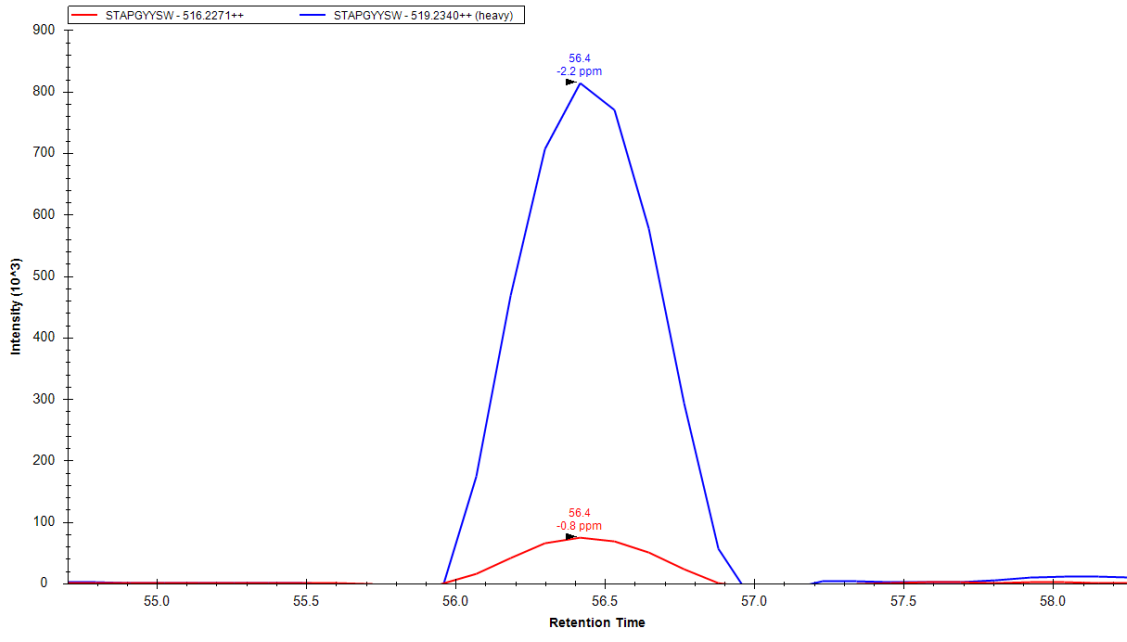


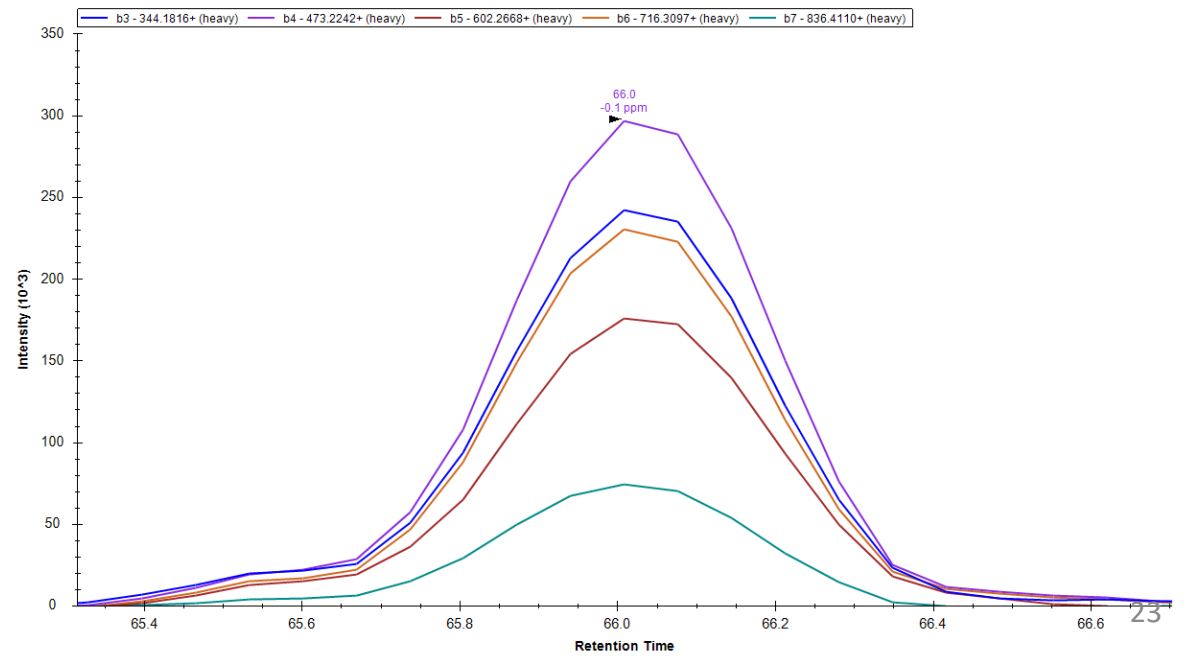
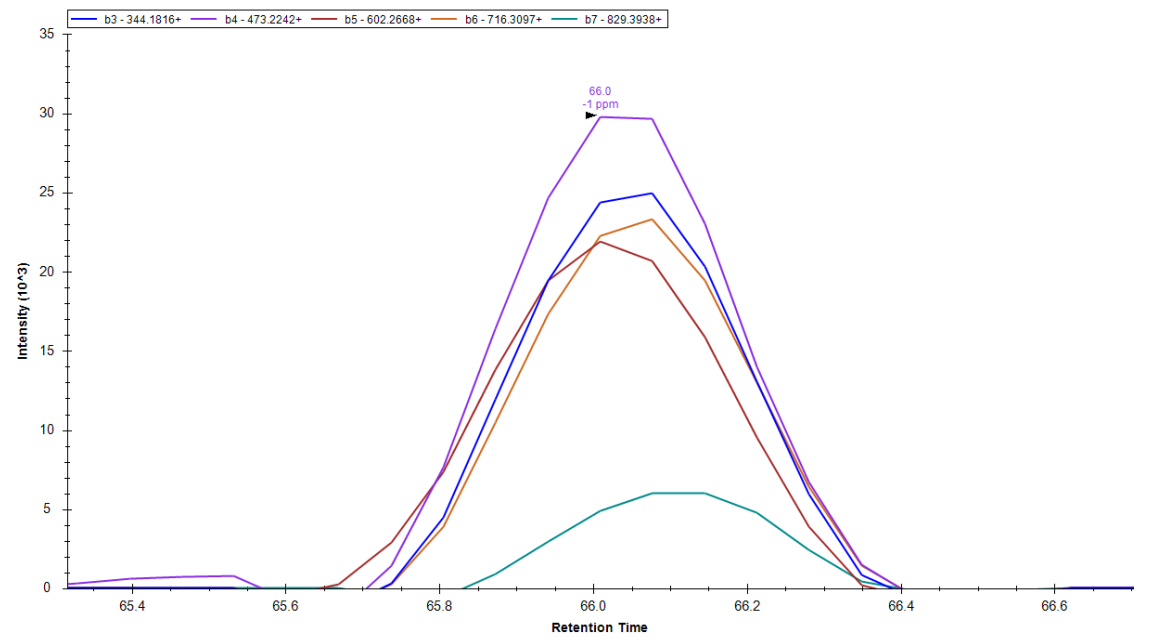
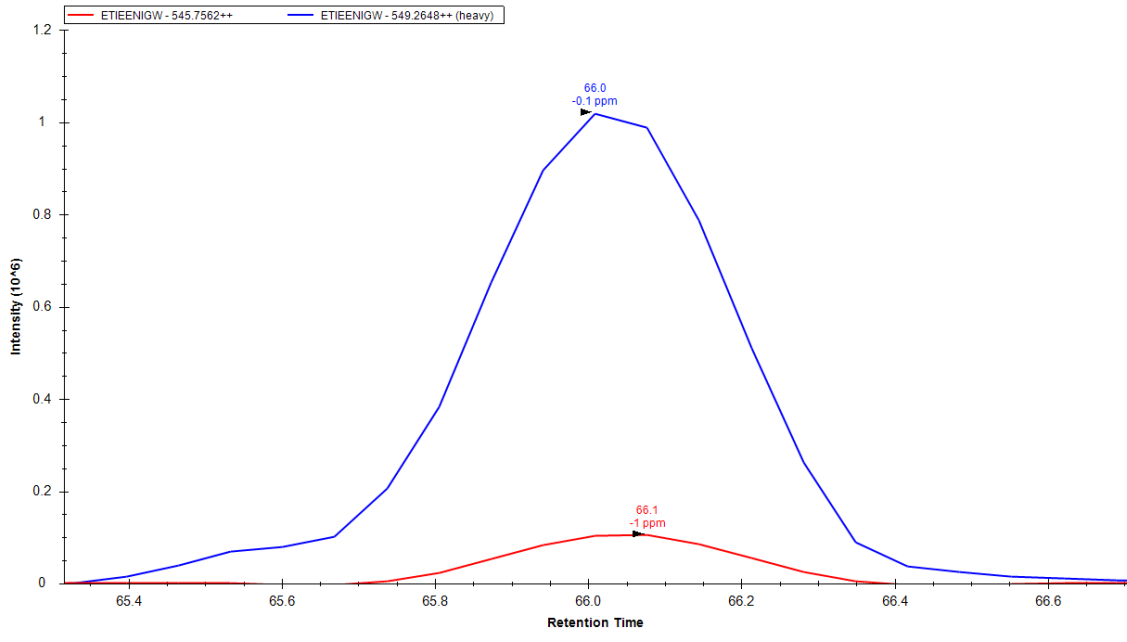


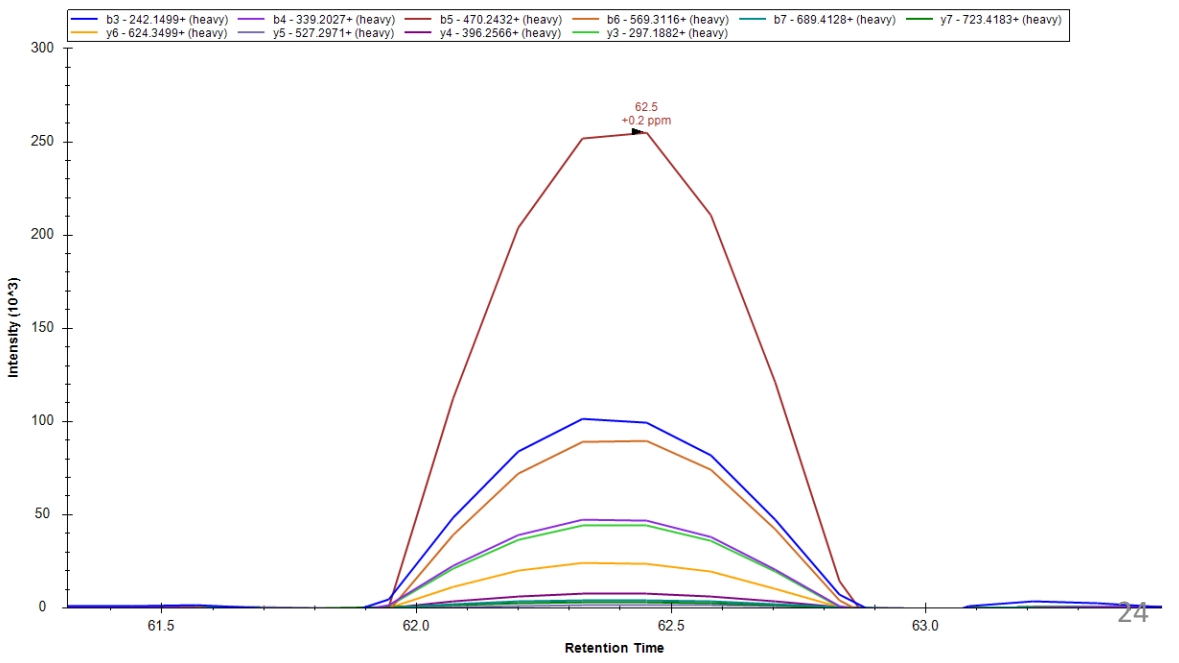
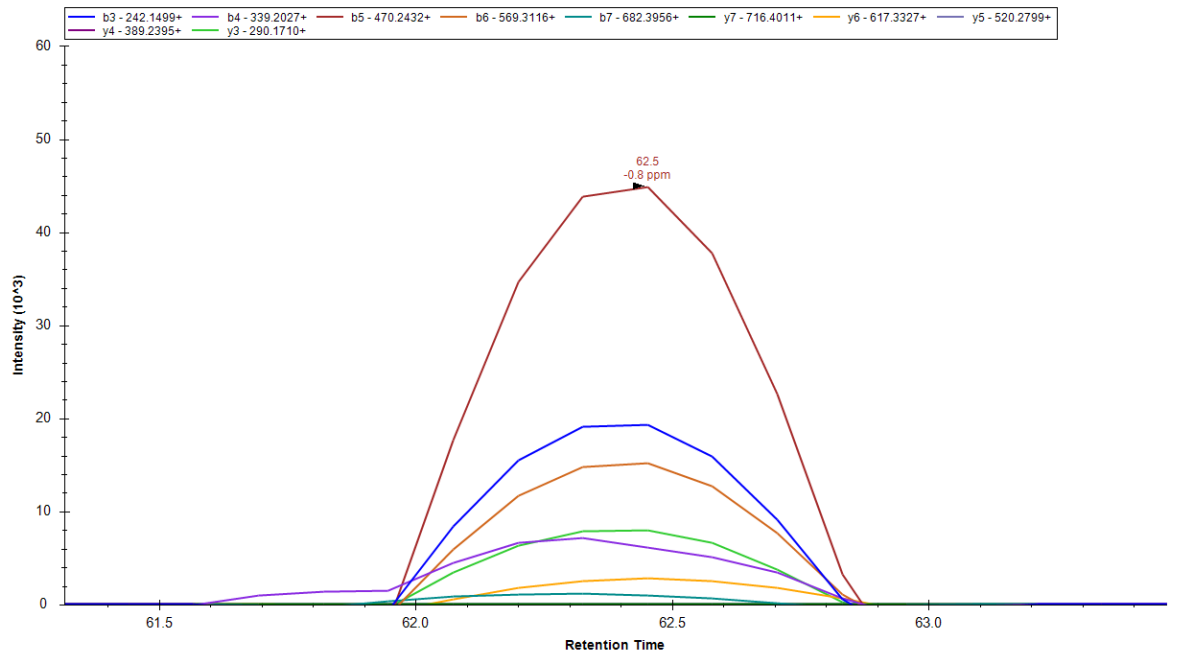
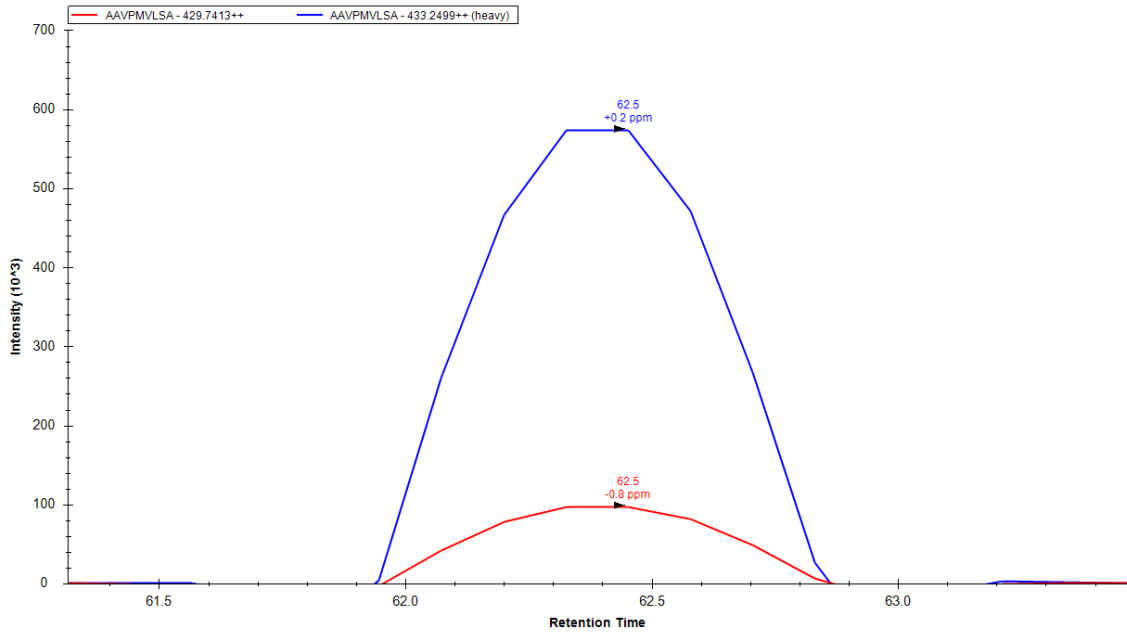




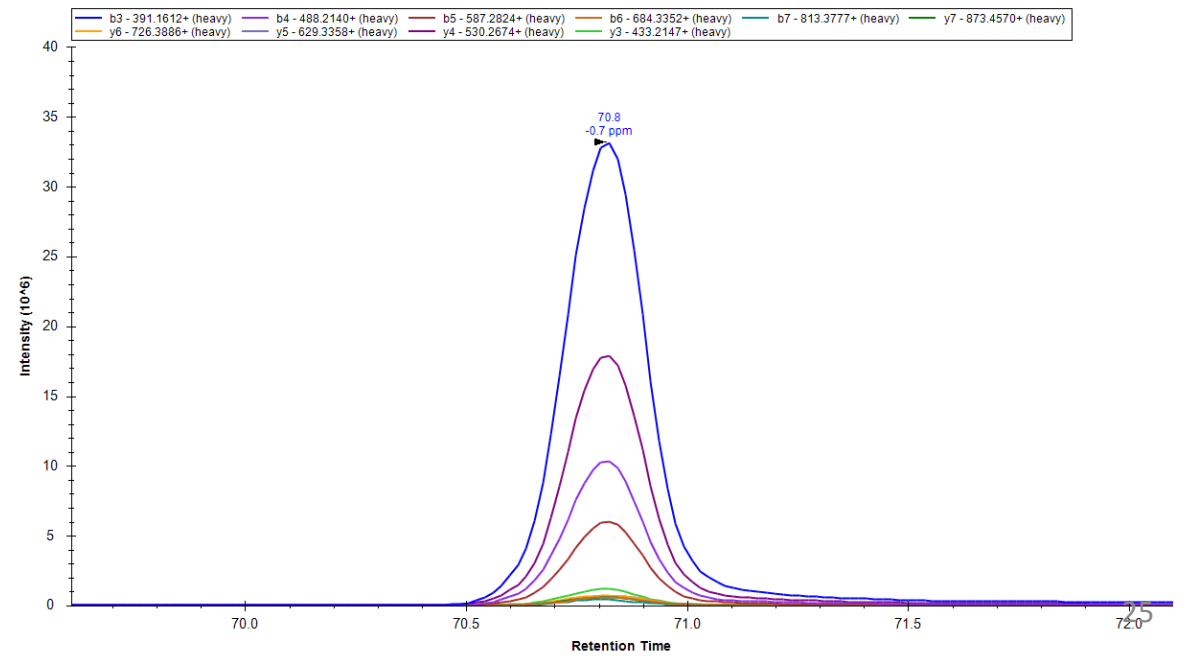
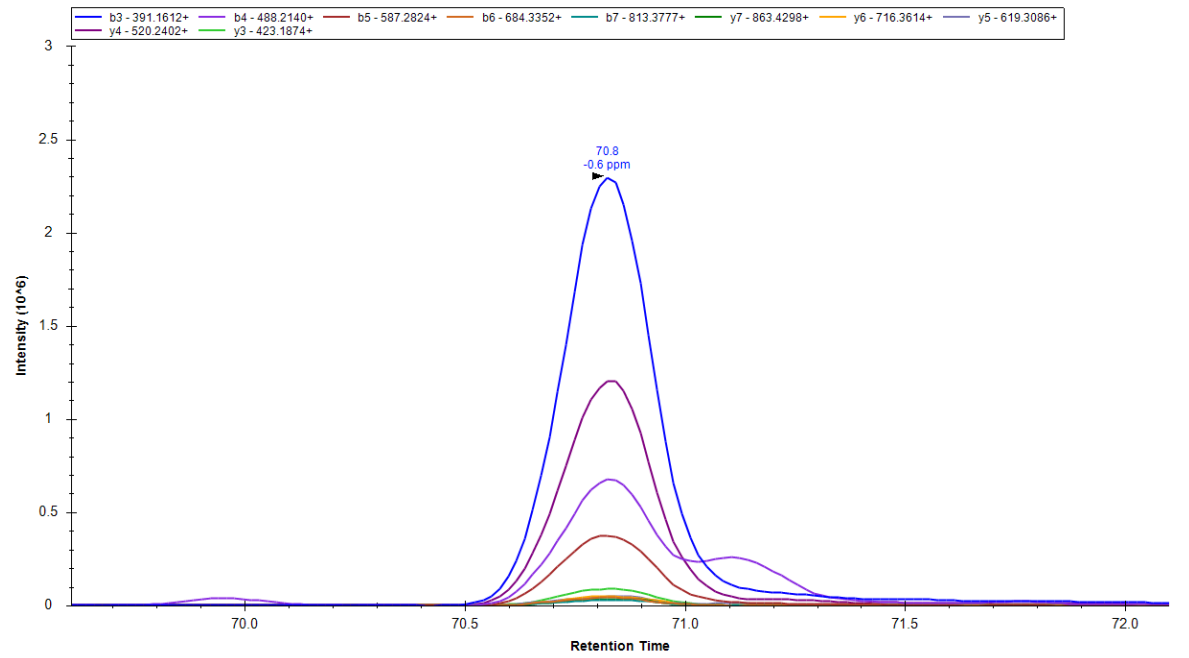
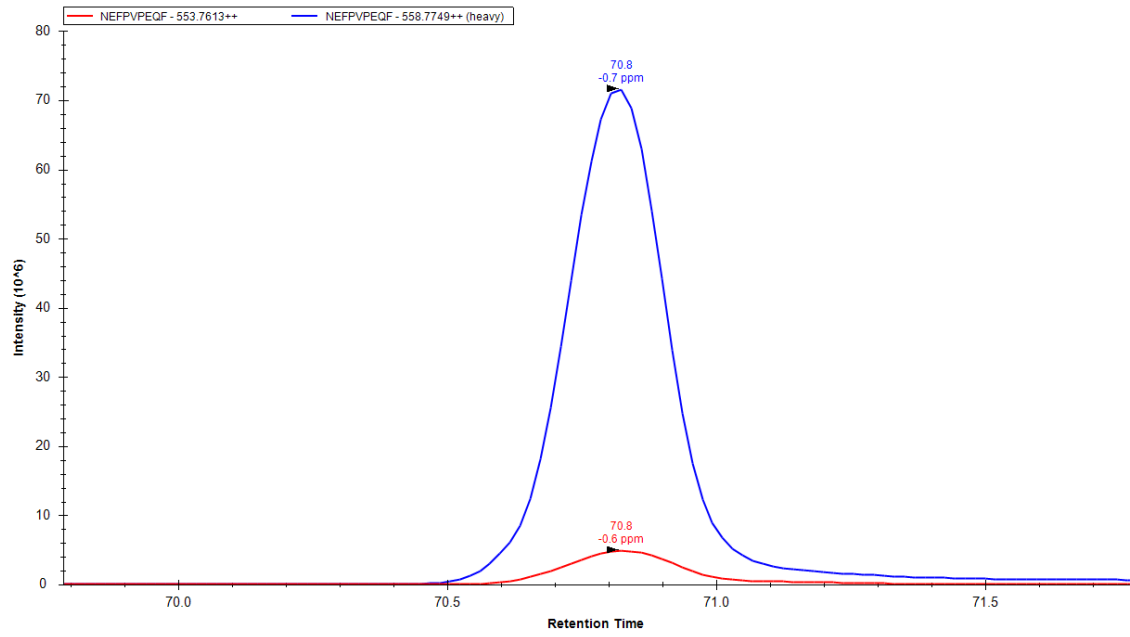


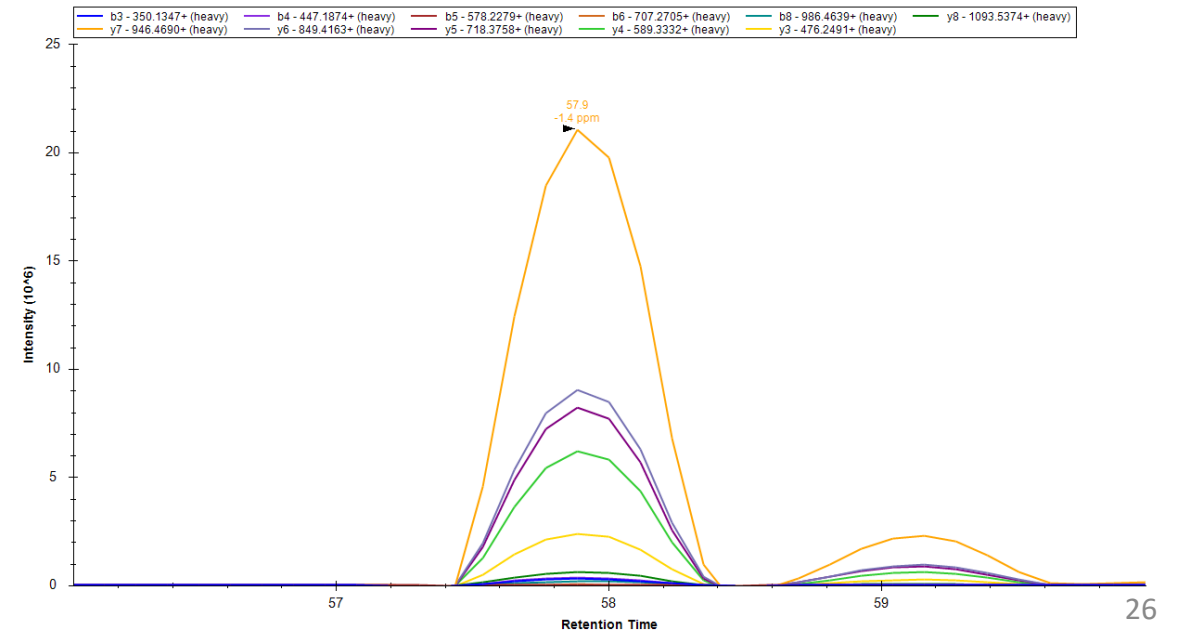
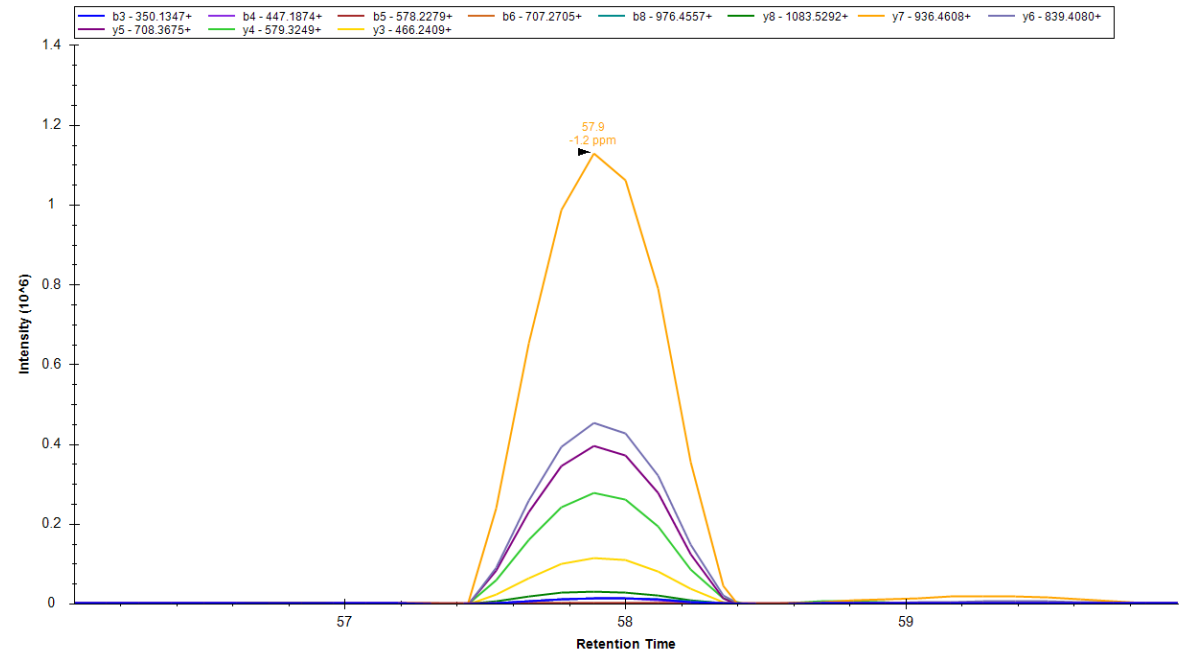
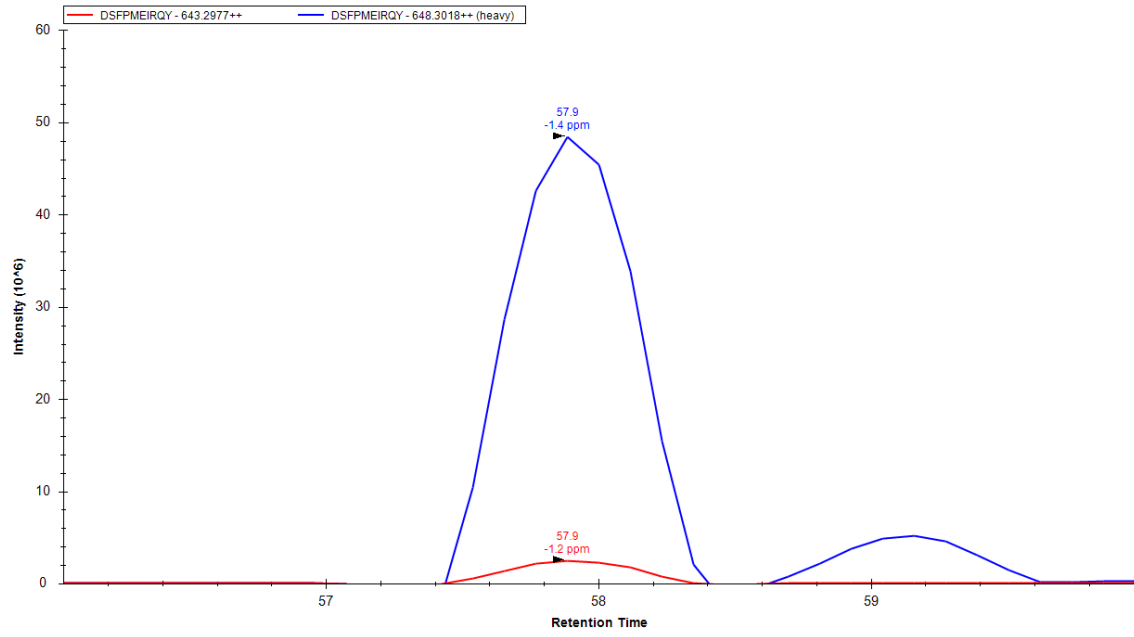


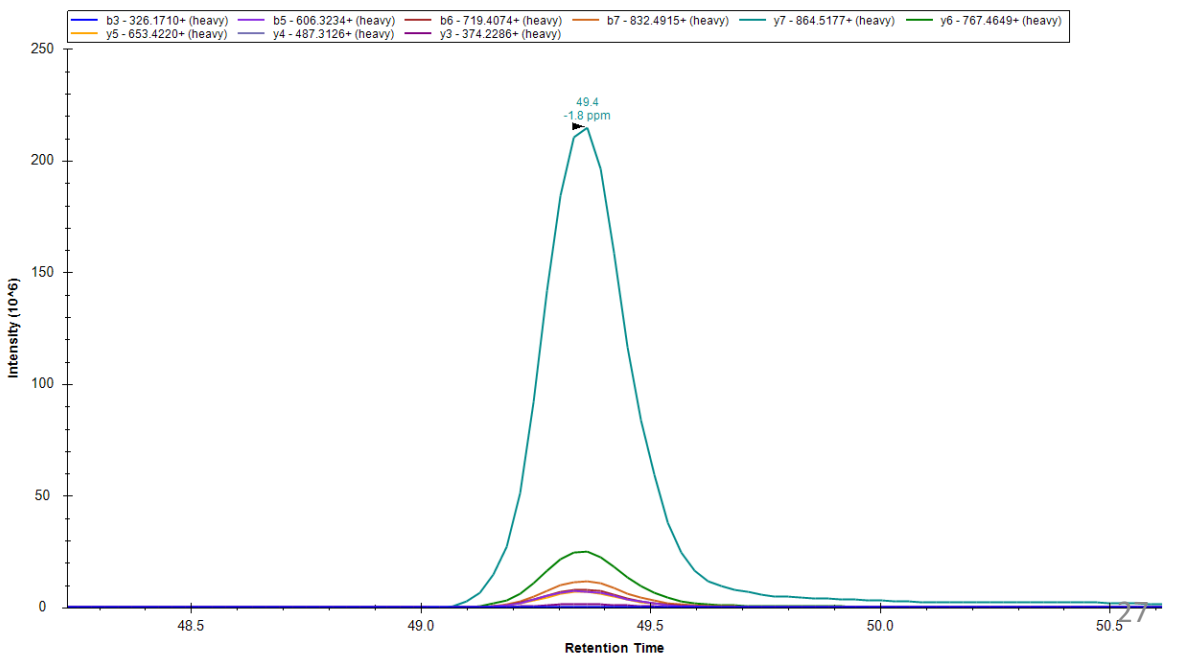
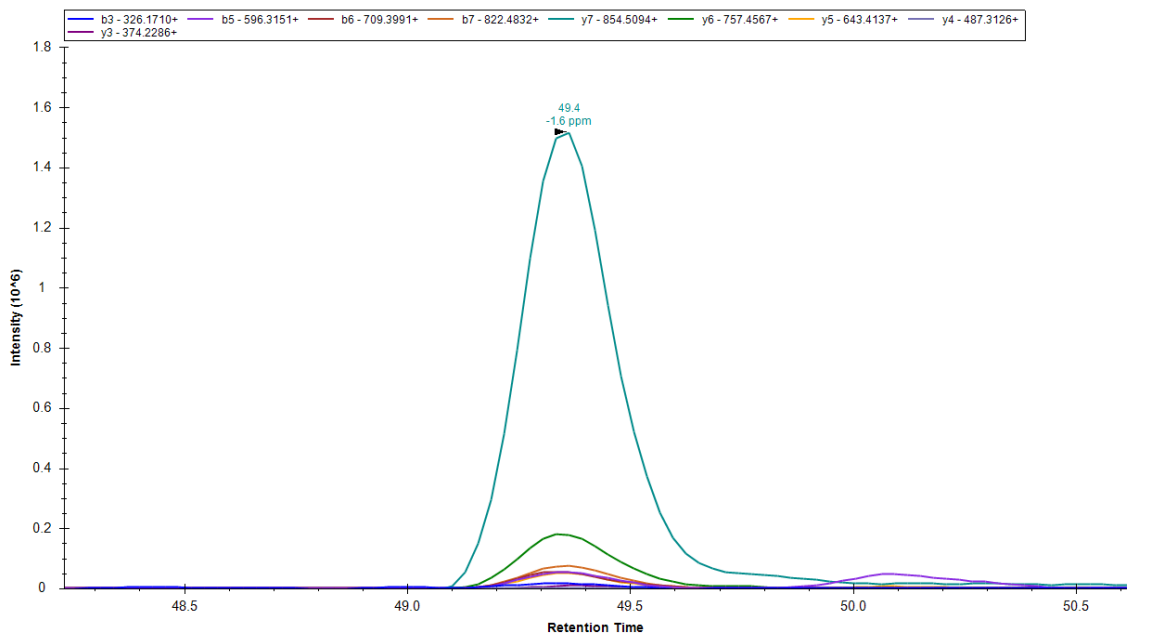
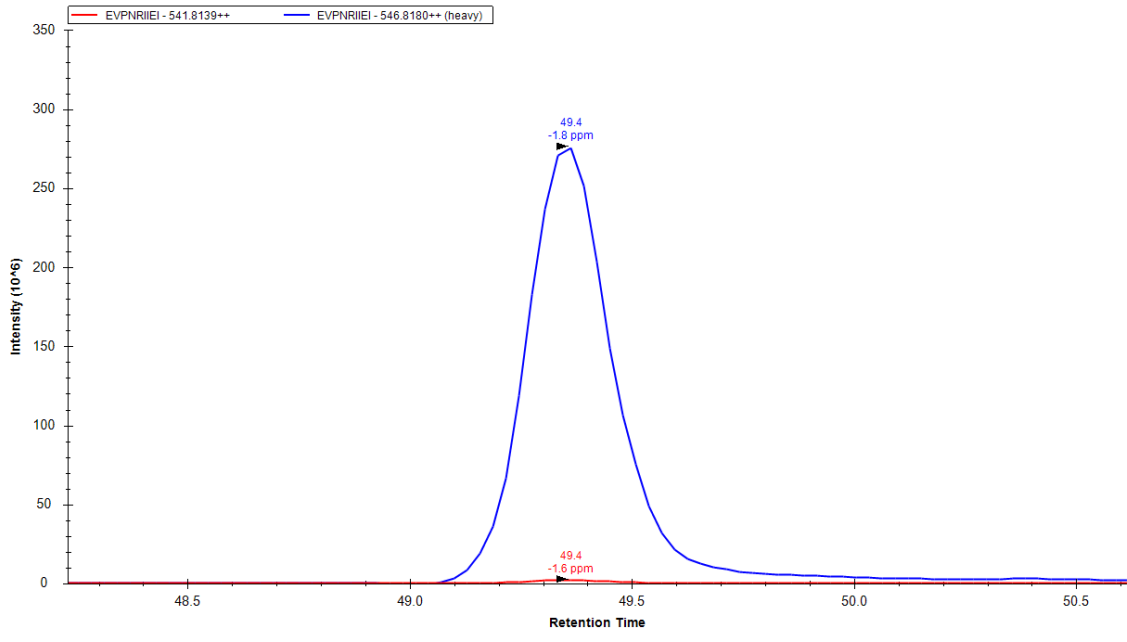


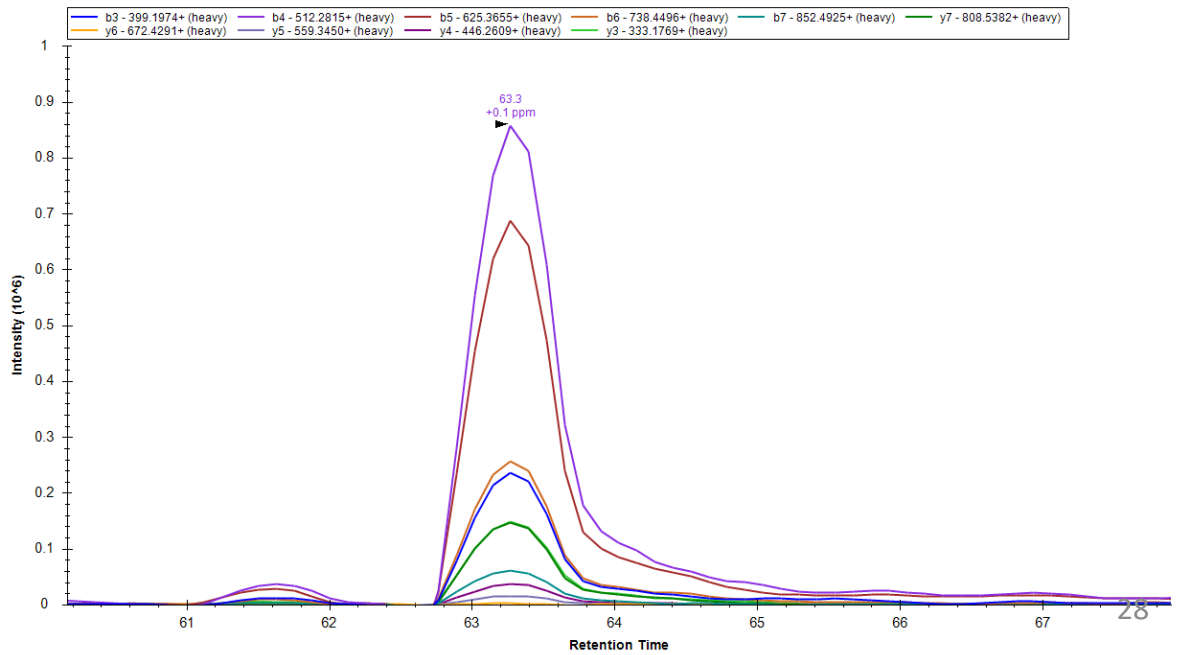
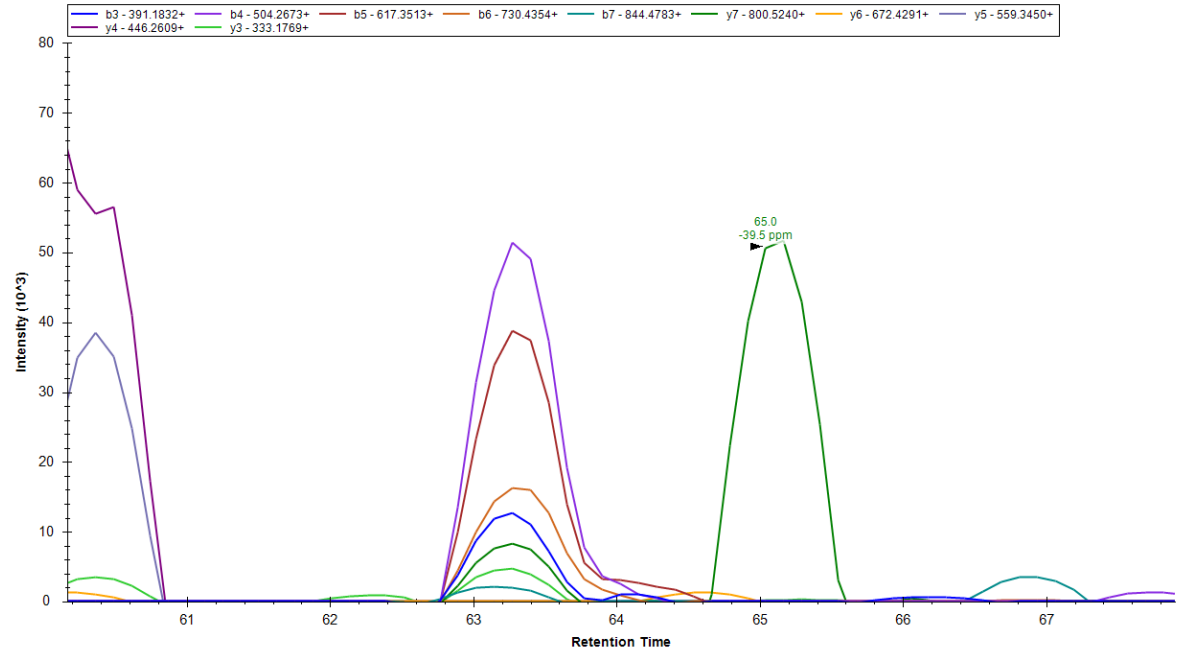
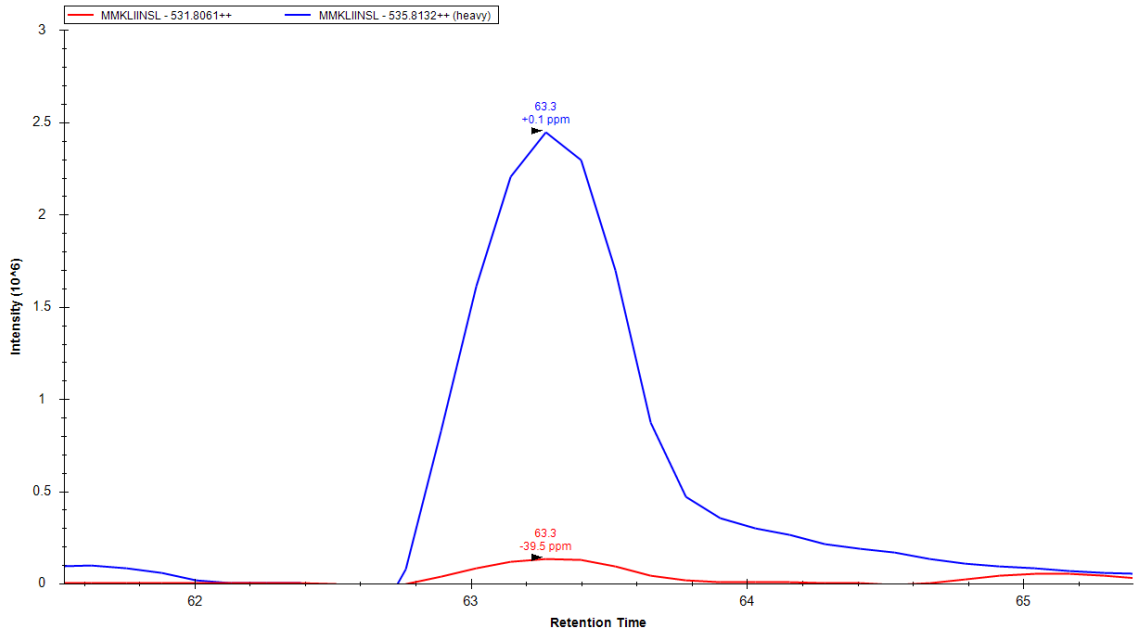


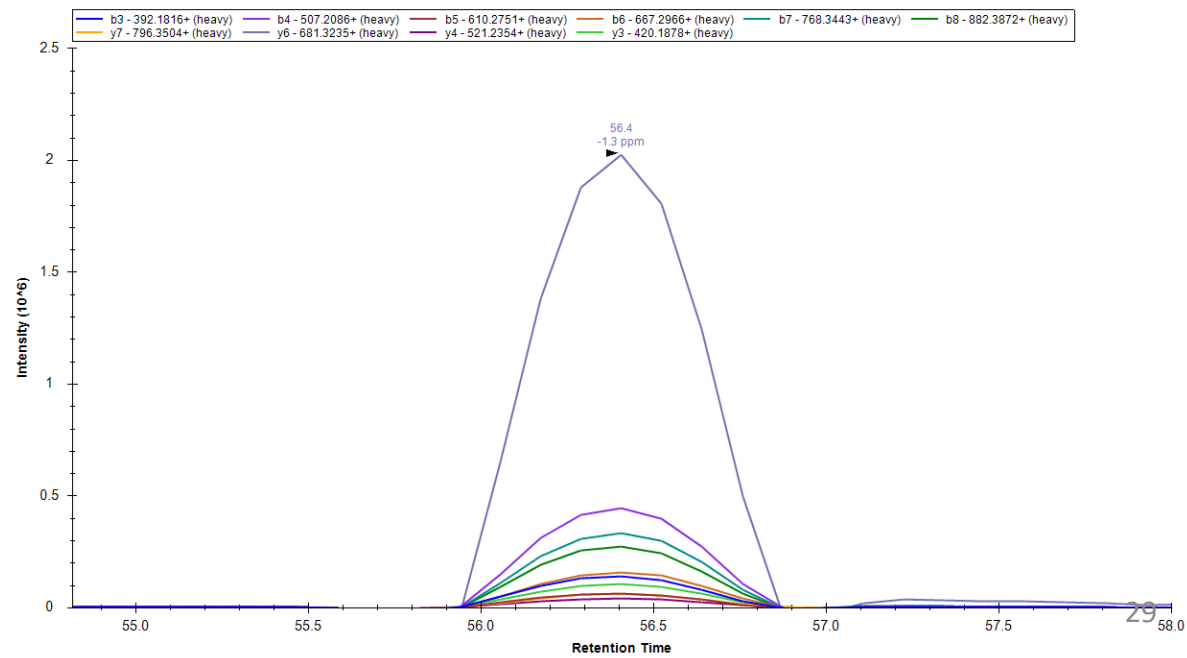
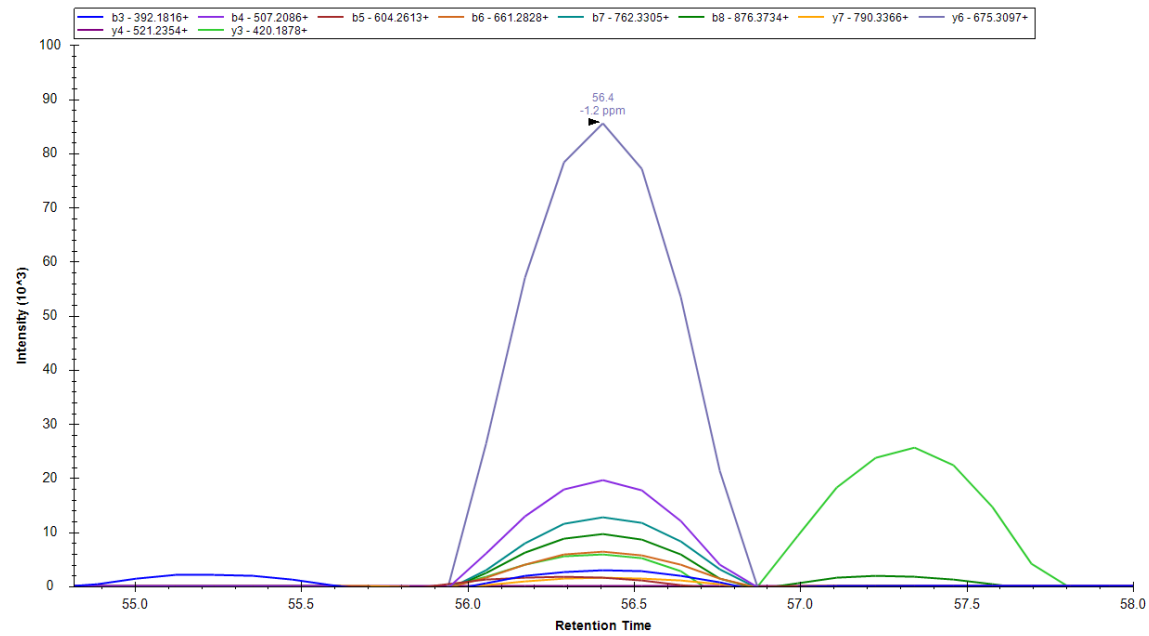
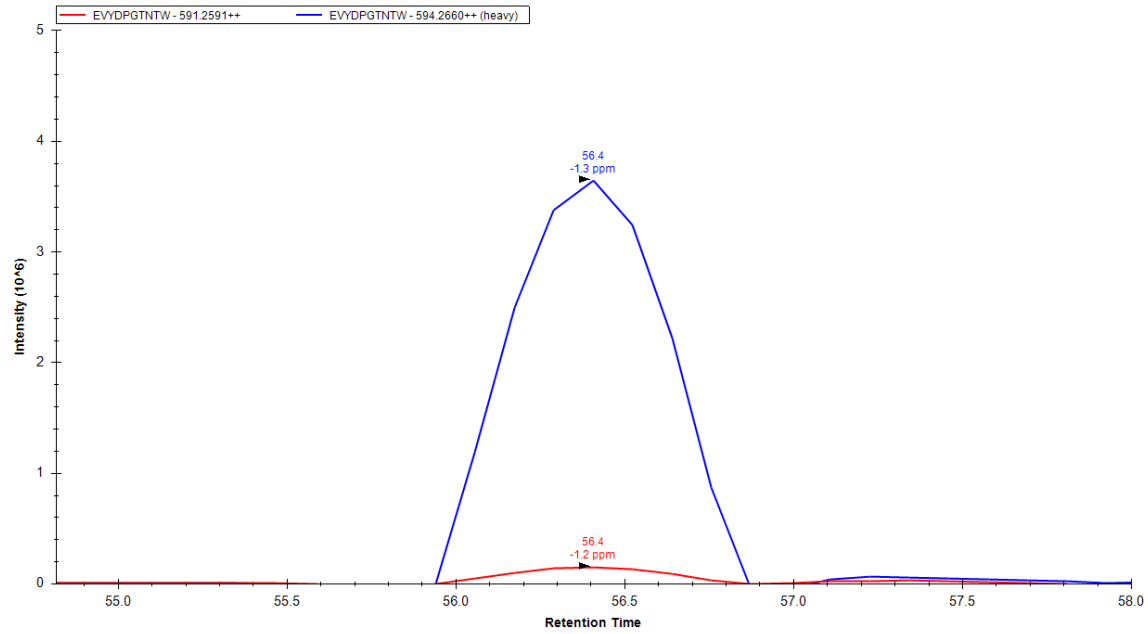


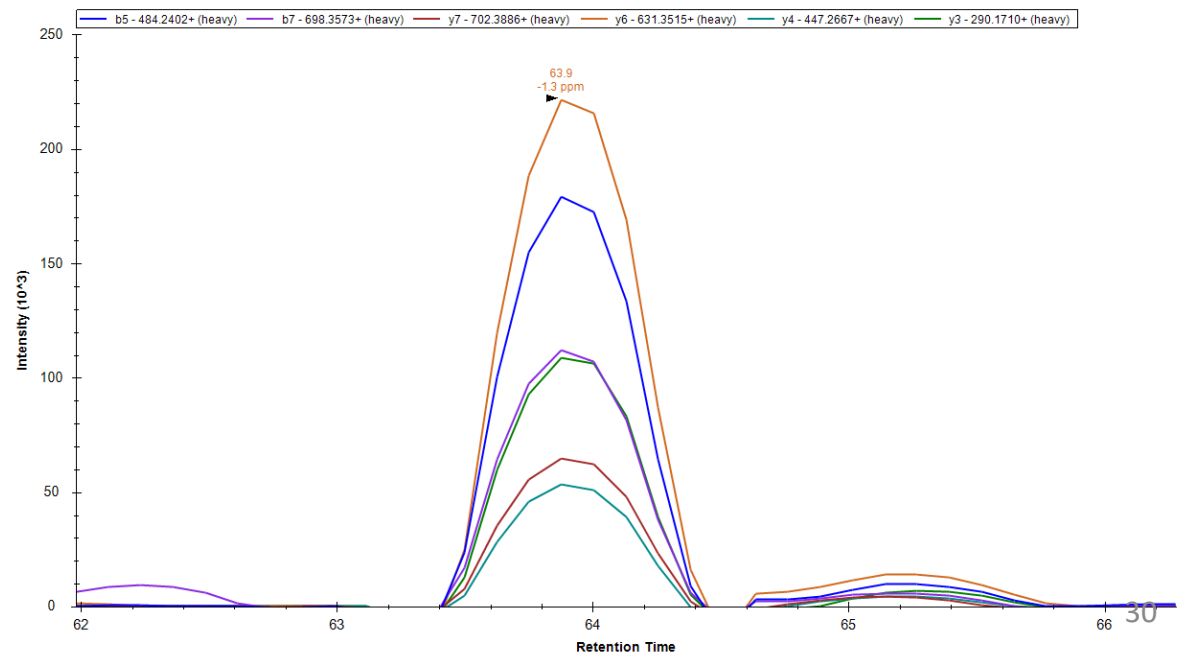
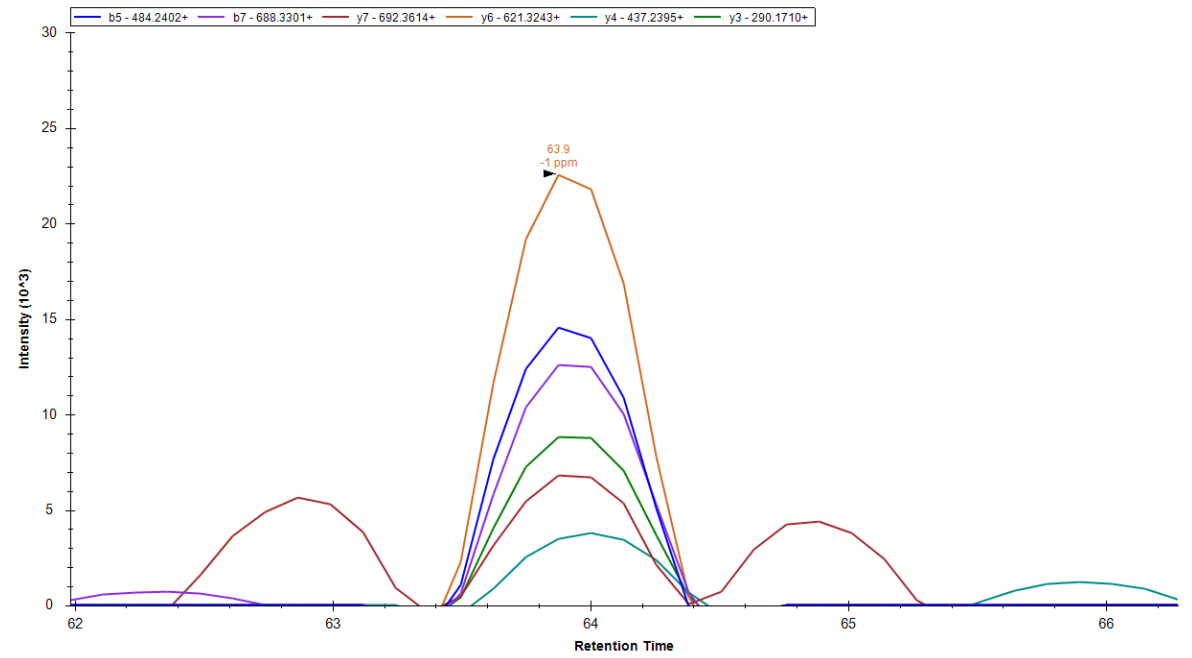
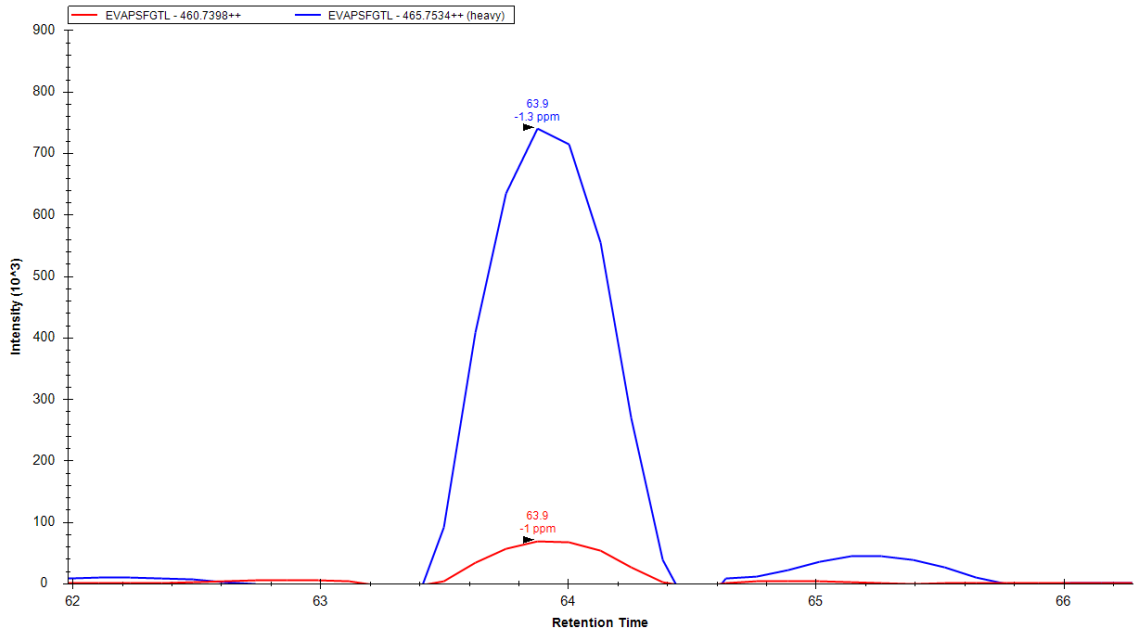


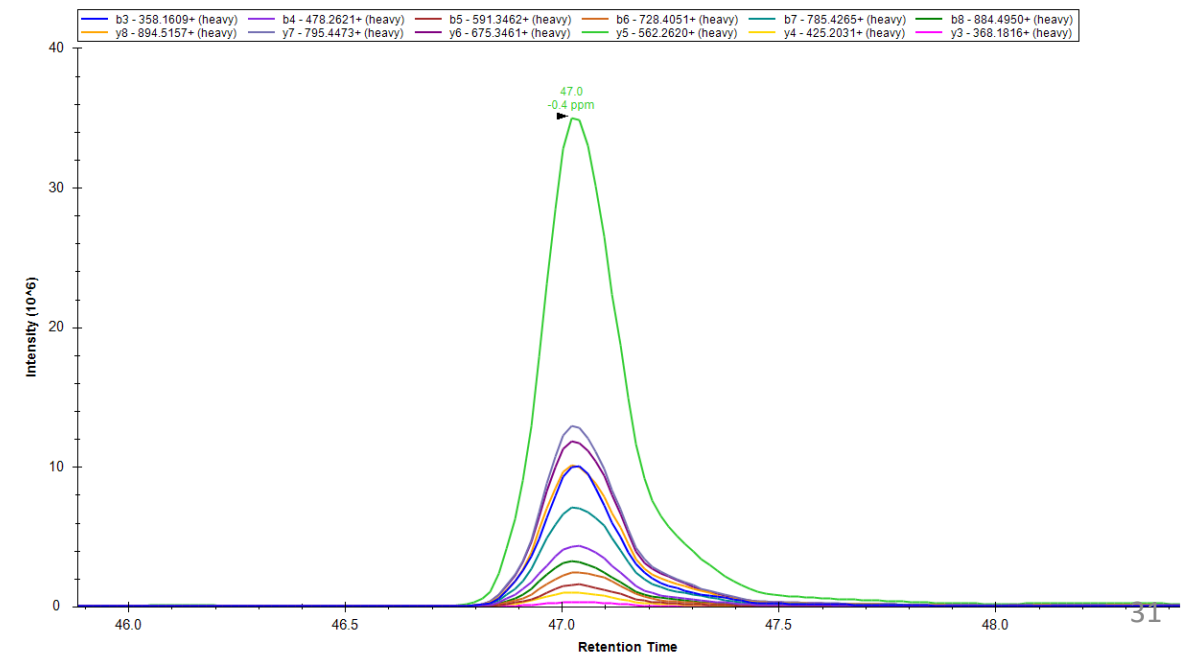
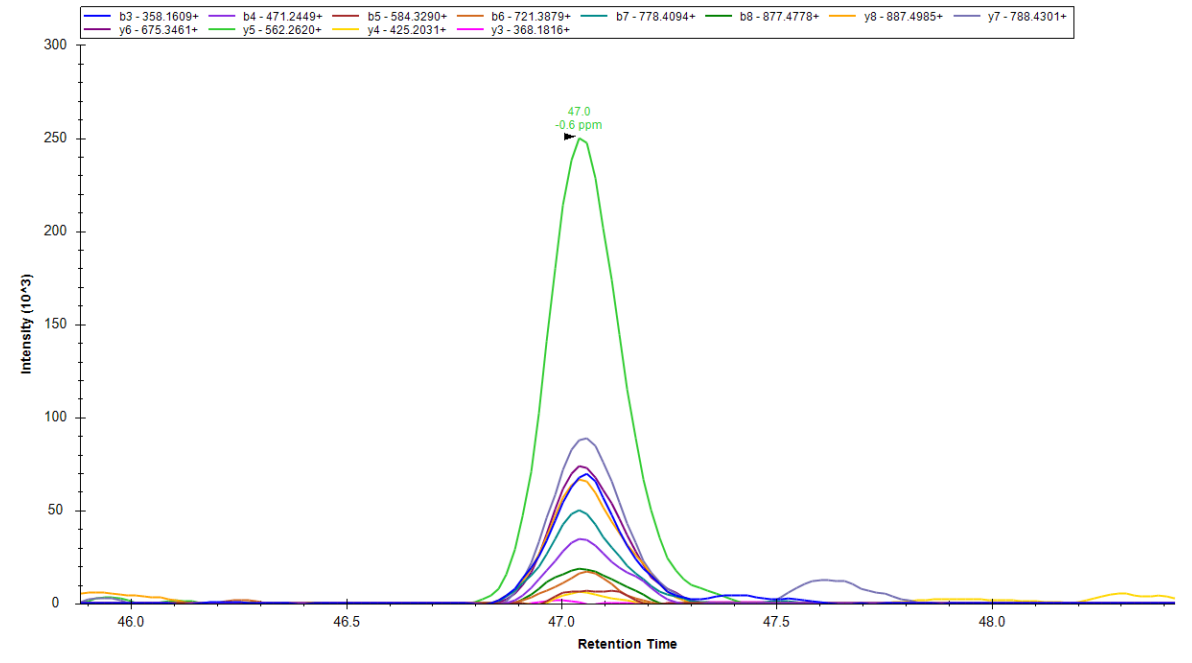
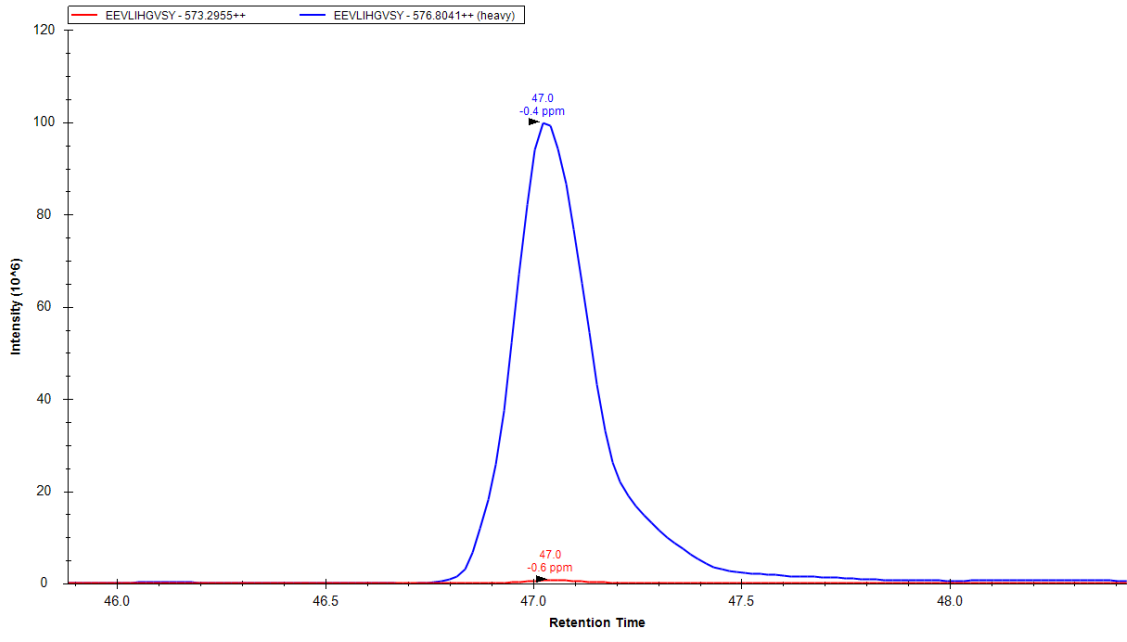


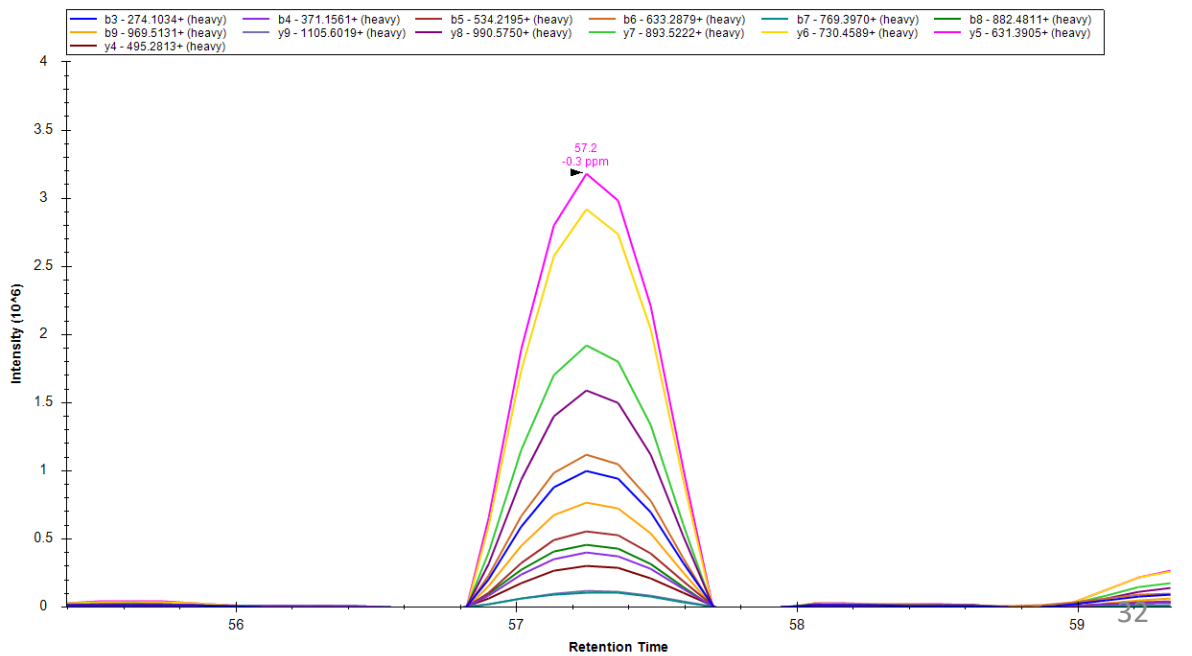
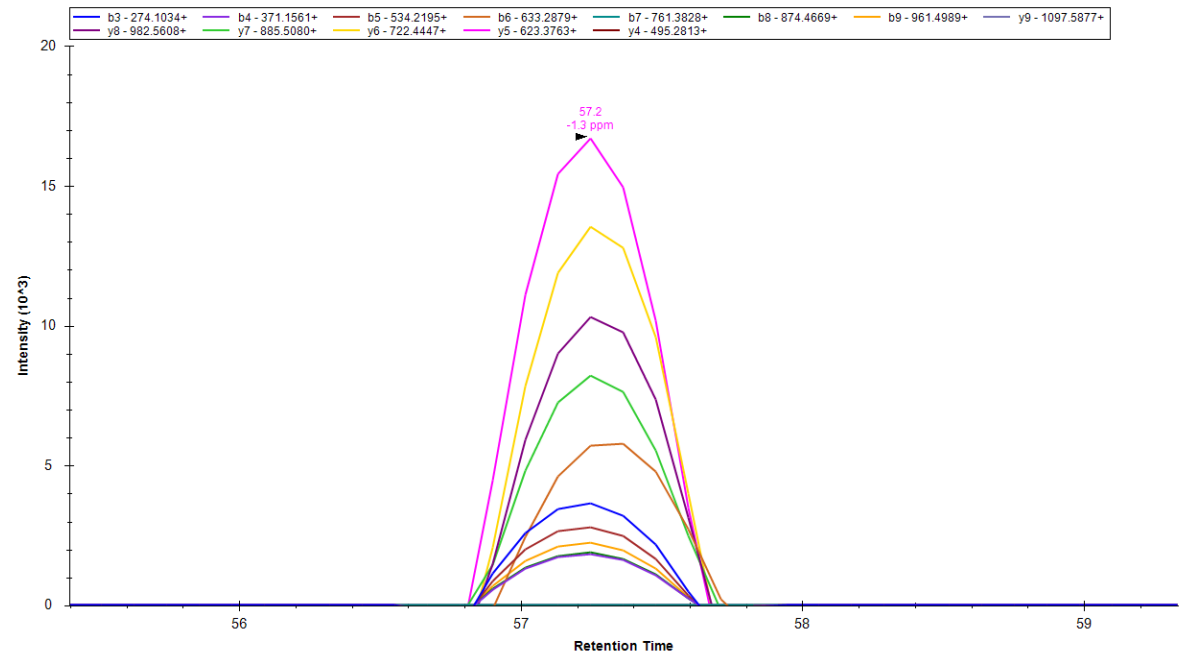
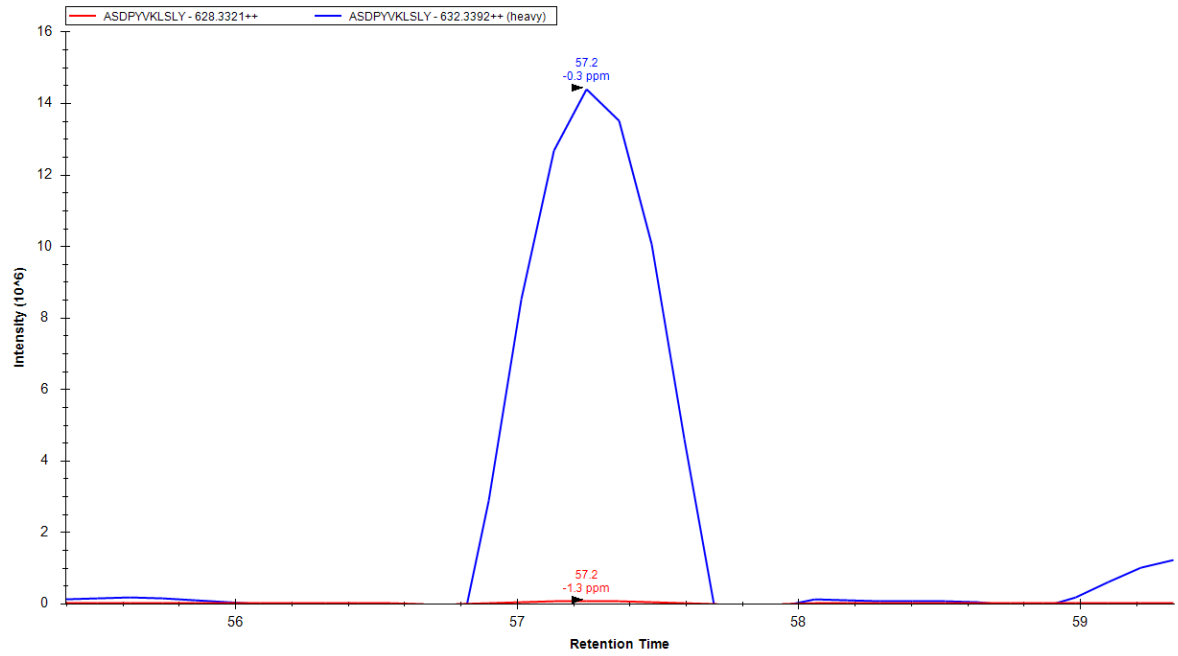




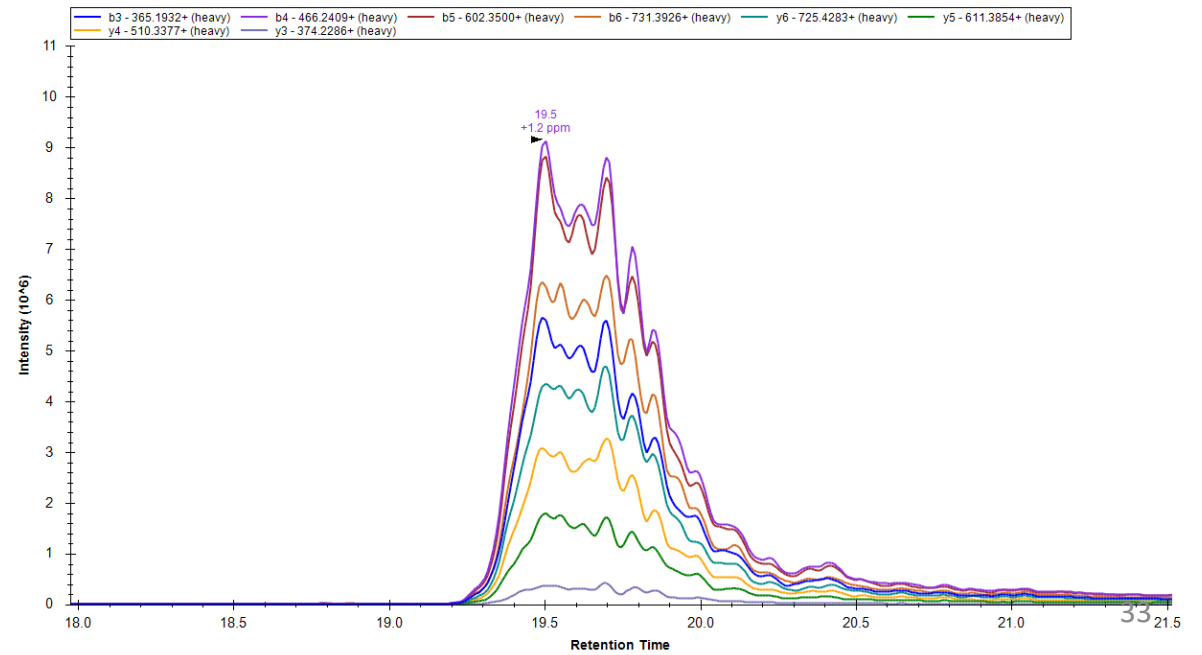
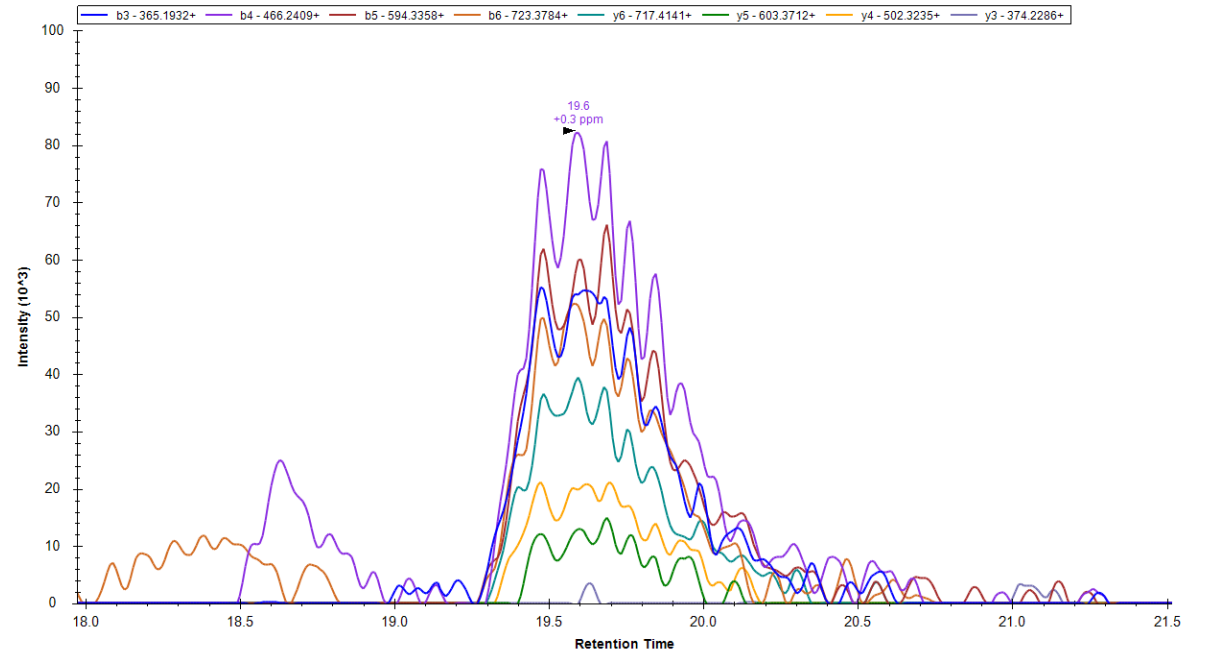
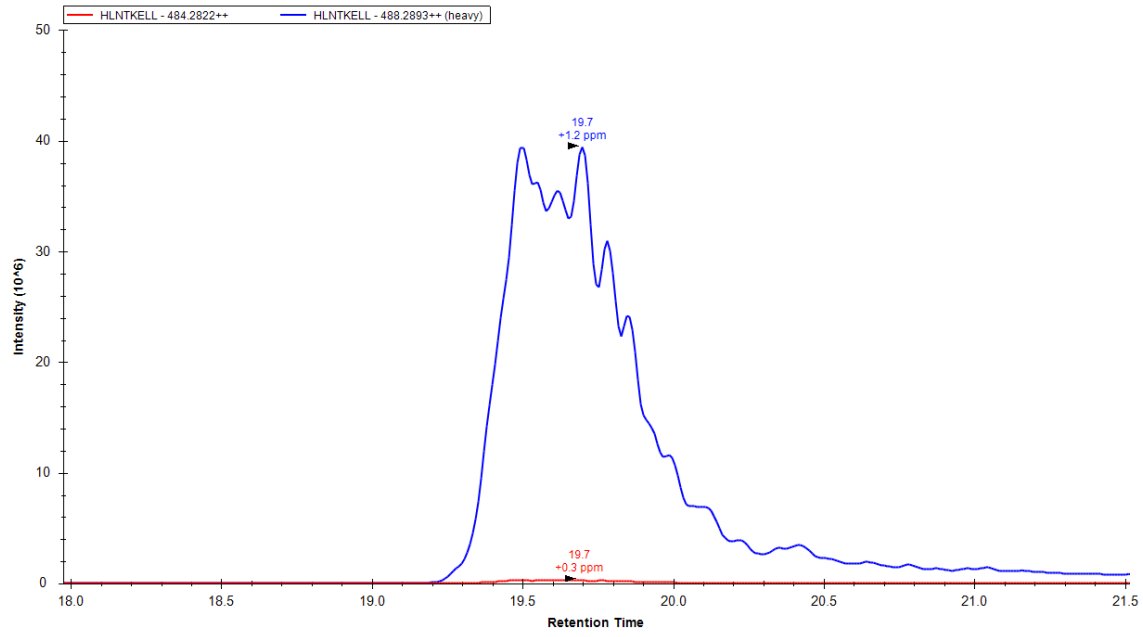


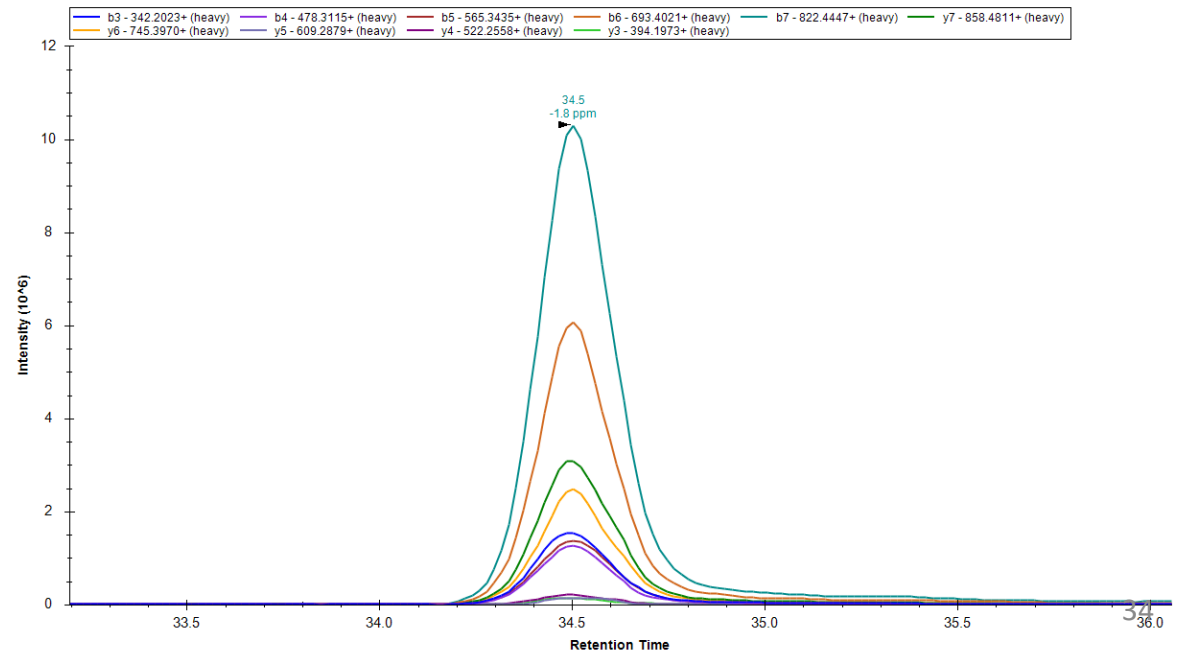
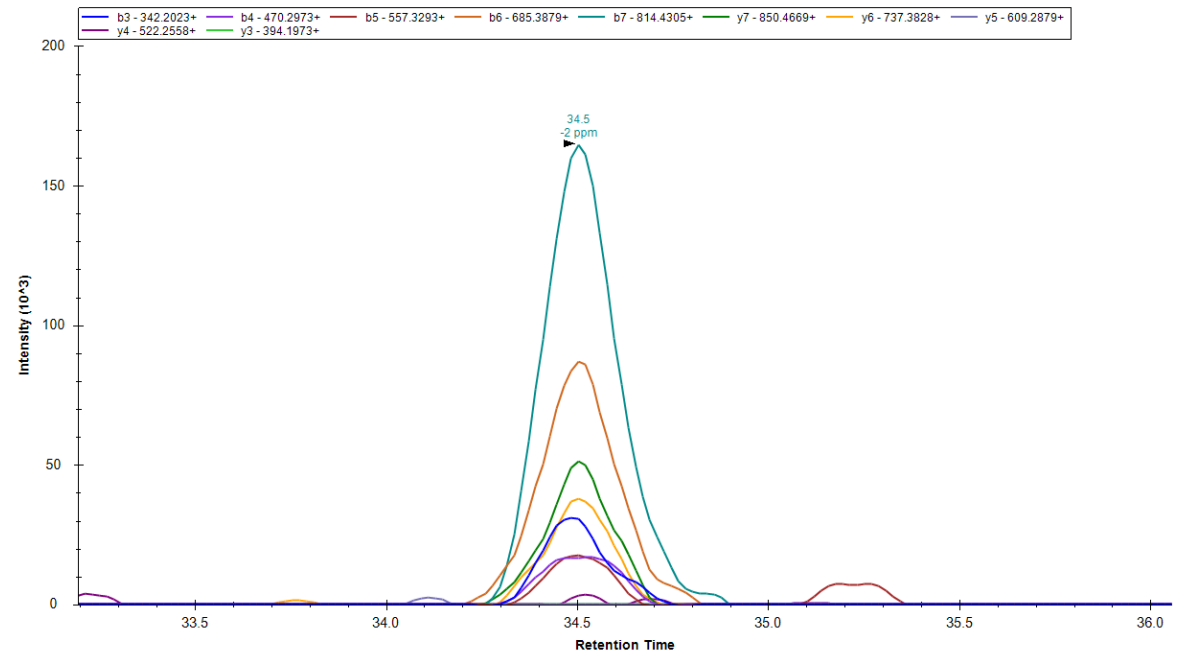
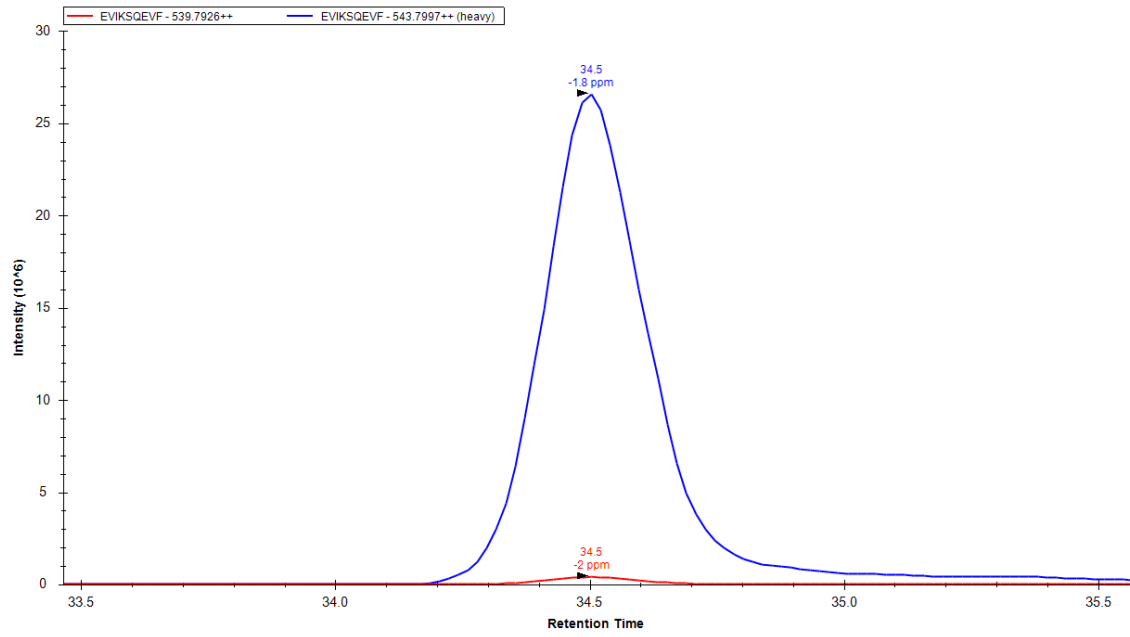


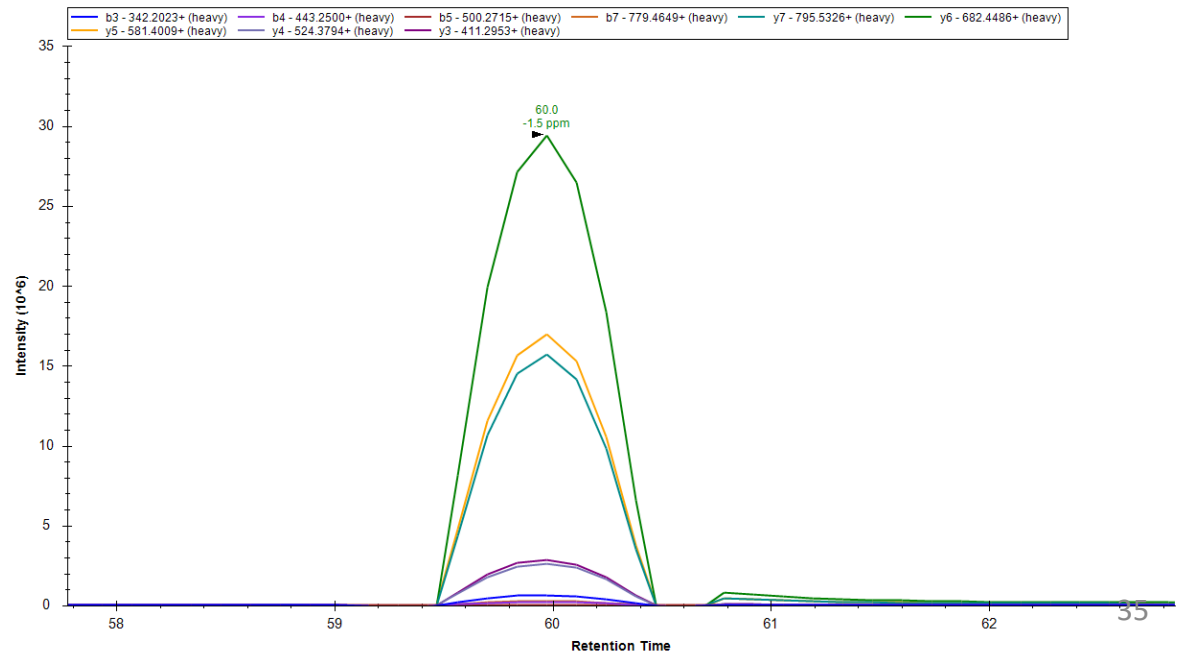
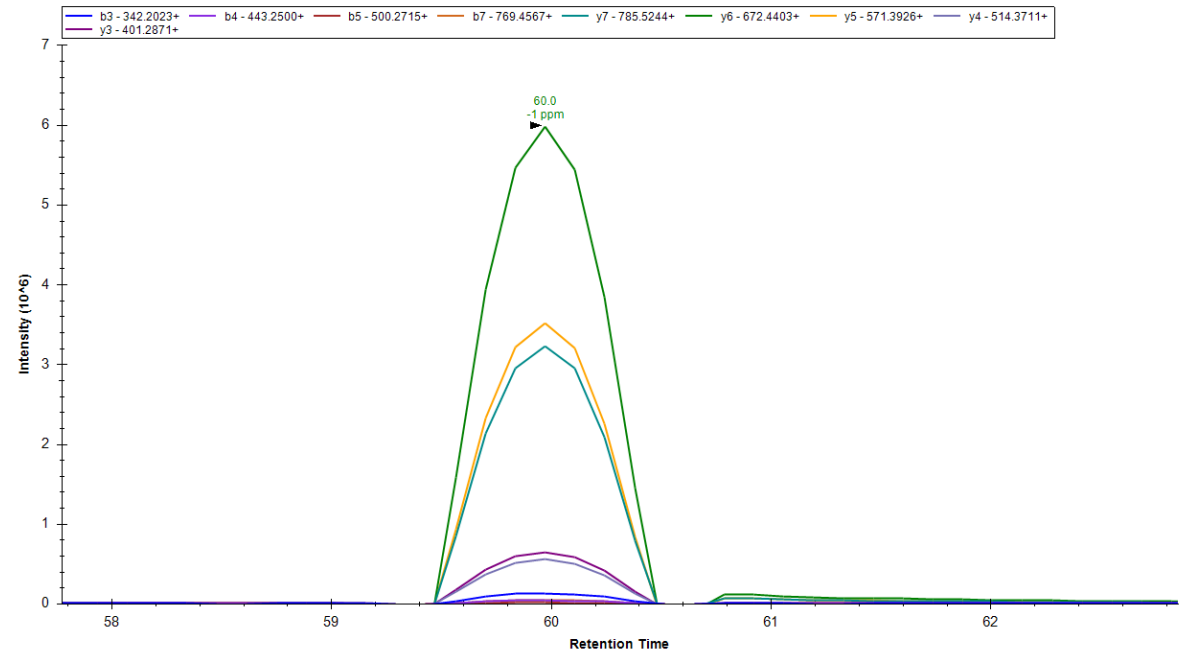
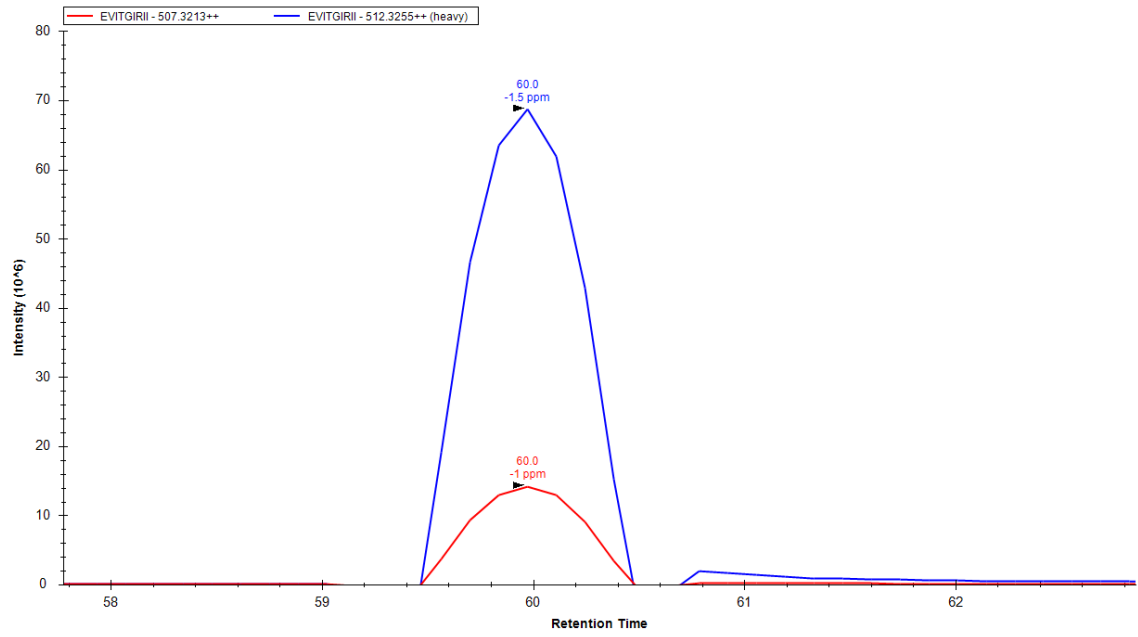


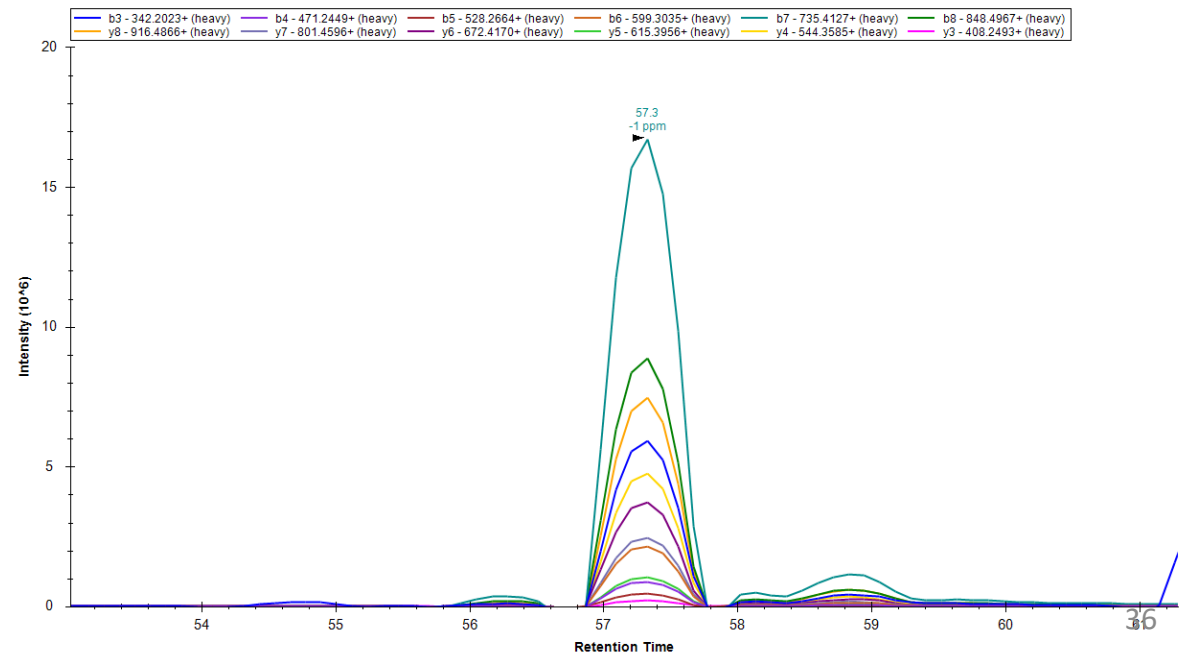
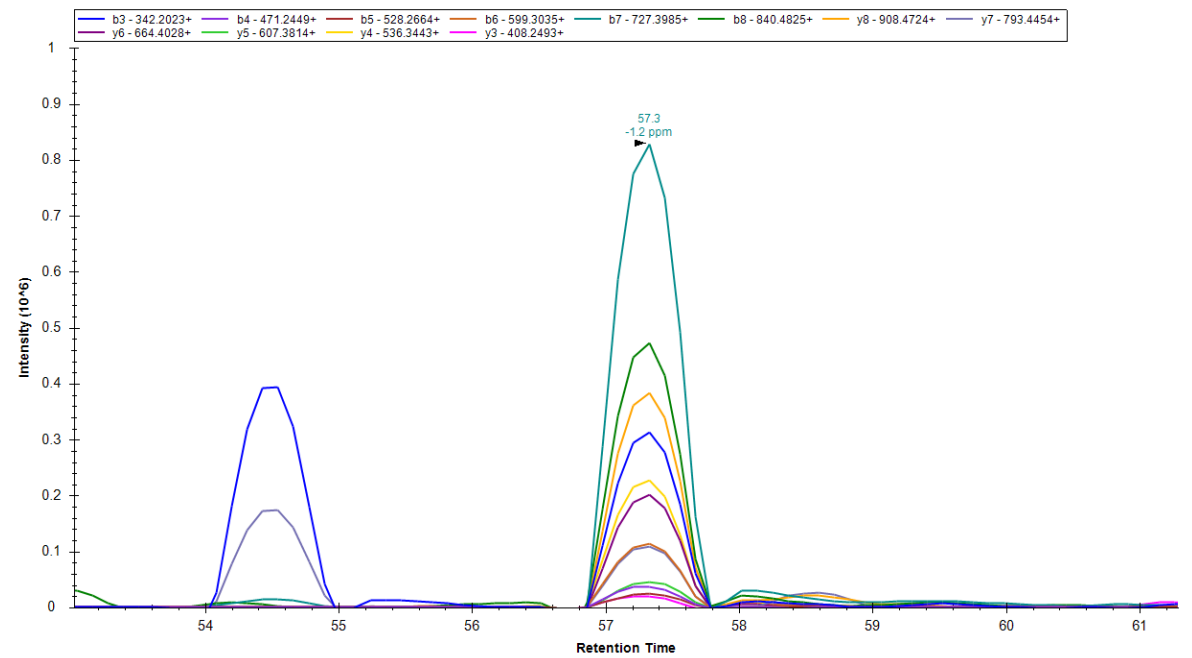
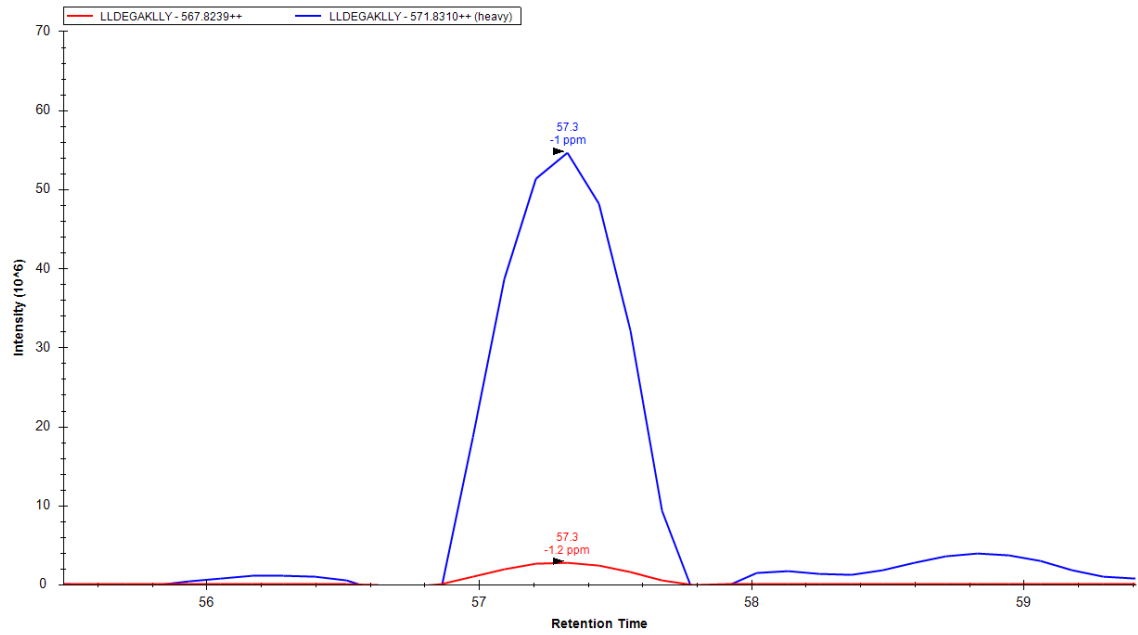


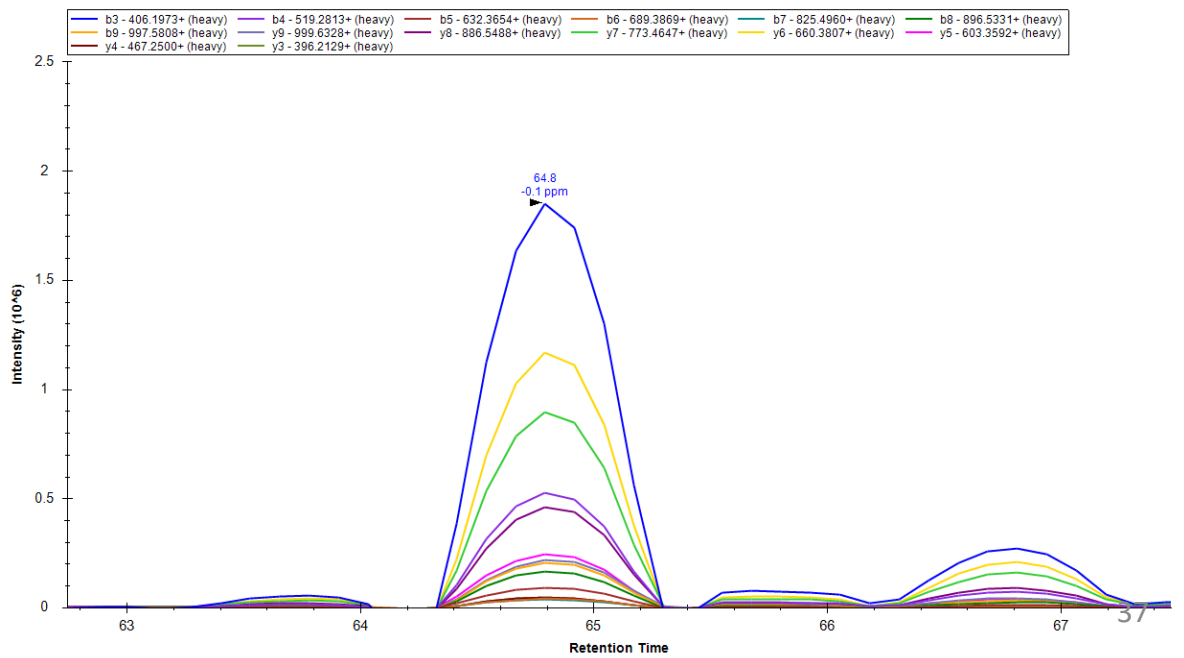
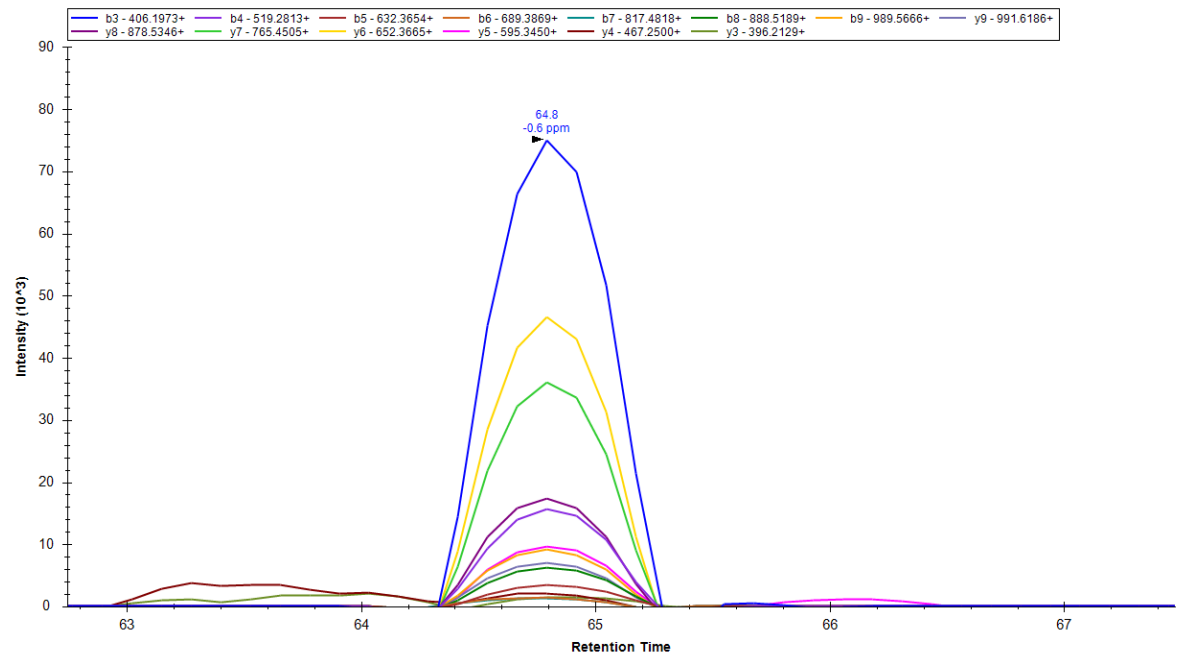
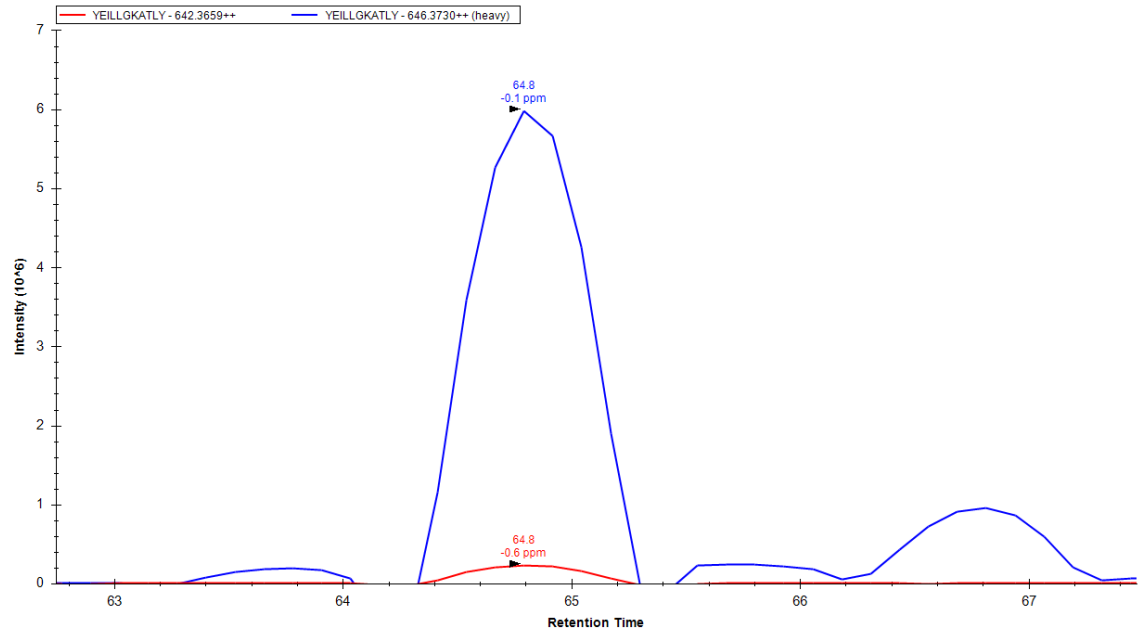






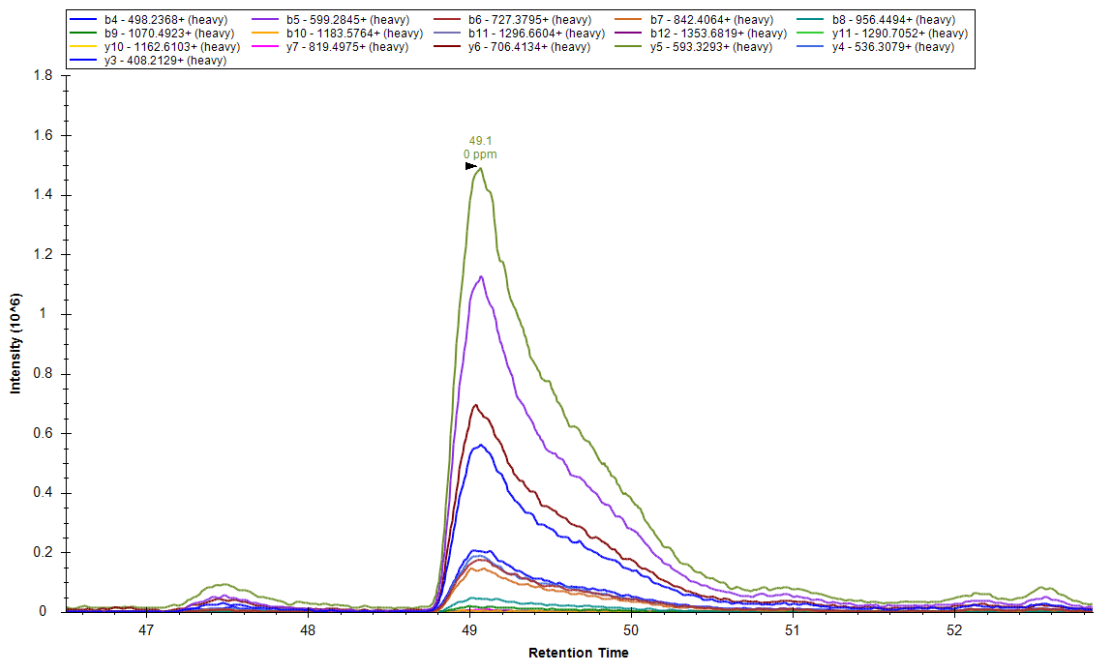
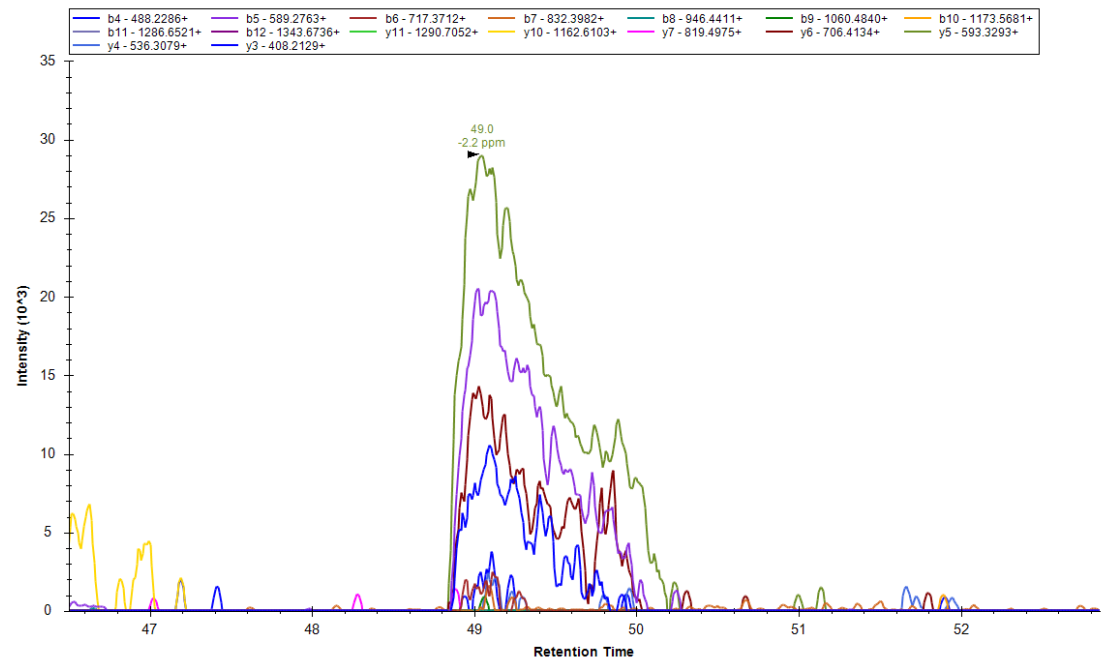
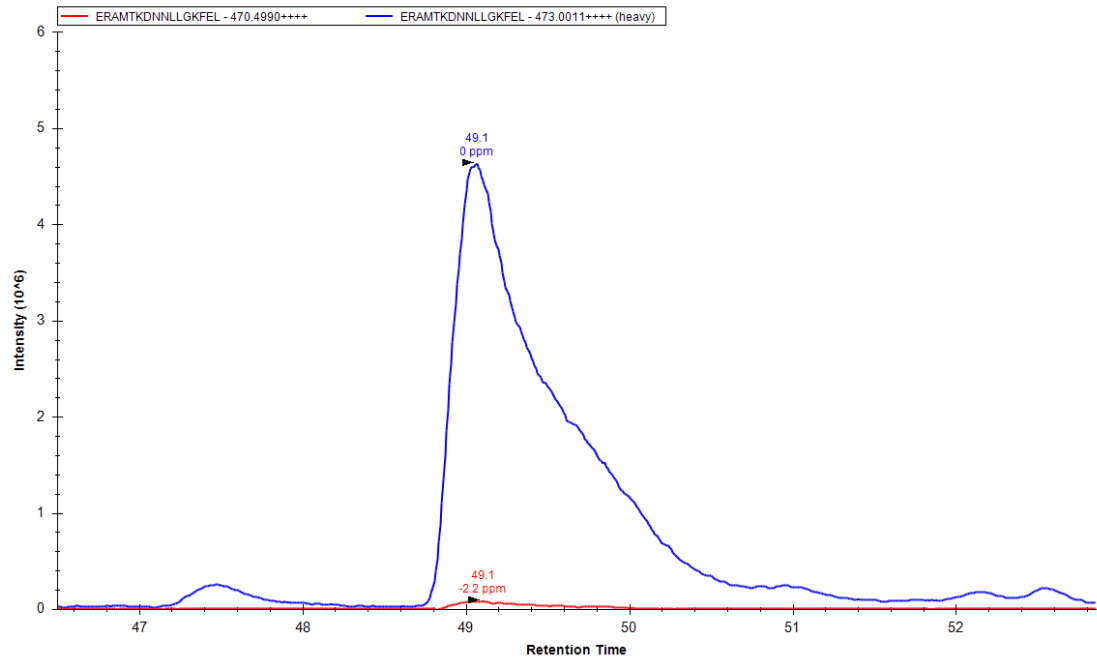


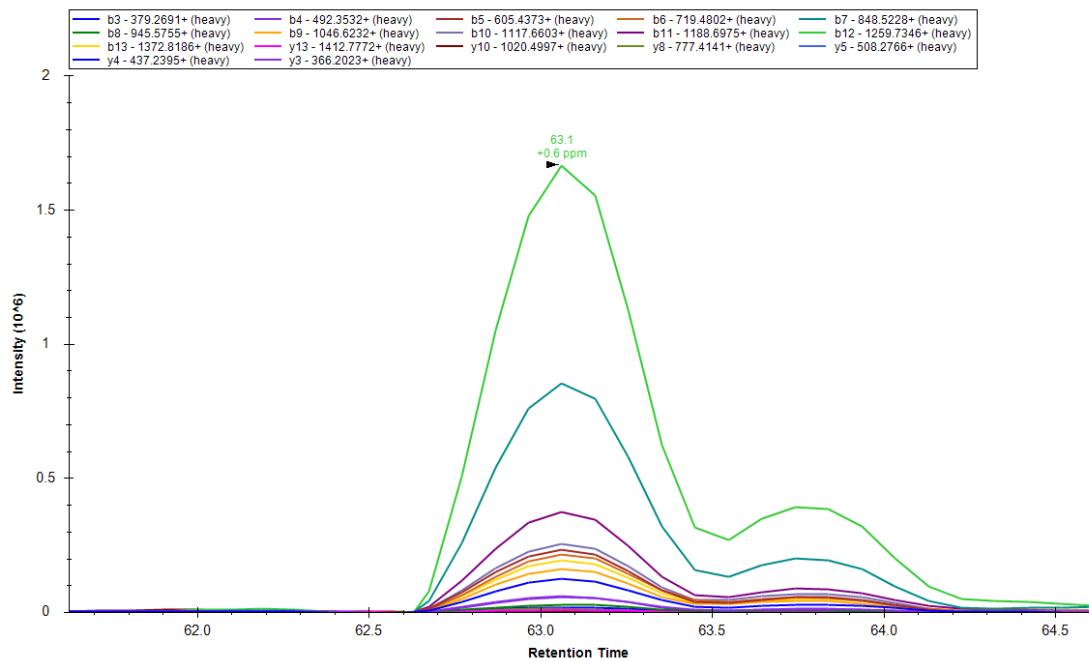
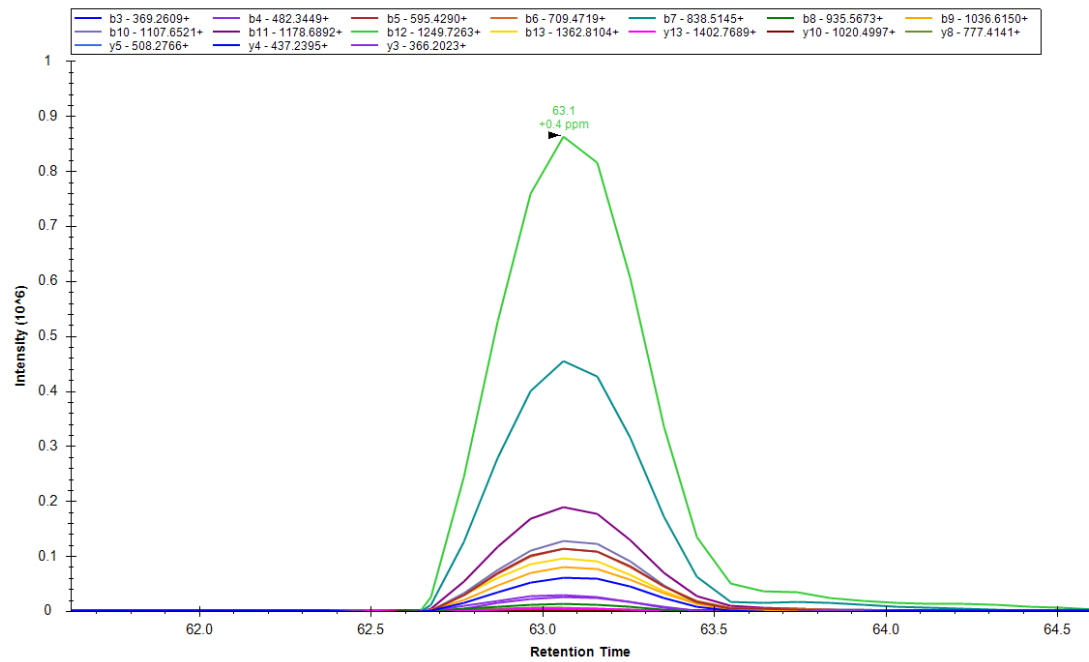
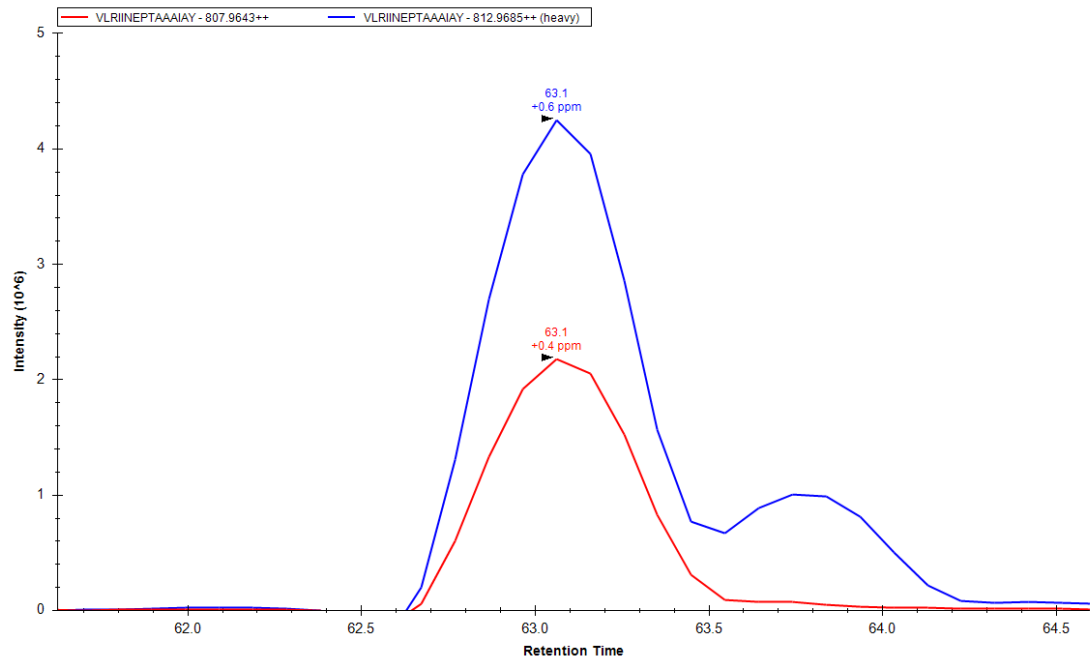




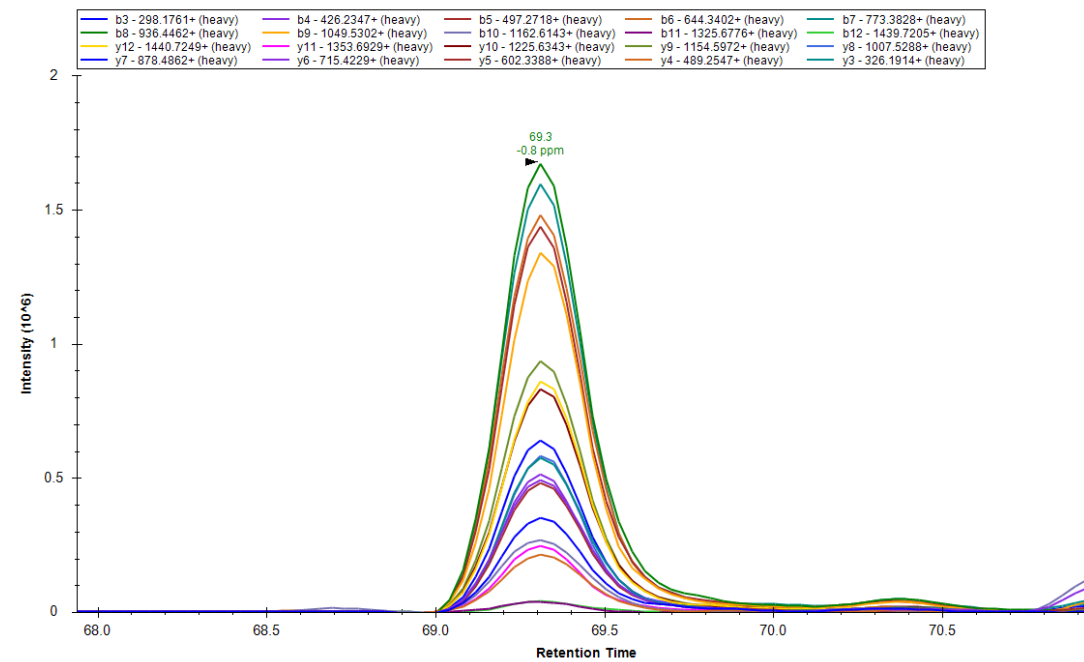
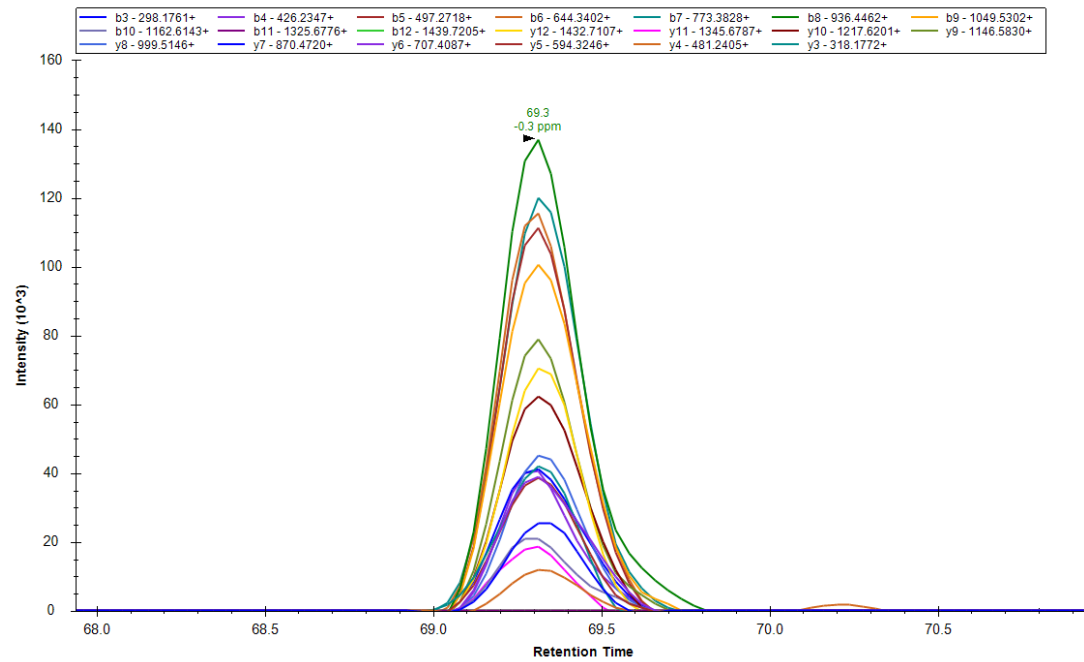
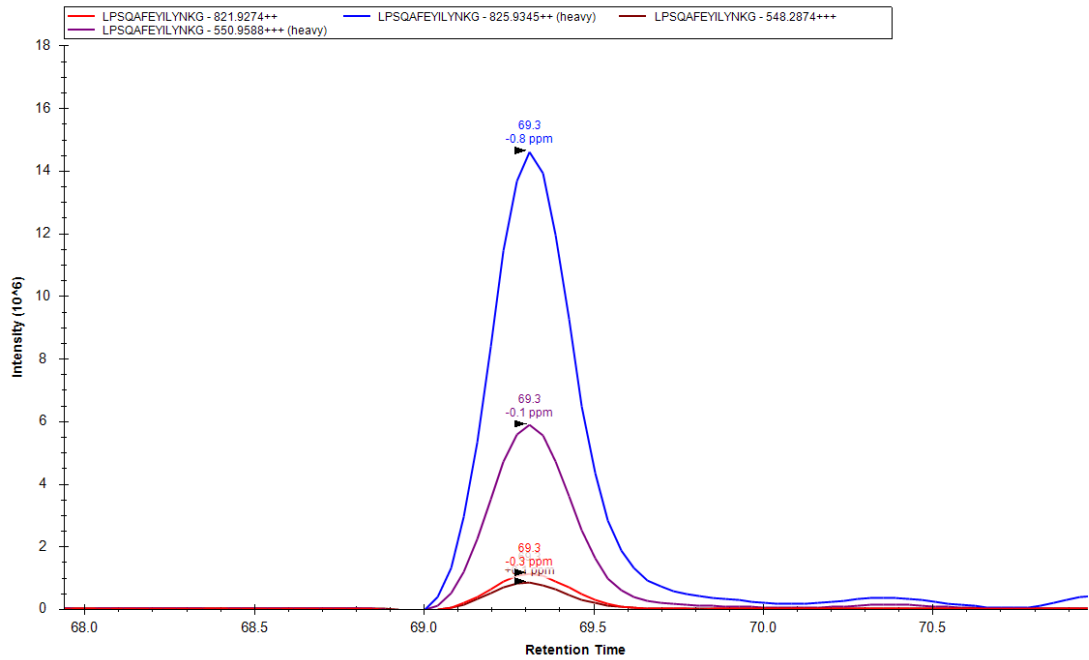
**Fig. S8**

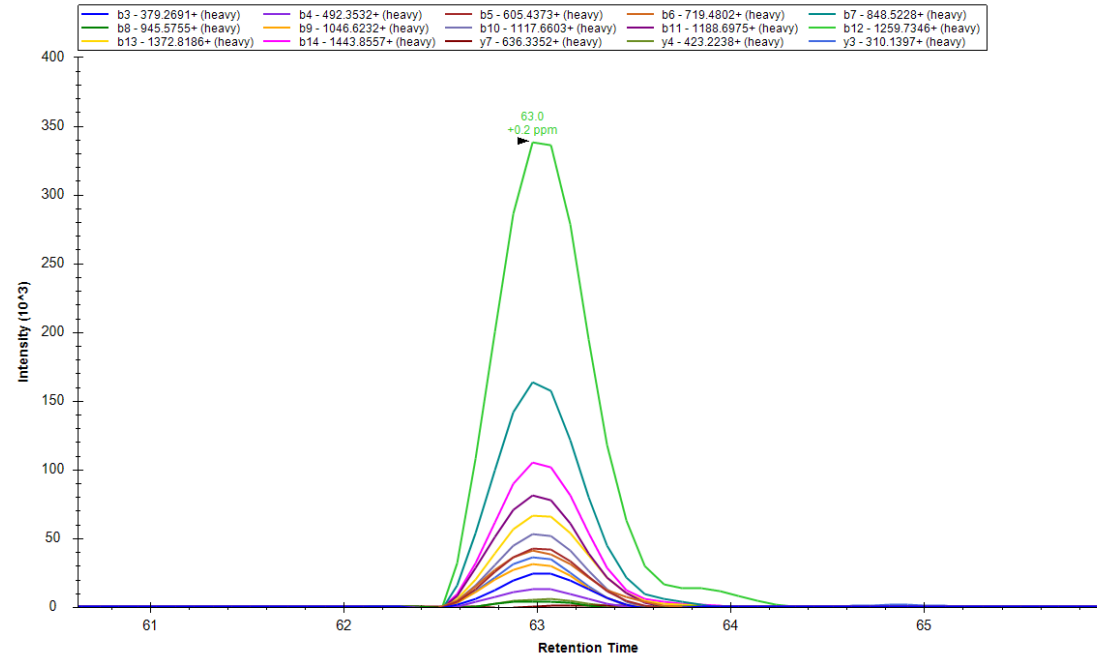
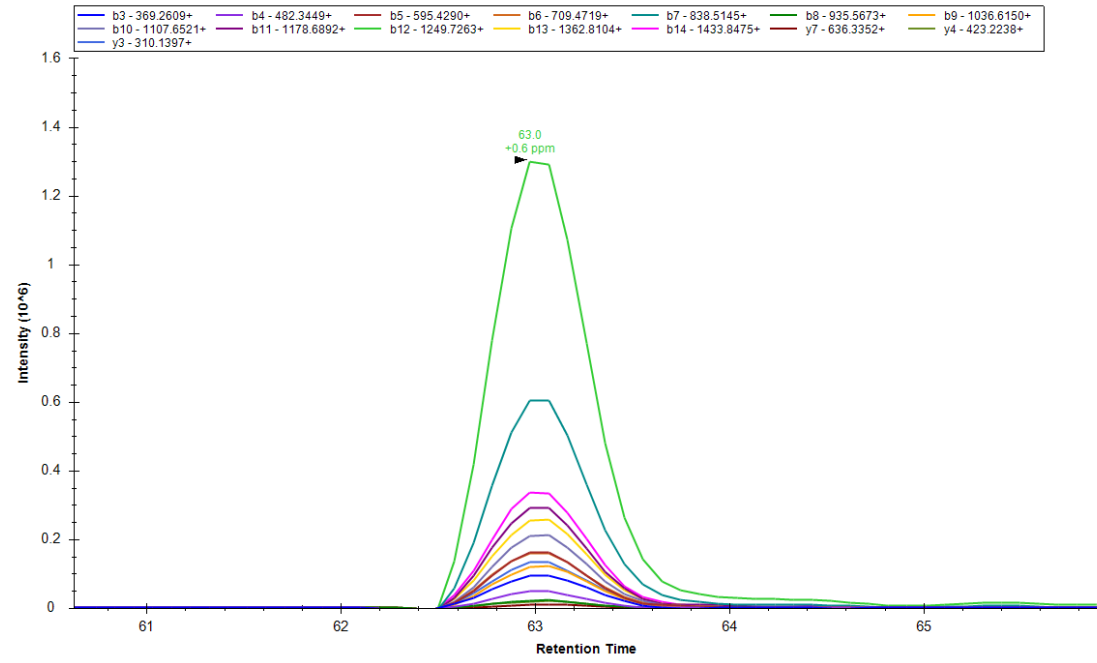
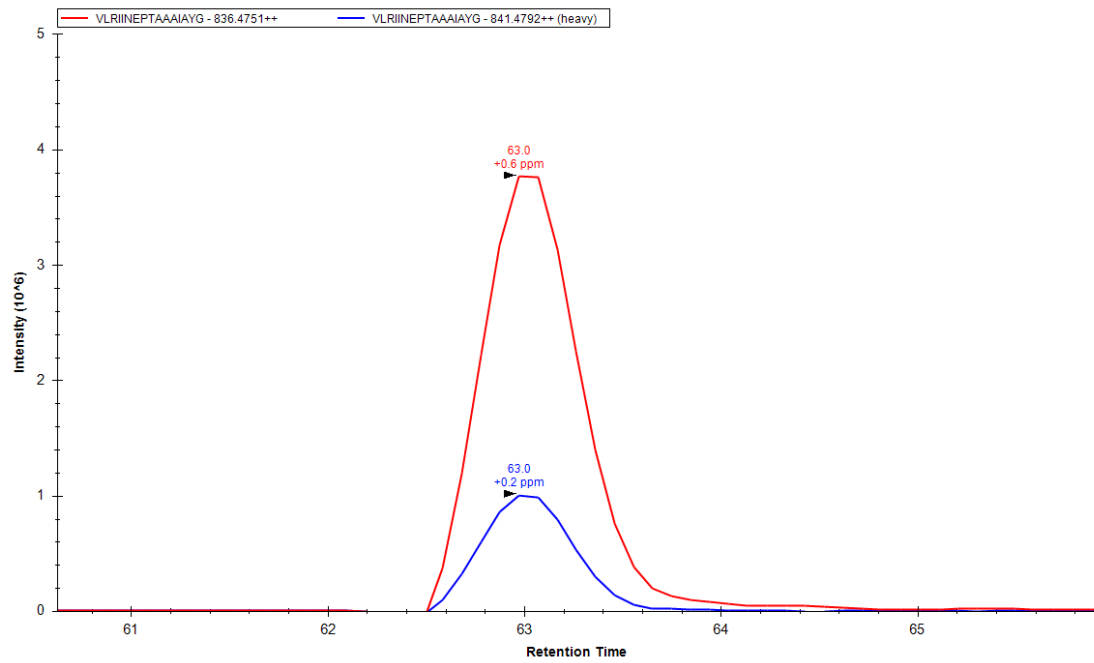
**Pages: 38 - 67**

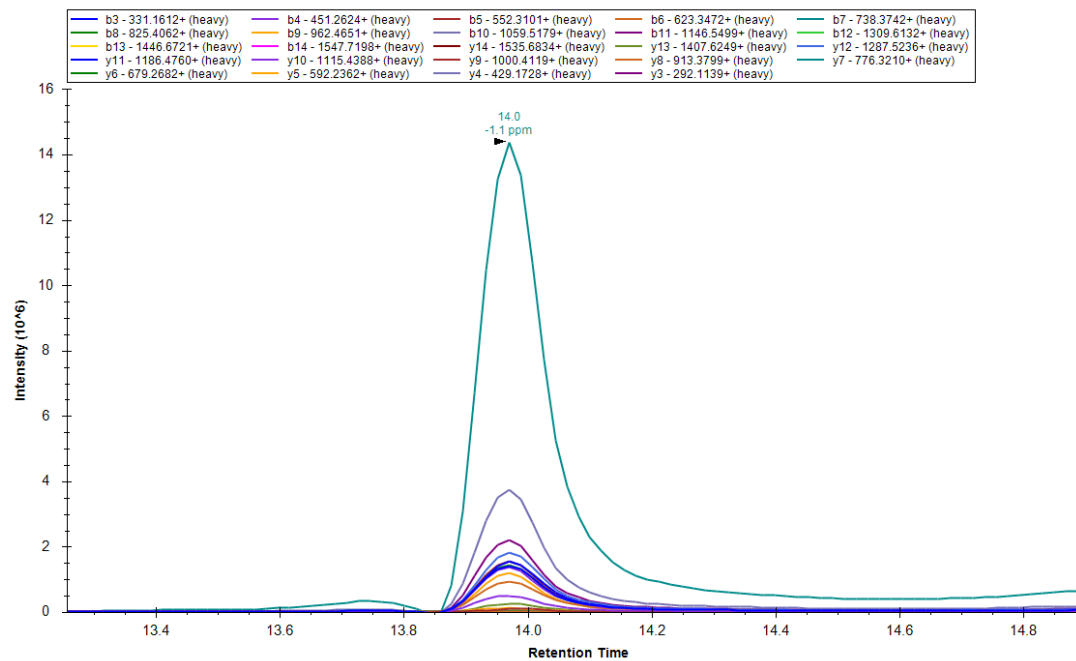
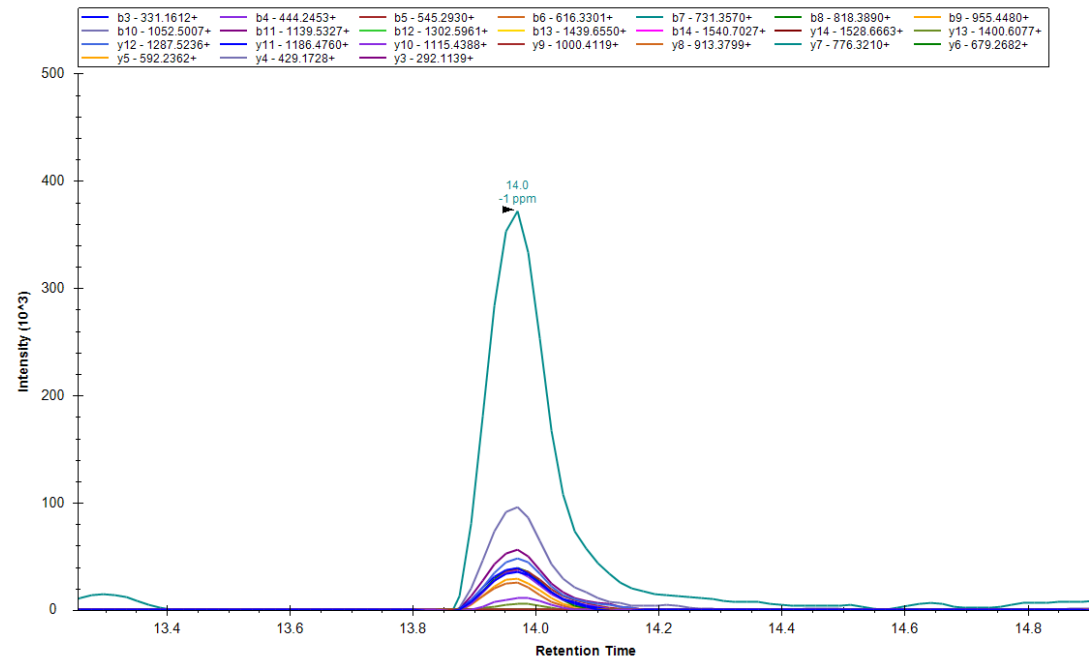
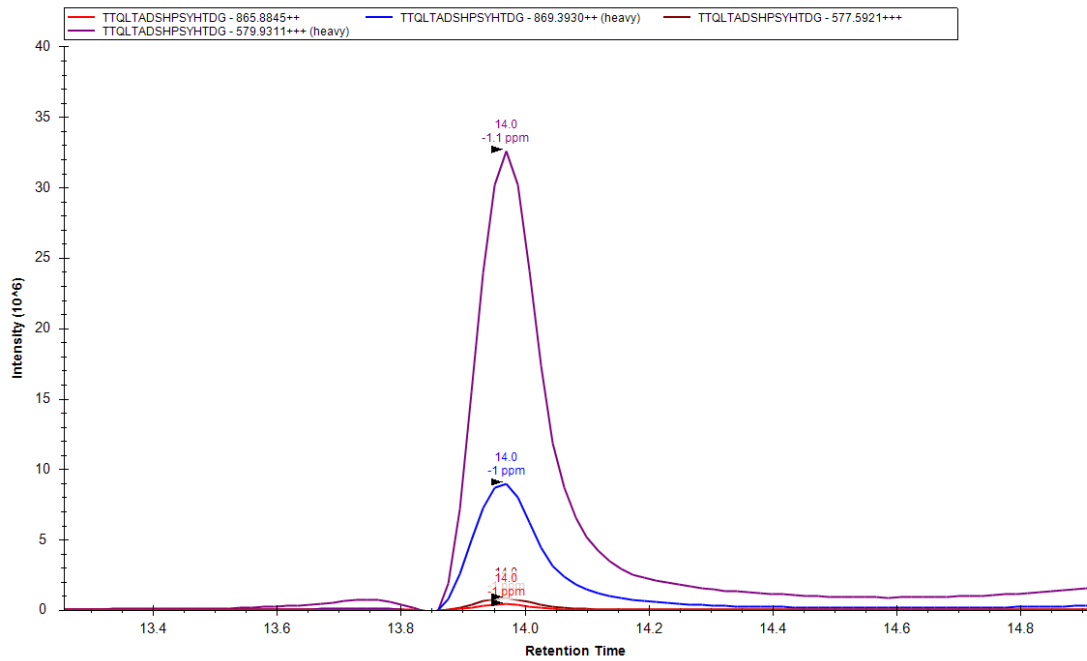


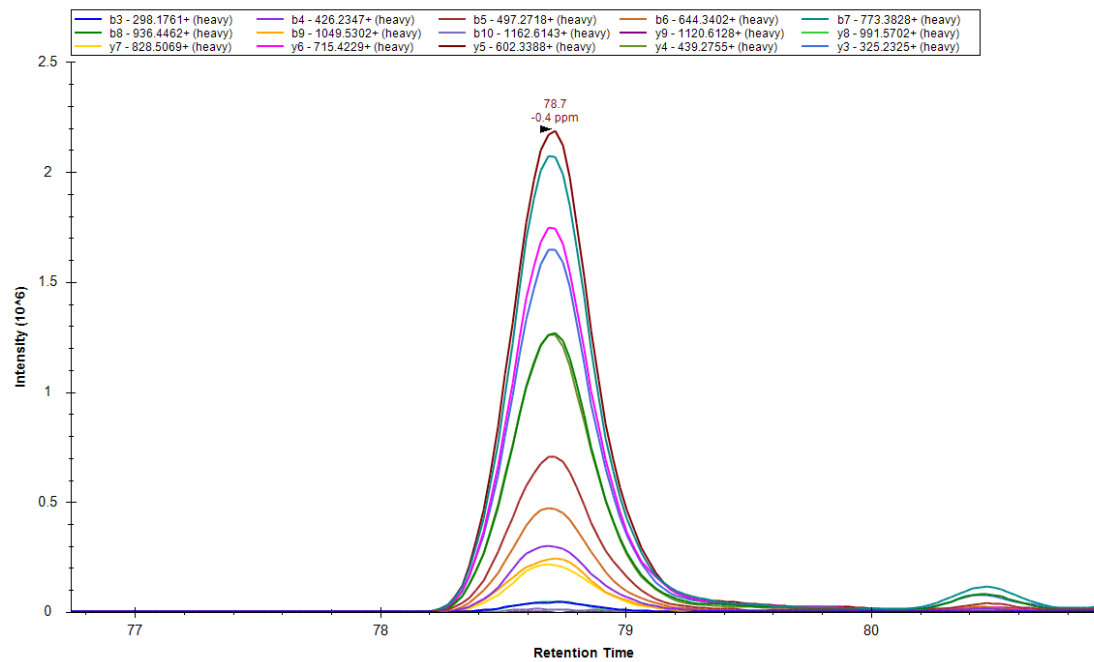
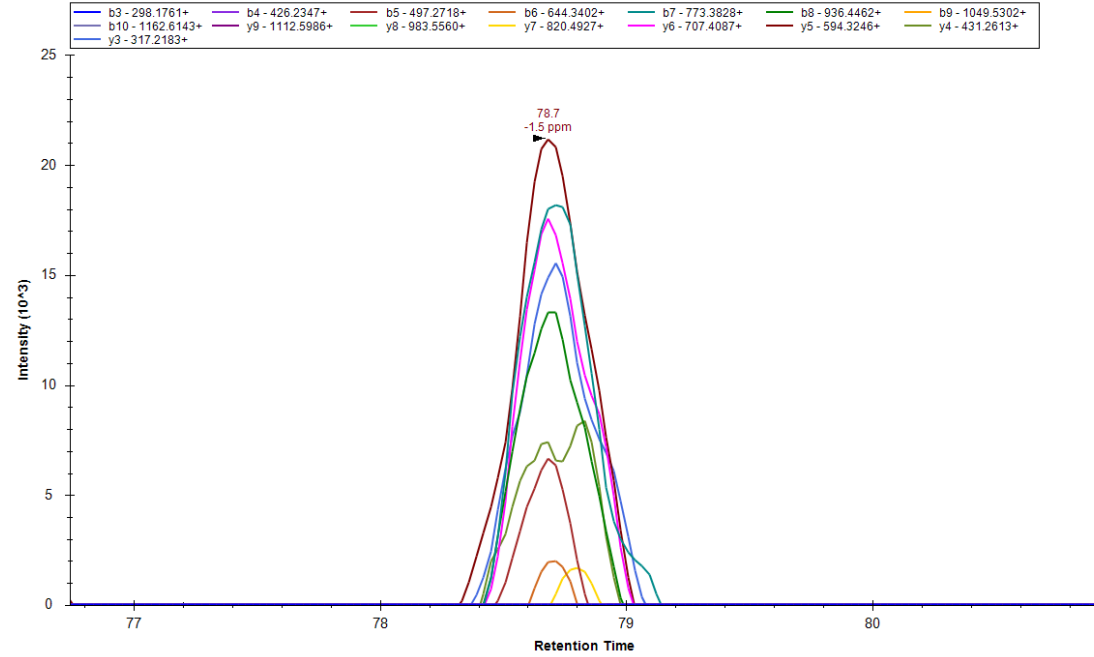
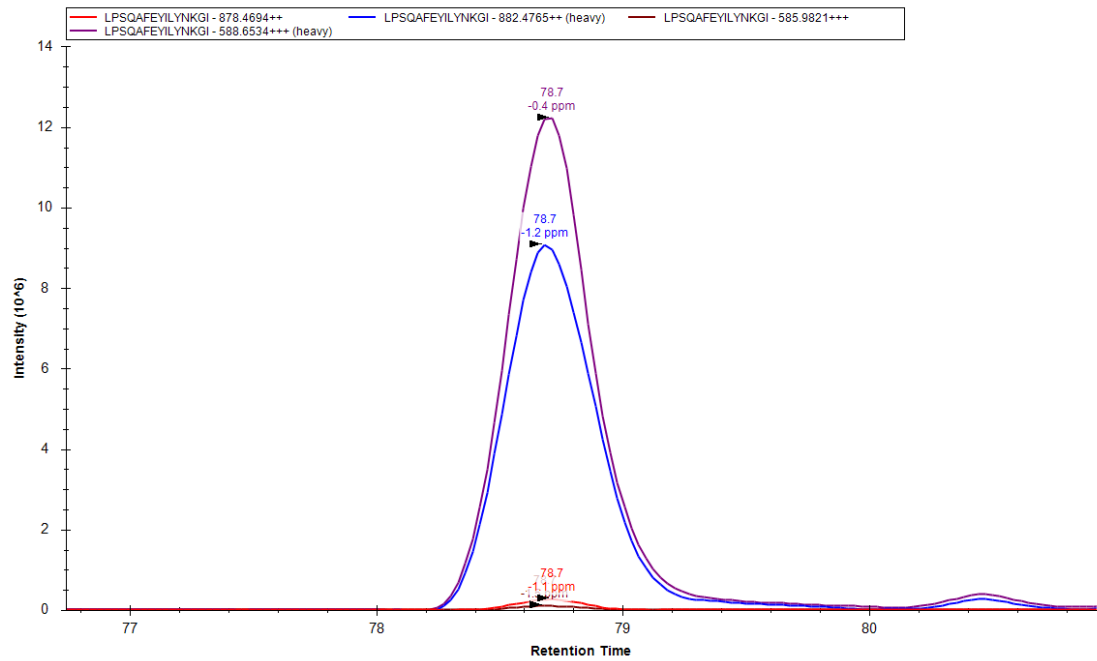


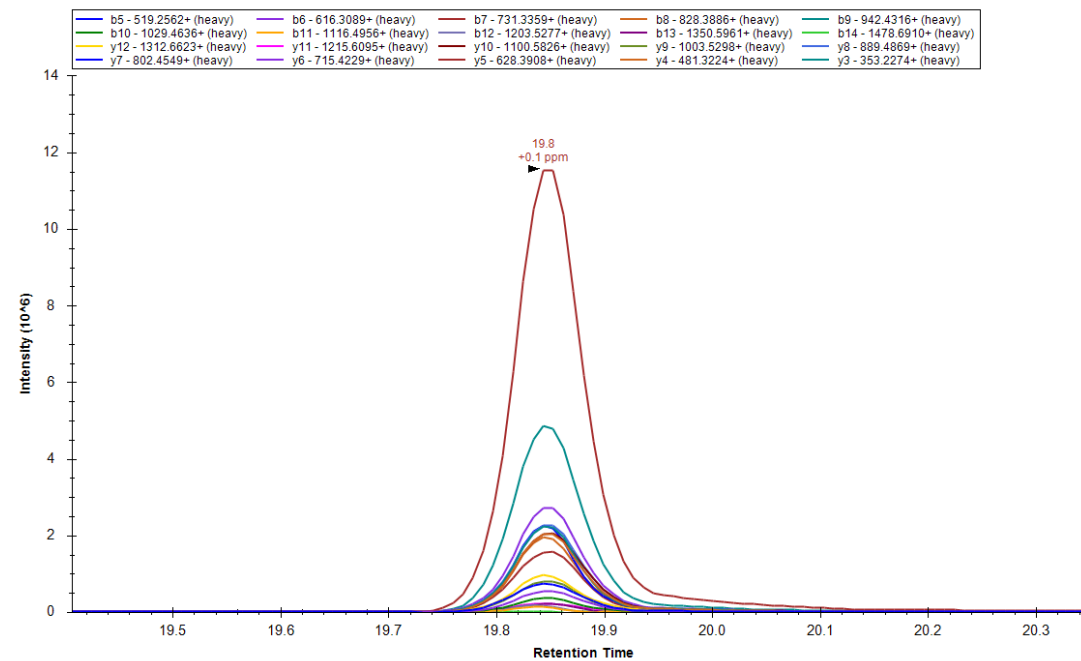
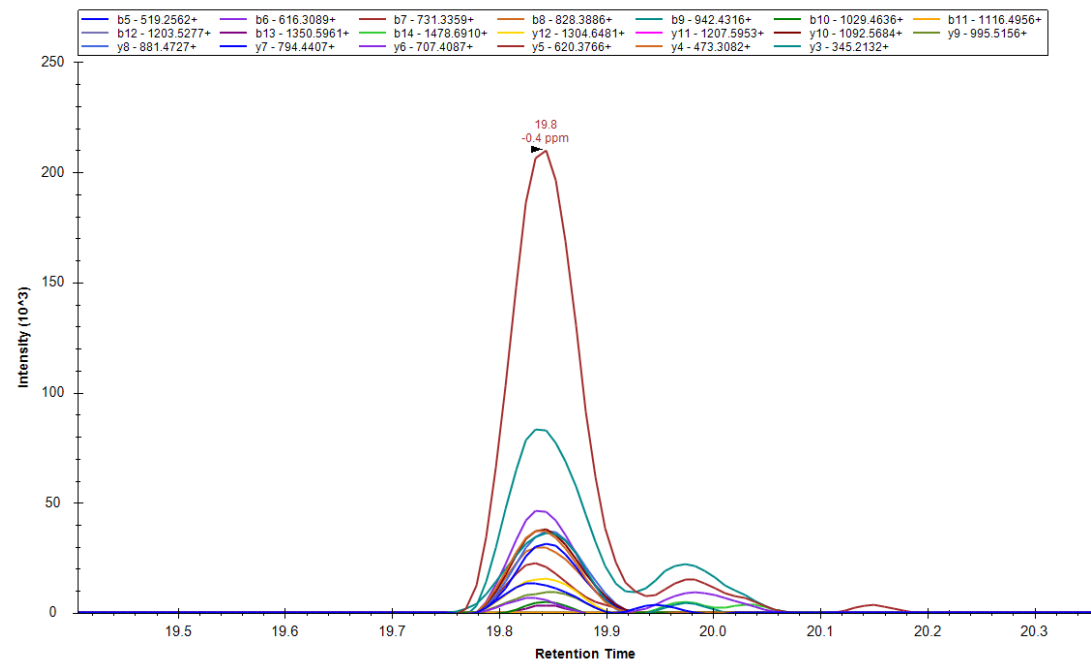
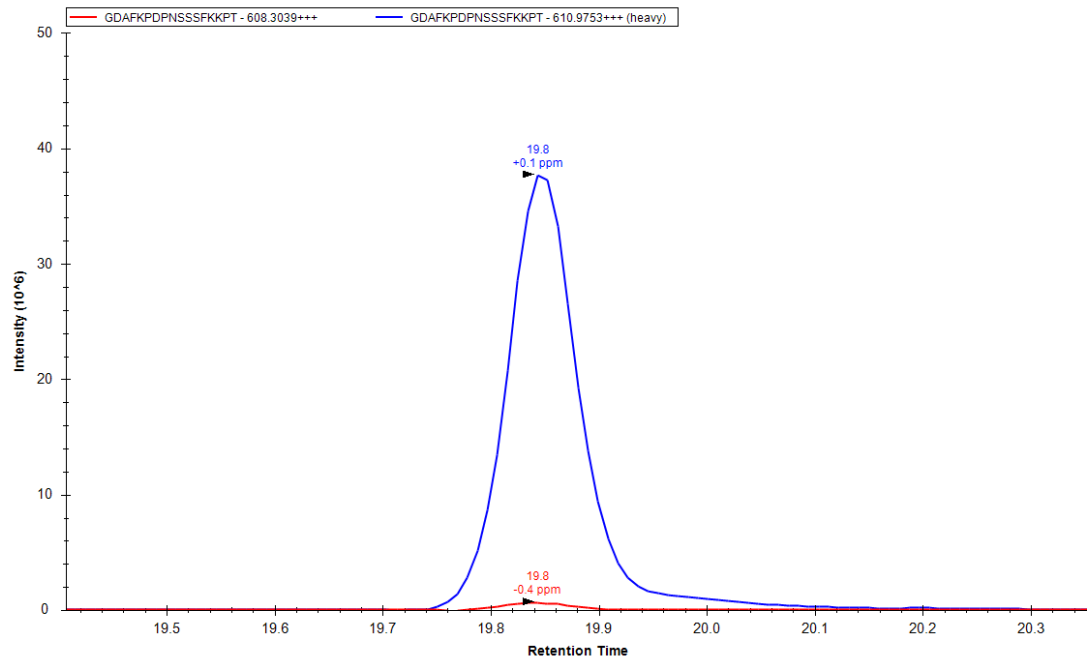


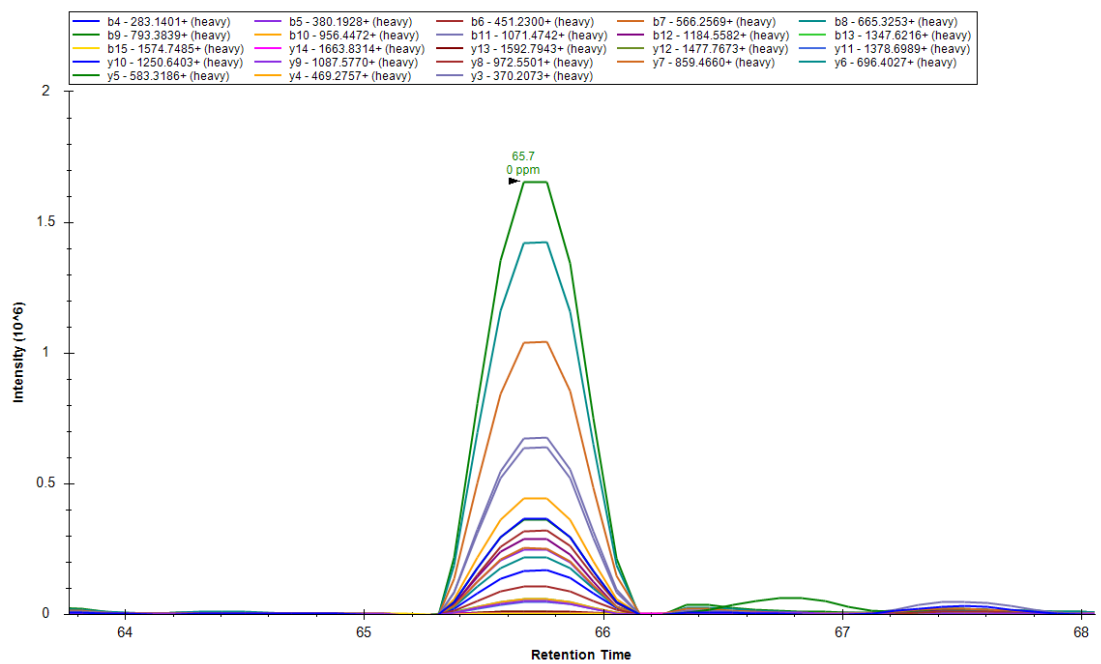
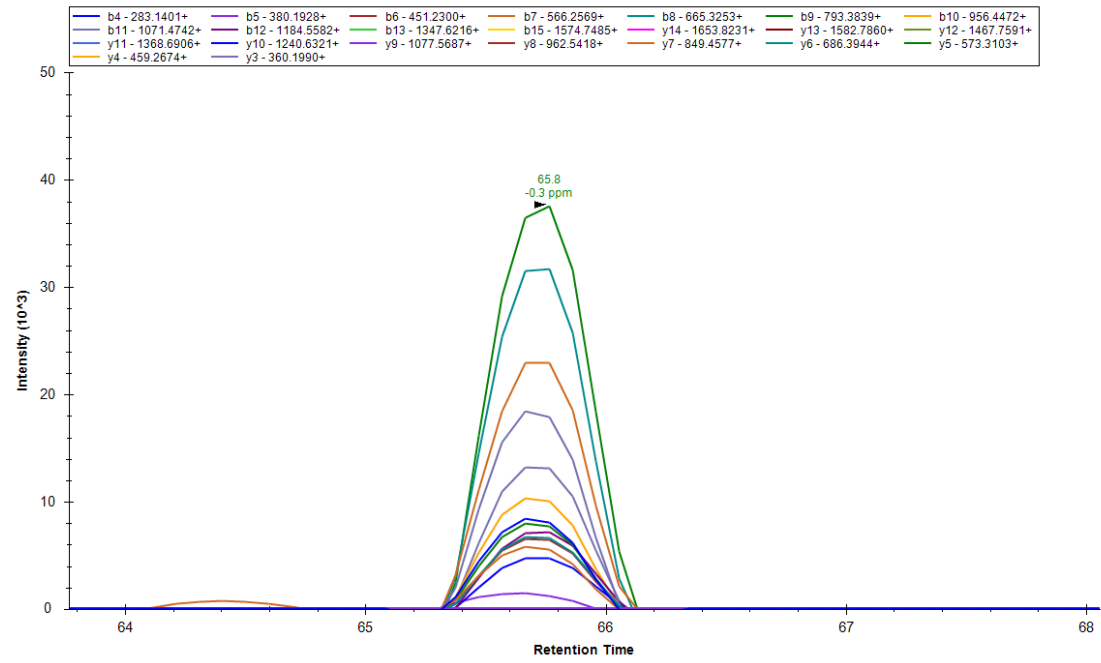
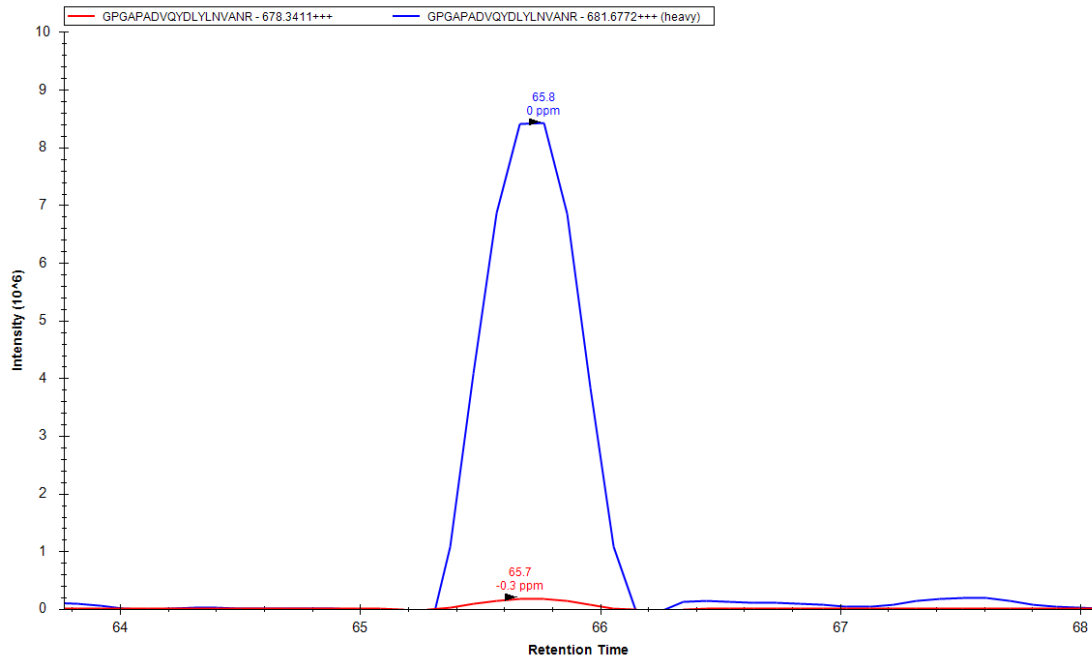


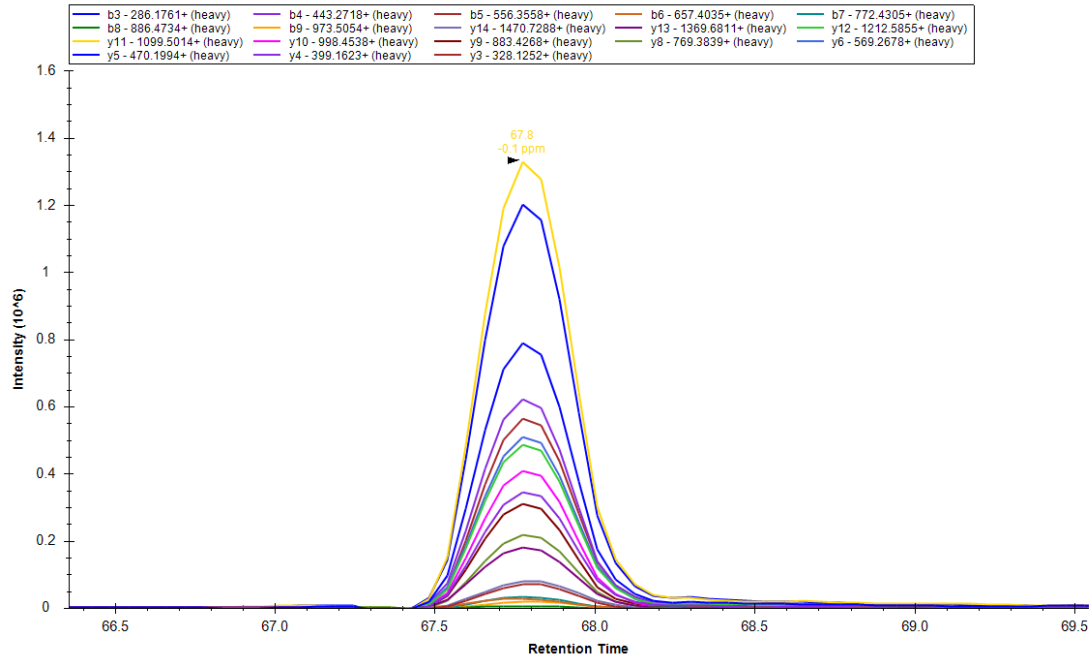
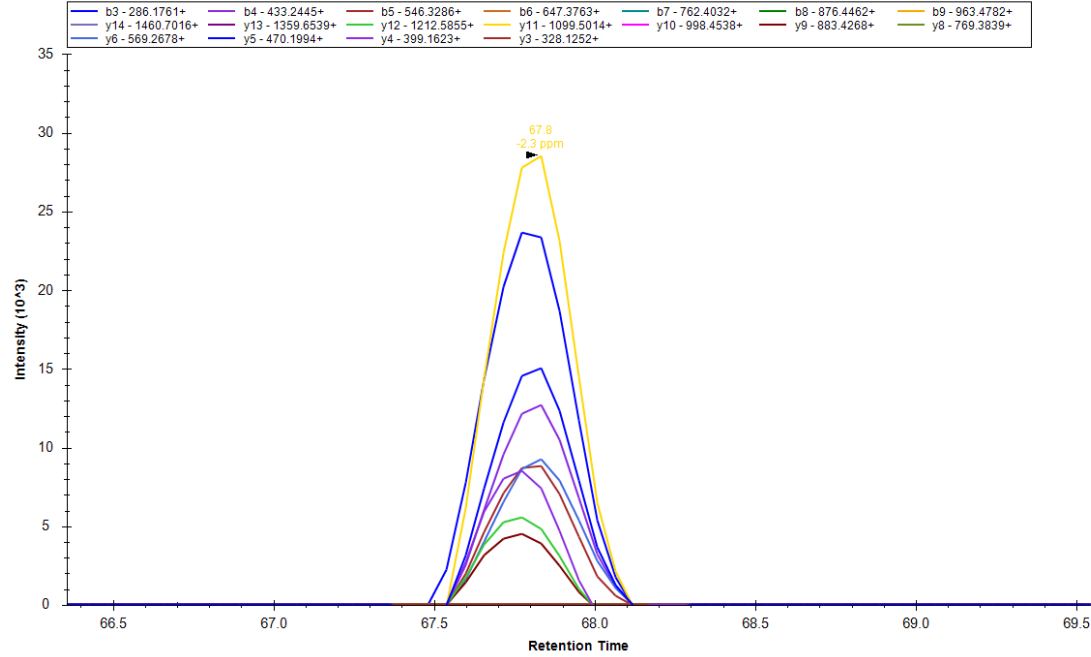
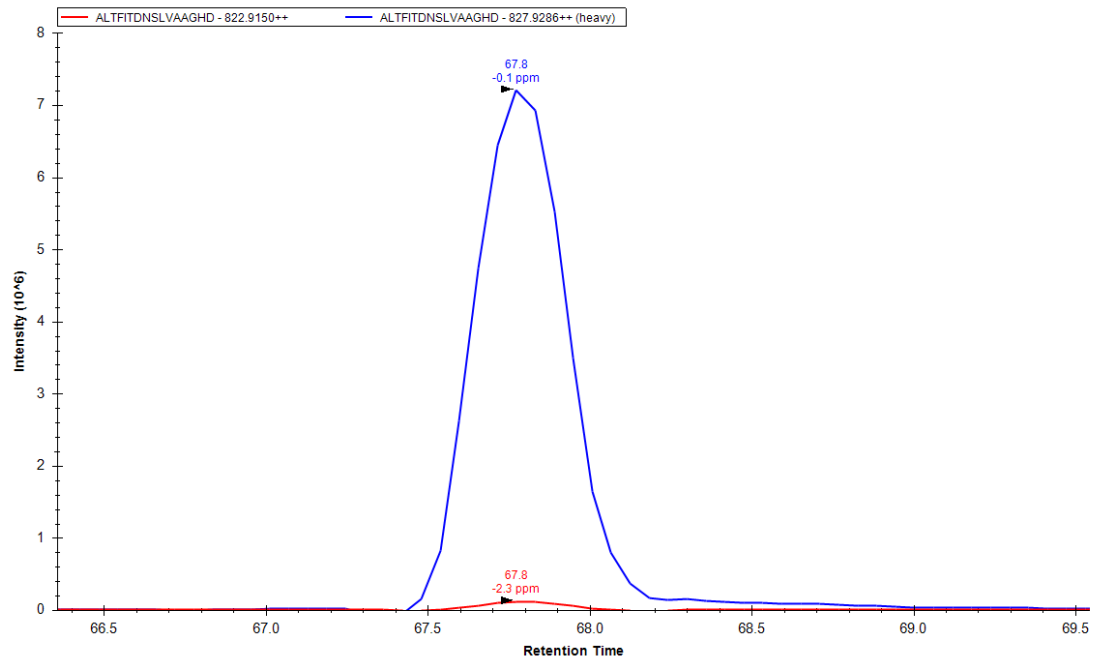


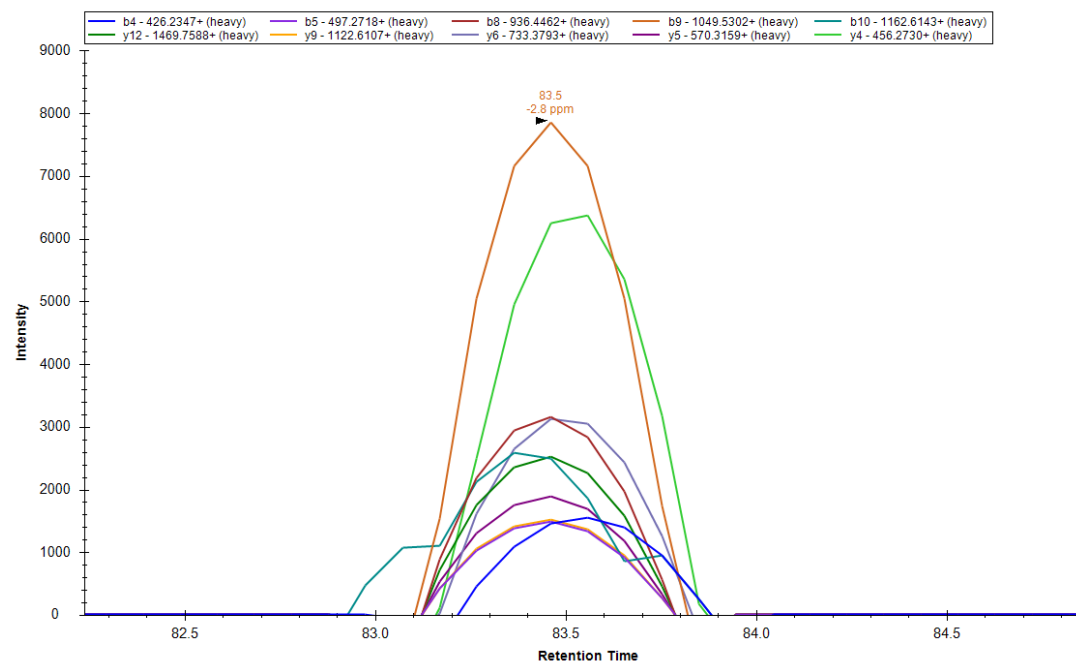
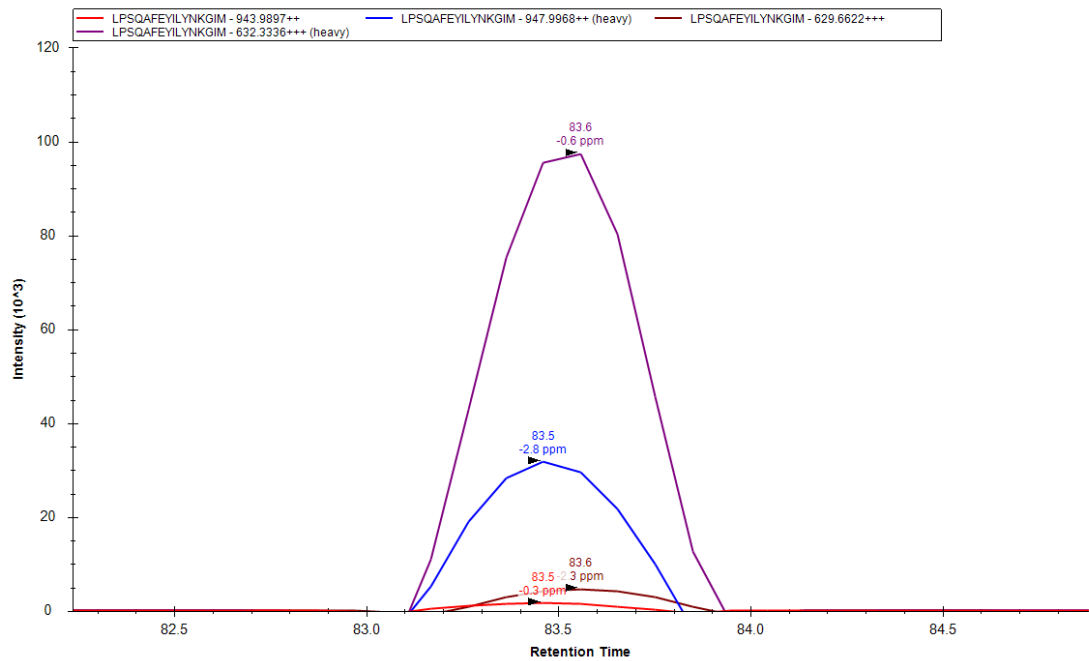






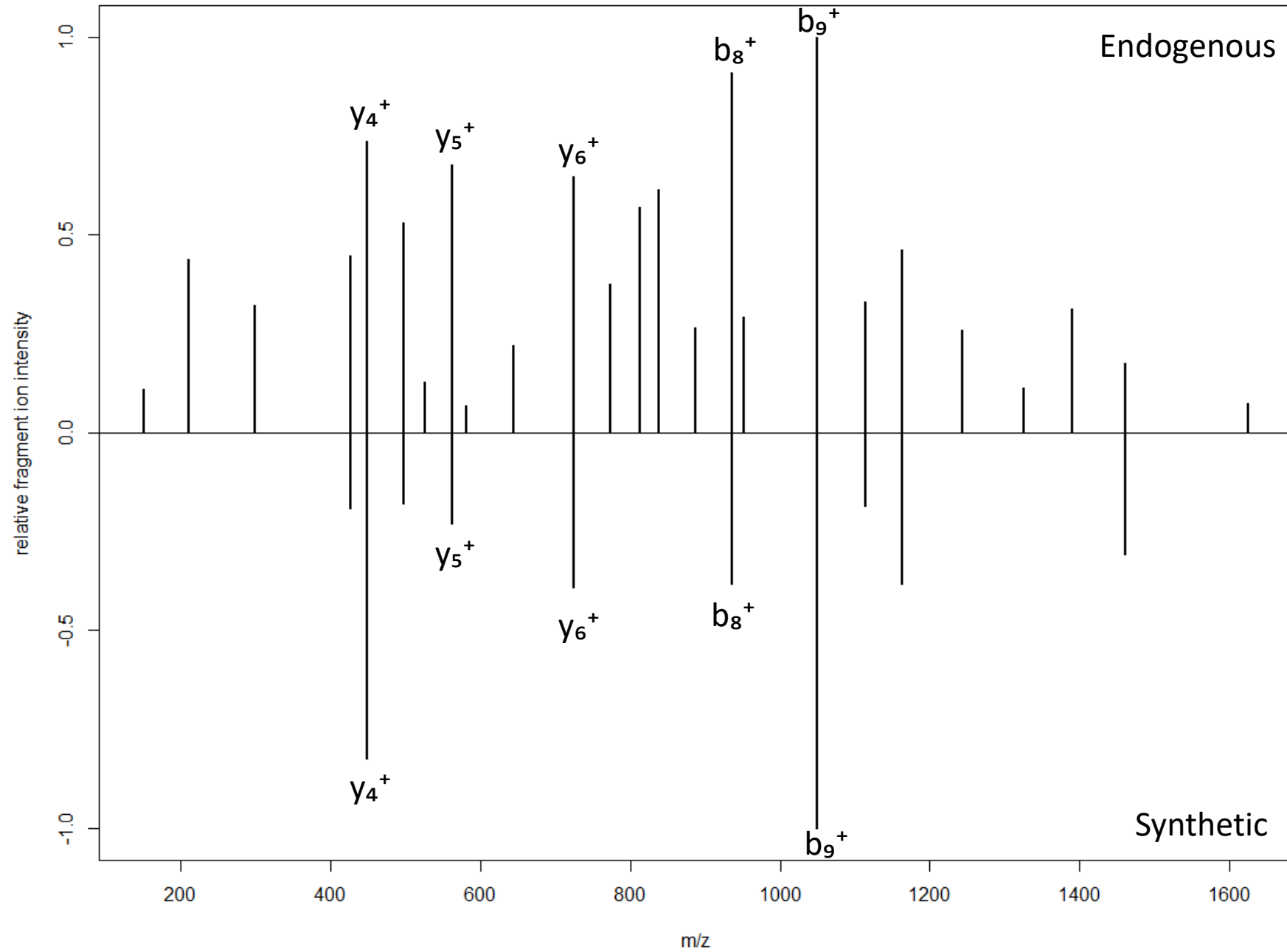


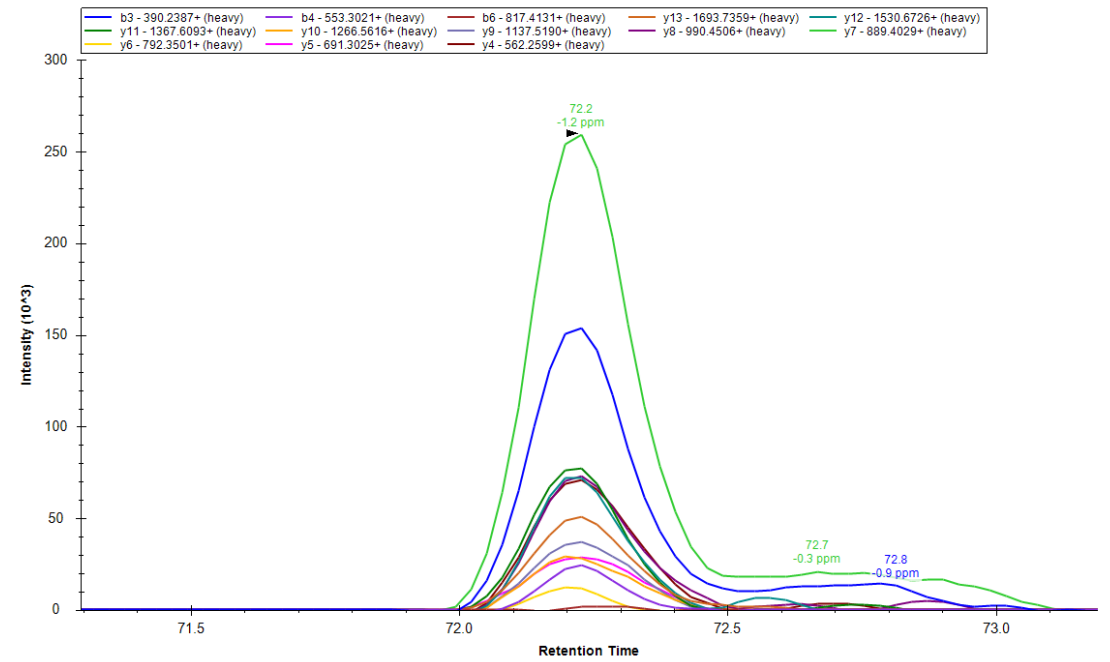
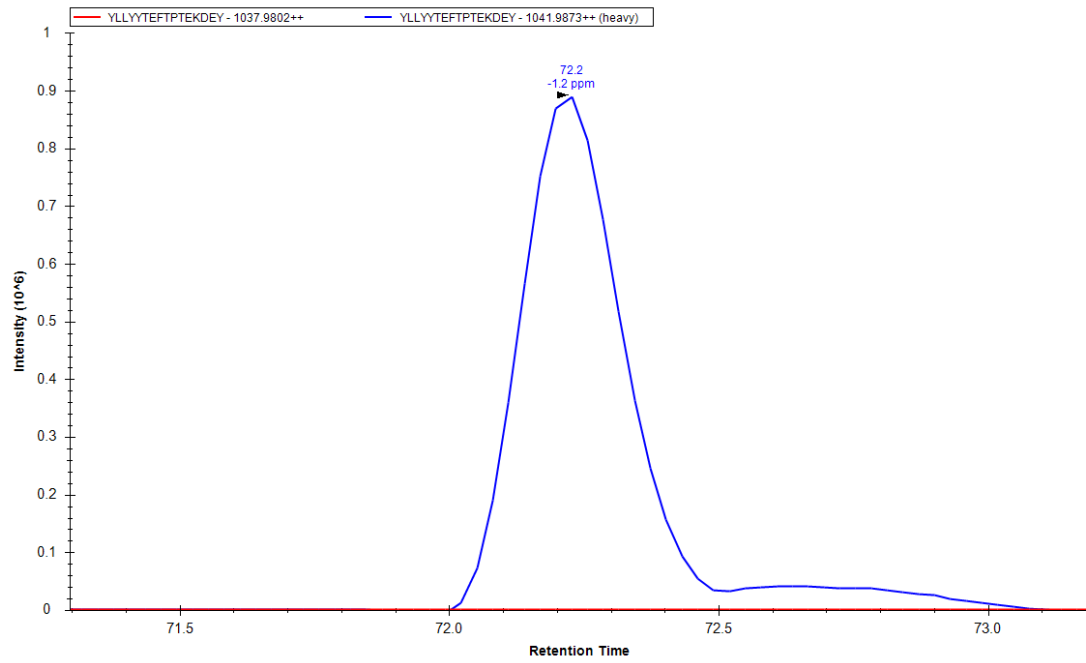




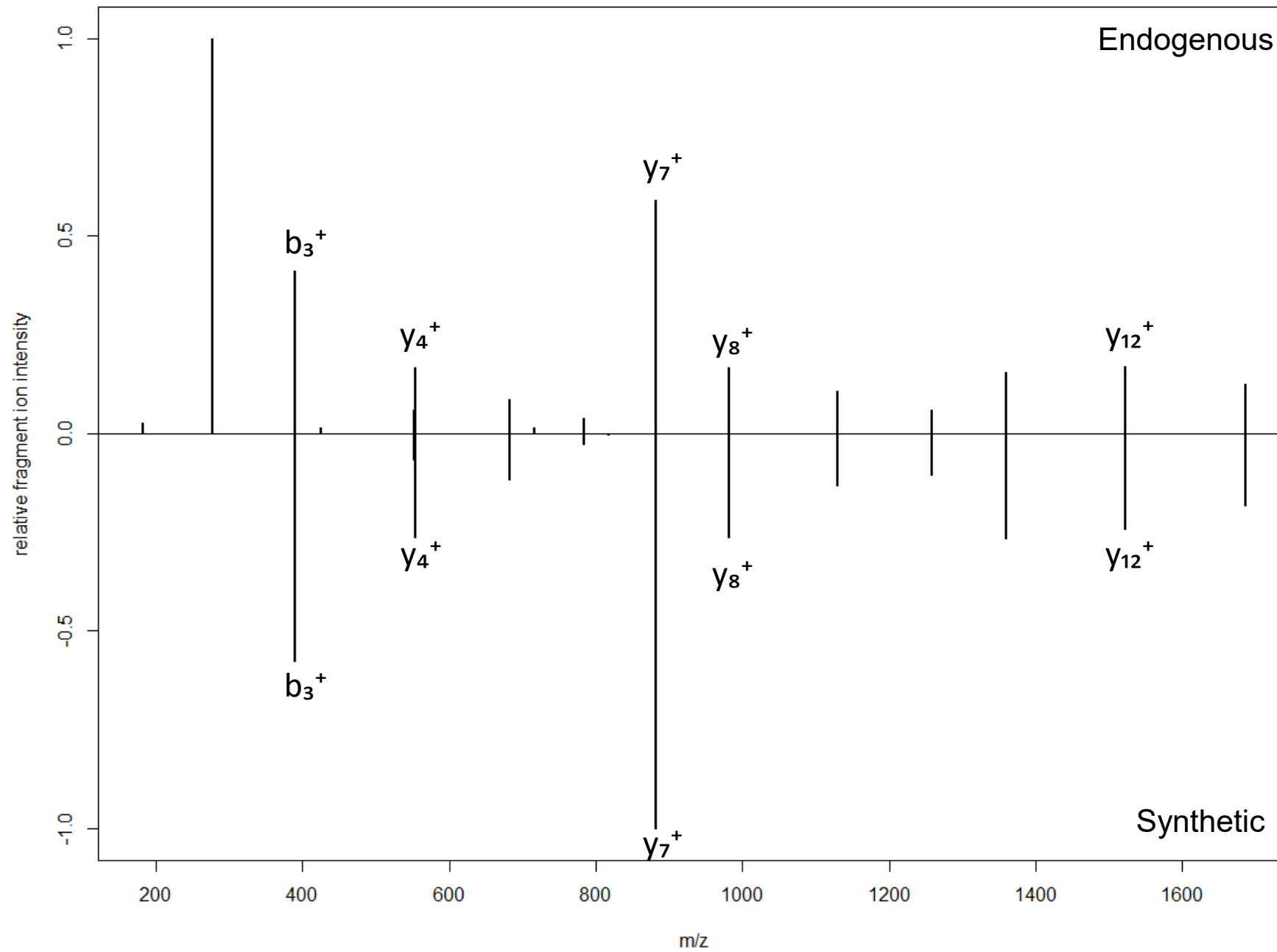


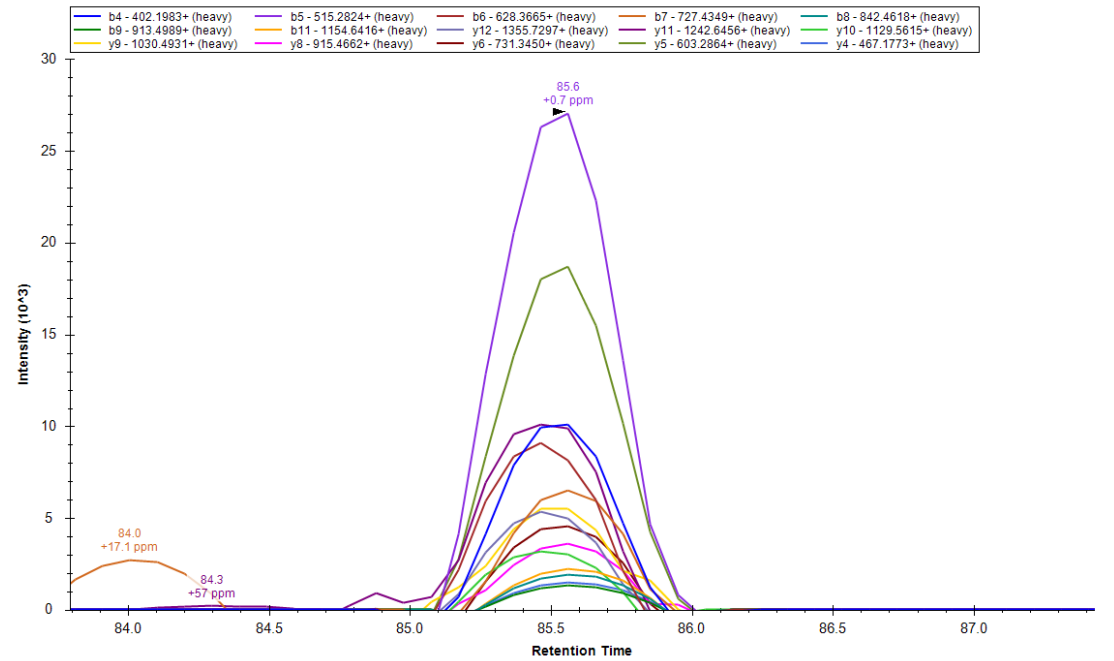
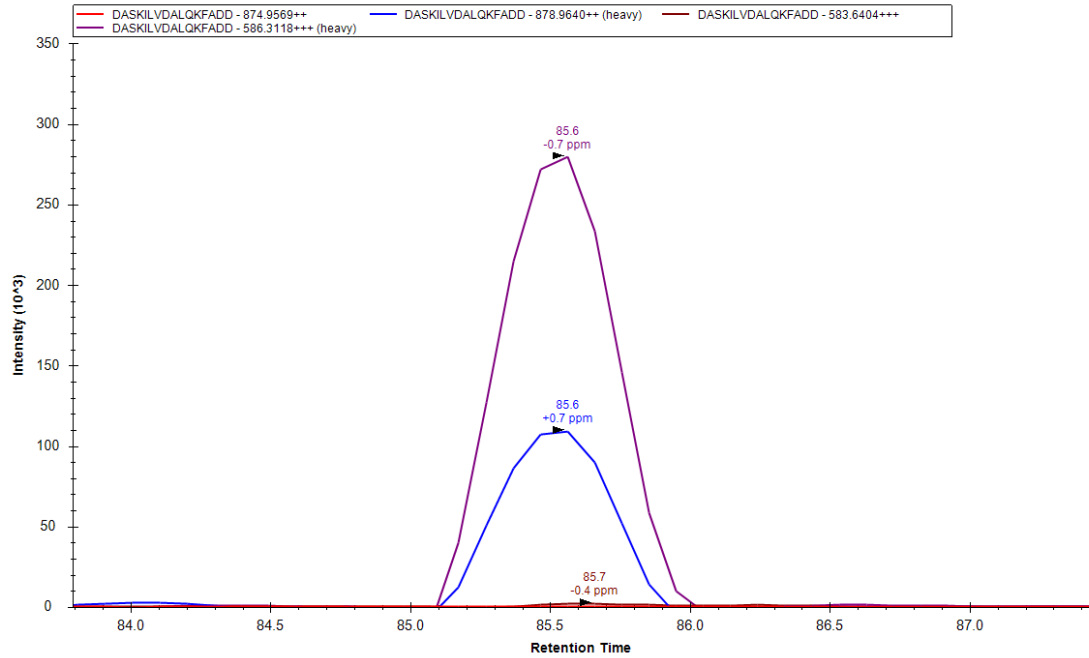
# LPSQAFEYILYNKGIM



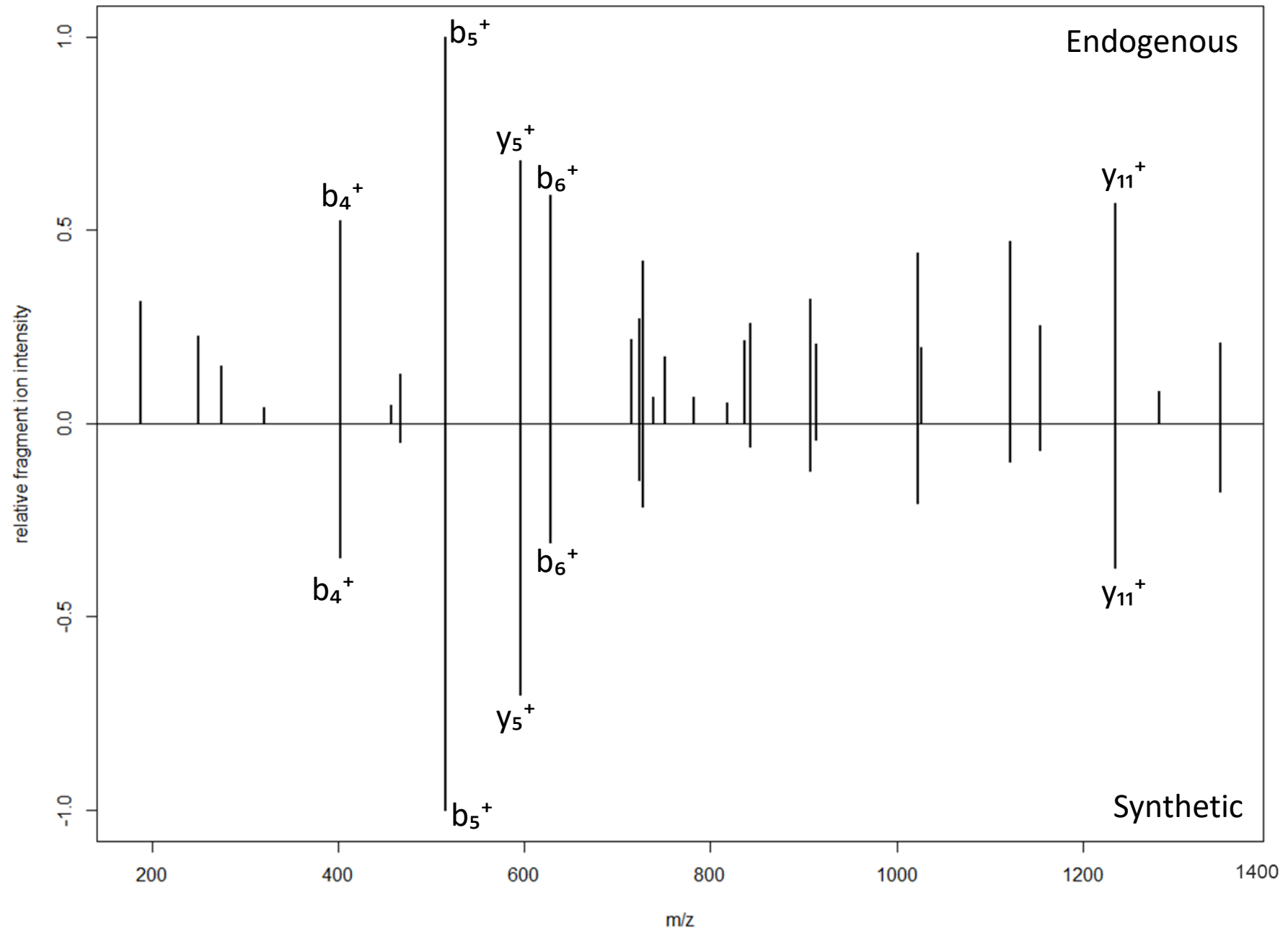


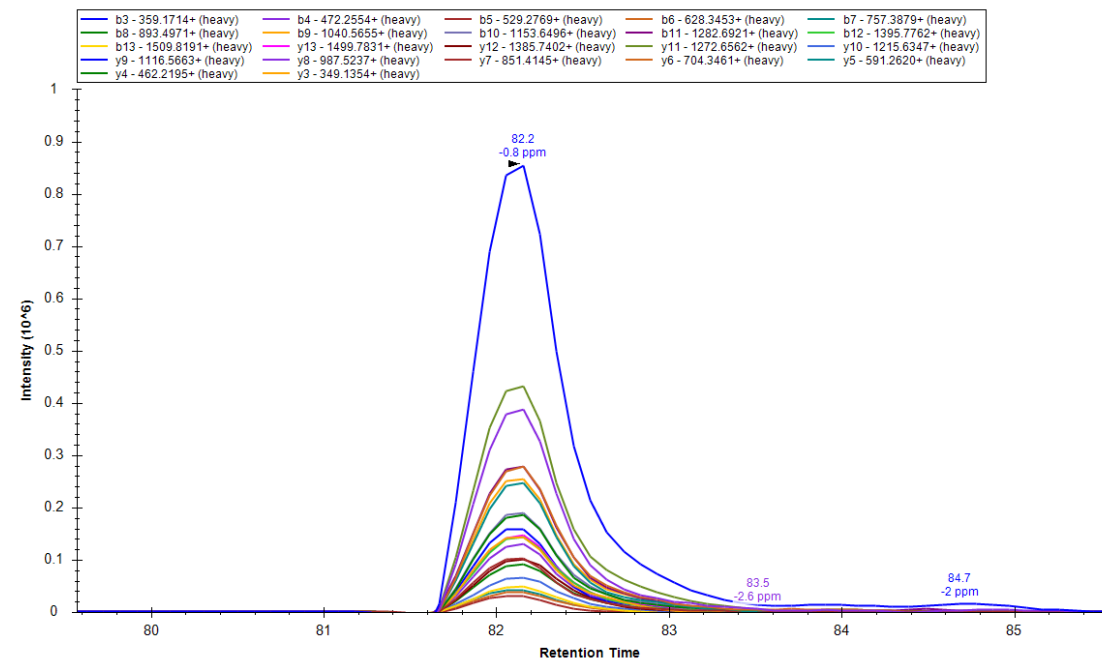
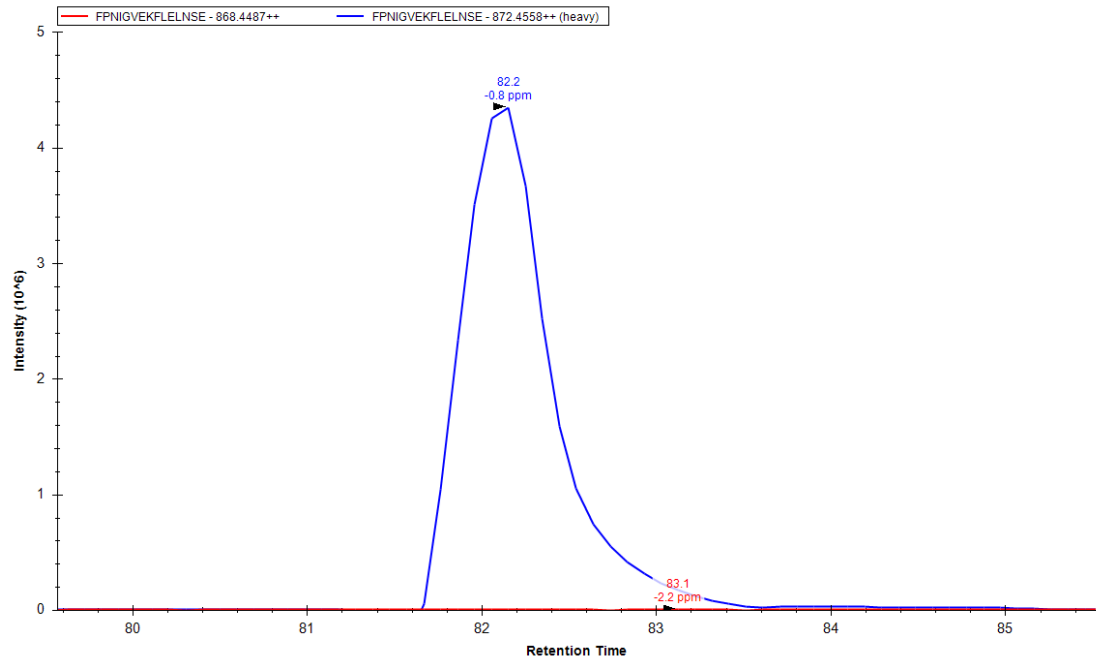
YLLYYTEFTPTEKDEY



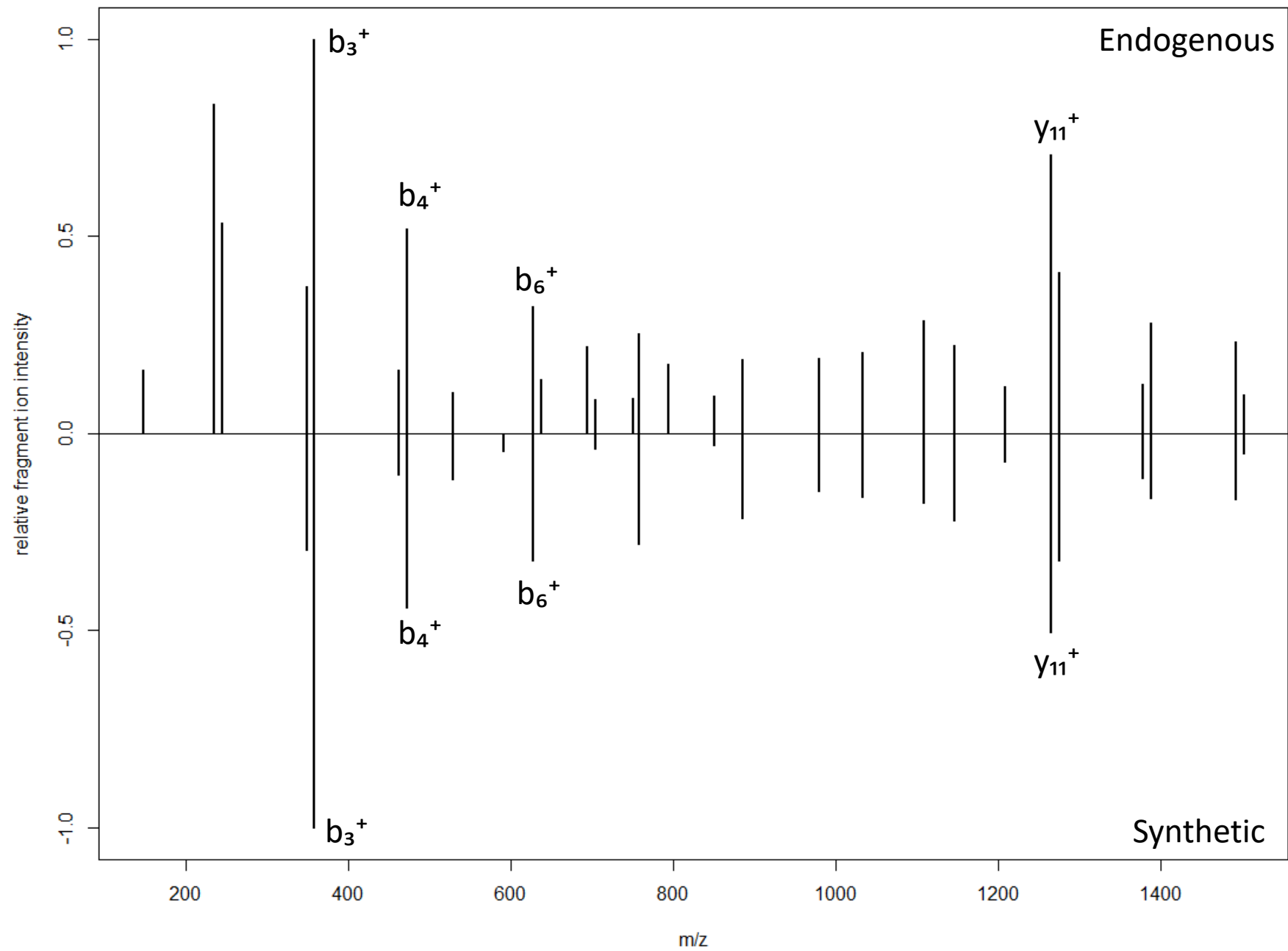


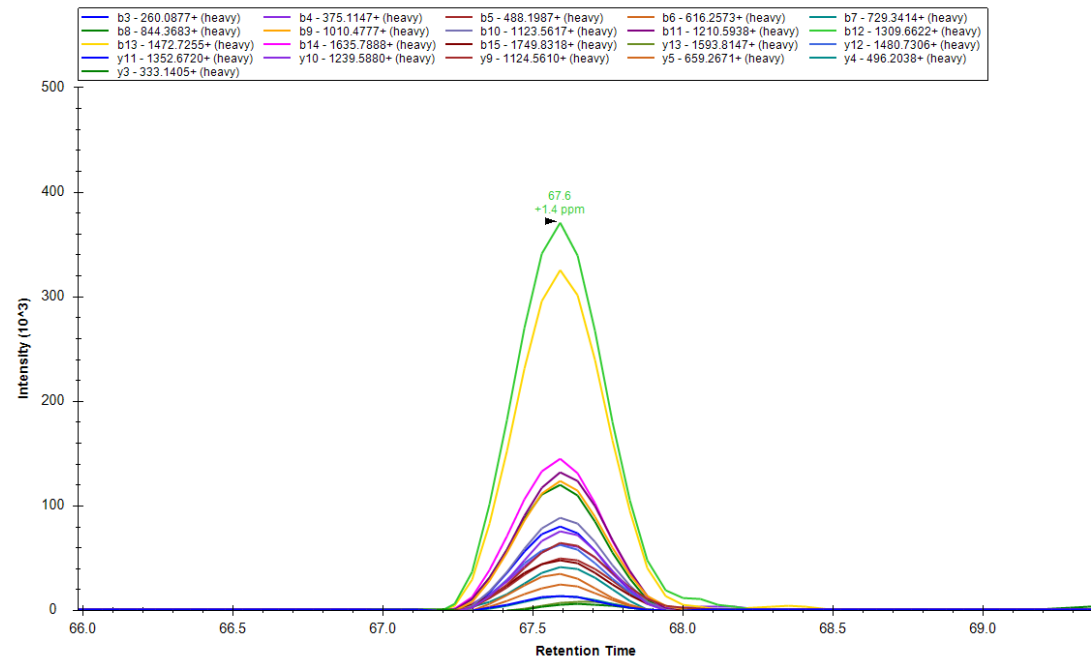
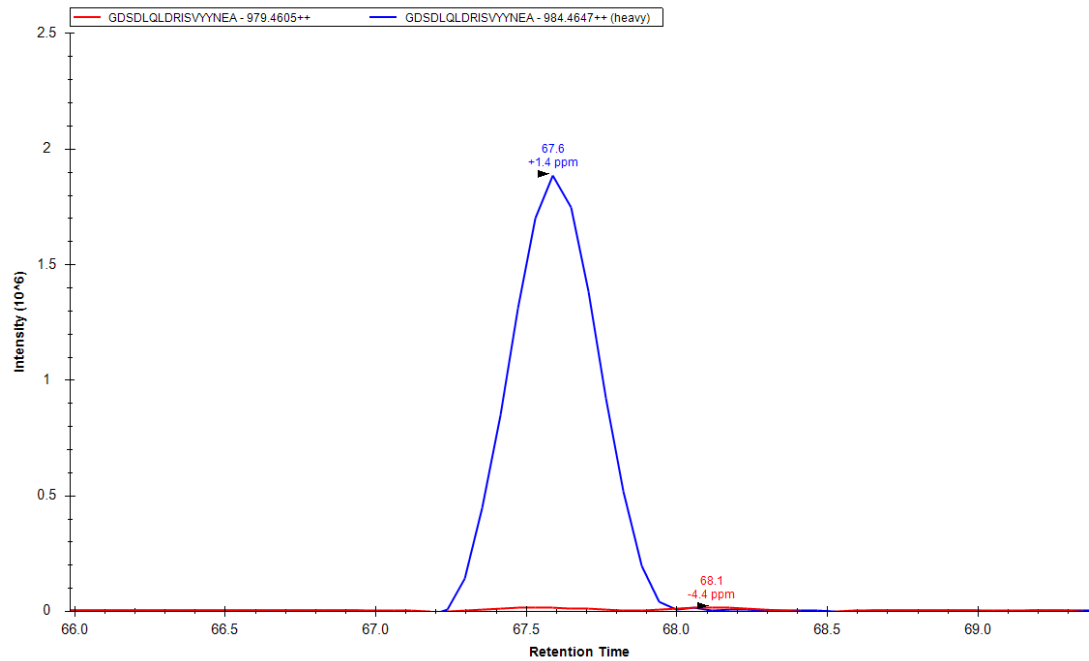
# DASKILVDALQKFADD





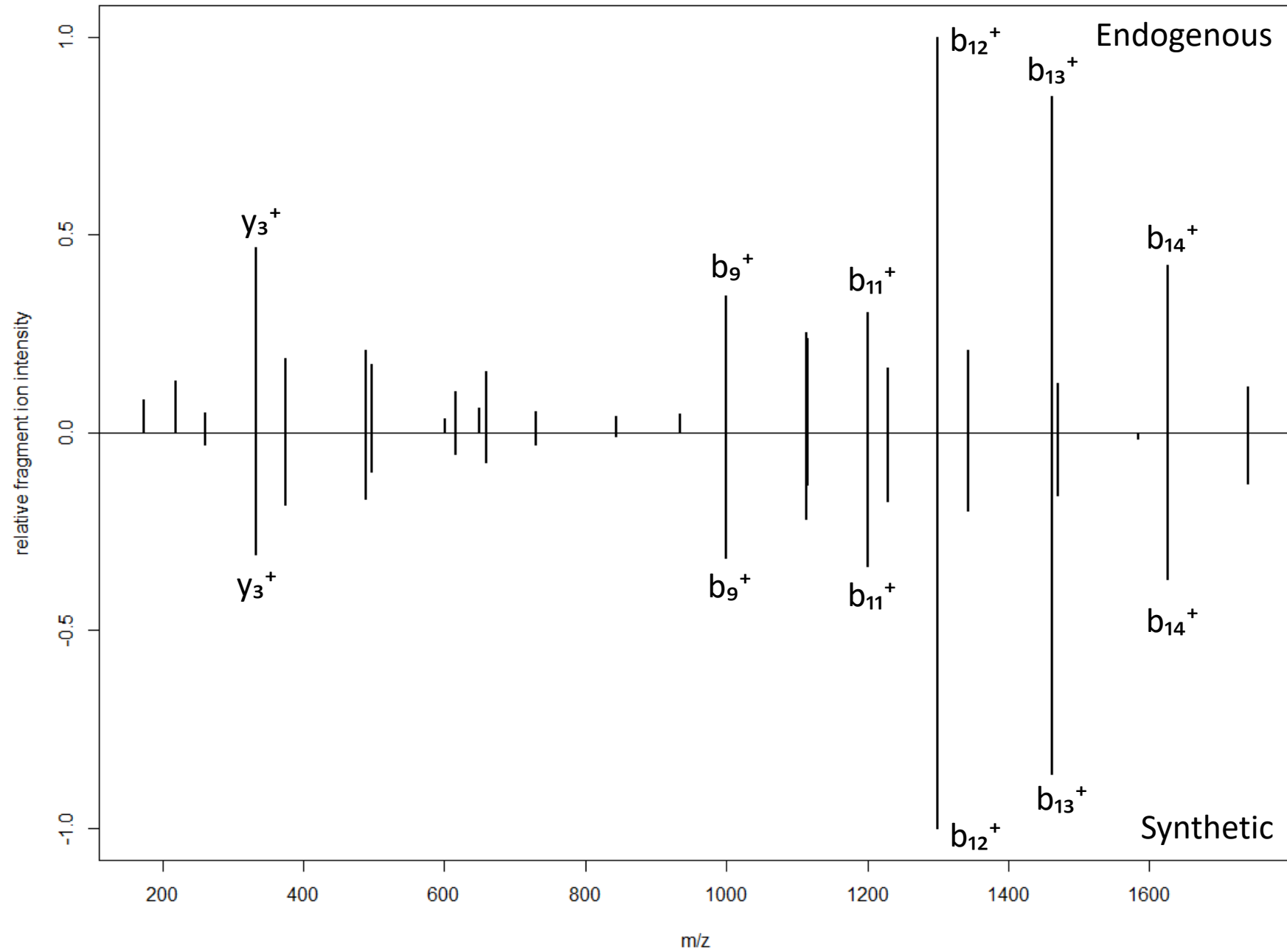
# FPNIGVEKFLELNSE

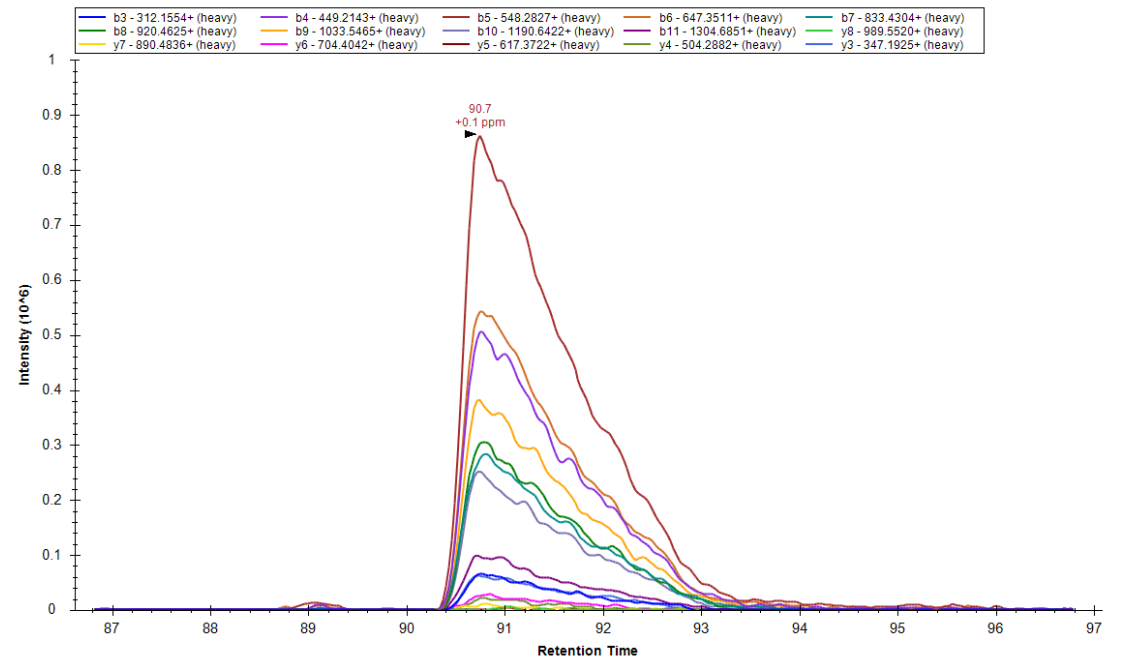
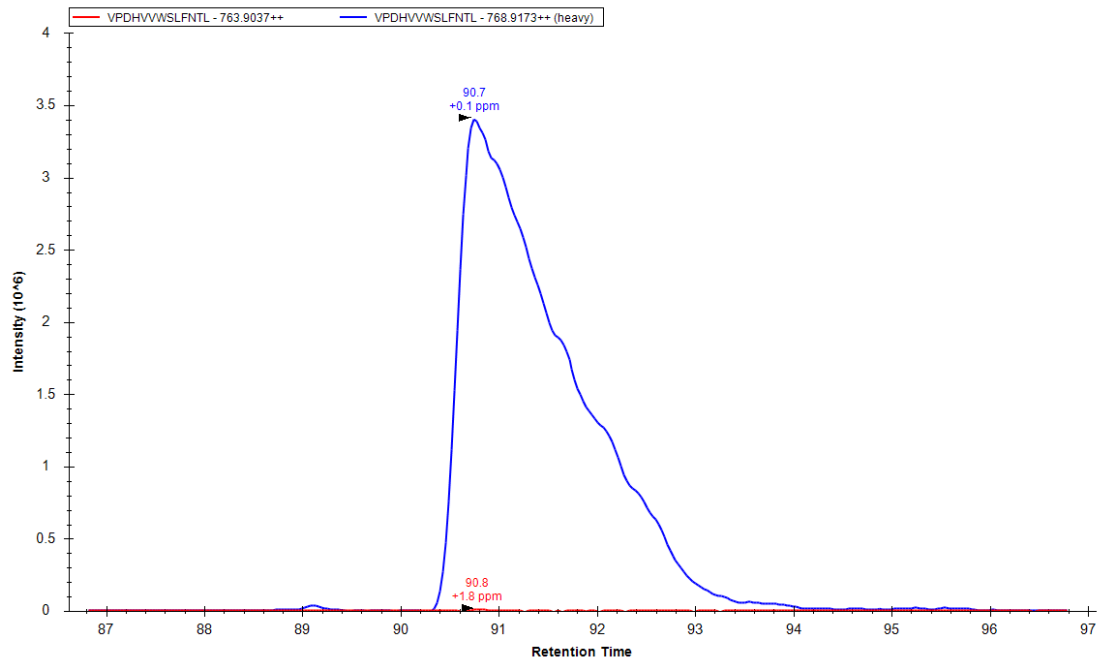




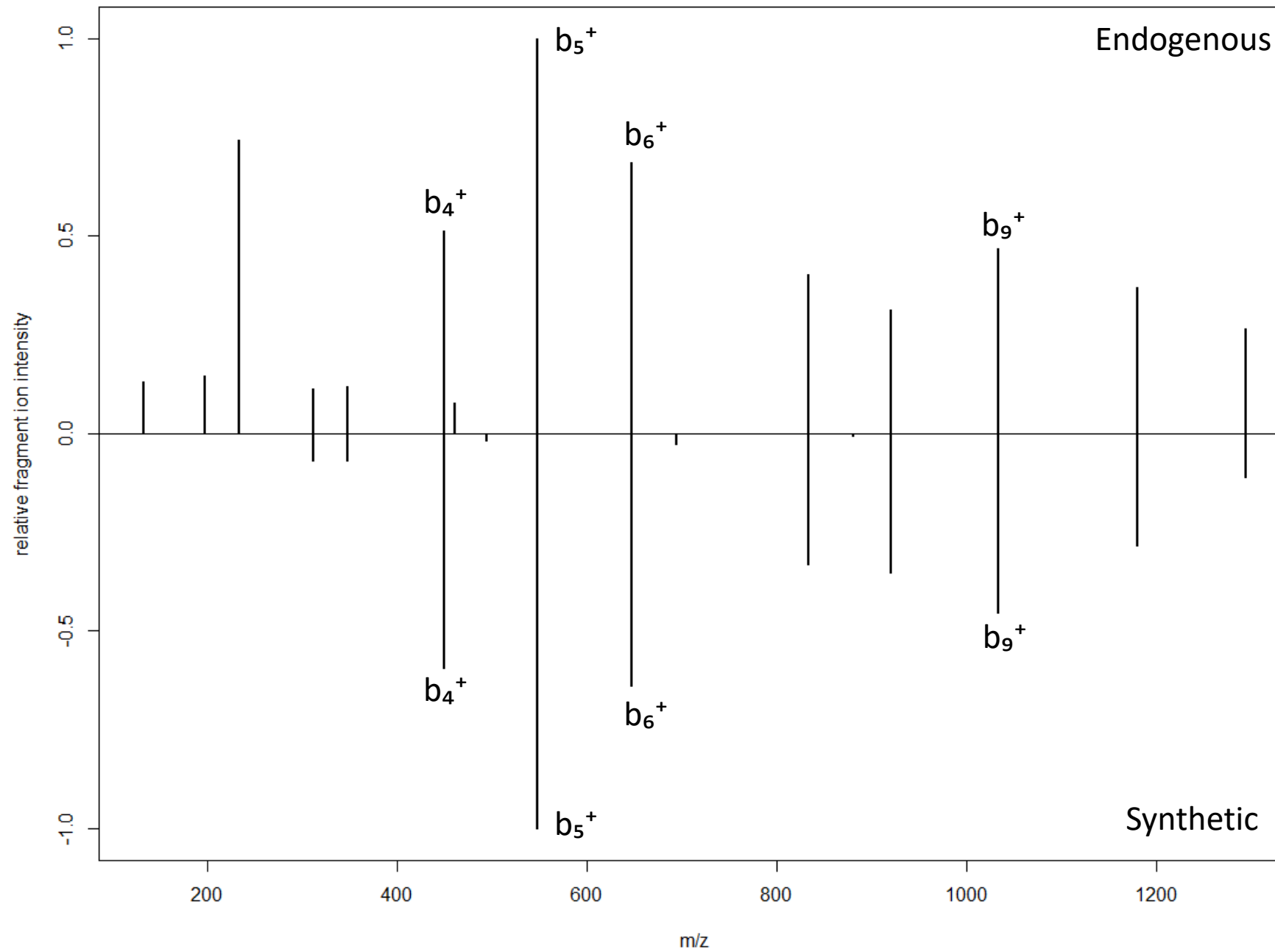


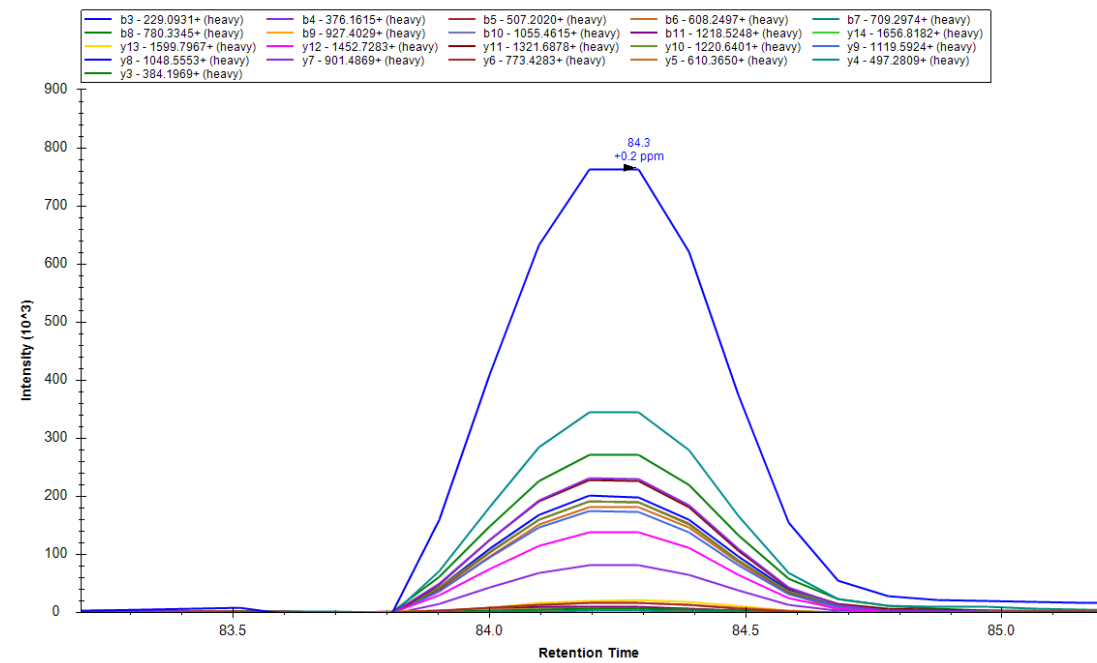
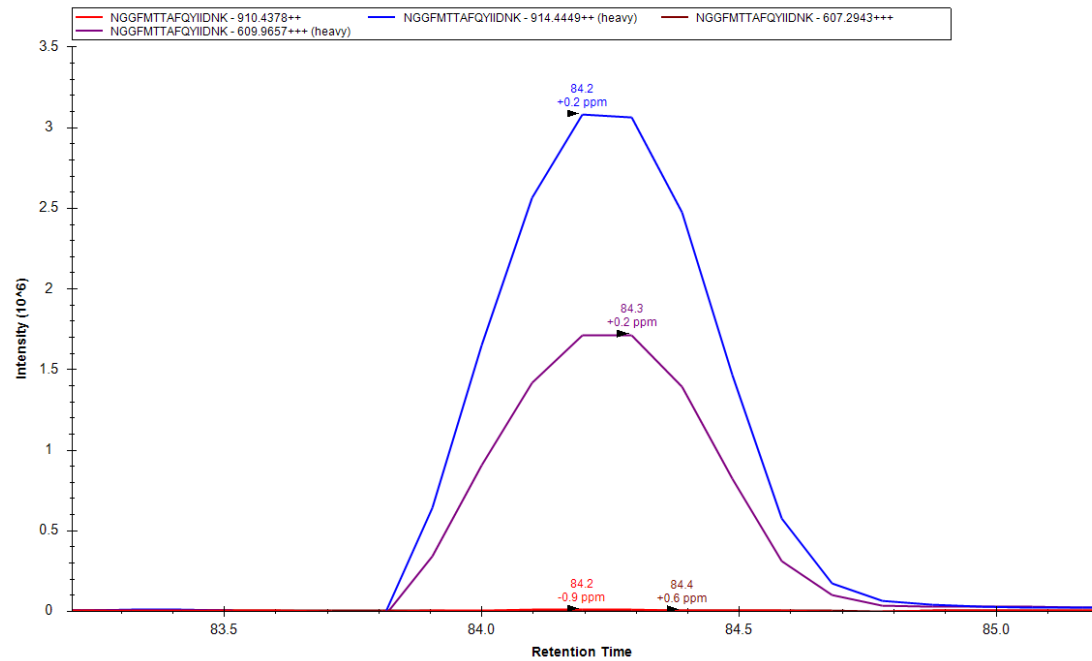
# GDSLQLDRISVYYNEA



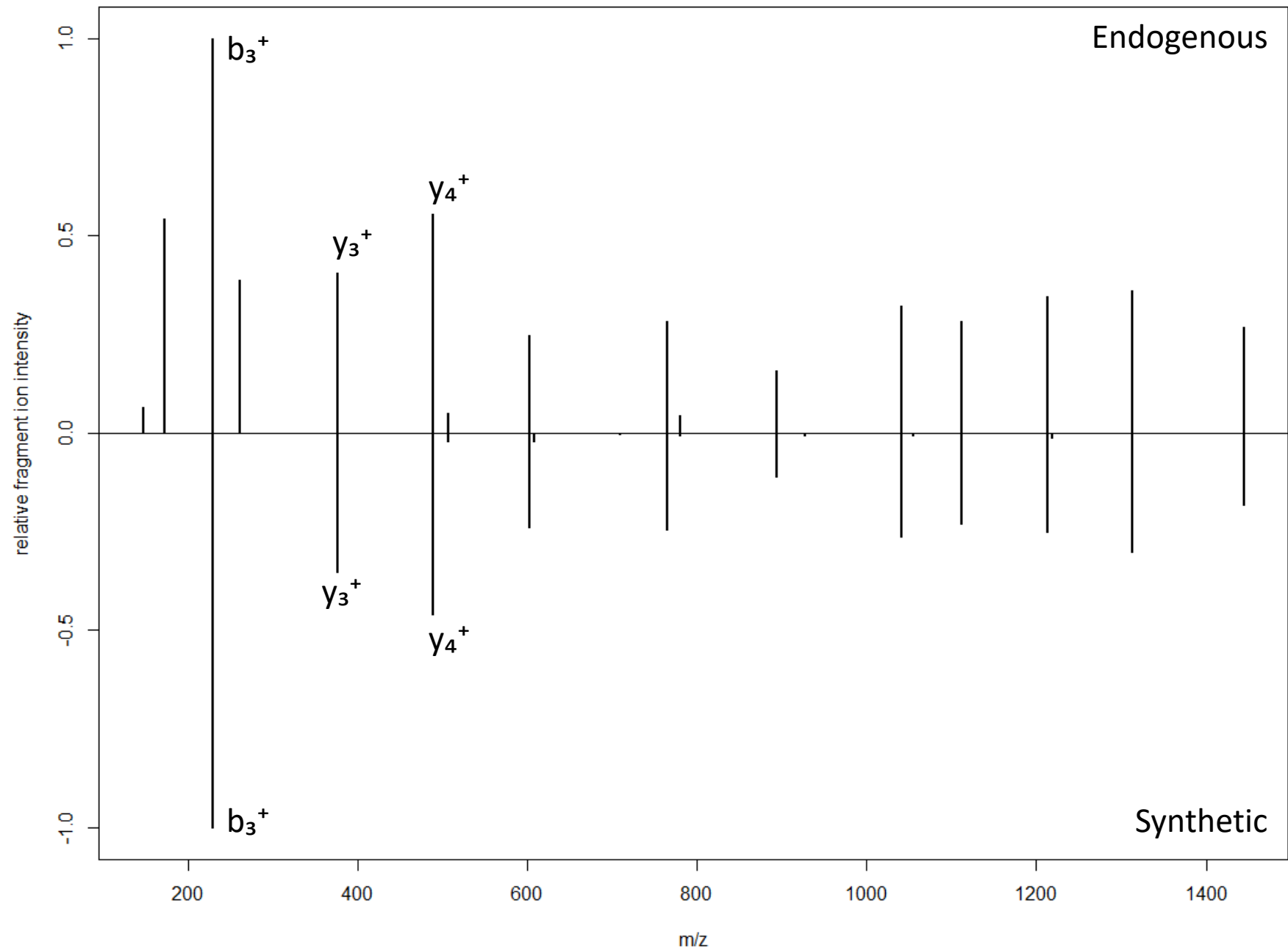


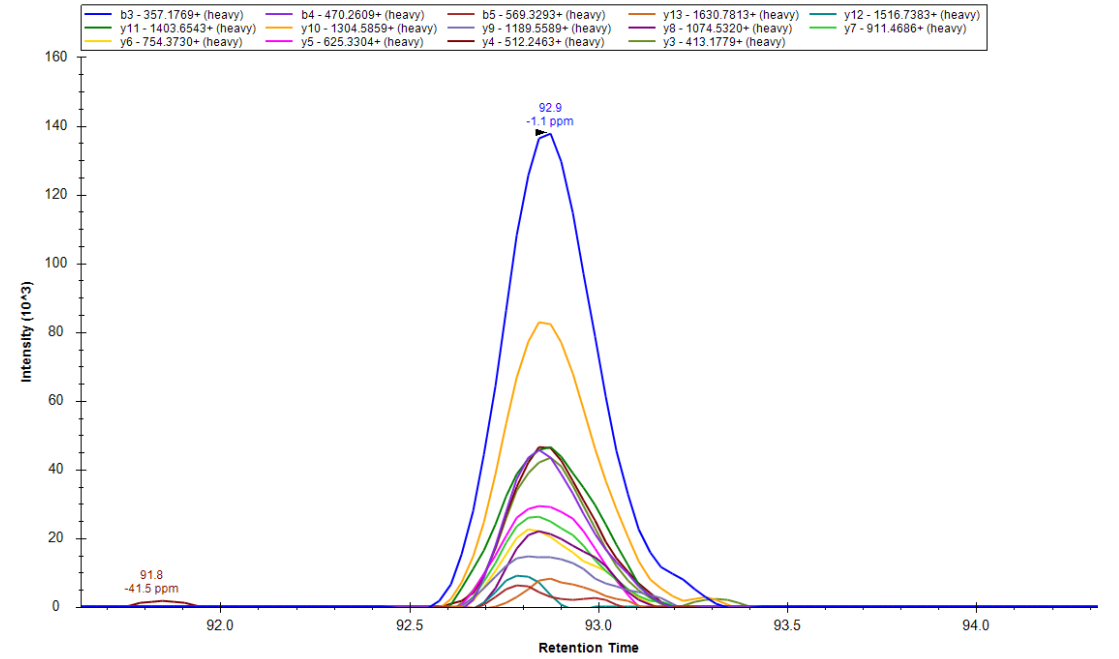
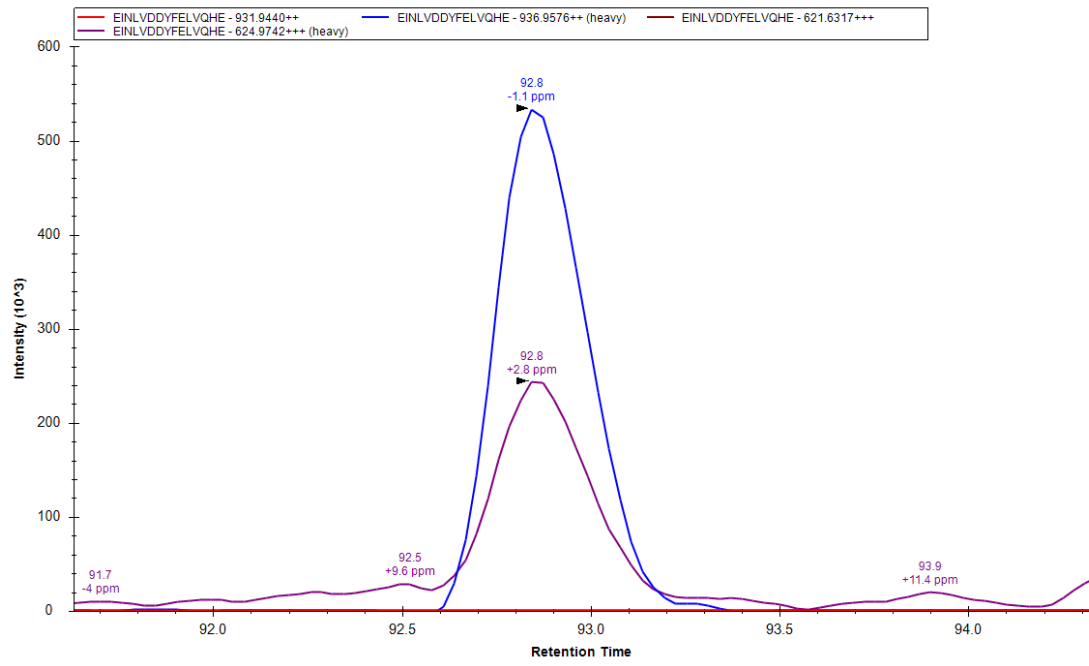
# VPDHVVWSLFNTL



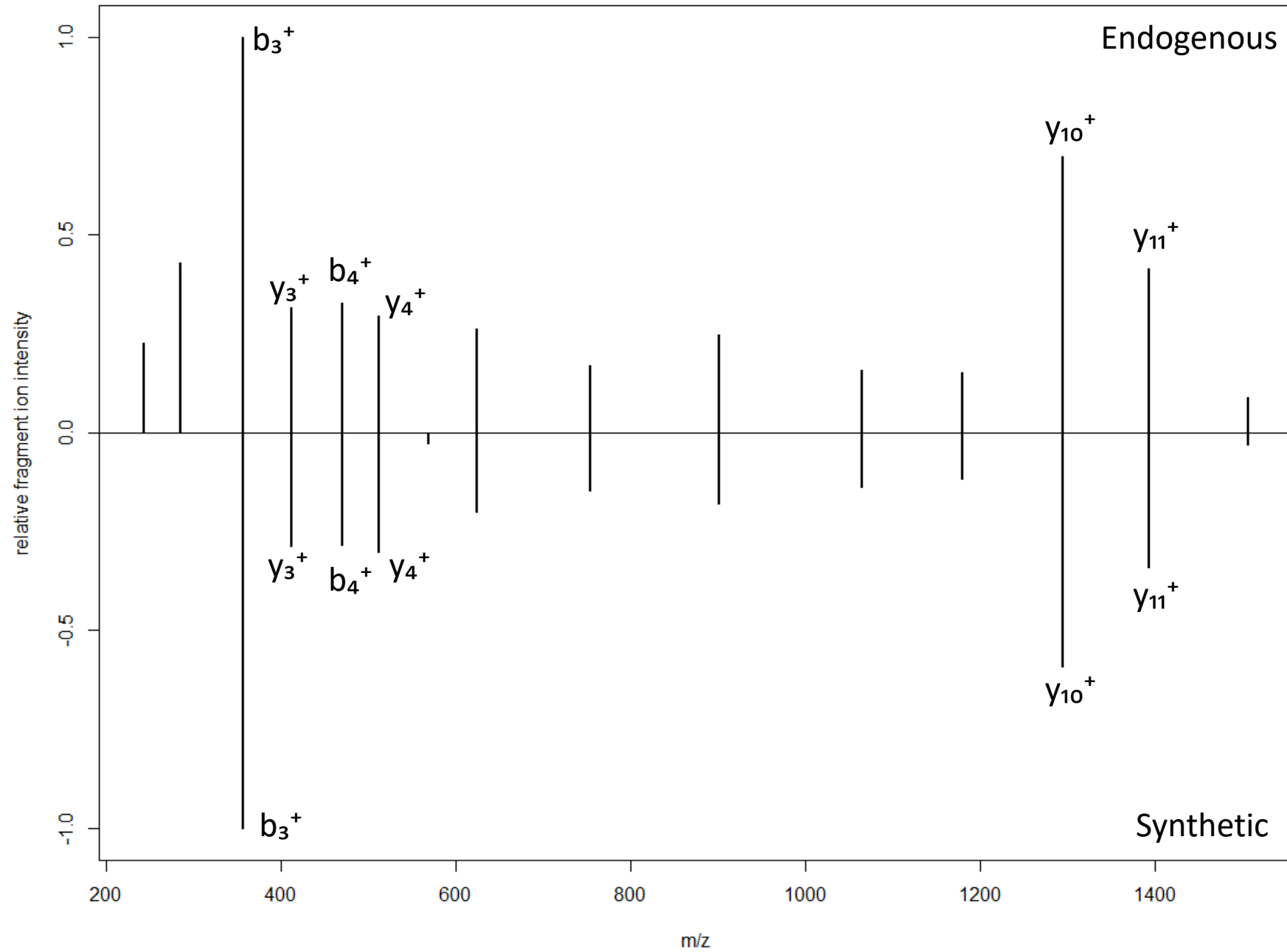


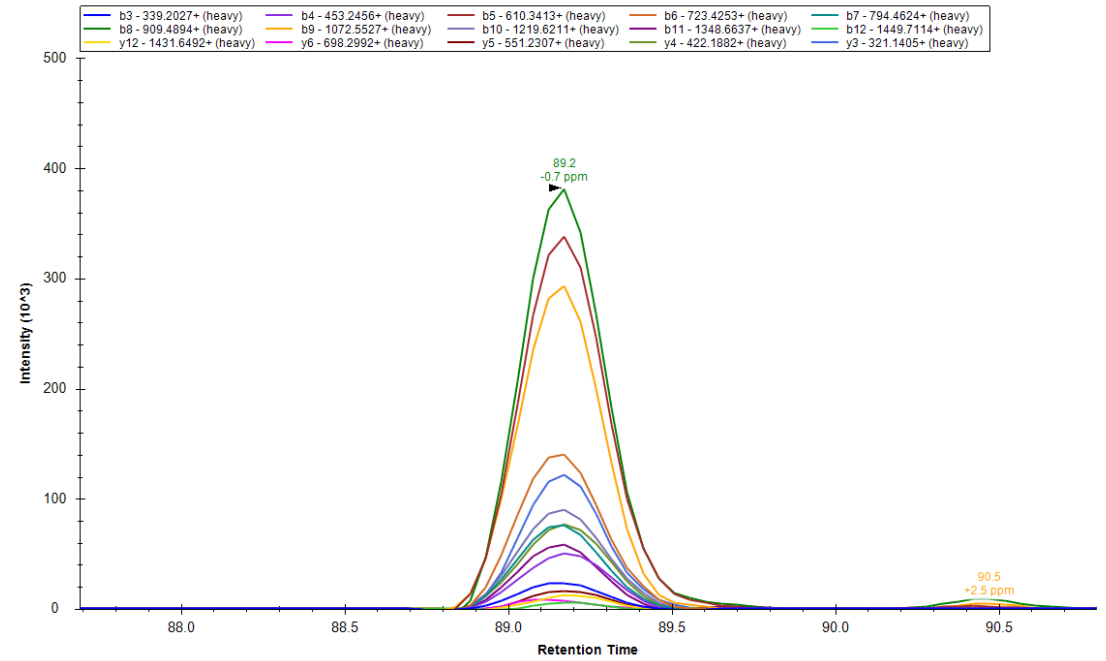
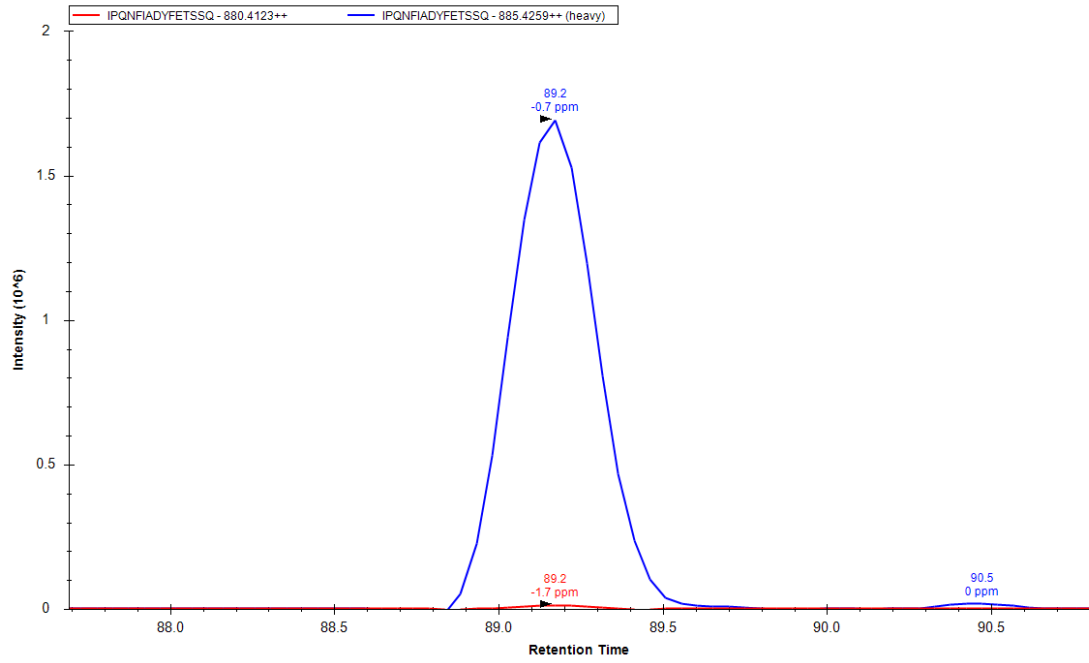
NGGFMTTAFQYIIDNK





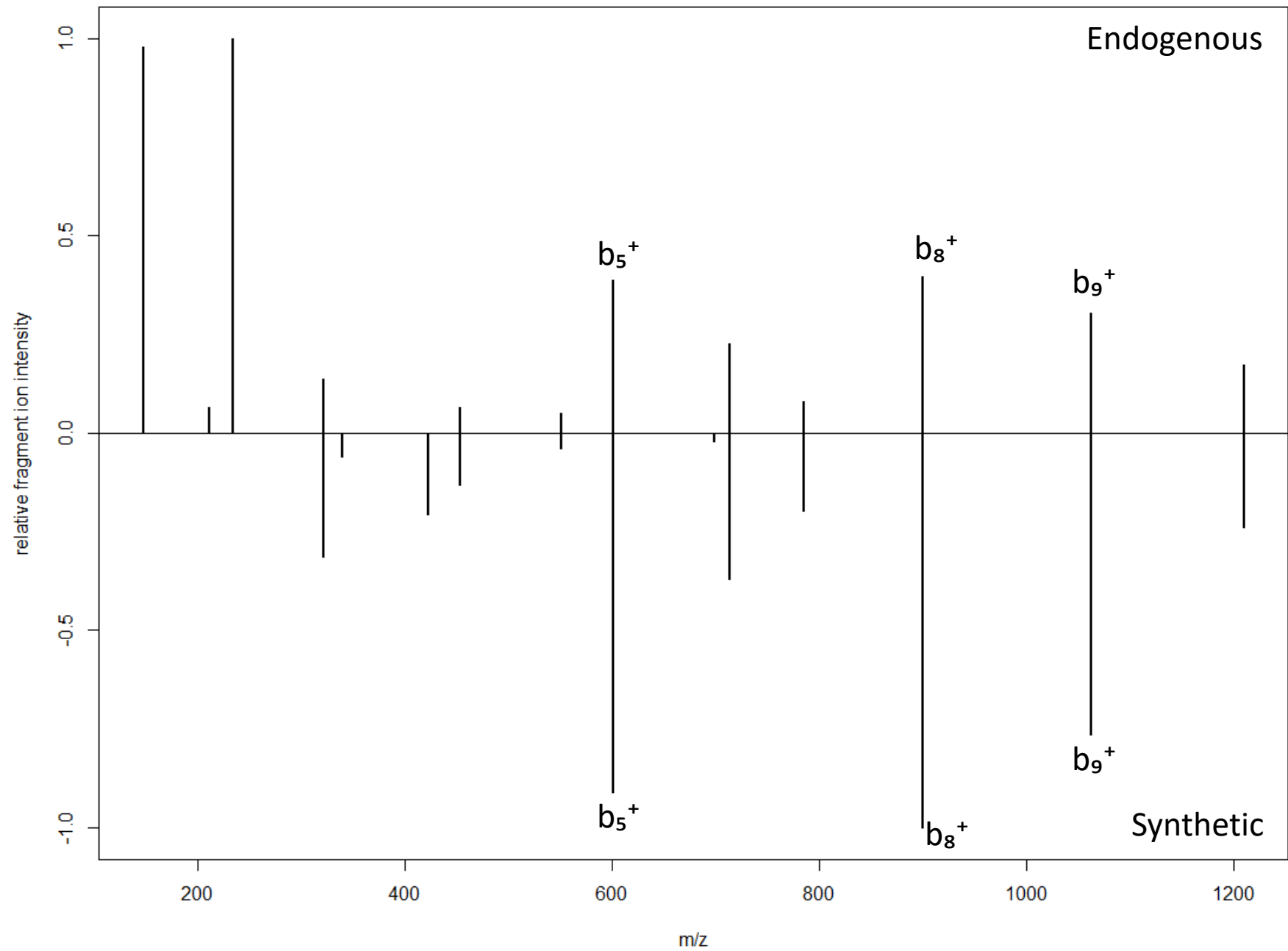
EINLVDDYFELVQHE

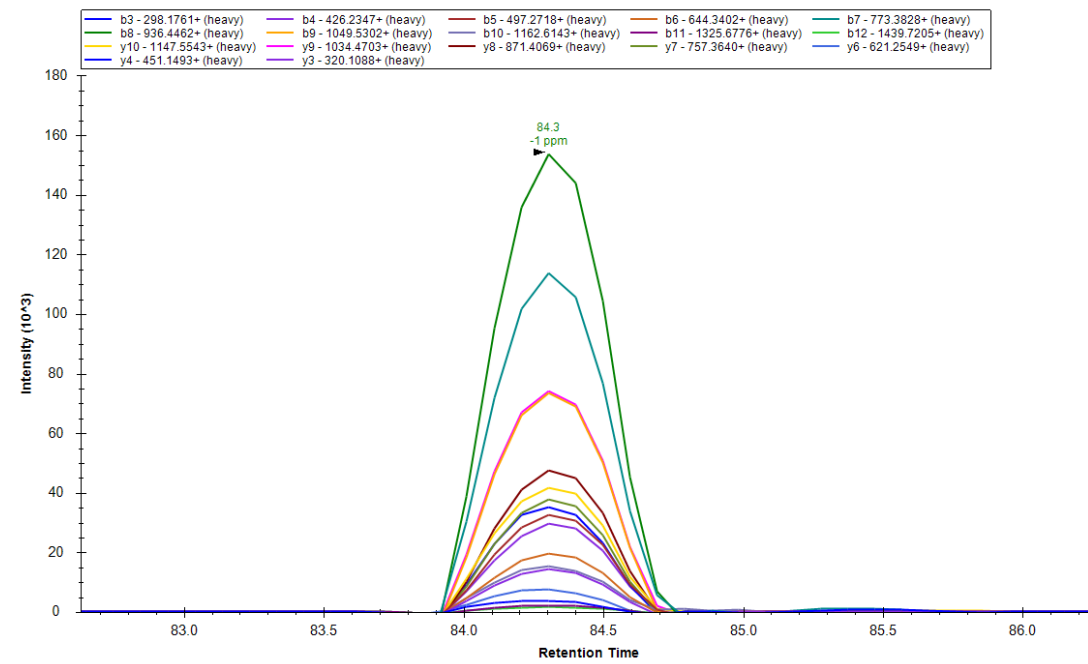
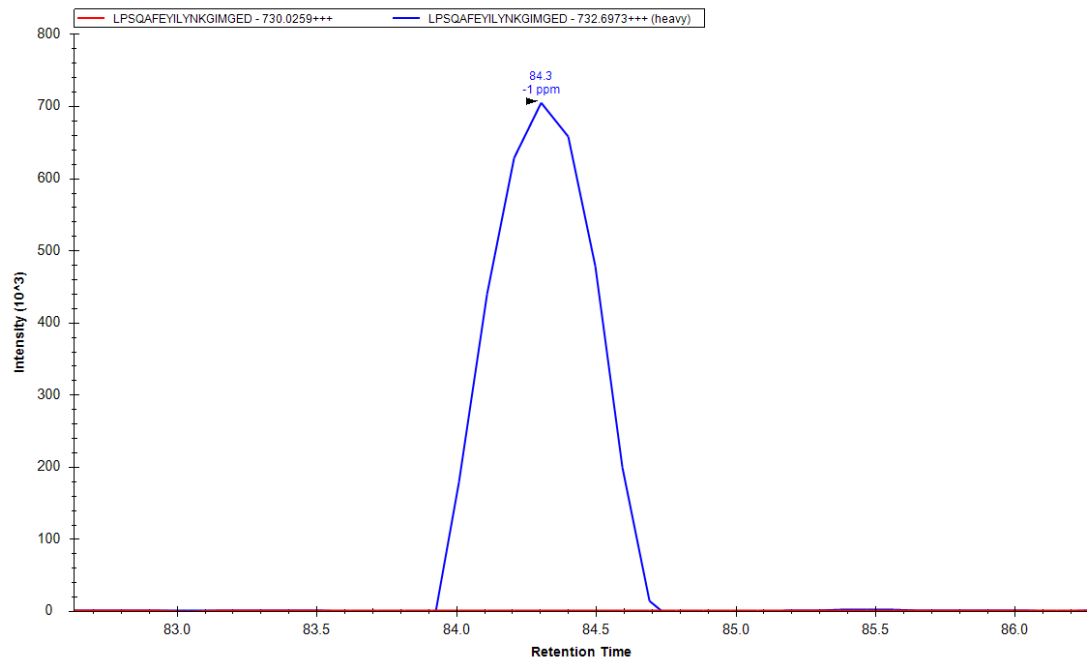




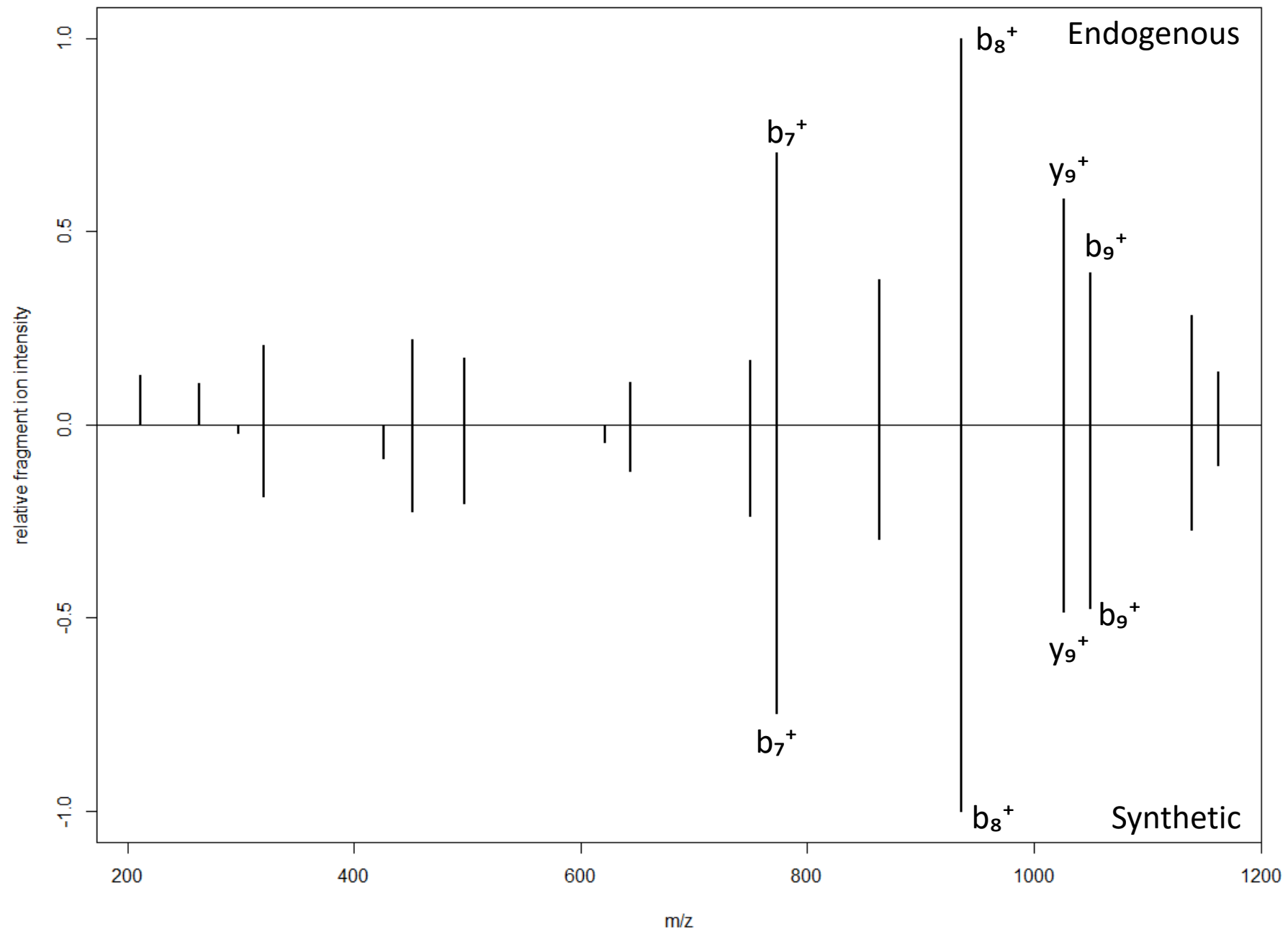


IPQNFADYFETSSQ



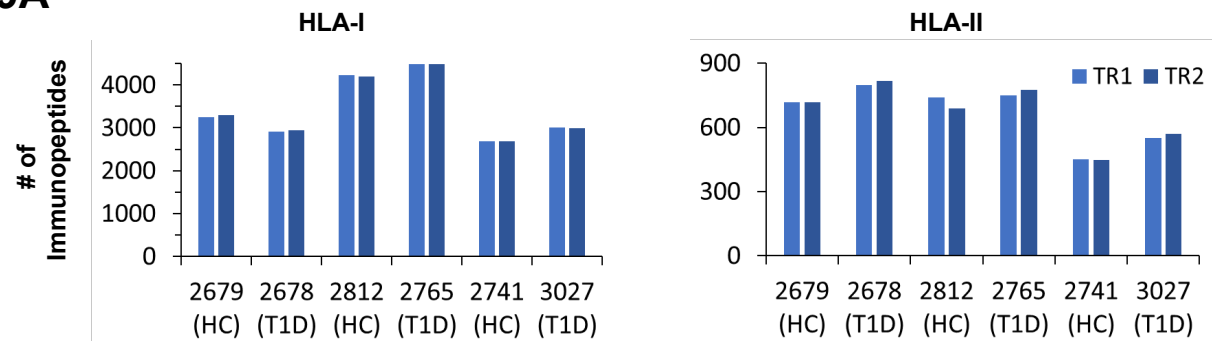


# LPSQAFEYILYNKGIMGED



**Fig. S9**

**S9A**



**S9B**

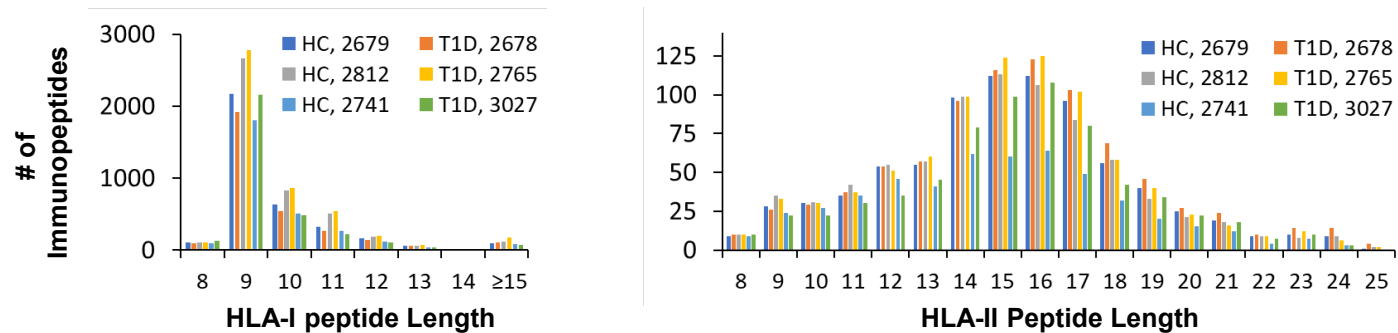
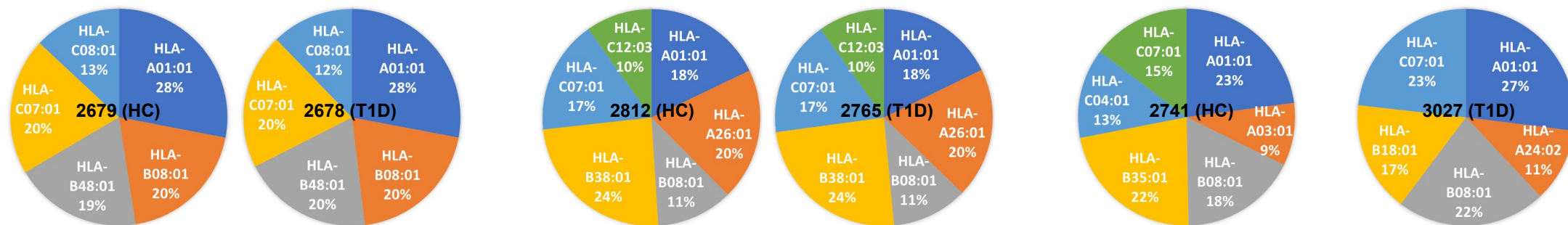
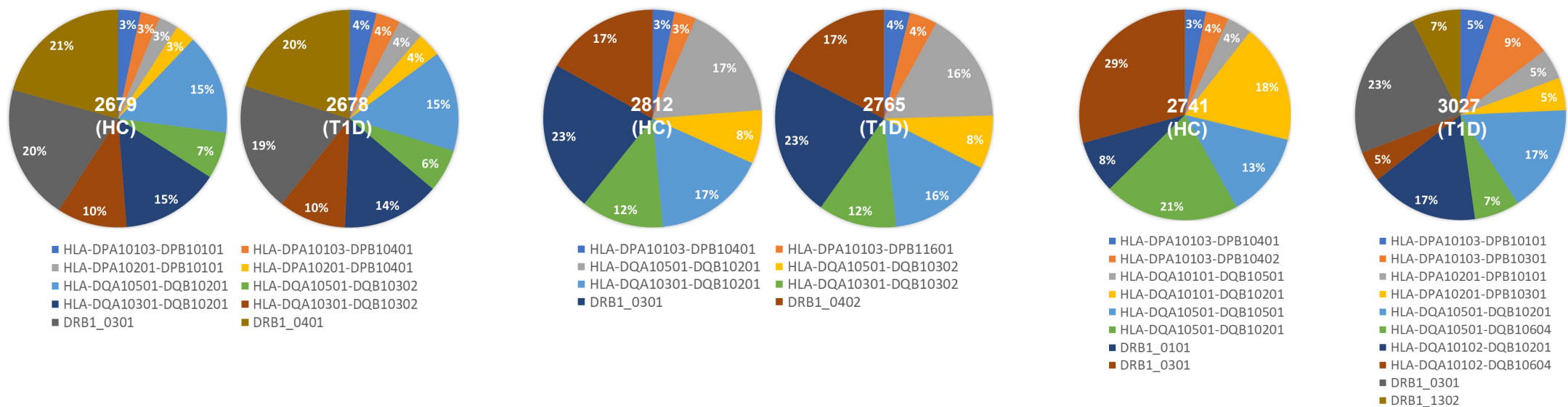


Fig. S10

S10A



S10B



**Fig. S11**

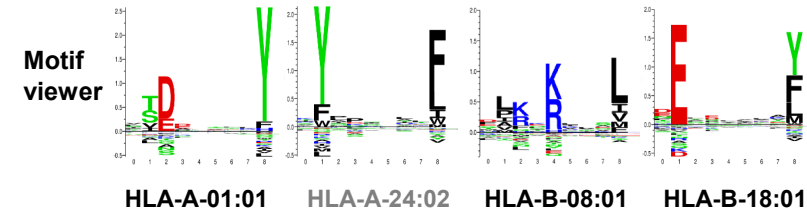
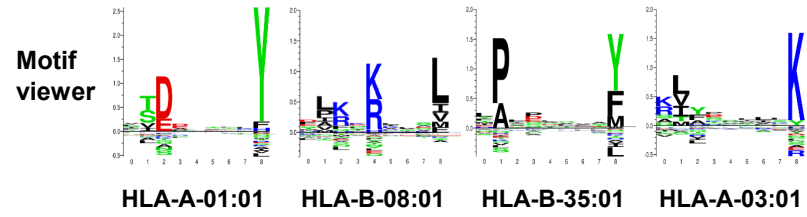
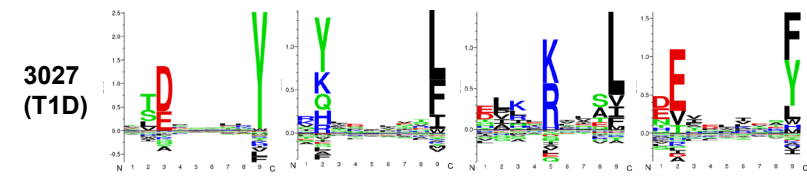
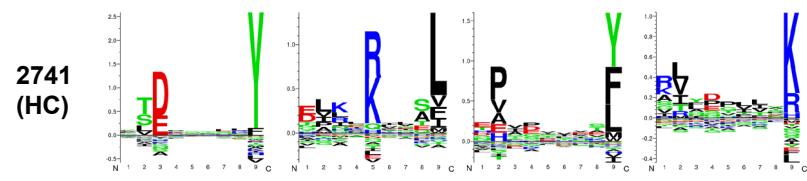
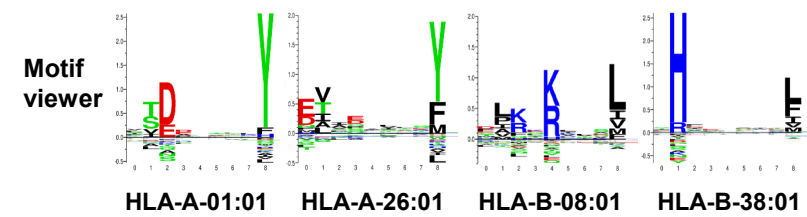
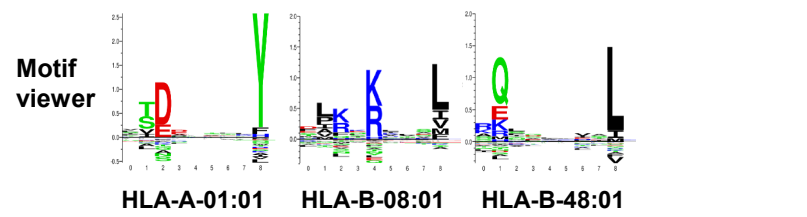
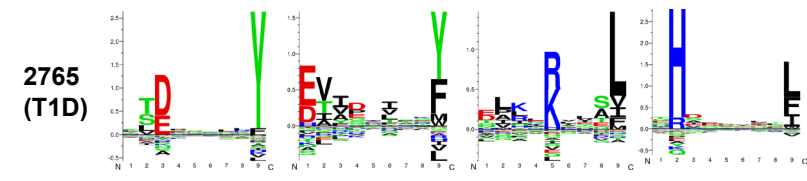
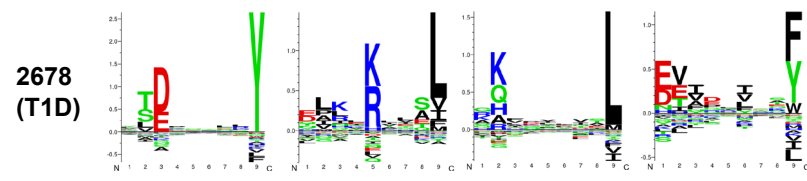
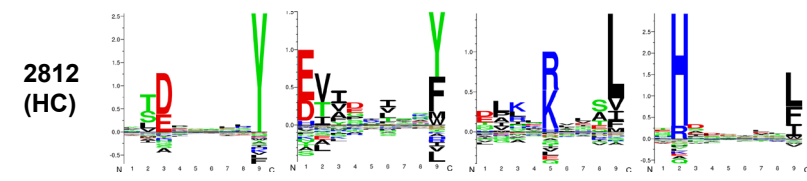
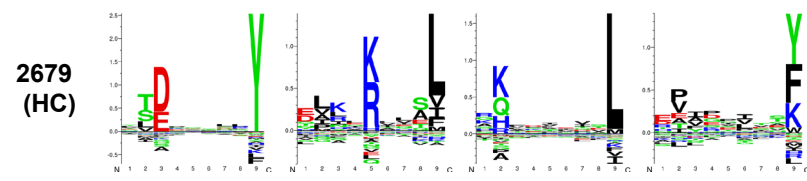
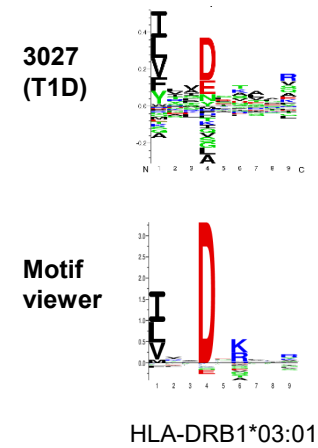
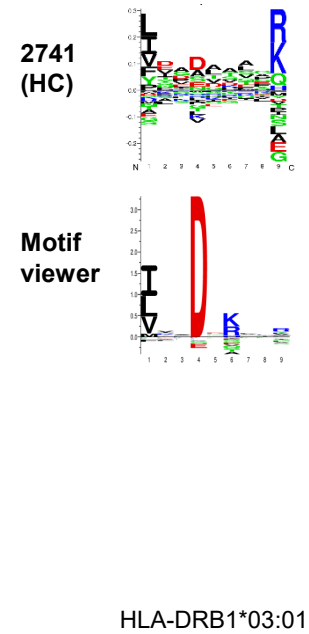
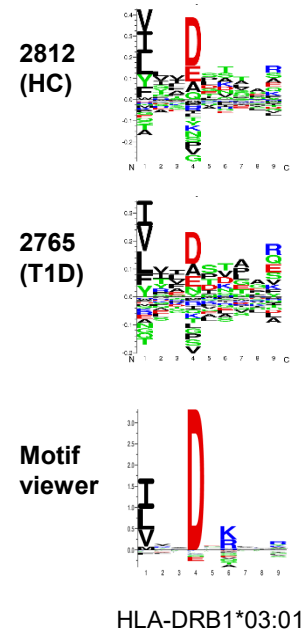
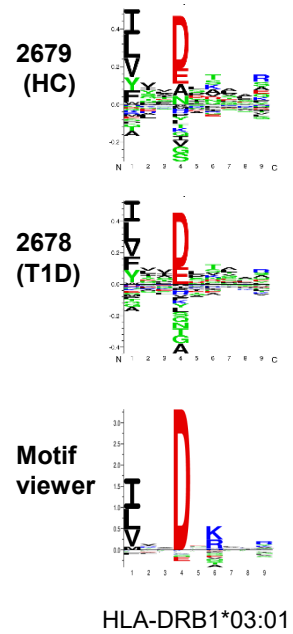


Fig. S12



**Fig. S13**

This study		IEDB	
Immunopeptide sequence	Antigen	Linear epitope	MHC ligand assay
<b>Identified during HLA-I immunopeptide enrichment</b>			
ETVWHLEEF	HLA-DPA1	Reported	Positive
QSEAGSHTLQWMY	HLA-B	Reported	Positive
<b>Identified during HLA-II immunopeptide enrichment</b>			
AQGALANIAVDKANLE	HLA-DRA	Reported	Positive
DDTQFVRFSDAASPREEPR	HLA-B	Reported	Positive
DDTQFVRFSDAASPRGEP	HLA-C	Reported	Positive
DDTQFVRFSDAASPRGEP	HLA-C	Reported	Positive
DDTQFVRFSDAASQKMEPR	HLA-A	Reported	Positive
DDTQFVRFSDAASQRMPEPR	HLA-A	Reported	Positive
DGTQFVRFSDAASPRTEPRAP	HLA-B	No information available	No information available
LANIAVDKANLEIMT	HLA-DRA	Reported	Positive
LRSWTAADTAAQITQR	HLA-B;HLA-C	Reported	Positive
LRSWTAVDTAAQISEQ	HLA-E	Reported	Positive
TQFVRFSDAASPRTEP	HLA-B	Reported	Positive
TQFVRFSDAASPRTEPR	HLA-B	Reported	Positive
TQFVRFSDAASQRMPEPR	HLA-A	Reported	Positive
VDDTQFVRFSDAASQKMEPRAPW	HLA-A	No information available	No information available