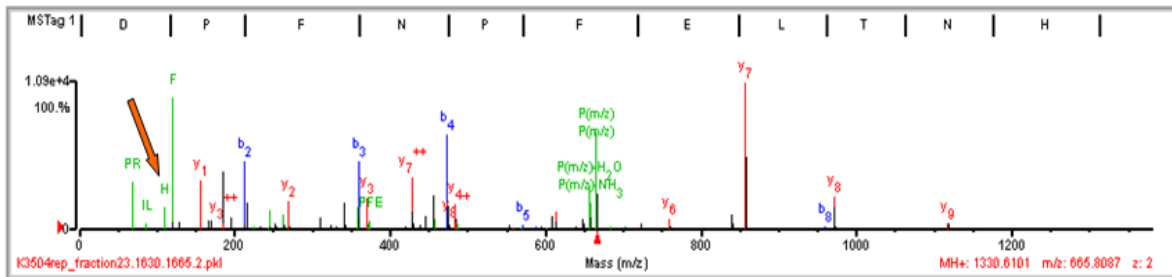


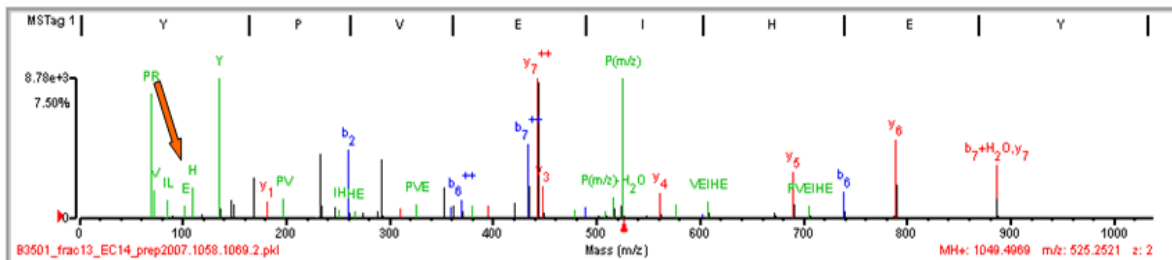
Utility of characteristic QTOF MS/MS fragmentation for MHC class I peptides.

Supporting Figure 1.

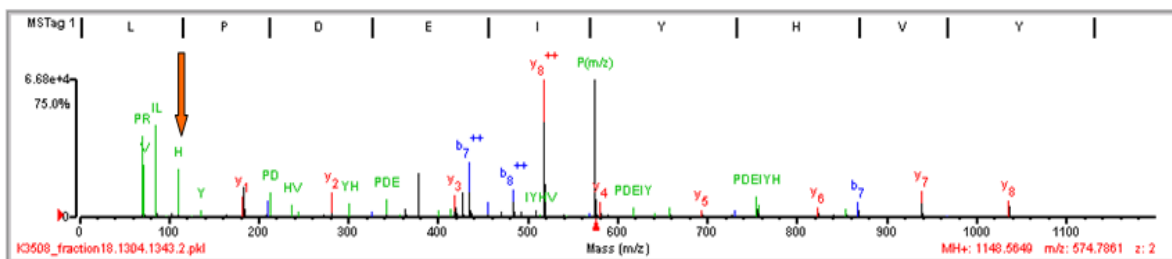
Immonium ion presence. Spectra of peptide MPDVDLHL from HLA allele B3501 where the Histidine immonium ion is clearly present inside the peptide at m/z 110.



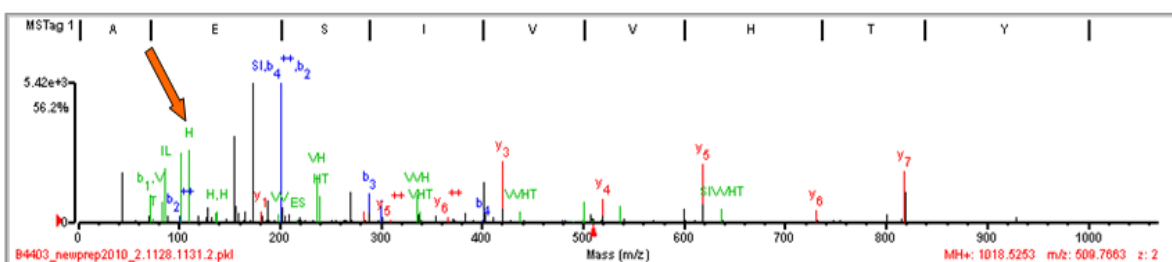
Peptide YPVEIHEY from HLA-B3501



Peptide DPFNPFELTNH from HLA-B3504



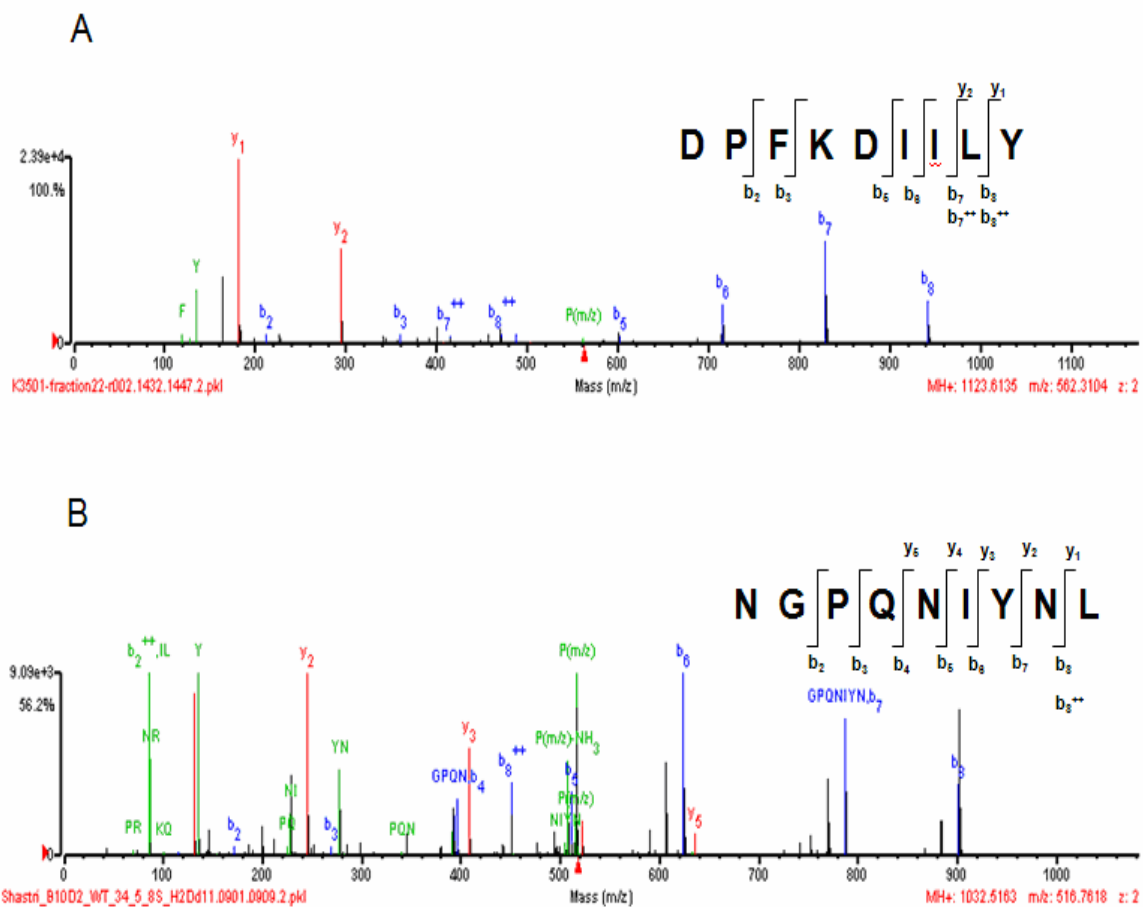
Peptide LPDEIYHVY from HLA-B3508



Peptide AESIVHTY from HLA-B4403

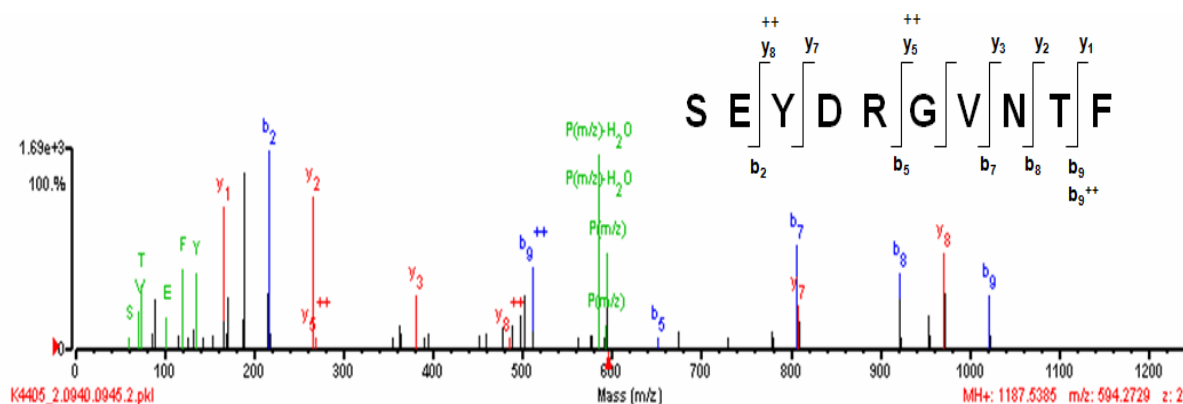
Supporting Figure 2.

Decreased proline cleavage. A. This peptide from HLA-B3501 shows the effect of basic amino acids close to N-terminus generating abundant b-type ions that affect both cleavage of N-terminus to proline in this case by impossibility of b_1 formation from protonated oxazolone, also the small peak C-terminal to proline is affected and is not evident. B. Peptide extracted from H2-Kd generate abundant b-type ions in gaseous phase collision by absence of basic residue inside the sequence. It is observed increased cleavage toward C-terminus in aliphatic residues as Leu, Ile and Tyr with Asn weak bond helping the process, all effects together lead to suppression of cleavage N-terminus to proline and C-terminal to proline.

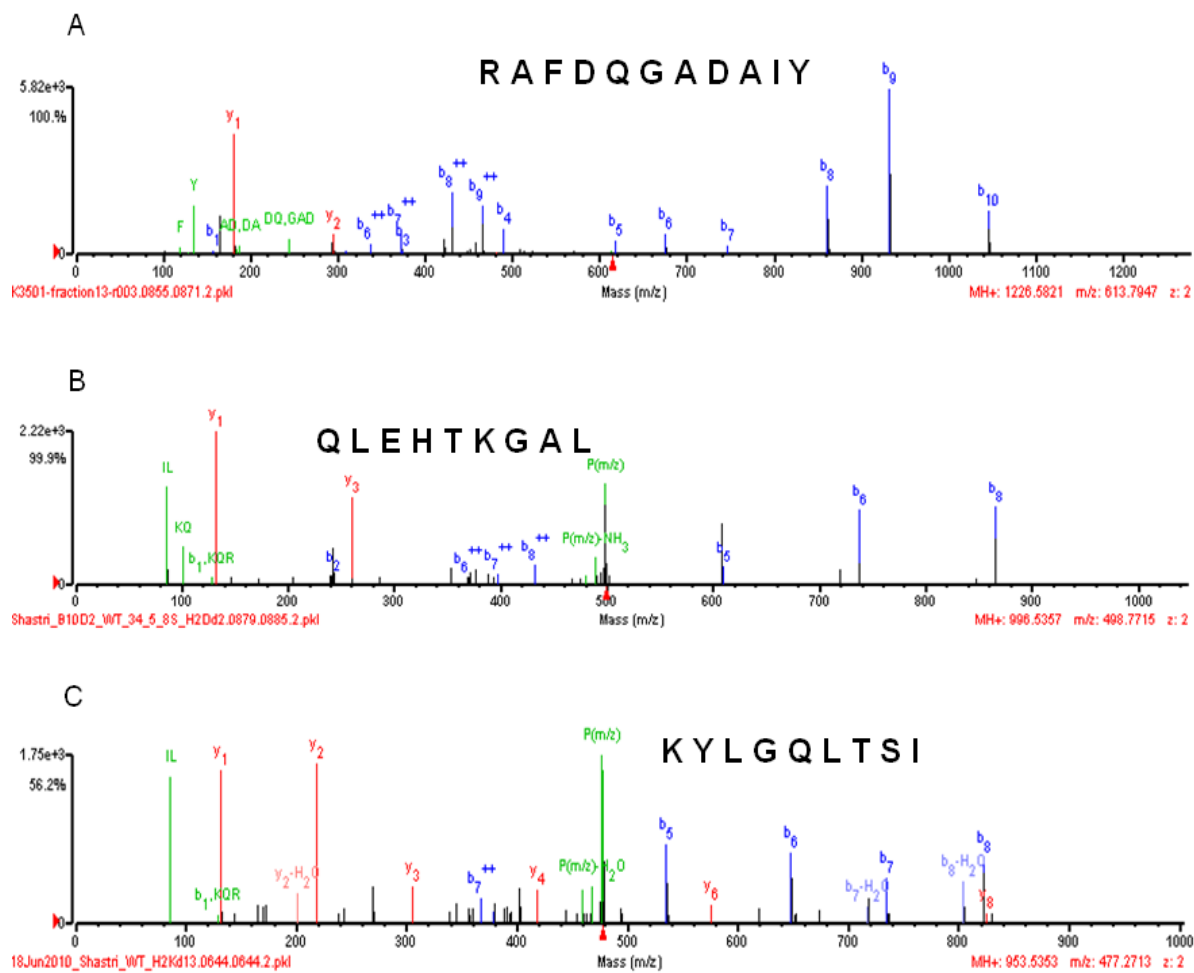


Supporting Figure 3.

Cleavage at Arginine residue is suppressed. Some sequences with peptides containing Arg showed absence or marked decrease in fragmentation around this basic residue. In this spectrum, the y-type mono-charged ions from cleavage surrounding Arg are absent; with only low intensity peaks for b_5 and y_5^{++} visible. When the basic residue is inside the sequence, it is more common to find spectral patterns with peak-wave behavior in the y-type and b-type series – an important clue when confirming validity of algorithm matched sequences.

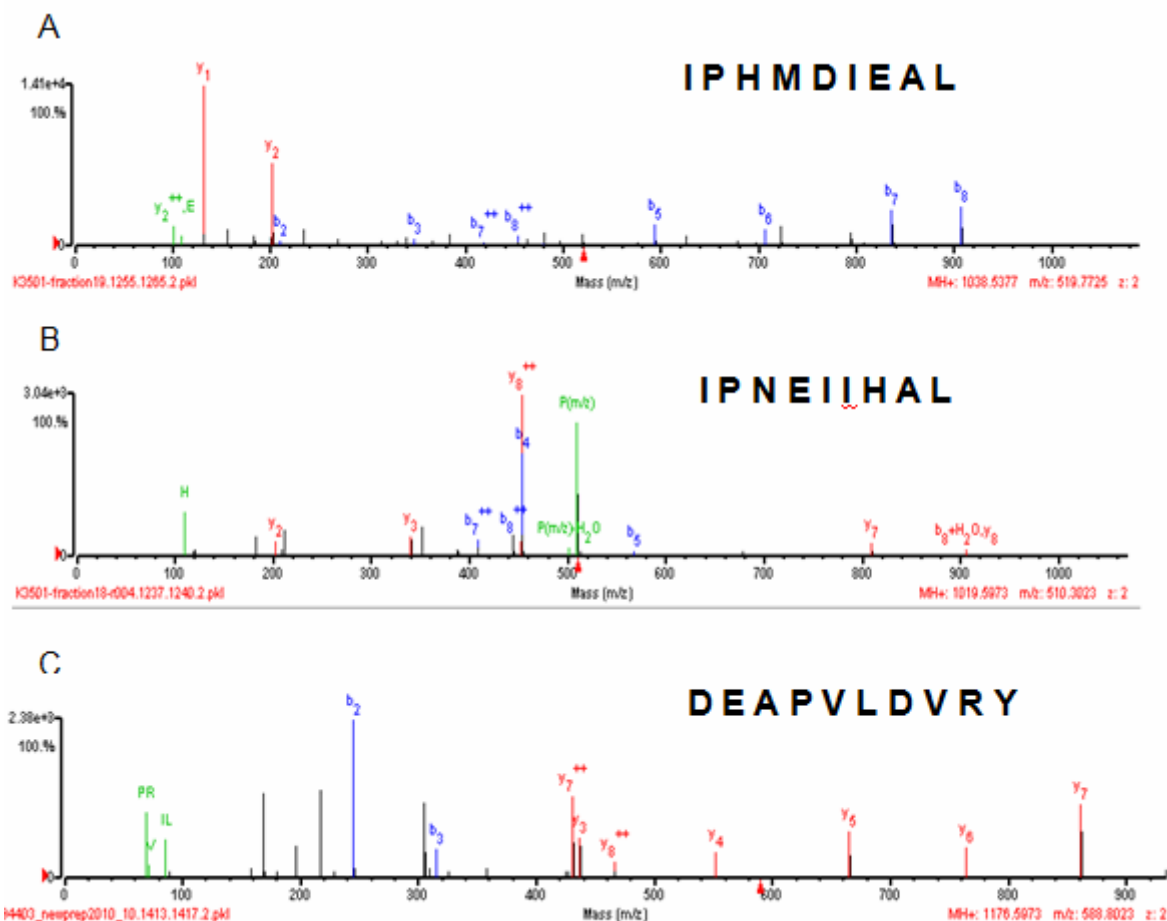


Supporting Figure 4. Position of basic amino residues enhance certain cleavages. In these three spectra (from HLA-B*3501, H2-Dd and H2-Kd), the increase of b-ions is seen by the basic residues proximity toward N-terminus, also non-polar and polar residues located at C-terminus produce intense peaks.

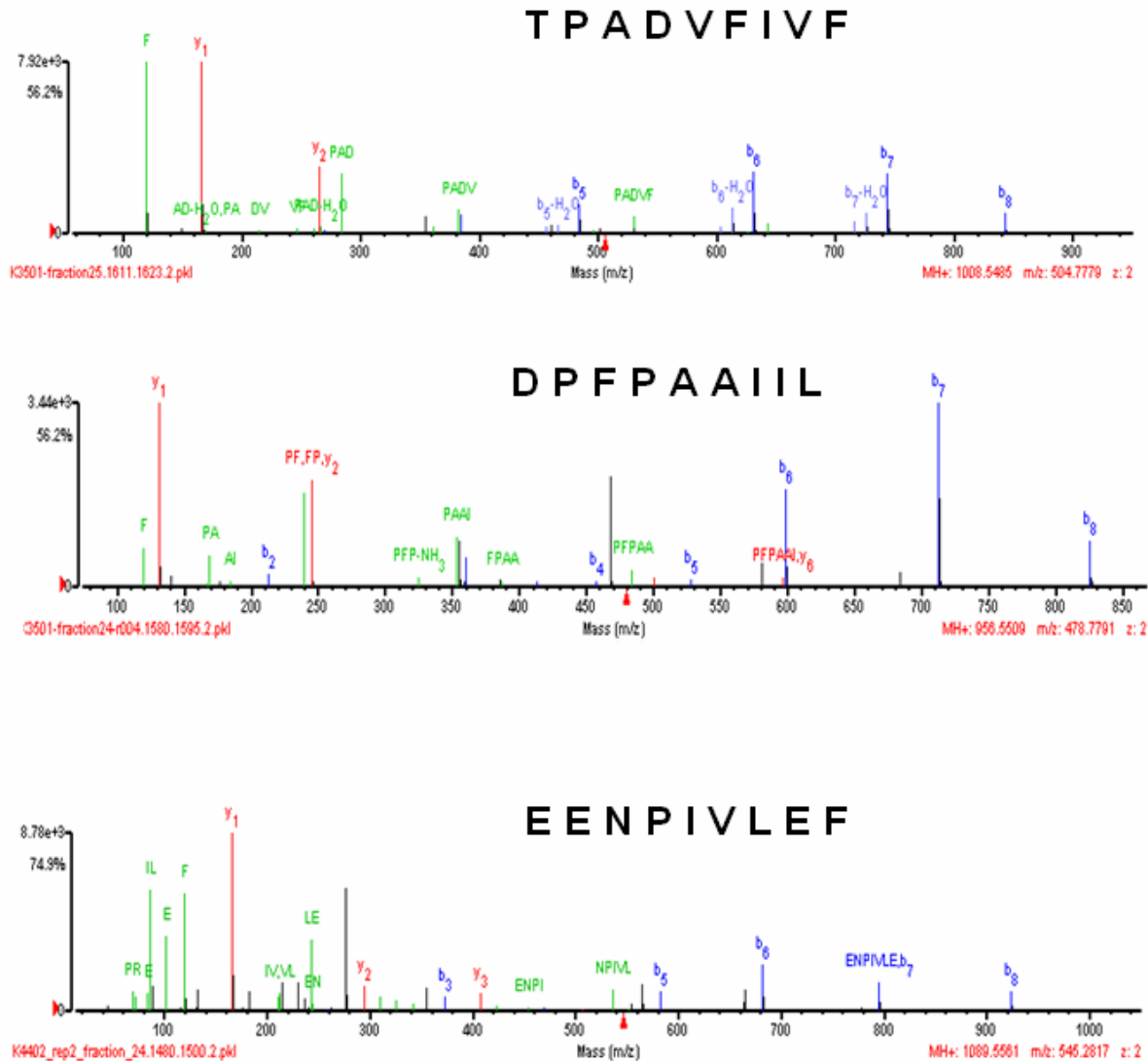


Supporting Figure 5. Presence of basic residues suppress cleavage of non-polar and polar residues.

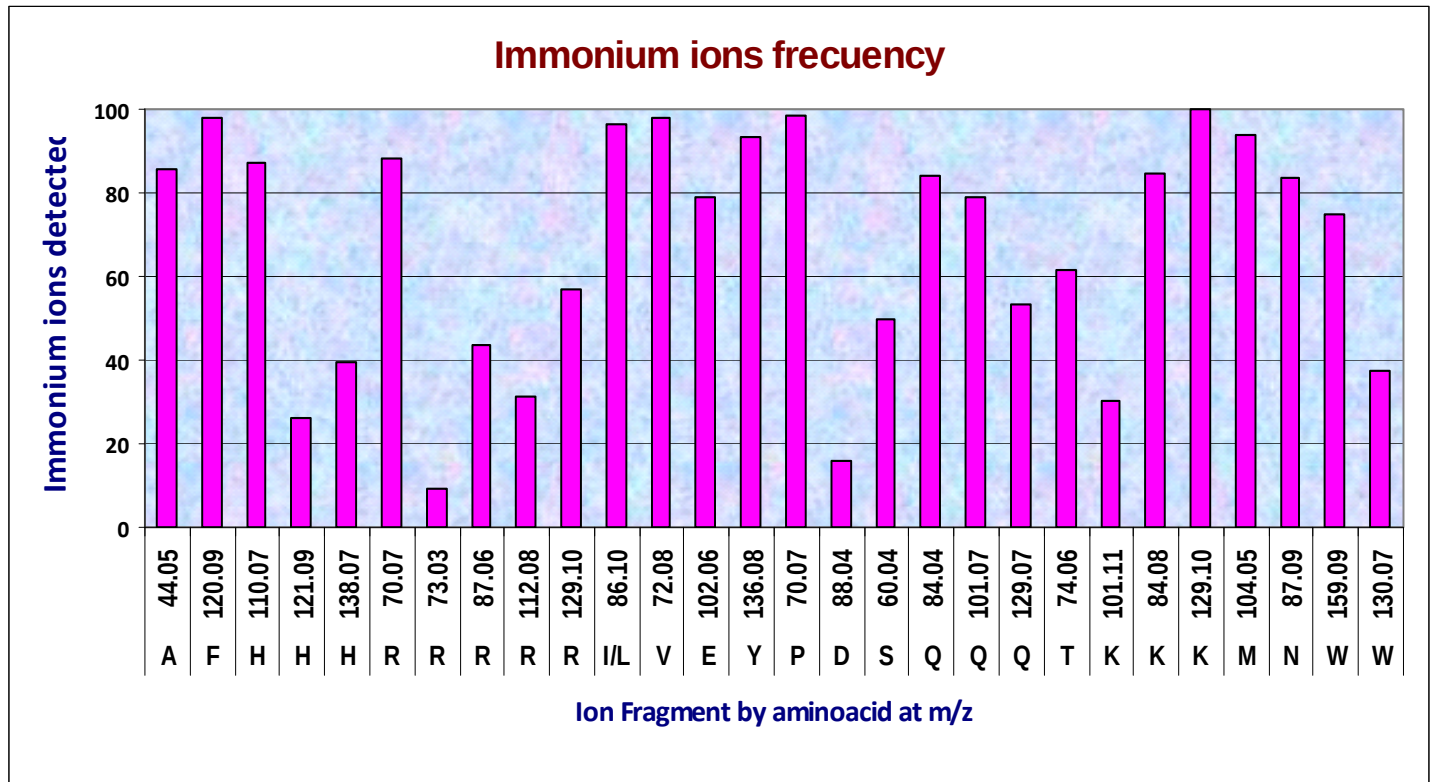
(A) Intense y-ions fragments are generated on C-terminus when basic residues are located close to N-terminus but (B) they show decrease in peak intensity when the basic amino acid approaches to C-terminus. (C) Presence of arginine close to C-terminus also suppress completely the common fragmentation y_1 and y_2 in this peptide extracted from HLA-B4403.



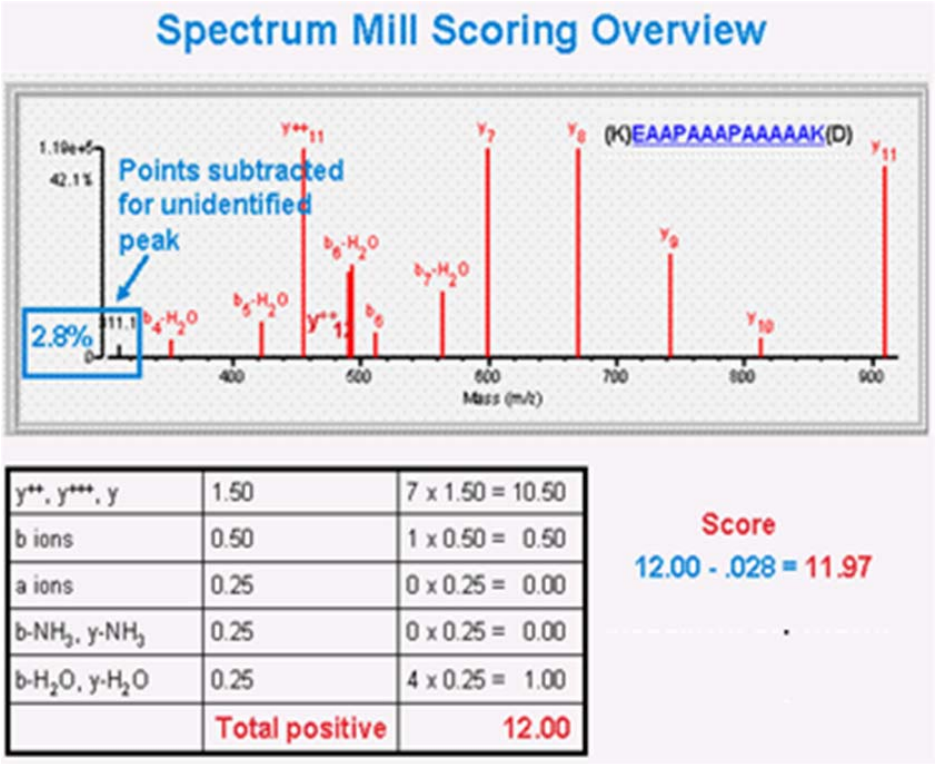
Supporting Figure 6. Non-polar and polar ions can lead CID fragmentation. These three spectra (from MHC class I peptides) show abundant generation of b-ion with intense y-type fragmentation in non-polar residues as Val, Ser, Thr, Tyr, Phe, Ala, Ile and Leu. In some cases even exceeding the enhanced fragmentation N-terminal to Pro which disappears in these spectra.



Supporting Figure 7. Immonium ions and related fragments in peptides from MHC class I. This graph shows the trend for immonium ions by the most common amino acids found in the type of peptides studied. Some of them are more likely to be present in the spectrum so they can be a good guide to confirm sequences. Residues such as aspartic acid generate very poor immonium peak signal.



Supporting Figure 8. Impact of scoring defined by the algorithm. The values assigned for the presence of y-type and b-type ions within the algorithm is defined by different appraises for each of them, thus giving preference when a tryptic peptide was detected in the mass spectrometer. This can unfairly modify the scoring and fail in the report of the most probable sequences given by the algorithm search when abundant b-ions are generated (as is common in MHC class I peptides) with basic residues located away of C-terminus or when these basic amino acids are absent. In this study we modified the scoring set of parameters for b-ions and y-ions in the algorithm equalizing the values for both at 1.0 units.



Supporting Table 1, 2, 3. Basic residues suppress cleavage in endogenous non-tryptic peptides. For different peptides extracted from MHC class I alleles with basic residues inside the main fragment peaks for y-type and b-type ions are detailed here with the following indicative forward-slashes into the sequence: blue b-type ions; red y-type ions; purple vertical lines both y- and b- ions. The trend is suppression of cleavage around basic residues with marked trend with arginine followed by histidine and last lysine. The presence of internal ions generated by Arg show few peaks but they tend to increase when residues as Asn and/or Gln are present maybe for the facile breaking bond presented by these amino acids..

Supporting Table 1

Peptides with Arginine				
#	MHC I	Protein source	Sequence & Cleavage	#internal ions generated
1	HLA-B4402	AP-2 complex subunit mu-1	G E V L I S R V Y	4
2	HLA-B4402	Vacuolar proton pump subunit F	N E I E D T F R Q F	7
3	HLA-B4402	Poly ADP-ribose] polymerase 1	E E L G F R P E Y	6
4	HLA-B4402	C-1-tetrahydrofolate synthase, cytoplasmic	S E L D L I S R L	7
5	HLA-B4402	Bardet-Biedl syndrome 2 protein	A E Q D L I R E L	2
6	HLA-B4405	Myeloid leukemia factor 2	A E G P P R L A I	5
7	HLA-B4405	Glutamyl-peptide cyclotransferase-like protein	E E L P L G R E L	7
8	HLA-B4405	Ser/Thr-protein phosphatase 2A 65 kDa regulatory subunit A alpha isoform	D E Q D S V R L L	9
9	HLA-B3508	Tubulin beta-2B chain	Y P D R I M N T F	12
	HLA-B3508	Microtubule-associated protein 1A	Y P D E R S F Q Y	5
11	HLA-B3508	Heterogeneous nuclear ribonucleoprotein H	L P Y R A T E N D I Y	12
14	H2Db	DNA-directed RNA polymerase II subunit RPB1	V T P F N I D R L	6
15	H2Db	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 1	I S G V N R Y Y V	6
16	H2-Db	RNA polymerase II subunit B1	V T P F N I D R L	2
17	H2-Db	Serine/threonine-protein kinase MRCK gamma	A A V I D Q E R L	12
18	H2-Db	Protein C10	S A P E N A V R M	5
19	H2-Db	H-2 class I histocompatibility antigen, D-B alpha chain precursor	F A Y E G R D Y I	7
	HLA-B0702	Staphylococcal nuclease domain-containing protein 1	S P A F S T R V L	12
	HLA-B0702	NACHT, LRR and PYD domains-containing protein 2	I P F S N P R V L	4
20	HLA-B0702	DNA-directed RNA polymerase II subunit RPB1	T P Q S N R P V M	7
21	HLA-B0702	Transcription elongation factor SPT4	L P Q G I V R E L	5
22	H2-Kd	Collagen alpha-2(I) chain	S F V D T R T L L	8
23	H2-Dd	Heterogeneous nuclear ribonucleoprotein K	S G P E R I L S I	7
24	H2-Dd	60S ribosomal protein L6	T G P L V I N R V	8
25	H2-Dd	Galectin-8	M G P G R T V V I	8
26	H2-Dd	Zinc finger CCH-type antiviral protein 1	A G P D R F V L L	6
27	H2-Dd	Protein YIPF3	V G P T Q R L L L	10
28	HLA-B3501	Cytochrome b-c1 complex subunit 9	R A F D Q G A D V I Y	3
29	HLA-B3501	Peptidyl-prolyl cis-trans isomerase A	A V D G E P L G R V S F	9
30	HLA-B3501	Actin-related protein 3	I P I A G R D I T Y	7
31	HLA-B3501	Protein tyrosine phosphatase type IVA protein 1	R P A P V E W T Y	3
32	HLA-B3501	DnaJ homolog subfamily B member	N P D D V F R E F	5
33	HLA-B3501	Actin-like protein 3	I P I A G R D I T Y	3
34	HLA-B3501	Tubulin beta-2A chain	Y P D R I M N T F	3
35	HLA-B3501	eIF4E-binding protein 1	T P G G T R I I Y	2

Supporting Table 2

Peptides with Histidine				
#	MHC I	Protein source	Sequence & Cleavage	#internal ions generated
1	HLA-B3501	Heterogeneous nuclear ribonucleoprotein M	I/P N/E// H A L	16
2	HLA-B3501	Nuclear nucleic acid-binding protein C1D	Y/P V/E//H/E Y	18
3	HLA-B3501	Spectrin beta chain, brain 1	Y/P N/V/N H N F	18
4	HLA-B3501	Methyltransferase-like protein 11A	L/P D/E I/Y H V Y	18
5	HLA-B3501	Regulator of nonsense transcripts 1	A/P V/E/V/T/H N F	15
6	HLA-B3501	Proline-rich protein BCA3	V/P E E/G G A T/H V Y	14
7	HLA-B3501	Thymocyte nuclear protein 1	E A Y P D H T Q/F	12
8	HLA-B3501	Cytochrome c oxidase subunit 6A1, mitochondrial	F/P W/G D G N H T L F	13
9	HLA-B3501	ASH2-like protein	T P H S E N F Y	18
10	HLA-B3501	26S proteasome non-ATPase regulatory subunit 7	L/P N H Q V I Y	18
11	HLA-B3501	Heat shock protein beta-1	D V N H F A P D E L	10
12	HLA-B3501	60S acidic ribosomal protein P0	V P H S V N G Y	20
13	HLA-B3501	Catalase	S A L E H S V Q Y	13
14	HLA-B3501	Transmembrane protein 50A	H A C G V W A T I	9
15	HLA-B3501	Dynein heavy chain 1, cytoplasmic 1	H V N W V V S E L	9
16	HLA-B3501	U5 small nuclear ribonucleoprotein 200 kDa helicase	Y A Q D E H L V T F	9
17	HLA-B3501	Ubiquitin-like modifier-activating enzyme 1	F/P N A I/E/H T L	15
18	H2-Db	DmX-like protein 1	V/A P/A/N S L V/H/A/F	11
19	H2-Db	Cyclin-dependent kinase inhibitor 1B	F/G P V/N/H E E L	11
20	H2-Db	Actin-related protein 2	Y/A G S/N/F/P/E/H I	10
21	H2-Db	Vacuolar protein sorting-associated protein 35	F/S E/E N H E P L	14
22	H2-Db	tRNA (uracil-5-)-methyltransferase homolog A	A A P/F/D T V/H I	14
23	H2-Db	Homeobox protein SIX5	A/G P/T/N V/H/L I	13
24	H2-Db	DmX-like protein 1	V/A P/A/N S L V/H/A/F	13
25	H2-Db	Cyclin-dependent kinase inhibitor 1B	F/G P V/N/H E E L	24
26	H2-Db	Putative methyltransferase	T/A I E/N/S W//H L	17
27	H2-Db	Malignant T cell amplified sequence 1	I/G I E/N/H Y L	27
28	H2-Db	G1/S-specific cyclin-D3	A A V I/A/H D F L	24
29	H2-Db	Cyclin-dependent kinase inhibitor 1B	F/G P V/N/H E E L	21
30	H2-Db	DmX-like protein 1	V/A P/A/N S L V/H/A/F	13
31	HLA-B4402	Proteasome subunit beta type-4 precursor	E E L L G/D/G/H S Y	13
32	HLA-B4402	Exportin-2	A E S I V V/H/T Y	17
33	HLA-B4405	Dolichyl-diphosphooligosaccharide--protein glycosyltransferase 63 kDa subunit precursor	A E L T P/H Q T F	18
34	H2-Kd	Chromodomain-helicase-DNA-binding protein 1	S/Y I G G/H/E G L	18
35	H2-Kd	Alpha-1,3-mannosyl-glycoprotein 2-beta-N-acetylglucosaminyltransferase	S/Y G T A V/T/H I	14

Supporting Table 3

Peptides with Lysine				
#	MHC I	Protein source	Sequence & Cleavage	#internal ions generated
1	HLA-B4402	Hemoglobin subunit gamma-1	K E F T P E V Q A S W	15
2	HLA-B4402	Eukaryotic translation initiation factor 3 subunit C	E E F E L L G K A Y	20
3	HLA-B4402	40S ribosomal protein S5	A E T P D I K L F	15
4	HLA-B4402	Uncharacterized protein C14orf119	S E P D F V A K F	32
5	HLA-B4402	Nucleolar protein 11	D E N S V I K S F	24
6	HLA-B4402	Prolyl 4-hydroxylase subunit alpha-1 precursor	A E I E I V K D L	10
7	HLA-B4402	Uncharacterized protein	S E P D F V A K F Y	24
8	HLA-B4402	DNA topoisomerase 2-alpha	A E I N N I I K I	12
9	HLA-B4402	Nuclear migration protein nudC	A E K L V I T Q T F	17
10	HLA-B4402	Targeting protein for Xklp2	E E L E K L Q Q Y	36
11	HLA-B4402	Protein SET	T E F E D I K S G Y	22
12	HLA-B3501	Bromodomain adjacent to zinc finger domain protein 1B	T A W L E I M T K Y	9
13	HLA-B3501	Alpha-actinin-4	N A F E V A E K Y	20
14	HLA-B3501	Cytochrome P450 51A1	M V G K T F T Y	11
15	HLA-B3501	Ubiquitin-conjugating enzyme E2 E1	Y P F K P P K V T F	11
16	HLA-B3501	7-dehydrocholesterol reductase	T P A G V V N K Y	24
17	HLA-B3501	Serine/threonine-protein kinase RIO2	F P V P K P I D Y	10
18	H2-Db	Adenosine kinase	S A V V D K D F L	29
19	H2-Db	Nucleolar GTP-binding protein 2 OS	A S L T N P F G K G A F I	14
20	H2-Db	Sorting nexin-14	I G P K N Y E F I L	30
21	H2-Db	Acyl-protein thioesterase 1	K A L I N P A N V T F	18
22	H2-Db	Protein flightless-1 homolog	A A L K L G Q E L	15
23	H2-Db	B-cell scaffold protein with ankyrin repeats	N S P F N S K F P A	17
24	H2-Db	Beta-2-microglobulin	I Q K T P Q V Q V Y	38
25	H2-Db	ELAV-like protein 1	G A V T N V K V I	16
26	H2-Db	Tropomodulin-1	S S I V N K E G L	47
27	HLA-B4405	RNA-binding protein 8A	A E Y G E I K N I	20
28	HLA-B4405	Transcription intermediary factor 1-beta	N E A F G D T K F	40
29	HLA-B0702	Ran-binding protein 9 -	L P K Q P P L A L	25
30	HLA-B0702	Sphingolipid delta(4)-desaturase DES1	F P N I P G K S L	17
31	HLA-B3508	Dihydrofolate reductase	F P E I D L E K Y	16
32	HLA-B3508	Anaphase-promoting complex subunit 7	L P S E I E V K Y	26
33	HLA-B3508	Squalene synthetase	M P A V K A I I Y	15
34	HLA-B3508	Hypoxanthine-guanine phosphoribosyltransferase	I P D K F V V G Y	39
35	HLA-B3508	CCR4-NOT transcription complex subunit 1	F P Q Y P D K E L	18

Supporting Table 4. No basic residues present influence the b-type ions generation in MHC class I peptides. As well as occur with histidine and lysine, the absence of basic residues produce abundant internal ions, also the trend to generate b-ions is strong with this sequences as shown in the last column with more than 50% of b-type ions comparing with y-ions in the spectrum.

Peptides without basic residues							
#	MHC I	Protein source	Sequence & Cleavage	#internal ions generated	Total y-type ions	Total b-type ions	b-type abundance (%)
1	HLA-B4402	Kanadaplin	E E N P I V I E I F	13	4	11	73
2	HLA-B4402	Heterogeneous nuclear ribonucleoprotein H	T E N D I I Y N I F I F	23	7	9	56
3	HLA-B3501	Spectrin beta chain, brain 1	E A I F I L N N Q E I I Y	34	6	11	65
4	HLA-B3501	Pre-mRNA-processing-splicing factor 8	S P I W I F/P I P I S I Y	14	6	10	63
5	HLA-B3501	Nuclear pore membrane glycoprotein 210 precursor	L P A E I F I E I V I L	16	3	6	67
6	HLA-B3501	Complement decay-accelerating factor precursor	F P V G T V I V I E I Y	14	4	6	60
7	HLA-B3501	Lymphokine-activated killer T-cell-originated protein kinase	D P I F I P A I W I I I L	17	4	12	75
8	HLA-B3501	T-complex protein 1 subunit eta	L I P I G D I W I A I T I Q I Y	32	10	14	58
9	HLA-B3501	Replication protein A 32 kDa subunit	M P L E I I D M N E I I F	16	4	9	69
10	HLA-B3501	Protein transport protein Sec23B	M P Q I F S T I V I E I Y	10	5	7	58
11	HLA-B3501	26S proteasome non-ATPase regulatory subunit 7	L P D I V I S I L I Q I E I F	8	4	6	60
12	HLA-B3501	60 kDa SS-A/Ro ribonucleoprotein	T P A D I W I F I V I F	9	2	8	80
13	H2-Db	Transmembrane protein 131	A S I L I V N I S P S I Y I L	24	6	13	68
14	H2-Db	Diphosphomevalonate decarboxylase	T A P V M I A I V I I	20	3	10	77
15	H2-Db	Receptor tyrosine-protein kinase erbB-2	G A I V I E N I P E I Y I L	12	5	9	64
16	H2-Db	40S ribosomal protein SA	V A I I E N P A I D I V S I V I I	46	5	14	74
17	H2-Db	Cyclin-A2	V S I L I L N I P/P E I T I L	29	7	8	53
18	H2-Db	U3 small nucleolar RNA-interacting protein 2	A A I L I L N T D I L I V	37	5	12	71
19	H2-Db	Sterol regulatory element-binding protein 2	A A I W Q N I P A I L I T I A I L	72	15	21	58
20	H2-Db	Transmembrane protein 167 precursor	S A I I F N I F Q I S I L	40	19	21	53
21	HLA-B3508	Microtubule-associated protein tau	S P Q L A I T I L A I D I E I V I S I A	28	6	16	73
22	HLA-B3508	Protein BTG1	W A D I P Y E I V I S I Y	27	9	17	65
23	HLA-B3508	RNA-binding protein 10	Y P W P D I V S I T I Y	15	5	9	64
24	HLA-B3509	Zinc finger and BTB domain	M P A I P E I V I S I Y	24	6	15	71
25	HLA-B3508	Cleavage and polyadenylation specificity factor subunit 1	V A D I P Y W I V I I I M	23	6	12	67
26	HLA-B3508	Transforming protein RhoA precursor	Y P D T D I V I I I L I M	12	6	6	50
27	HLA-B4405	Elongation factor 2	L E I P E E L I Y Q I T I F	12	7	9	56
28	HLA-B0702	Ataxin-2-like protein	V I P I Q I S I G V I P A I L	20	7	12	63
29	HLA-B0702	Myc-associated zinc finger protein	A V I A I P V I A I S A I L	18	6	9	60
30	HLA-B0702	Rhomboid domain-containing protein 2	Y I P A I S I A I G T I S I L	16	9	15	63

Supporting Table 5. Presence of glutamine and asparagine increase internal ions and fragmentation inside peptide. These data show the impact of these two residues in the cleavage in MHC class I. First always there are y-type and or b-type ions generated from this residues in the listed peptides and second the internal ions found in the QTOF analysis are generated in more than 40% of the time from these residues as may be detailed in the last column. The presence of several residues inside the sequence increase the internal ion generation.

Generation Internal Ions from Asn and/or Glu						
#	Peptide sequence	y-type and b-type ions from Glu and or Asn	Residue analyzed	Total Internal ions presented	Internal ions generated from Glu and/or Asn	Influence Glu and/or Asn in generation internal ions (%)
1	IGIENIHYL	y ₄ , y ₅ , b ₄	N	34	17	50
2	FGPVNHEEL	y ₄ , y ₅ , b ₄ , b ₅	N	39	17	44
3	IGPKNYEFL	b ₄ , b ₅	N	33	16	48
4	YVVDNIDHL	y ₄ , y ₅	N	10	4	40
5	SLGKNPTDAYL	y ₆ , y ₇ , b ₄ , b ₅	N	38	16	42
6	VSVANVDLL	b ₅	N	19	9	47
7	TSVENHEFL	y ₄ , y ₅	N	19	8	42
8	GALENAKAEI	y ₄ , y ₆ , b ₄	N	11	7	64
9	RAIENIDTL	b ₄ , b ₅	N	15	9	60
10	GAVTNVKVI	y ₄ , y ₅ , b ₄ , b ₅	N	13	6	46
11	SLPTNLIHL	y ₅	N	9	6	67
12	LPDVSLQEF	y ₂ , b ₆ , b ₇	Q	24	14	58
13	FPDKPITQY	y ₁ , y ₂ , b ₇ , b ₈	Q	33	20	61
14	KEFTPEVQASW	y ₄ , y ₅ , b ₄ , b ₅	Q	42	17	40
15	DYQALRTSI	y ₄ , y ₅ , b ₄ , b ₅	Q	20	10	50
16	GYLGQVTTI	y ₄ , b ₅	Q	22	10	45
17	AALQNAVAF	b ₄ , b ₅	N,Q	18	16	89
18	GGIQNVGHI	y ₄ , y ₅ , y ₆ , b ₃ , b ₄	N,Q	18	14	78
19	FQIVNPHLL	y ₄ , y ₅ , y ₇ , y ₈ , b ₂ , b ₄	N,Q	36	29	81
20	SEDNPQTLLF	y ₄ , y ₆ , b ₄ , b ₅ , b ₆	N,Q	17	13	76
21	SALQNAESDRL	y ₆ , y ₇ , y ₈ , b ₃ , b ₄ , b ₅	N,Q	29	19	66
22	VSLNLRQVL	y ₃ , y ₄ , y ₆ , y ₇ , b ₃ , b ₆ , b ₇	N,Q	27	18	67
23	AEINNIKI	y ₄ , y ₅ , y ₆ , b ₄ , b ₅ , b ₆	N,N	19	10	53
24	NSIRNLDTI	y ₅ , b ₄ , b ₅	N,N	41	29	71
25	ASVLNVNHI	y ₂ , y ₃ , y ₄ , y ₅ , b ₄ , b ₅	N,N	37	30	81

Supporting Table 6. Summary of internal ion generation with presence and absence of residues. This data is a summary of tables I,II, III and IV in terms of internal ions presence There is a clear trend for arginine to yield less ions comparing with the other basic residues or when they are absent.

<i>Summary data presence internal ions per peptide</i>			
Basic residue	Peptides analyzed	fragments by peptide	
		Mean	STDV
No presence	30	23	13
Arginine	35	7	3
Histidine	35	15	4
Lysine	35	21	10

Supporting Table 7. False discovery rate.

Calculations with decoy database search									
Algorithm parameters	Allele	Algorithm Score	#Sequences	False Positive	True Positives	False Negatives (manually validated)	True Negatives	FDR %	True FDR %
Default (y-ions 3X to b-ions)	H2-Kb	>7.0	542	111	431	21	90	25.8	19.9
	H2-Db	>7.0	1897	506	1391	50	456	36.4	31.6
y-ions and b-ions equal	H2-Kb	>7.0	513	70	443	12	58	15.8	12.7
	H2-Db	>7.0	1920	444	1476	40	404	30.1	26.6

Formulation:

Sequences: Total of assignments over score threshold

False positive: passing reverse search assignment

True positive: passing assignments – false positives

False Negatives: False positives passing manual validation assignments

True negatives: False Positives – False negatives

*FDR%: $100 * \text{False positives} / (\text{\#sequences} - \text{False positives})$*

*True FDR%: $100 * \text{True negatives} / (\text{\#sequences} - \text{True negatives})$*

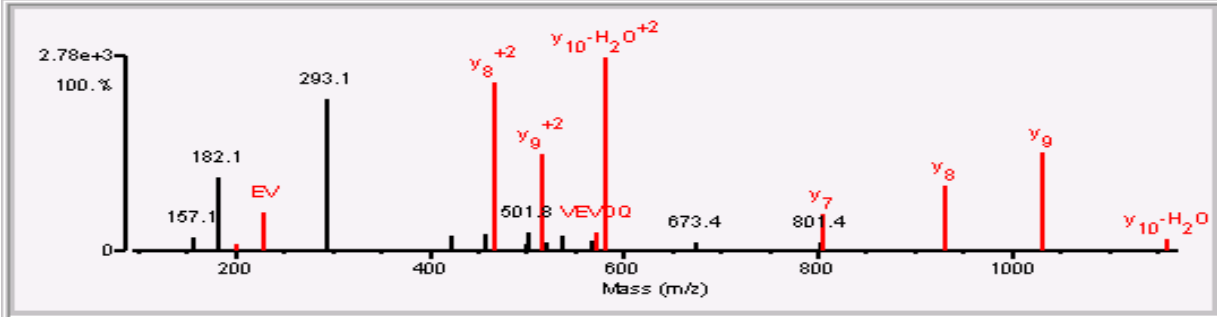
Practical application example:

In the search of the correct sequence for one peptide extracted from MHC class I molecule for the Allele HLA-B3501 we found the following report from the mass spectrometer algorithm:

Rank	Score	SPI (%)	BCS	# Unmatched Ions	Sequence	MH* Calculated (Da)	MH* Error (Da)	MH* Error (ppm)	Protein MW/pI (Da)	Species	Accession #	Protein Name
1	12.42	67.6	3	14/25	(N)YFVQTVEVDQL(I)	1340.6738	-0.0120	-9.0	68361.7/5.98	HUMAN	Q6TFL4	Kelch-like protein 24 - Homo sapiens (Human) [MASS=68361]
2	11.01	73.6	2	15/25	(M)YFVPPPYELSE(S)	1340.6414	0.0203	15.2	307693.0/6.23	HUMAN	Q9NT68	Teneurin-2 - Homo sapiens (Human) [MASS=307689]
3	9.94	85.6	4	14/25	(Y)AKELREGVEY(T)	1340.6850	-0.0233	-17.4	124714.3/4.87	HUMAN	O60518	Ran-binding protein 6 - Homo sapiens (Human) [MASS=124713]
4	9.91	89.7	4	11/25	(F)SPEGRLYQVEY(A)	1340.6487	0.0131	9.8	29484.0/7.58	HUMAN	P25789	Proteasome subunit alpha type-4 - Homo sapiens (Human) [MASS=29484]
4	9.91	89.7	4	11/25	(F)SPEGRLYQVEY(A)	1340.6487	0.0131	9.8	27399.6/6.35	HUMAN	P60900	Proteasome subunit alpha type-6 - Homo sapiens (Human) [MASS=27399]

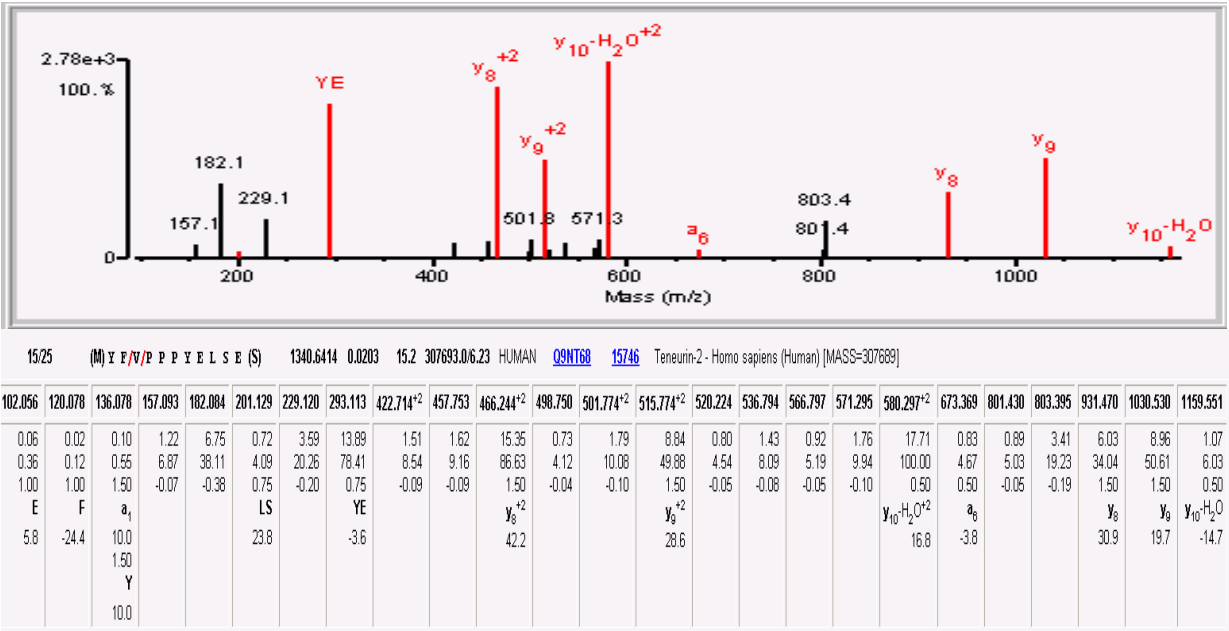
A. The first and second options are suspicious by the high unmatched ions and few internal ions content detected in the spectra. When there are no basic residues, we have found generation of internal ions greater than 7 ions , in this example the spectra shows for the first and second sequence options only four internal ions and one internal ion, respectively. Then both sequences can be discarded.

1.YFVQTVEVDQL



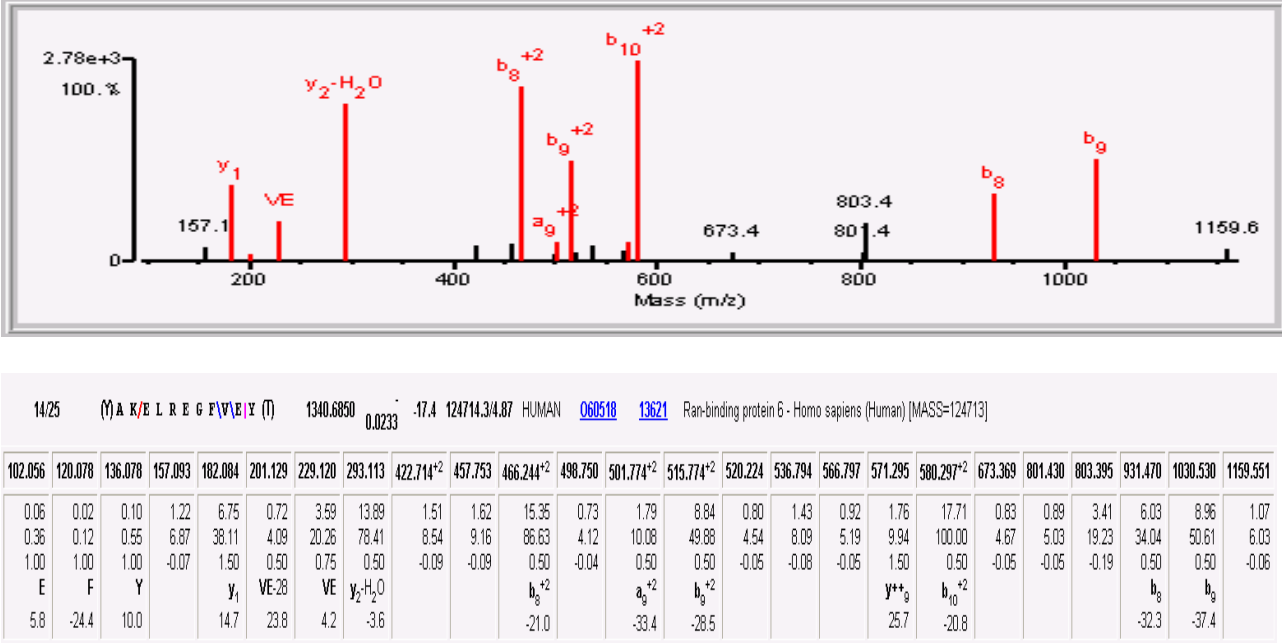
14/25	(N)YFVQTVEVDQL(I)										1340.6738	-0.0120	-9.0	68361.7/5.98	HUMAN	Q6TFL4	8477	Kelch-like protein 24 - Homo sapiens (Human) [MASS=68361]										
102.056	120.078	136.078	157.093	182.084	201.129	229.120	293.113	422.714 ⁺²	457.753	466.244 ⁺²	498.750	501.774 ⁺²	515.774 ⁺²	520.224	536.794	566.797	571.295	580.297 ⁺²	673.369	801.430	803.395	931.470	1030.530	1159.551				
0.06	0.02	0.10	1.22	6.75	0.72	3.59	13.89	1.51	1.62	15.35	0.73	1.79	8.84	0.80	1.43	0.92	1.76	17.71	0.83	0.89	3.41	6.03	8.96	1.07				
0.36	0.12	0.55	6.87	38.11	4.09	20.26	78.41	8.54	9.16	86.63	4.12	10.08	49.88	4.54	8.09	5.19	9.94	100.00	4.67	5.03	19.23	34.04	50.61	6.03				
1.00	1.00	1.50	-0.07	-0.38	0.75	0.75	-0.78	-0.09	-0.09	1.50	-0.04	-0.10	1.50	-0.05	-0.08	-0.05	0.75	0.50	-0.05	-0.05	1.50	1.50	1.50	0.50				
E	F	a ₁			TV	EV				y ₈ ⁺²			y ₉ ⁺²				VEVDQ	y ₁₀ -H ₂ O ⁺²			y ₇	y ₈	y ₉	y ₁₀ -H ₂ O				
5.8	-24.4	10.0			23.8	4.2				7.4			-2.8				39.1	-11.2			-24.7	-3.8	-11.7	-42.6				
		1.50			EV-28	23.8																						
		Y																										
		10.0																										

2.YFVPPPYELSE



B. Third option gave one sequence with two basic residues close to N-terminal justifying the presence of abundant b-ions in the spectra but the high error in ppm for this b-type ions with majority of them over 20ppm (last line of report) forces to discard it:

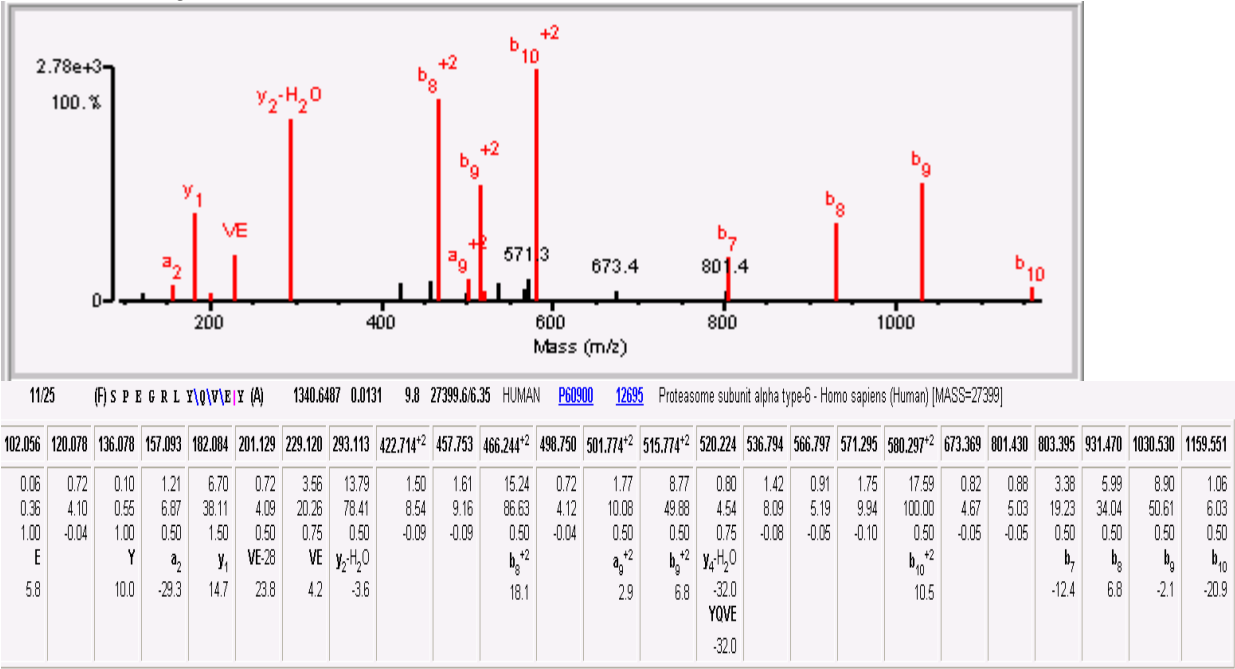
3.AKELREGFVEY



C. Finally the fourth option seems comfortable by the presence of Arg closer to N-terminus and supporting the abundance in b-ions which principal fragments present error less than 20. Also two internal ions match with the

trend of few internal ions when Arg are inside the peptide. Then this sequence is accepted as valid and additionally it contains the motif characteristic of the HLA-B3501. The reason of the poor score for this peptide compared with the first three is focused on the difference in value assignment to b-ion and y-ions by the algorithm. If the values were equal for both ions in the calculations the score the fourth option would have been 17.5 units, one excellent score in the scale for this algorithm.

4.SPEGRLYQVEY



Escobar et al. SUPPORTING MATERIAL PART 2.

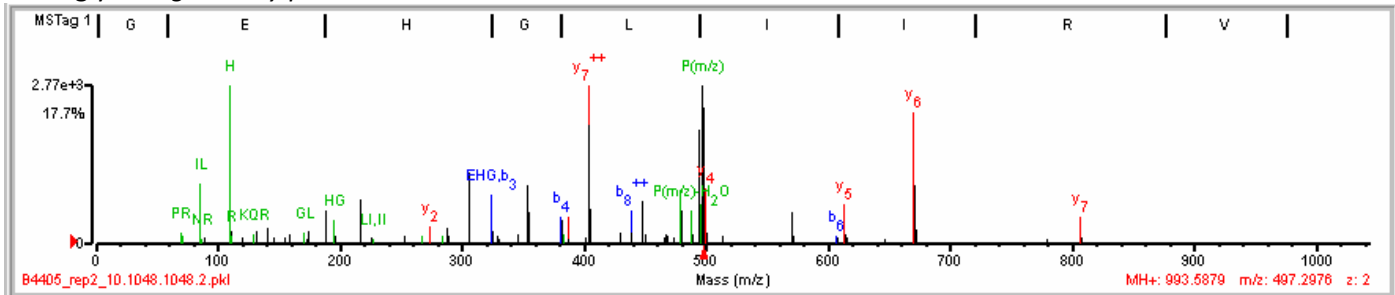
Utility of characteristic QTOF MS/MS fragmentation for MHC class I peptides.

A. QTOF spectra MHC class I peptides containing Histidine. y-ions, b-ions and internal ions C-terminus to His

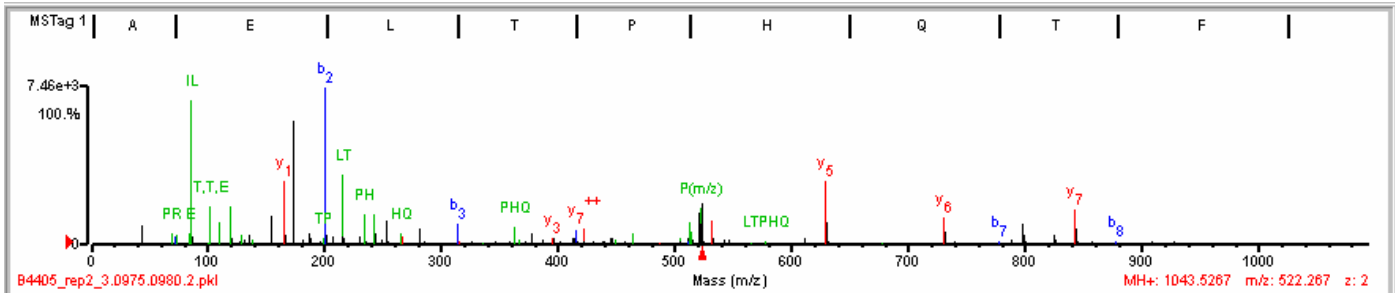
(Please note the software design does not default display all ion signals in the follow spectra. Complete or specific count of ions may require additional manipulations.)

HLA-B4405

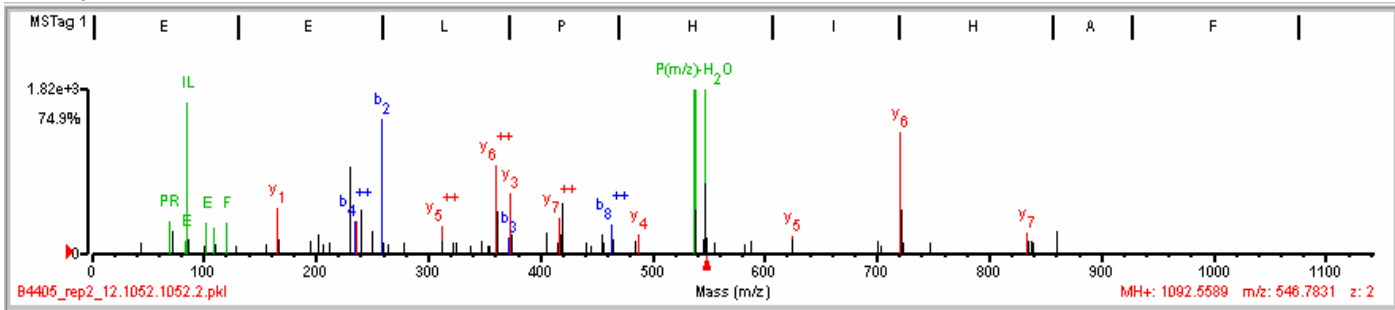
Strong y6. *Arg and Gly presence.*



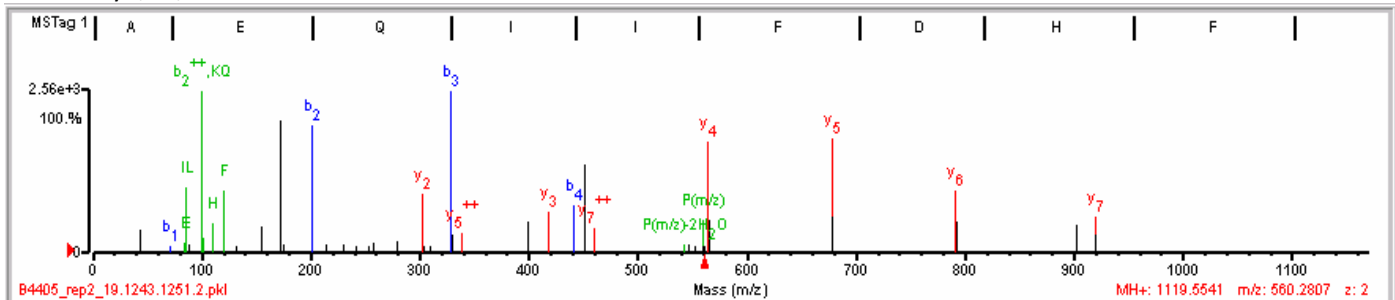
Weak y3, PH



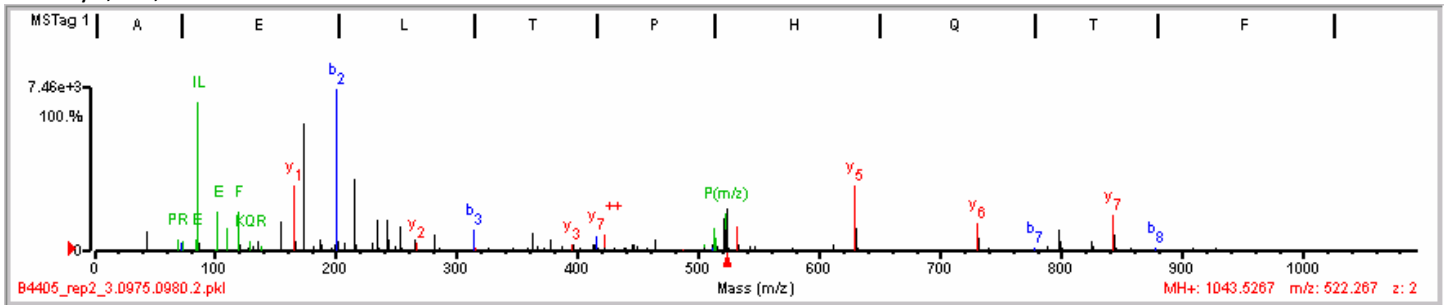
Weak y2, PHIH, LPH



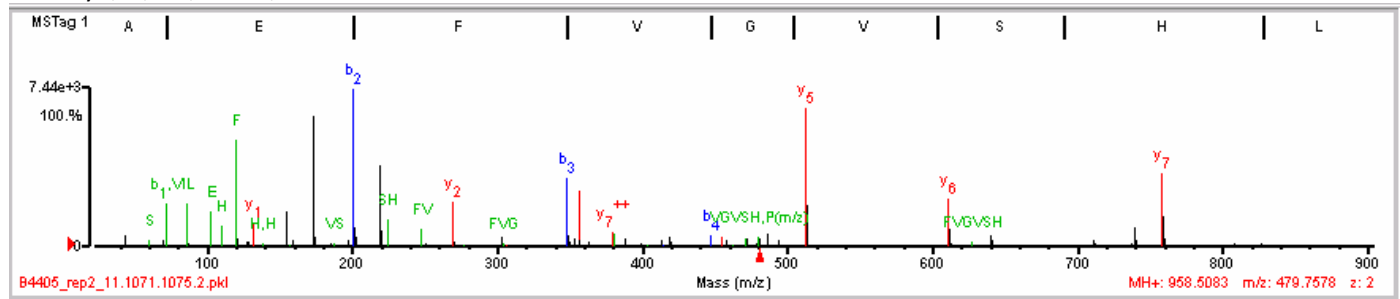
Absence y1, b8, FDH



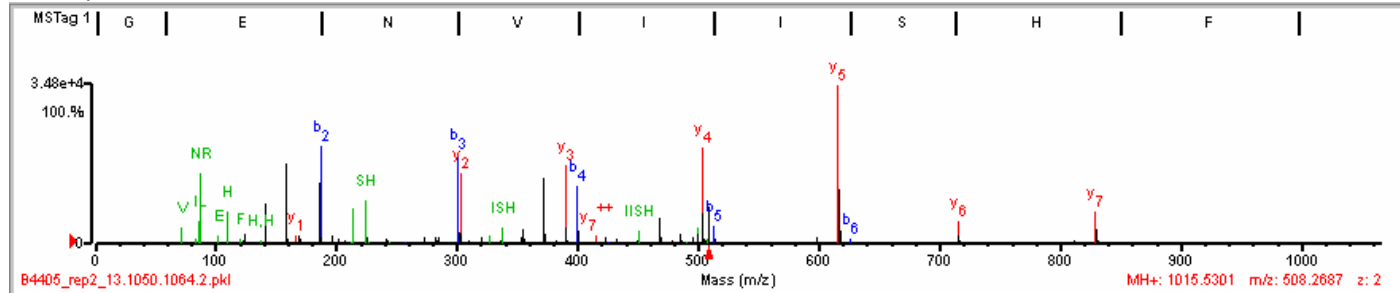
Weak y3, PH, HQ



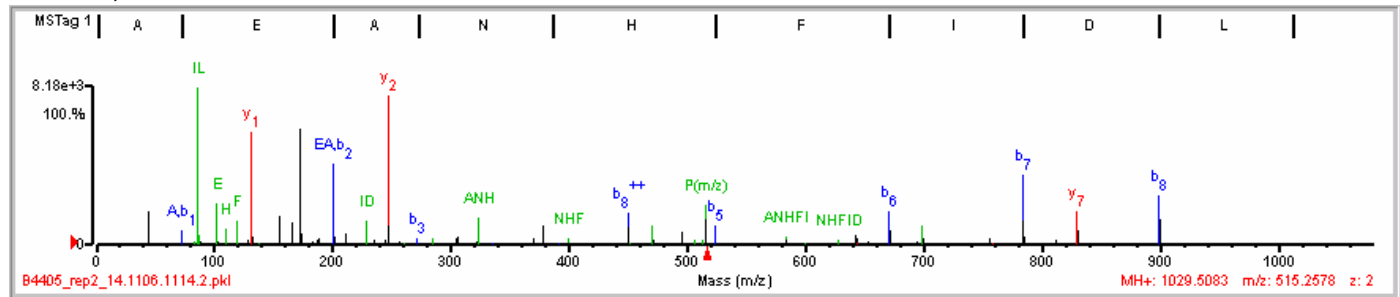
Weak y1, H, SH, VGSH, FVGVS



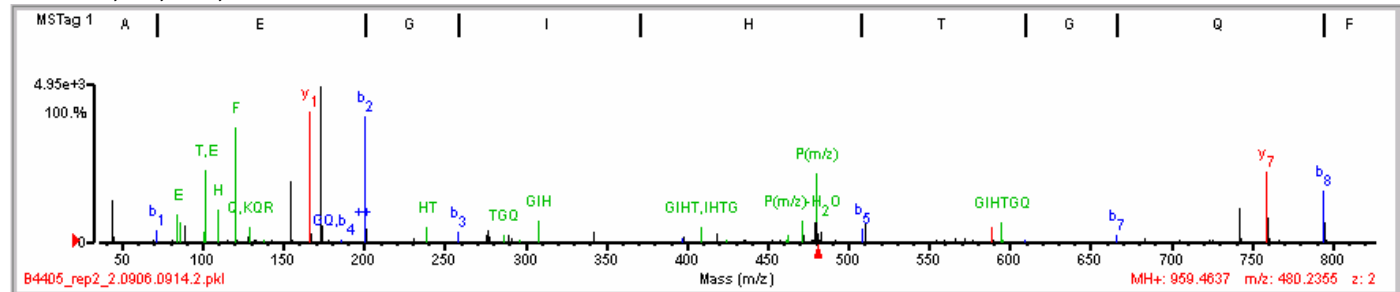
Weak y1, SH, ISH, IISH



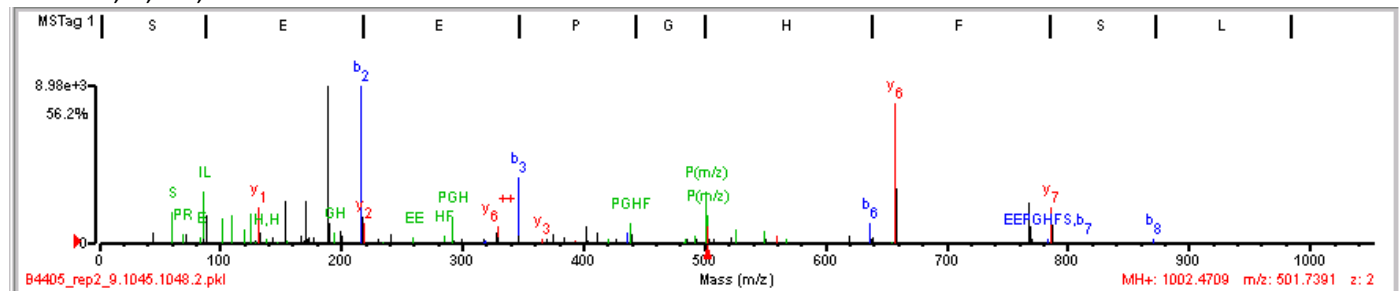
Weak b5, ANH



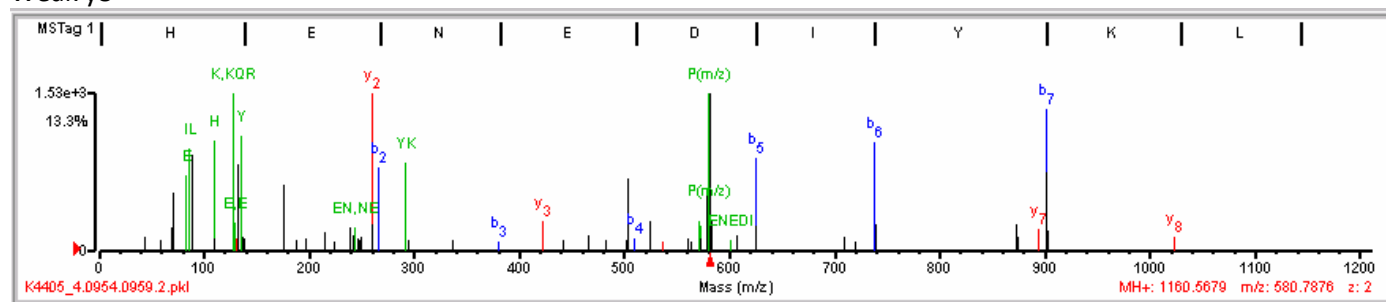
Weak b5, HT, GIH, GIHT



Weak b6, H, GH, PGH

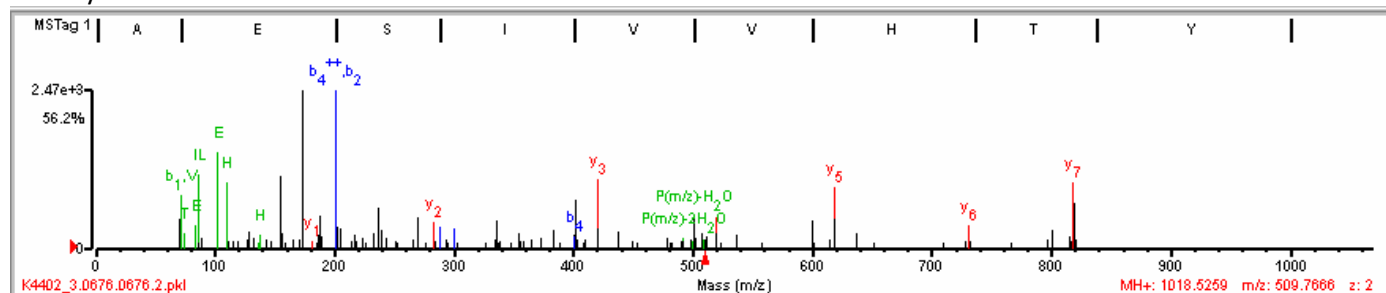


Weak y8

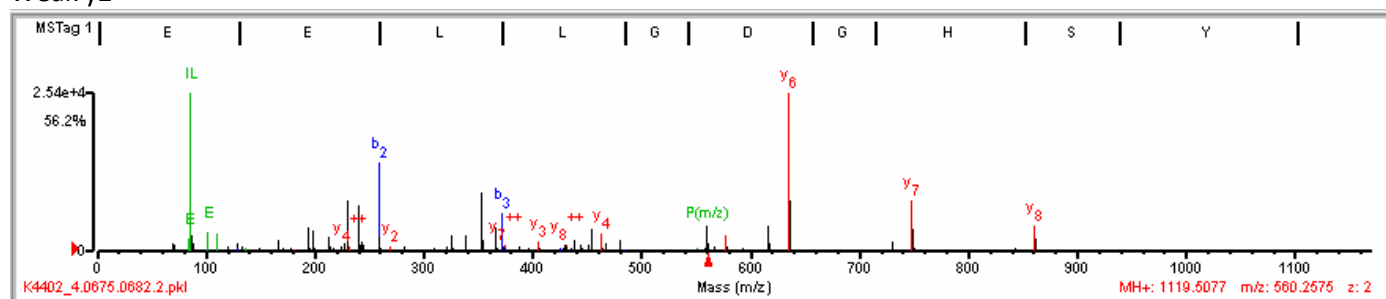


HLA-B4402

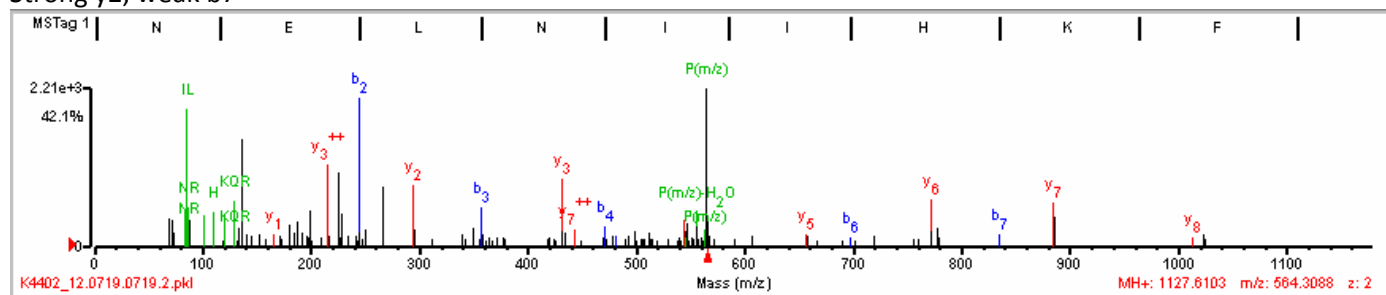
Weak y2



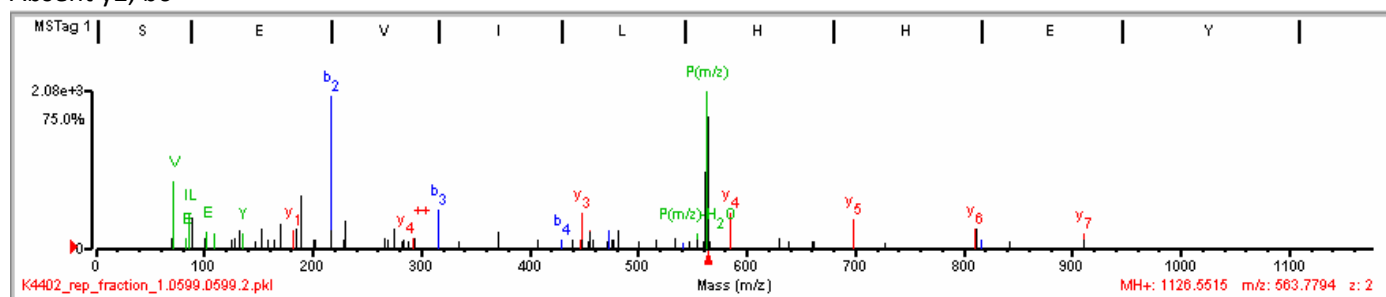
Weak y2



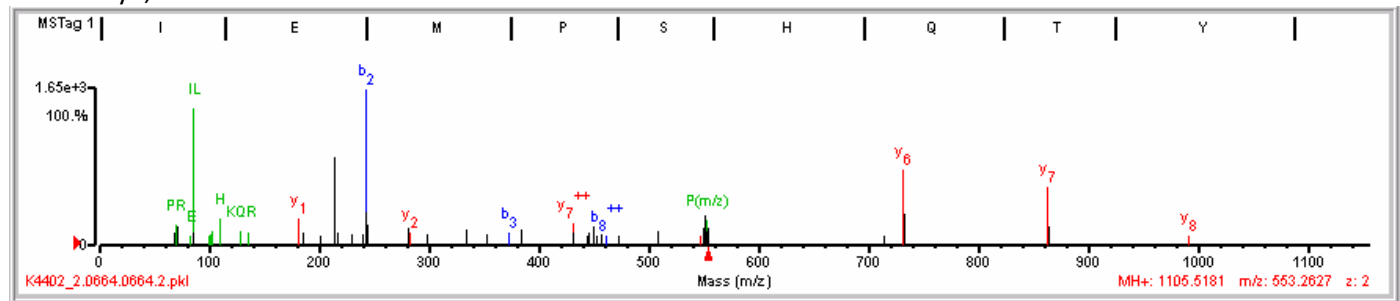
Strong y2, weak b7



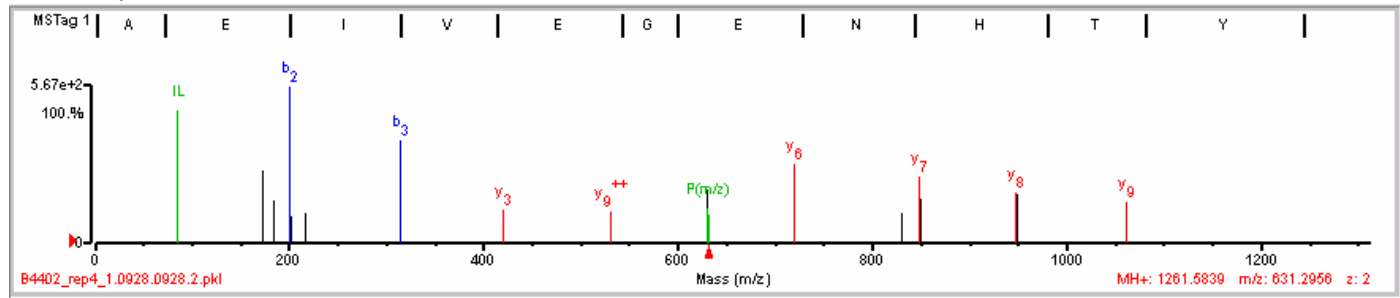
Absent y2, b6



Absence y3, b6

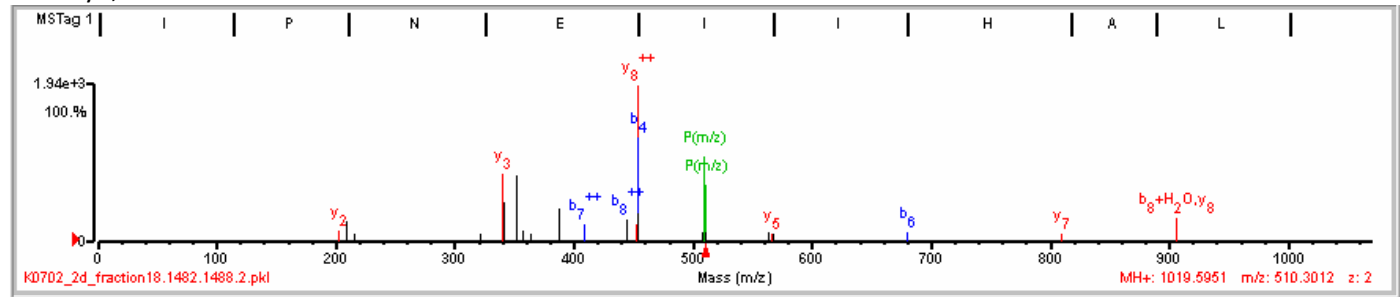


Absence y2, b9



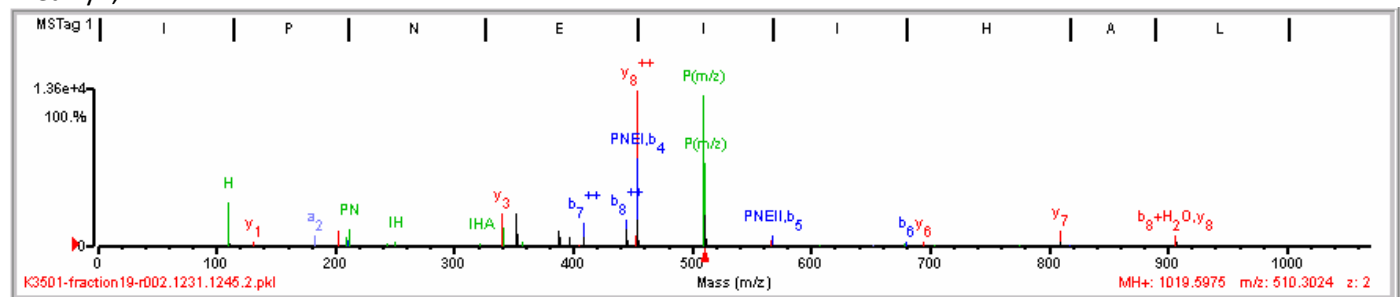
HLA B0702

Weak y2, b7++

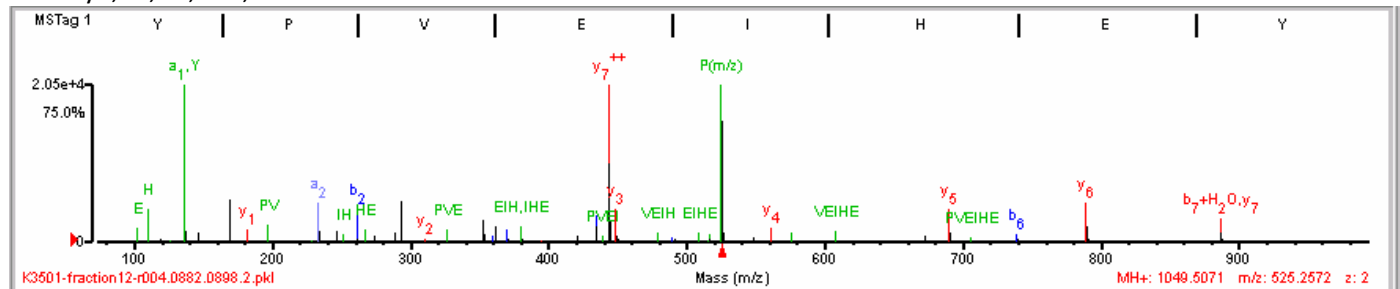


HLA-B3501

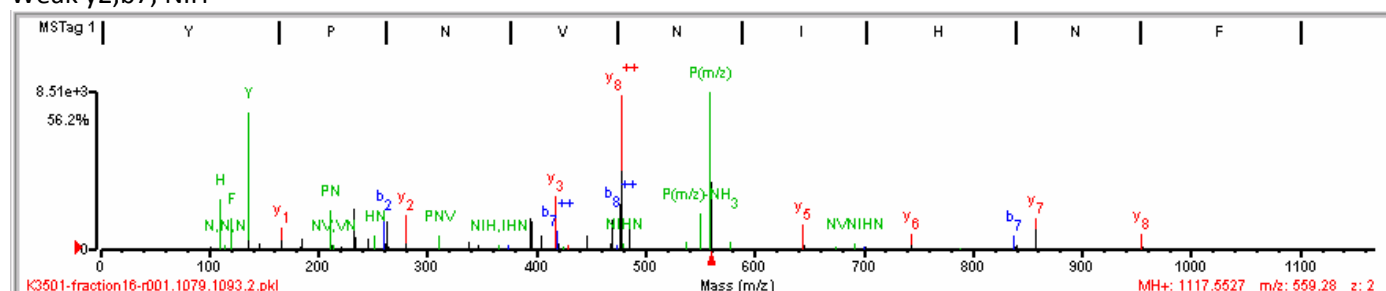
Weak y2, IH



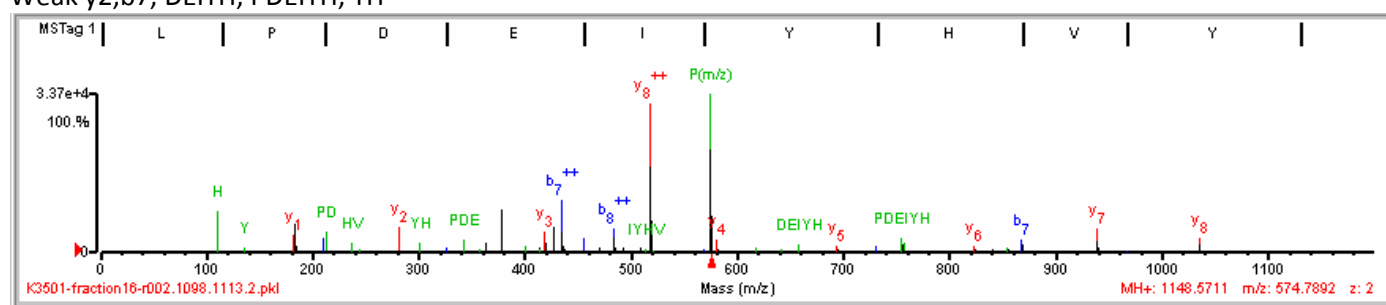
Weak y2,b6, IH, EIH, VEIH



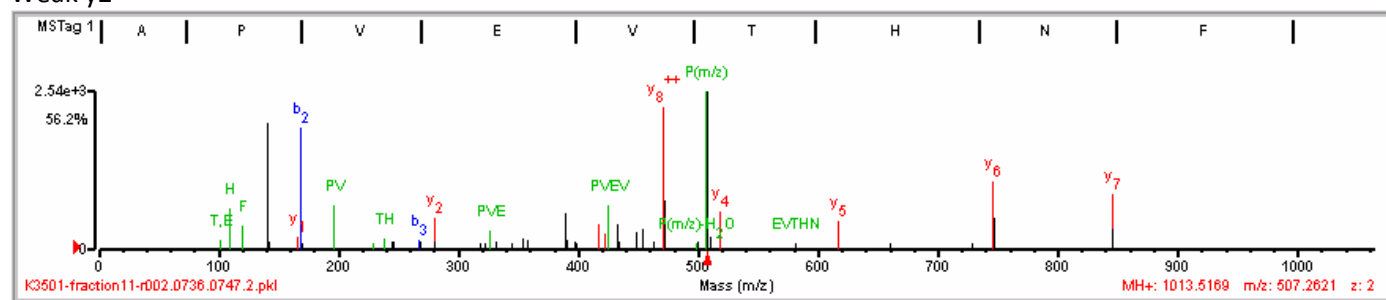
Weak y2,b7, NIH



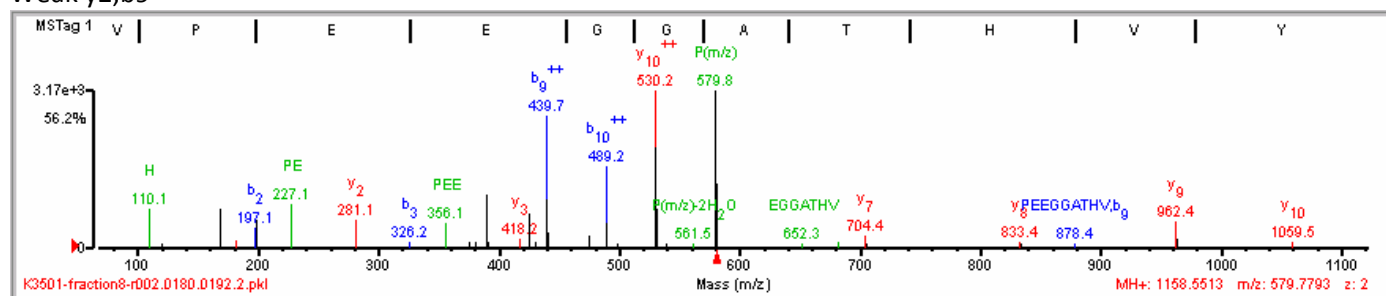
Weak y2,b7, DEIYH, PDEIYH, YH



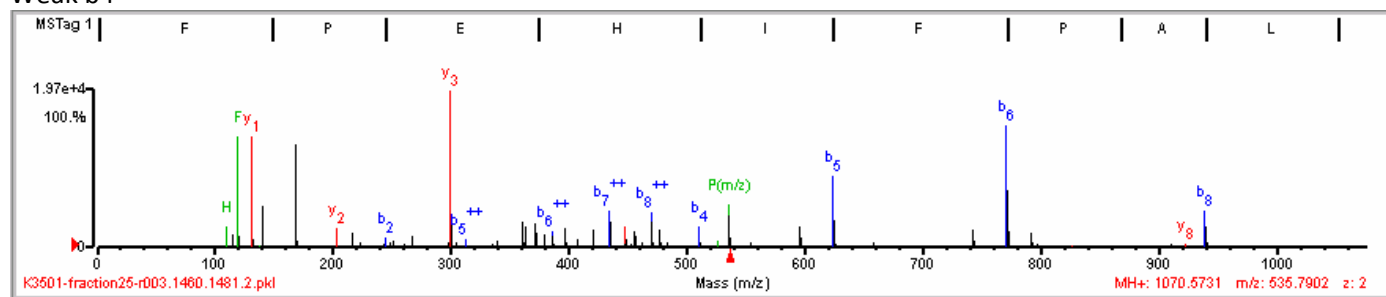
Weak y2



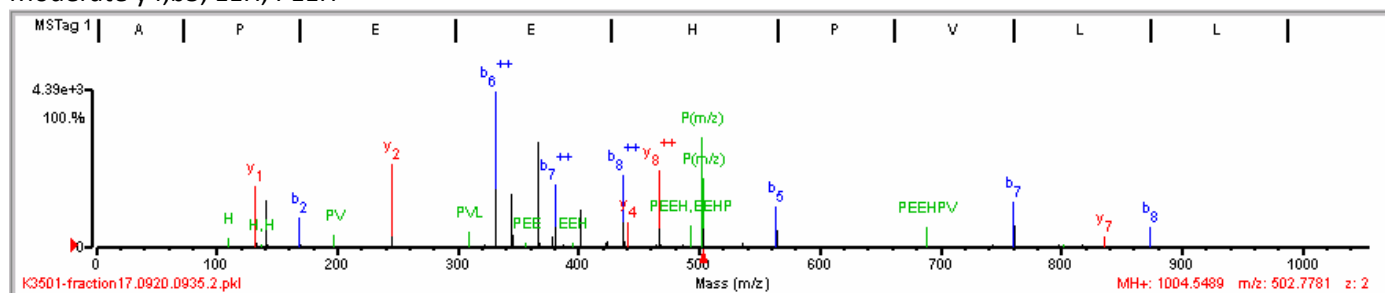
Weak y2,b9



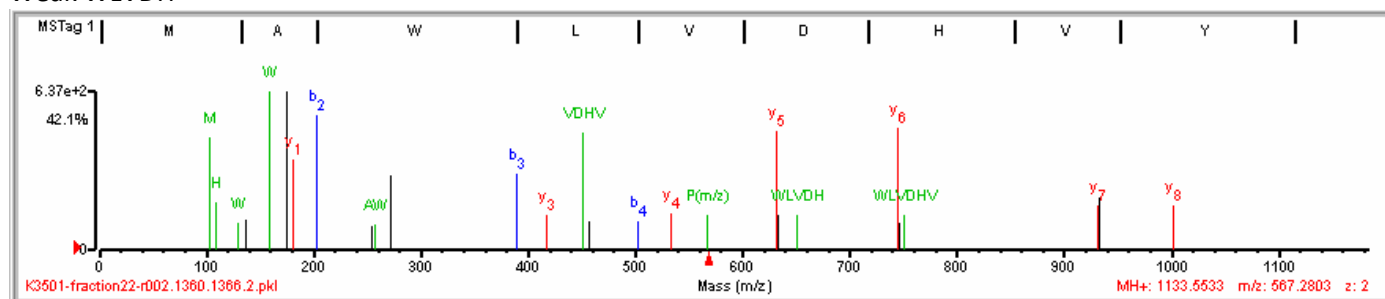
Weak b4



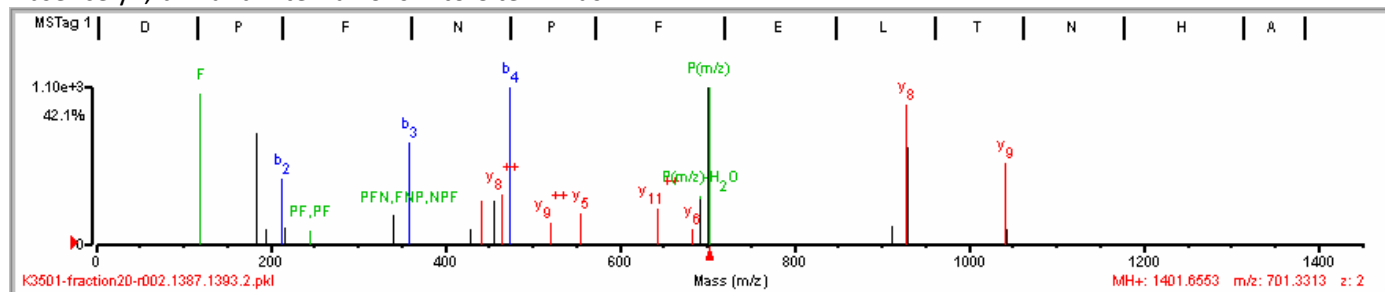
Moderate y4,b5, EEH, PEEH



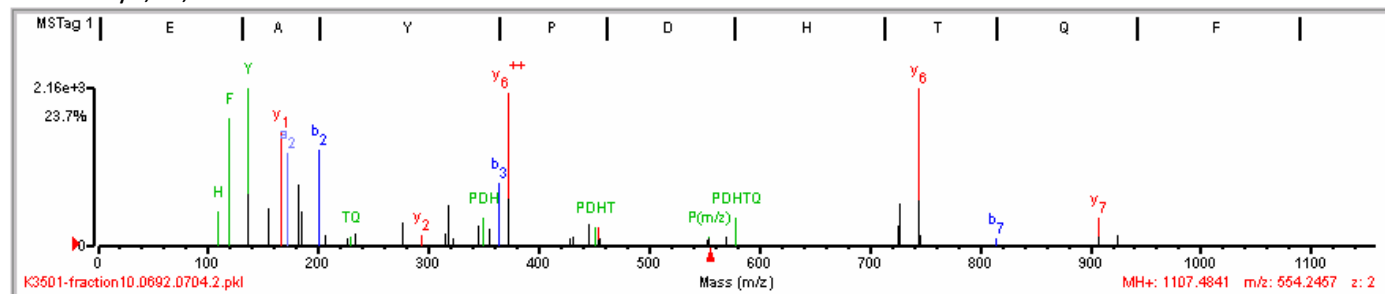
Weak WLVDH



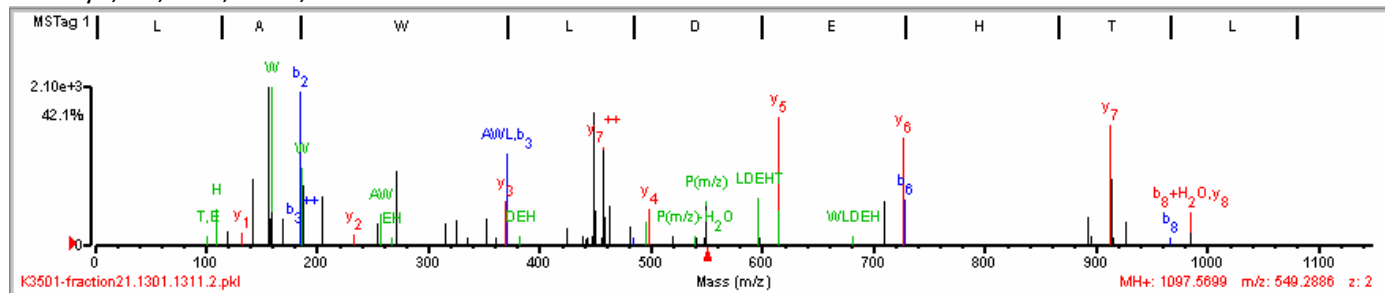
Absence y1, b11 and internal ions H to C terminus



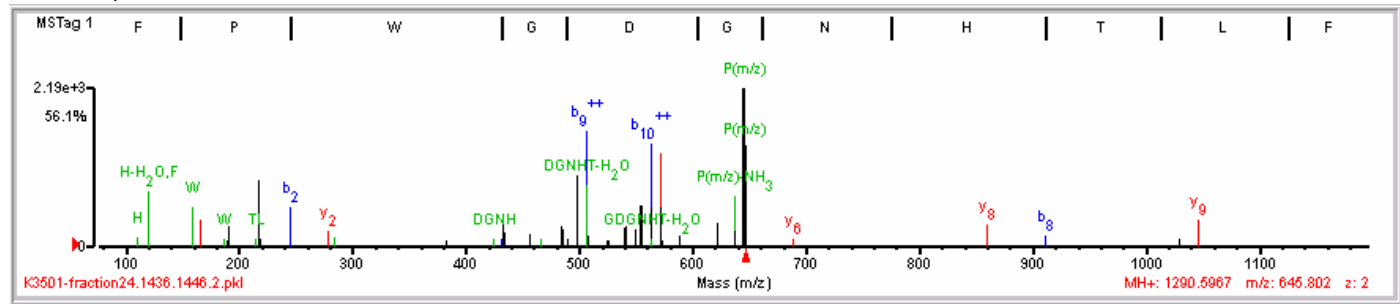
Absence y3,b6, PDH



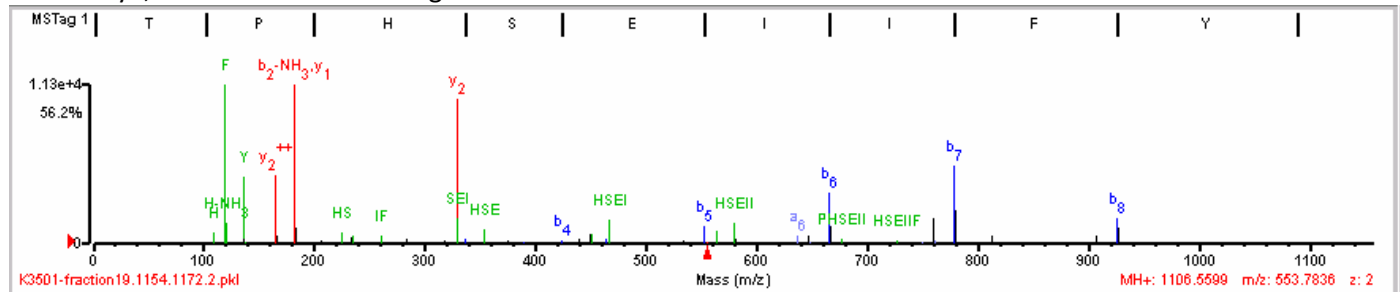
Weak y2, EH, DEH, LDEH, WLDEH



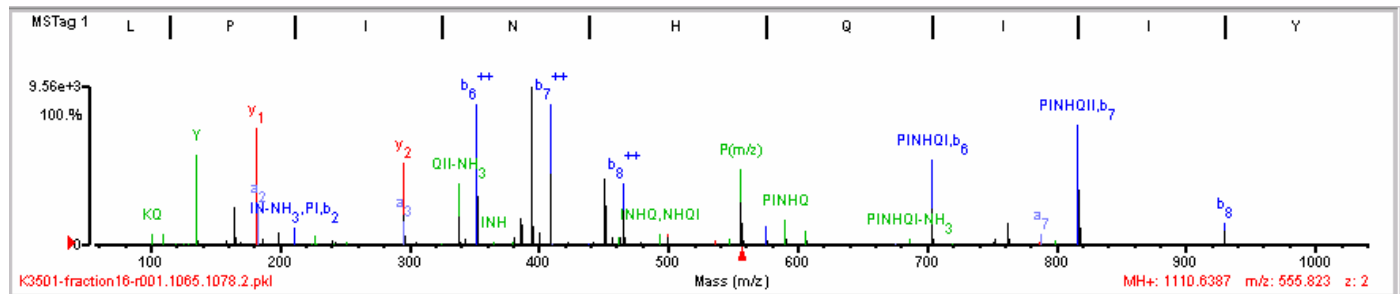
Weak b8, DGNH



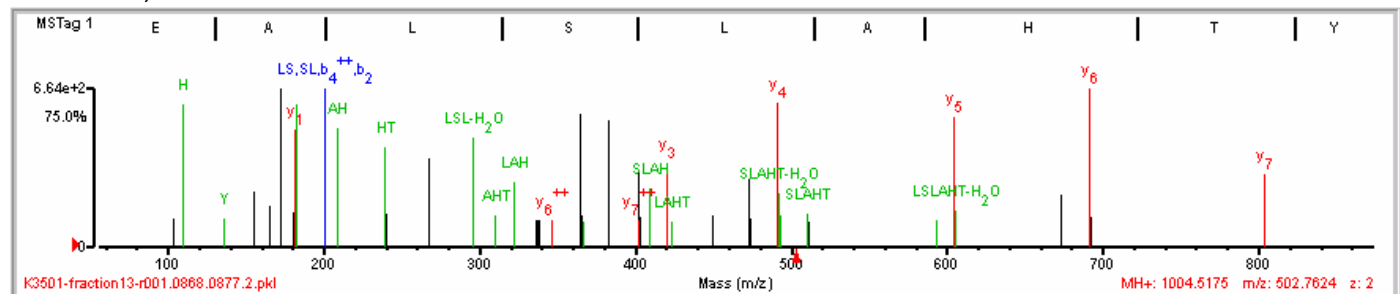
Absence y6,b3 and internal ions fragments H to C terminus



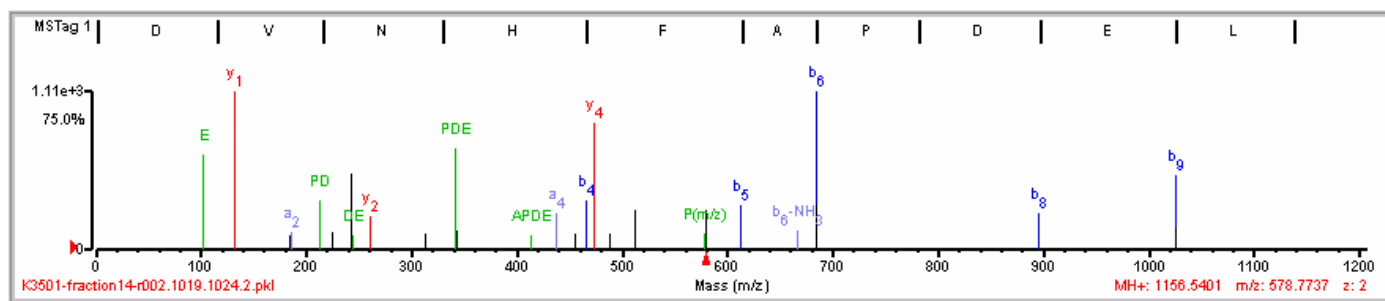
Weak INH



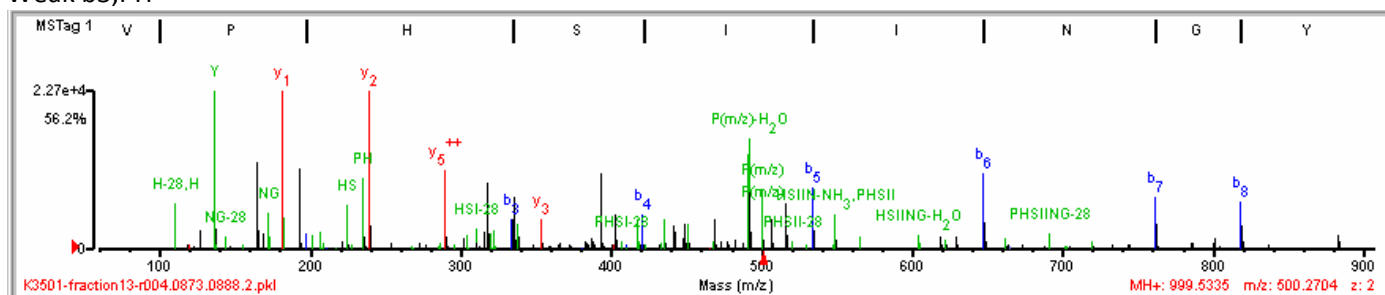
Weak SLAH, LAH



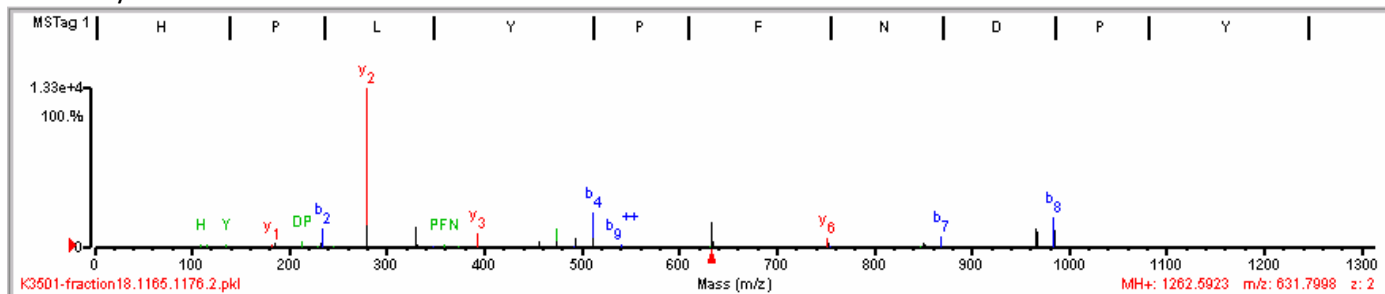
Weak b4



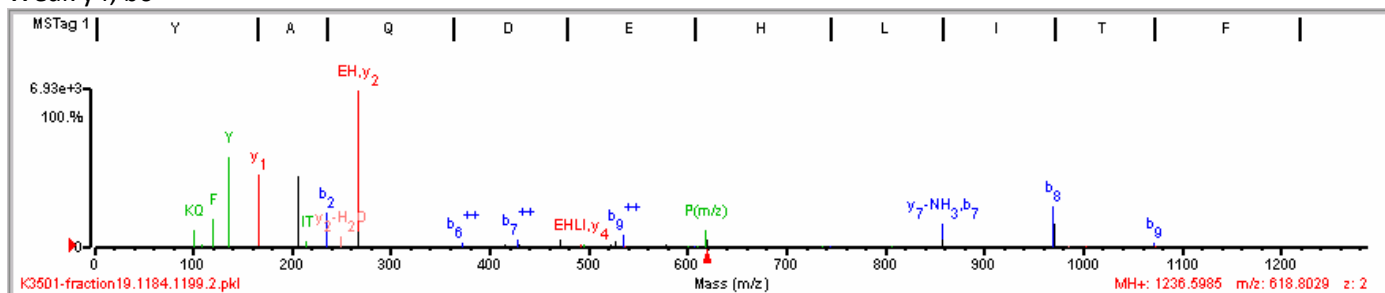
Weak b3,PH



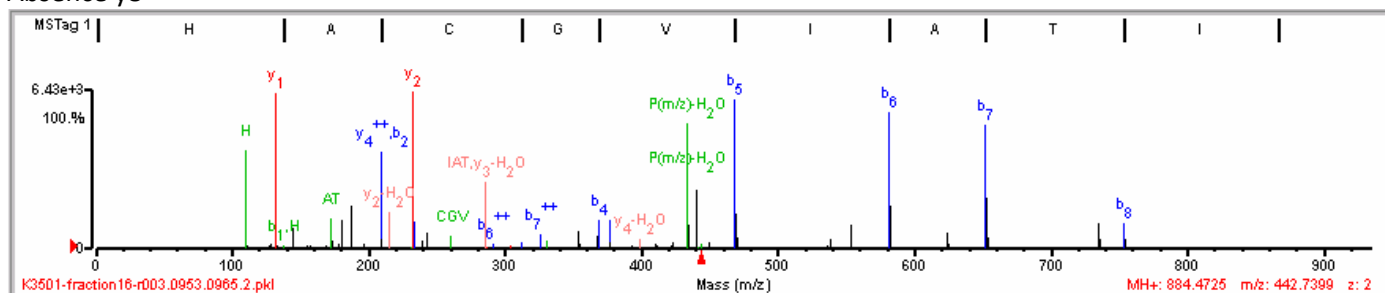
Absence y9



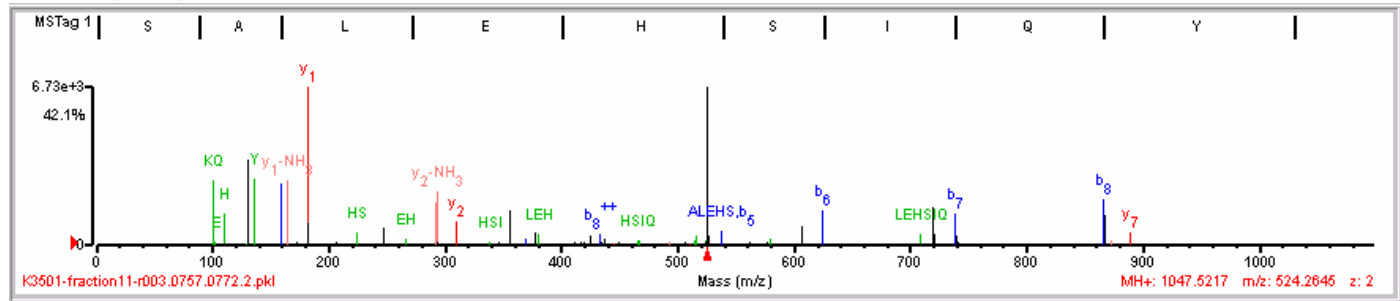
Weak y4, b6++



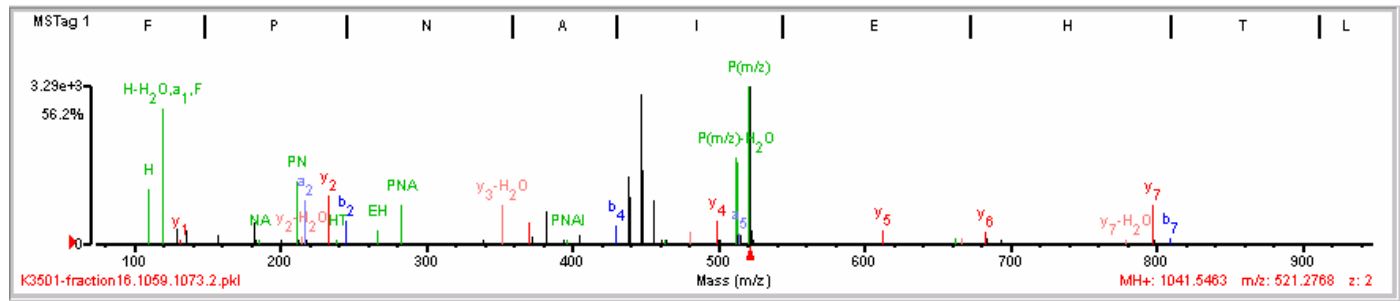
Absence y8



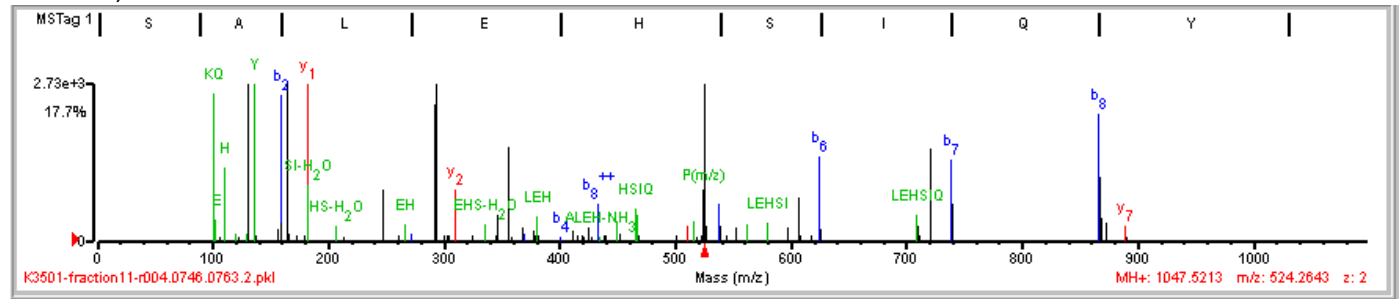
Weak b5, LEH, EH



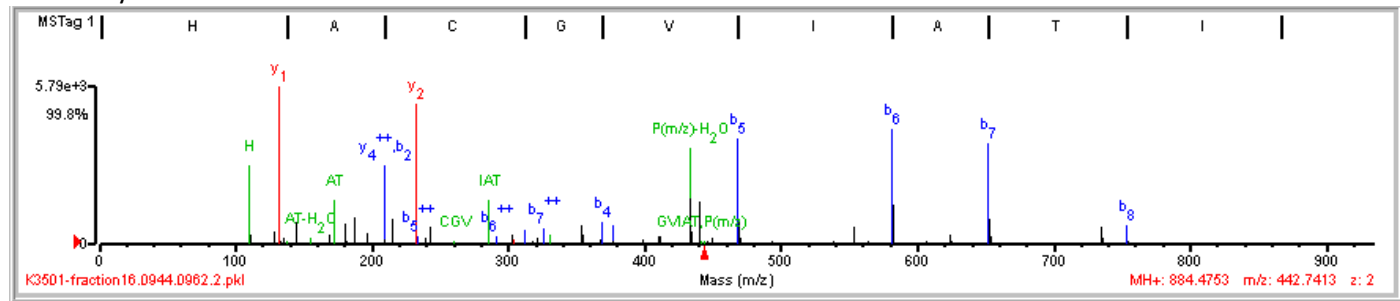
Weak EH



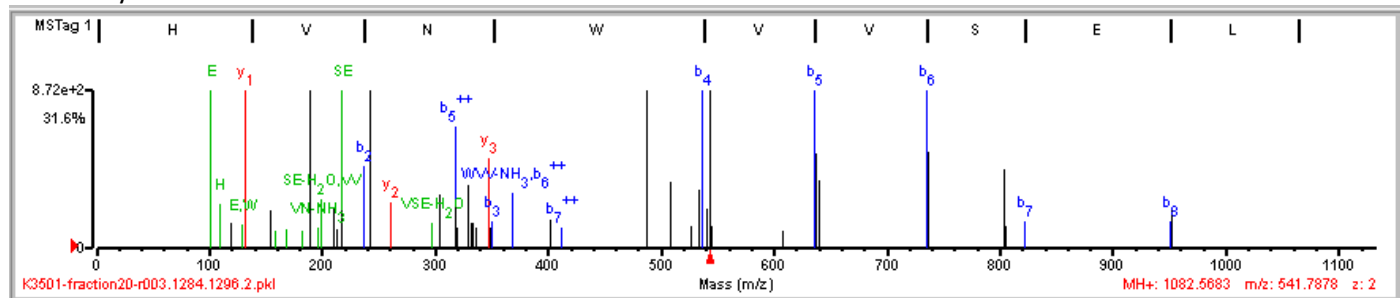
Weak EH, LEH



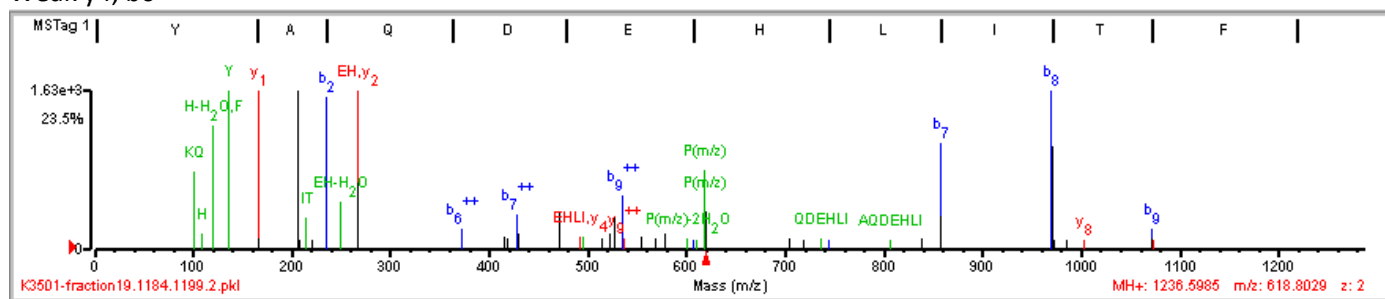
Absence y8



Absence y8

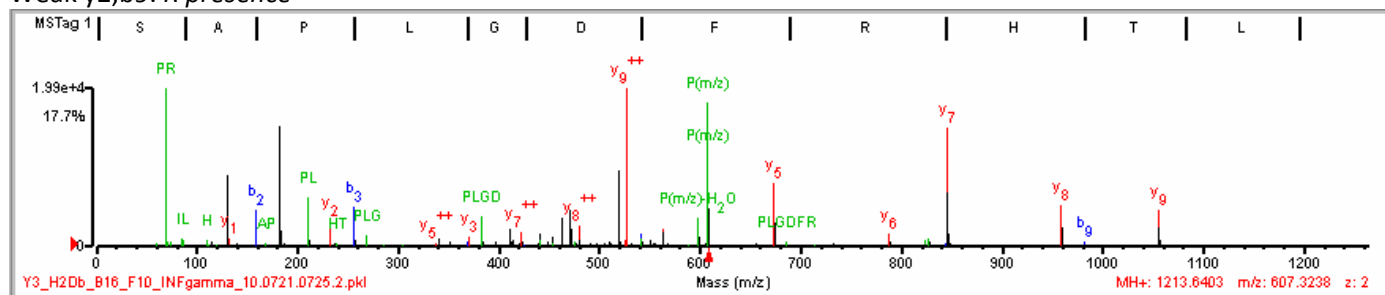


Weak y4, b6++

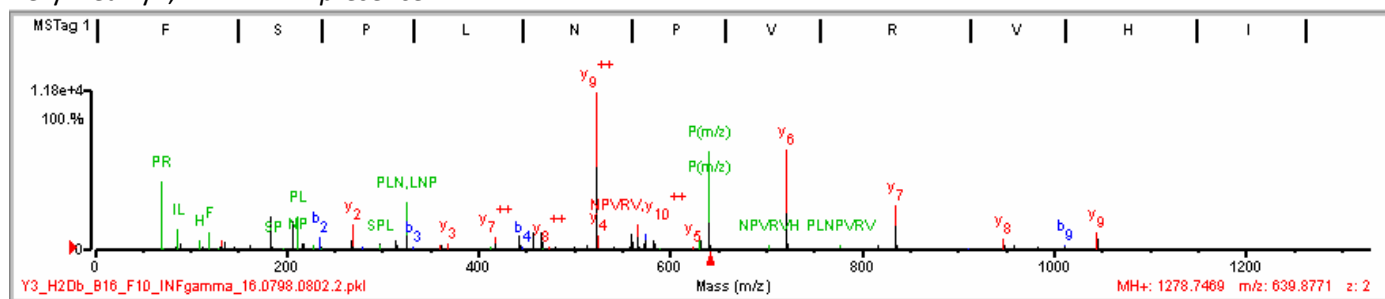


H2-Db

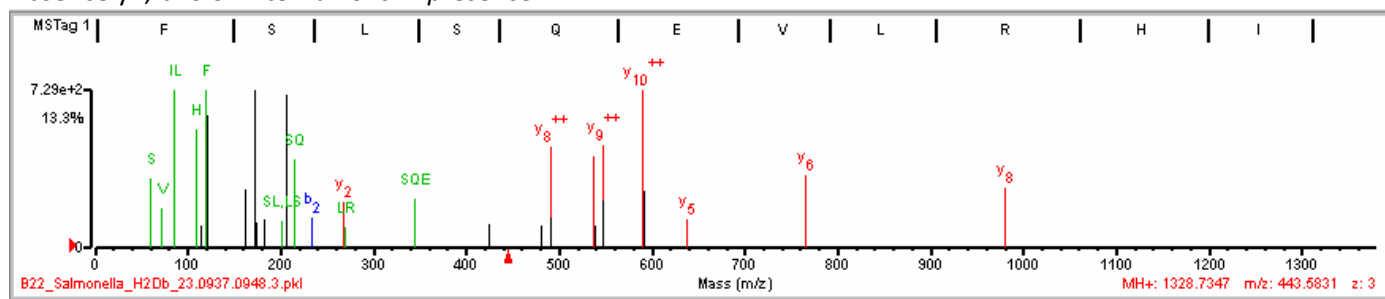
Weak y2,b9. *R* presence



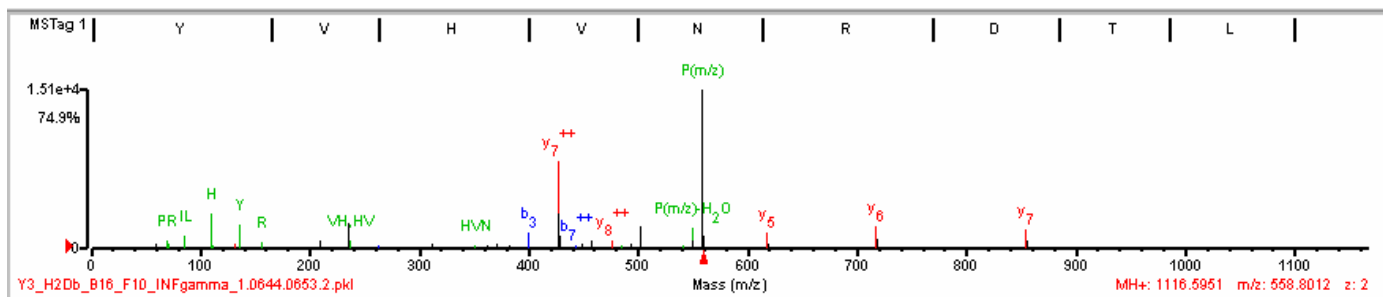
Very weak y1, NPVRVH. *R* presence



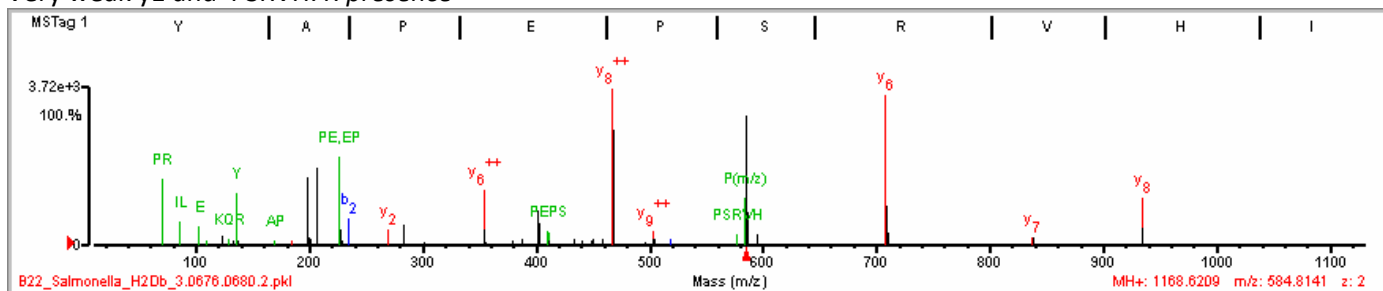
Absence y1, b10 or internal ions. *R* presence



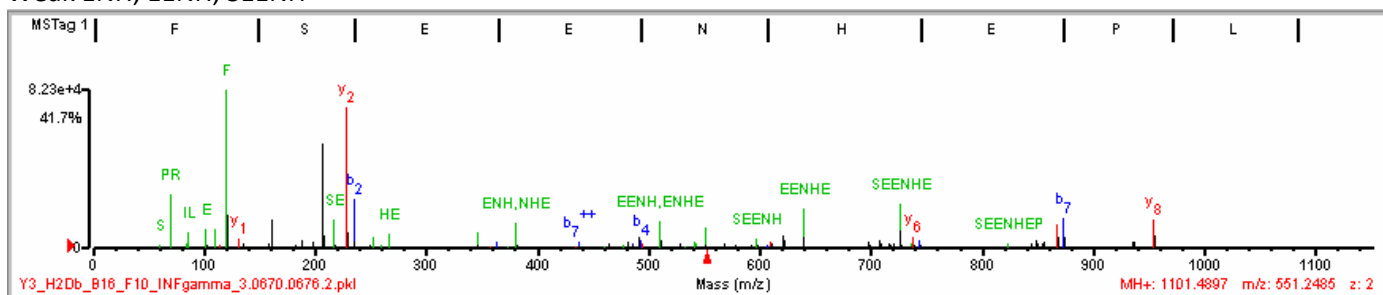
Weak y6, b3. *R* presence



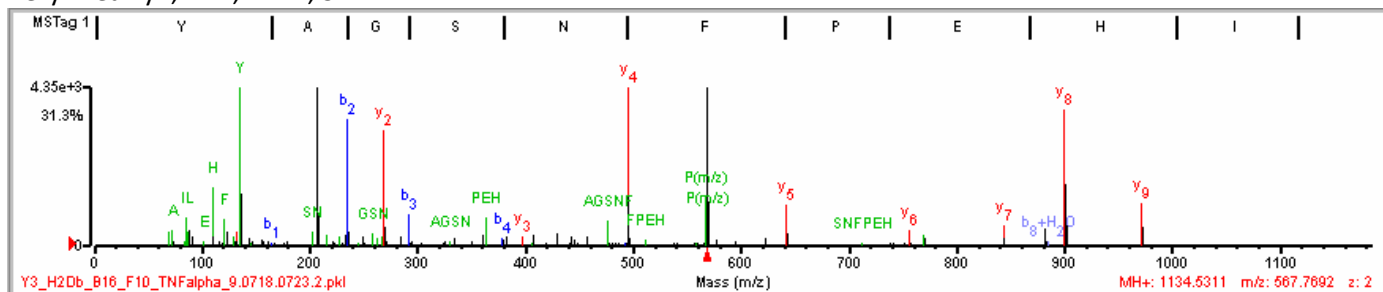
Very weak y1 and PSRVH. *R* presence



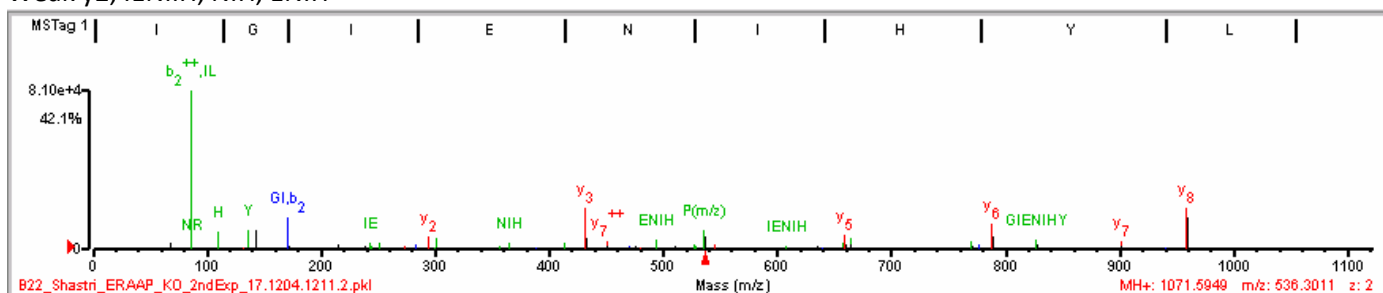
Weak ENH, EENH, SEENH



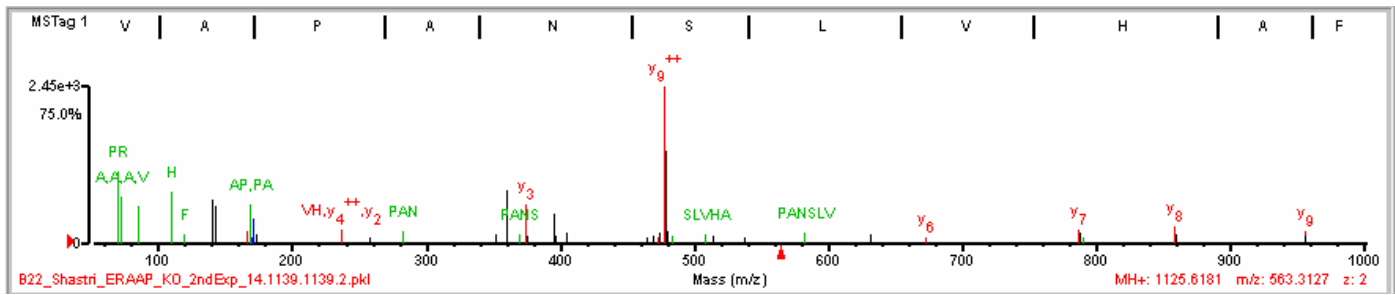
Very Weak y1, PEH, FPEH, SNFPEH



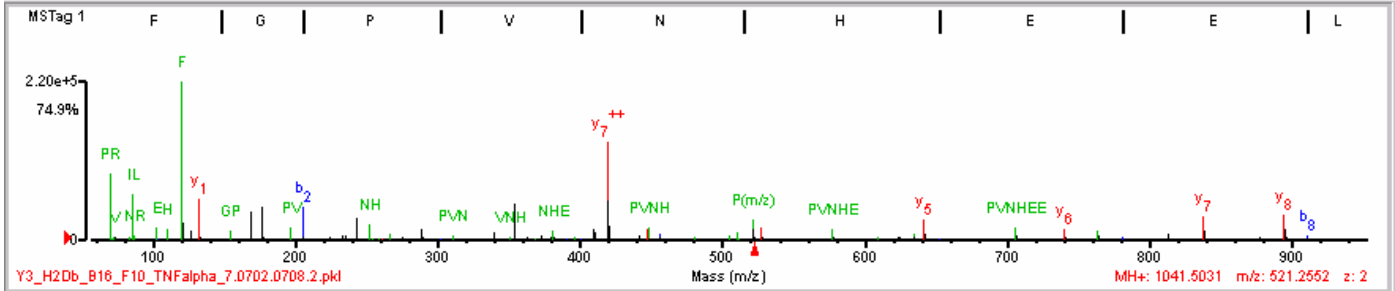
Weak y2, IENIIH, NIH, ENIH



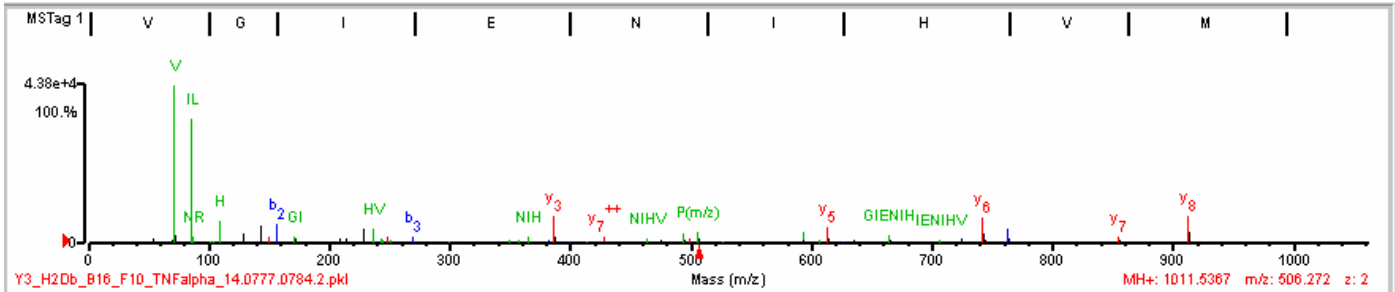
Weak y2



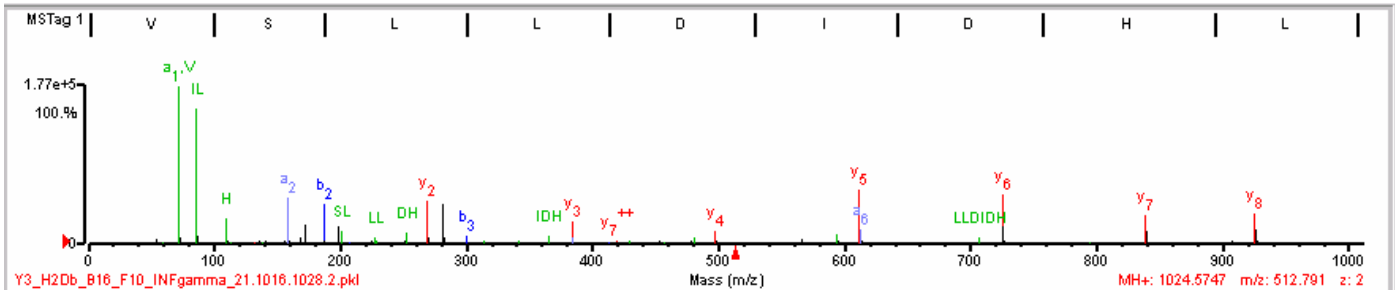
Weak NH, VNH, PVNH



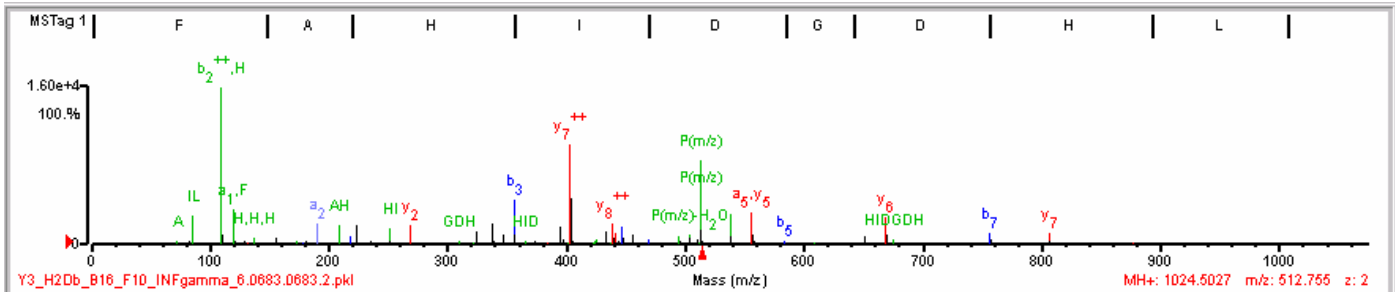
Weak y2, b7, NIH, GIENIH



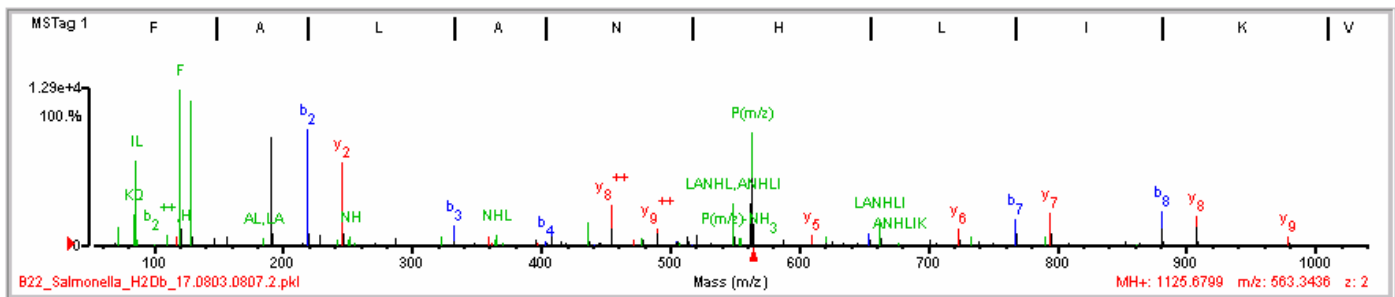
Weak DH, IDH, LLDIDH



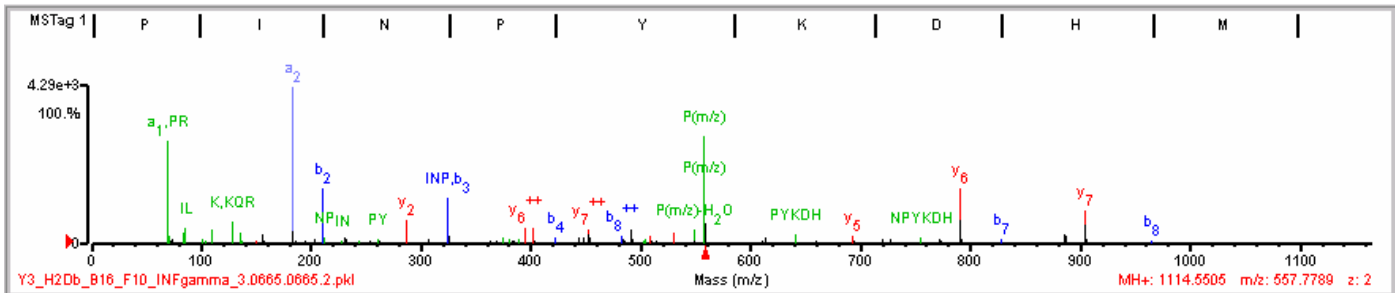
Weak H, AH, HI, GDH, HIDGDH



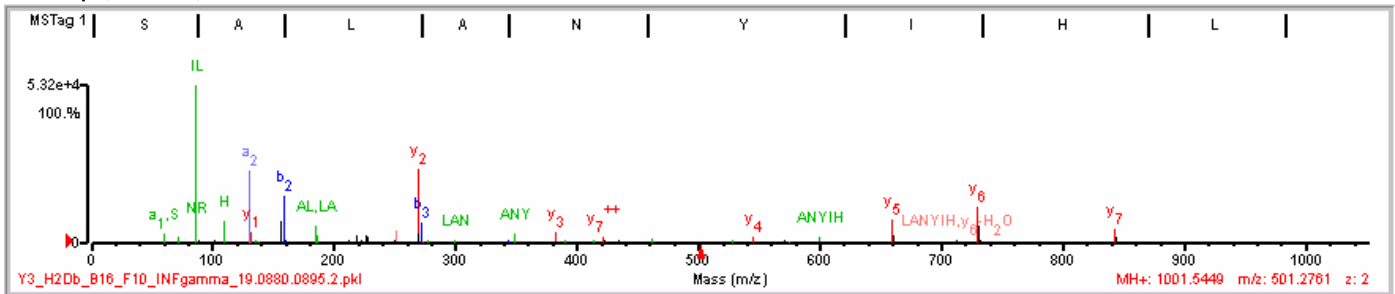
Weak NH



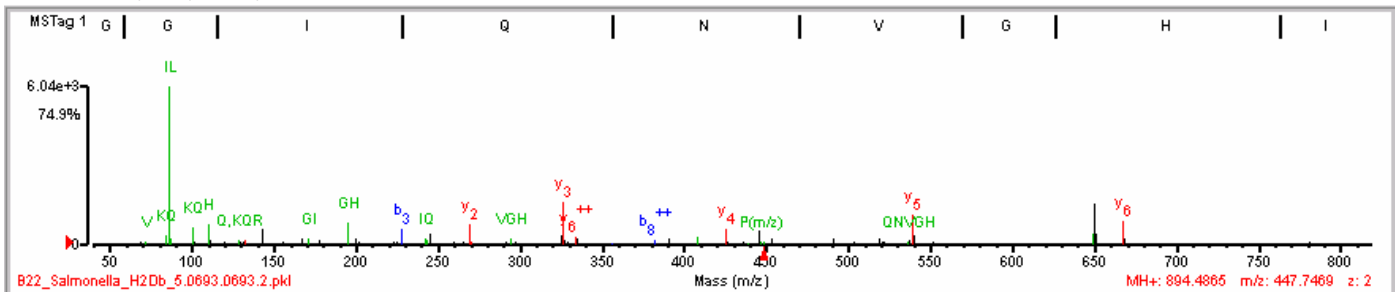
Weak PYKDH, NPYKDH



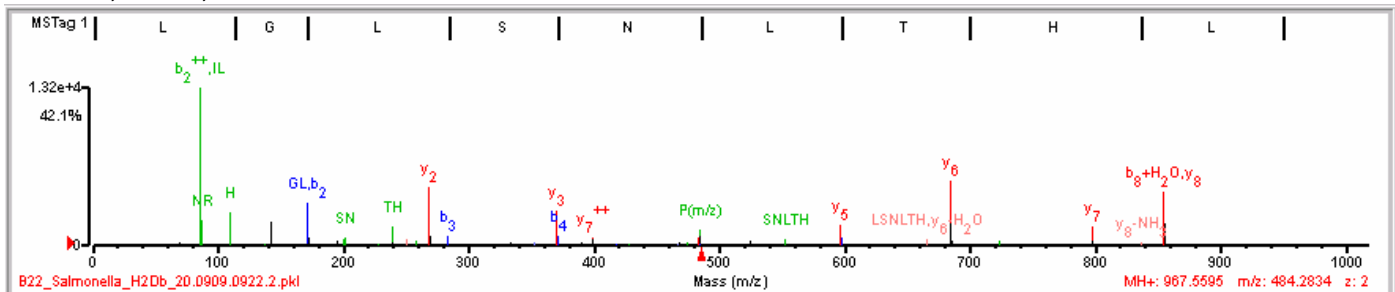
Weak y1, ANYIH, LANYIH



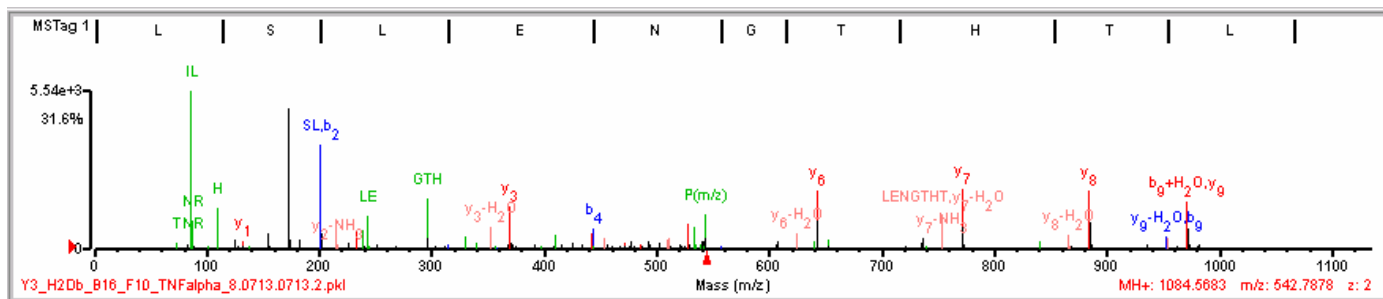
Weak b8++, GH, VGH, QNVGH



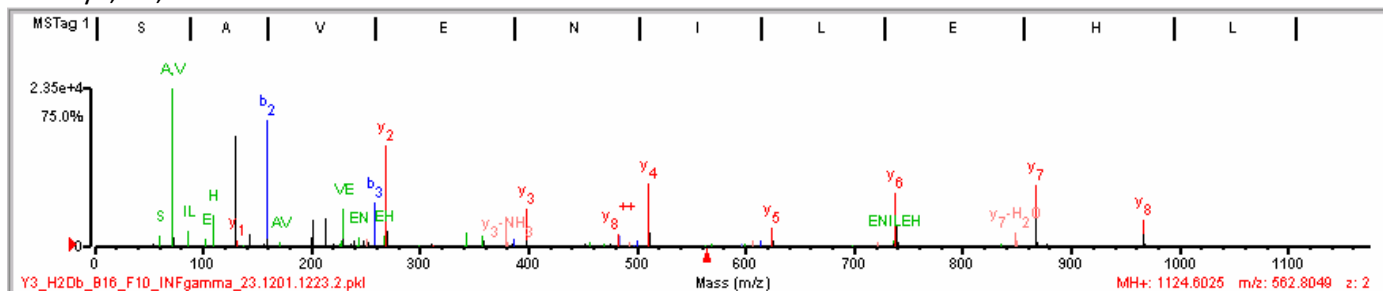
Weak TH, SNLTH, LSNLTH



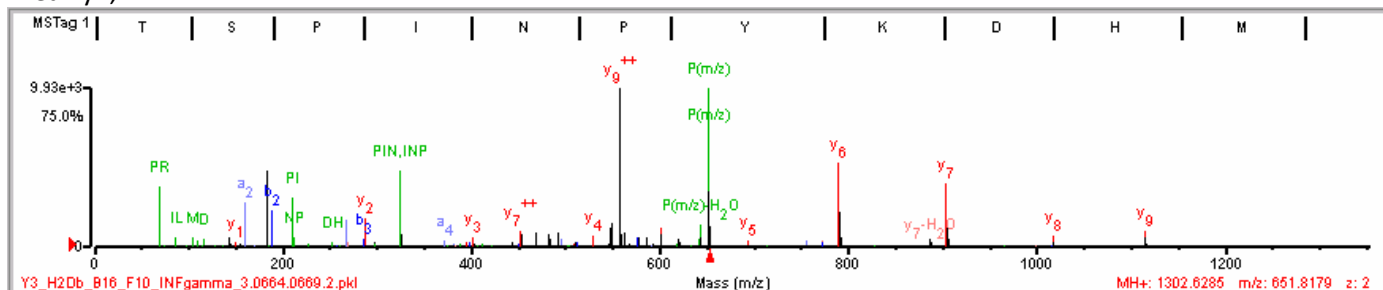
Weak y2, GTH



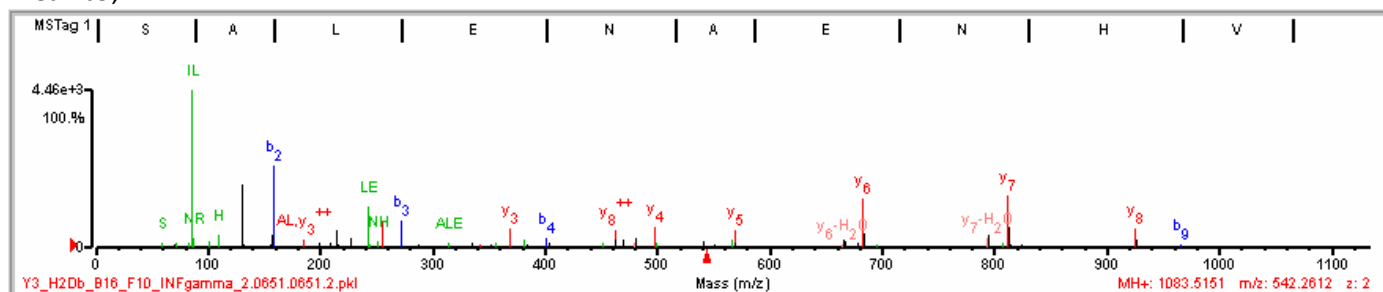
Weak y1, EH, ENILEH



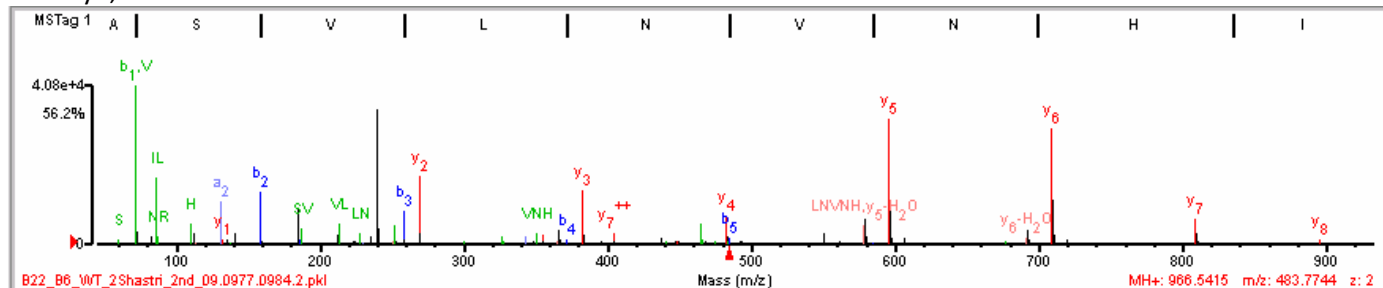
Weak y1, DH



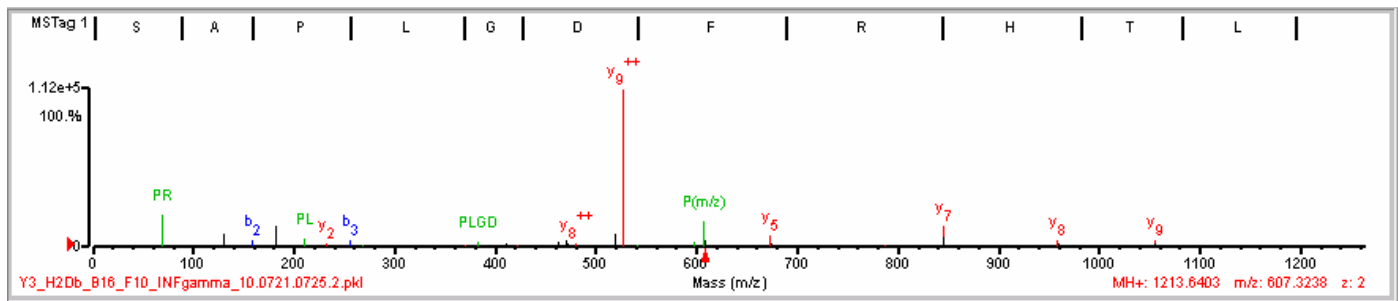
Weak b9, NH



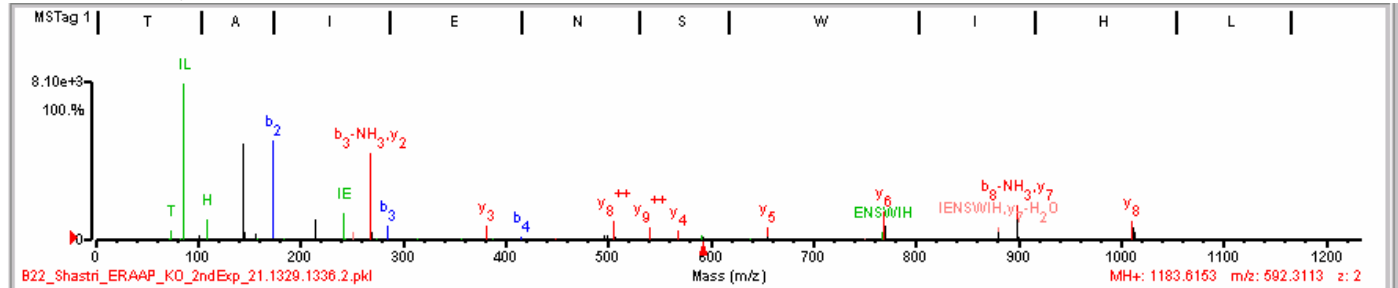
Weak y1, VNH



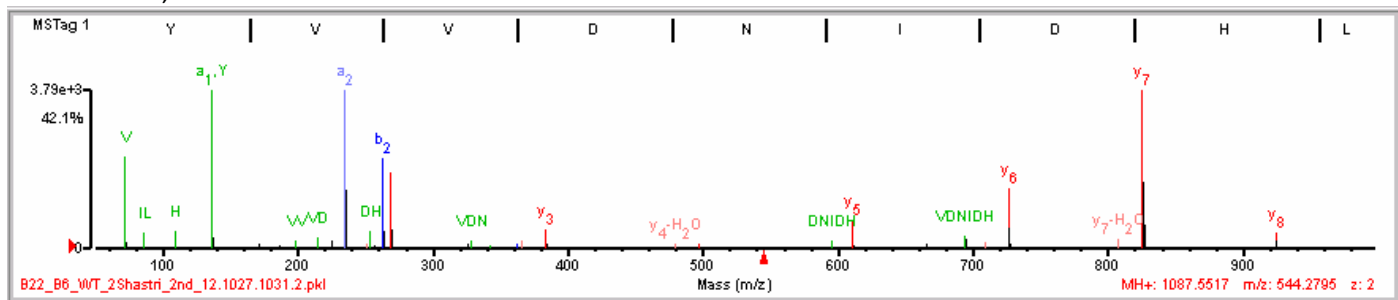
Weak y2



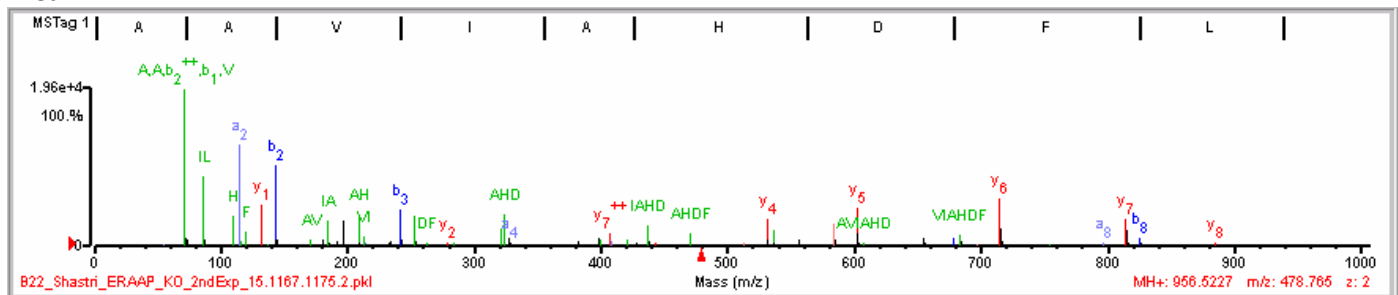
Weak ENSWIH, IENSWIH



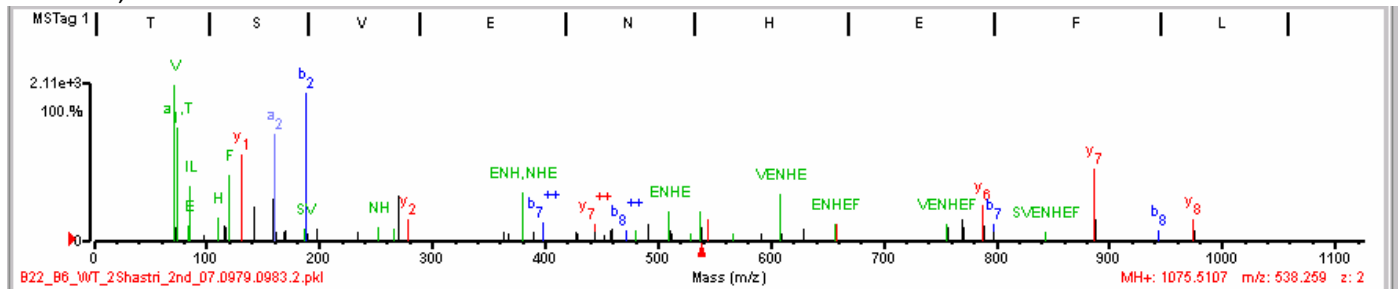
Weak DNIDH, VDNIDH



Weak AH



Weak NH, ENH



Mass spectrum plot showing relative intensity (0 to 1.03e+3) versus Mass (m/z) (0 to 1100). The base peak is at m/z 47. The spectrum is labeled with various peaks and their corresponding amino acid sequences: S-H₂O, PR; A, A, A, V, H; b₂⁺⁺, IL; AP, PA; b₂; y₉⁺⁺; y₃; NSLV-H₂O; PANS-H₂O; SLVHA; PANSLV; y₆; APANSLVH; y₇; PANSLVH; y₈; y₉. The x-axis is labeled "Mass (m/z)" and the y-axis is labeled "Relative intensity". The plot is titled "B22_Shastri_ERAAP_KO_2ndExp_14.1139.1139.2.pkl". The bottom right corner displays "MH+: 1125.6181 m/z: 563.3127 z: 2".

MS/MS spectrum of the protein B22_Shastrin_ERAAP_KO_2ndExp_01.1329.1336.2.pkl. The x-axis represents the mass-to-charge ratio (m/z) from 0 to 1200. The y-axis represents the relative intensity, with a maximum of 1.32e+3. The spectrum shows several characteristic peaks, including the precursor ion at m/z 592.3113 (labeled y₈). Other labeled peaks include b₂, y₂, y₃, y₄, y₅, y₆, y₇, and y₈. The spectrum is identified as B22_Shastrin_ERAAP_KO_2ndExp_01.1329.1336.2.pkl.

MS/MS spectrum of the protein B22_Shasti_ERAAP_KO_2ndExp_17.1204.1211.2.pkl. The x-axis represents the mass-to-charge ratio (m/z) from 0 to 1100. The y-axis represents the relative intensity from 0.0% to 10.0%. The spectrum shows several peaks, with the base peak at m/z 536.3011. The peaks are labeled with their corresponding amino acid sequences: IL, NR, E, Y, GL, b₂, IE, NI, y₂, y₃, NIH, GIEN, IYH, ENIH, GIENI, b₅, P(m/z), P(m/z)-H₂O, IENIH, y₅, ENIH, y₆, IENIH, GIENIH, y₇, b₈, and b₈+H₂O, y₈. The spectrum is a color-coded plot with peaks labeled in green, blue, red, and black.

MS/Tag 1 | A | A | V | I | A | H | D | F | L |

8.28e+3
42.1%

AAb₂⁺⁺, b₁, V
b₂
y₁
H
F
IA
AH
b₃
AHD
IAH
AHDF
IAHD
y₄⁺⁺
y₇
y₄-H₂O
y₅
IAHDF, y₆-H₂O
y₆-NH₂
y₆-H₂O
y₇-NH₂
b₈
y₈

ANH₃⁺, ANH₃⁺
D
H, H
AV
V
DAV, MA
AMAb₄

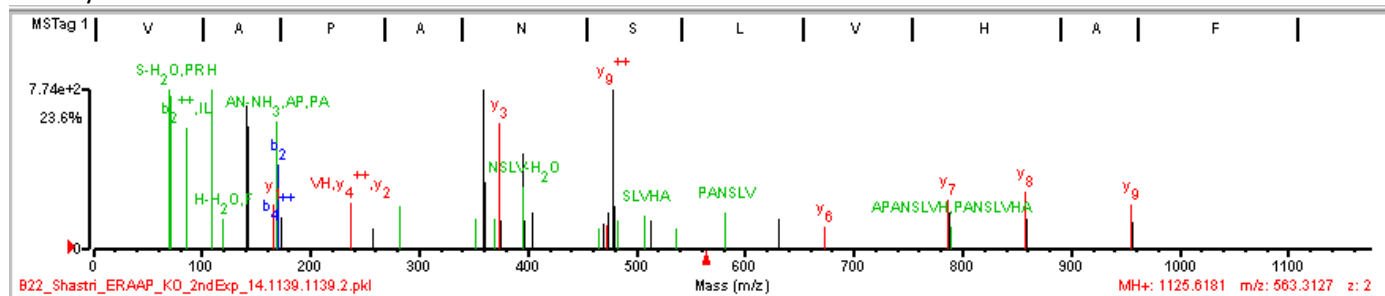
B22_Shastri_ERAAP_KO_2ndExp_15.1167.1175.2.pkl

Mass (m/z)

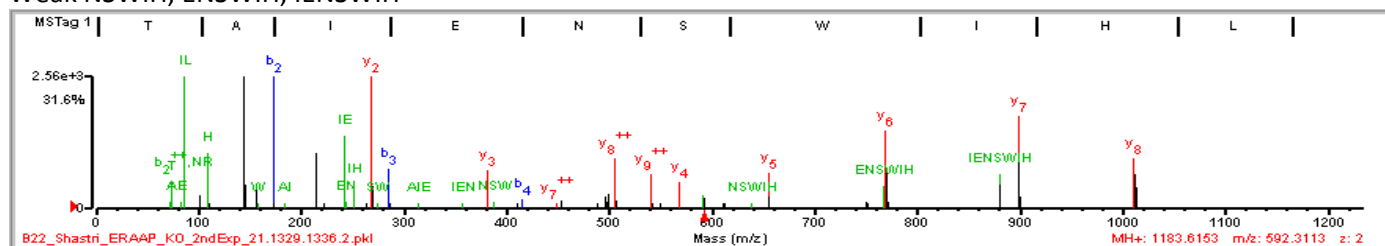
MH⁺: 956.5227 m/z: 478.765 z: 2

Mass spectrum plot showing relative intensity (Y-axis, 0 to 1.03e+4) versus Mass (m/z) (X-axis, 0 to 900). The spectrum displays various peaks labeled with amino acid sequences (PR, IL, E, H, GP, PV, NH, PVN, VNH, NHE, y7⁺⁺, y8⁺⁺, y5, y7, y8, b2, b7, b8) and fragmentation types (P(m/z), NHEE, PVNHE, GPNHEE). The base peak is at m/z 434 (y7⁺⁺). The plot is titled "B22_Shastreri_ERAAP_KO_2ndExp_6.1023.1030.2.pkl".

Weak y2

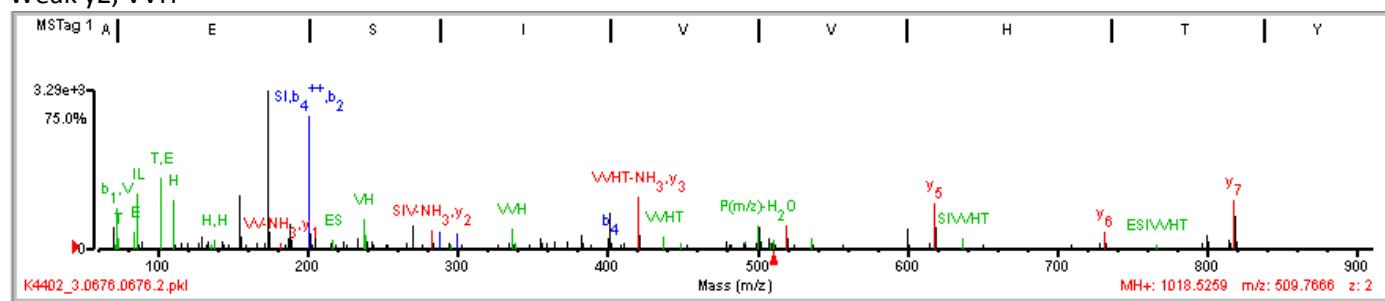


Weak NSWIH, ENSWIH, IENSWIH



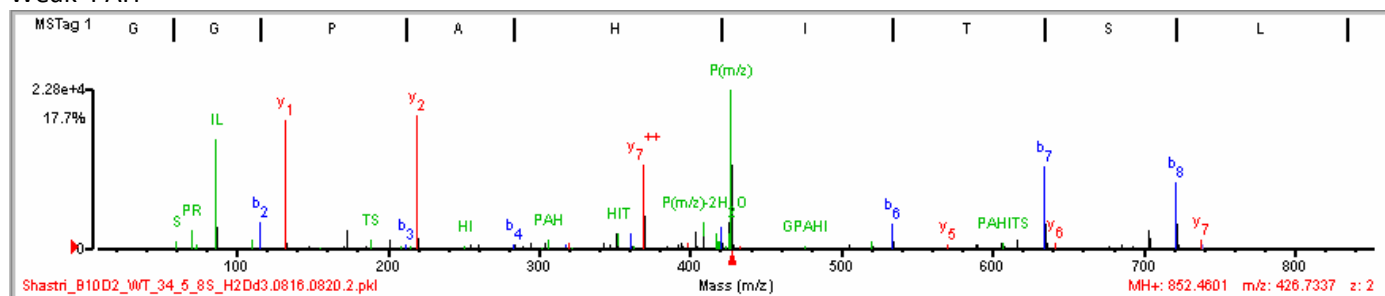
HLA-B4402

Weak y2, VVH

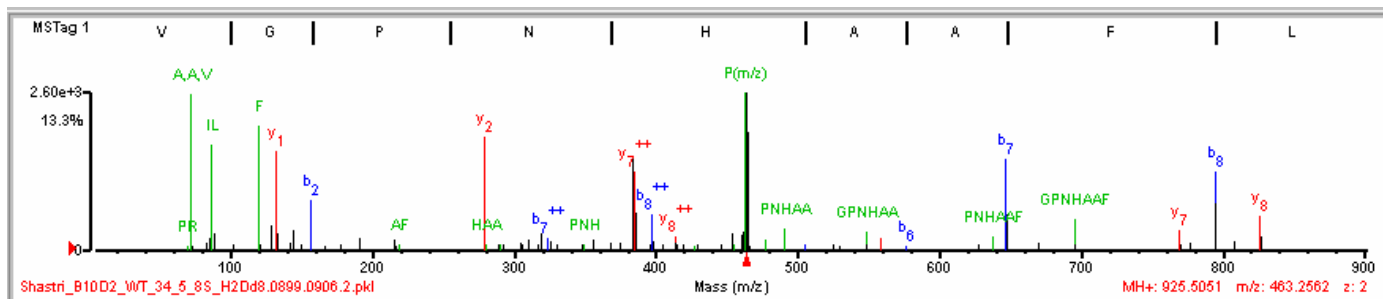


H2-Dd

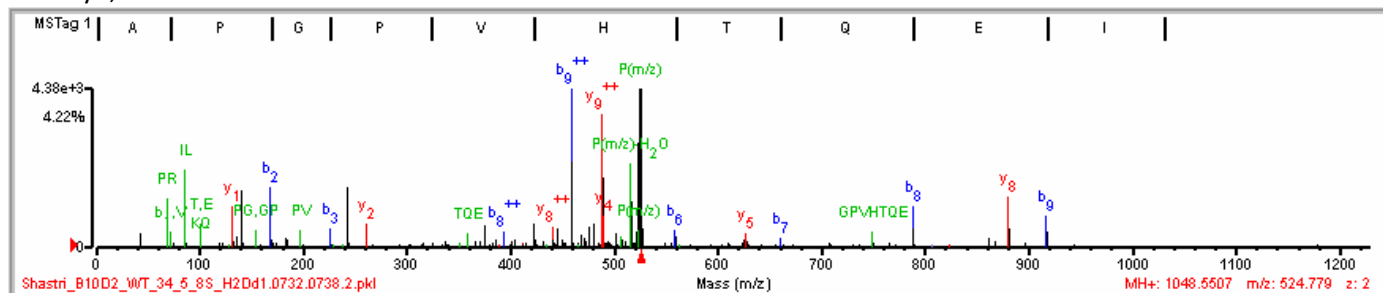
Weak PAH



Weak PNH

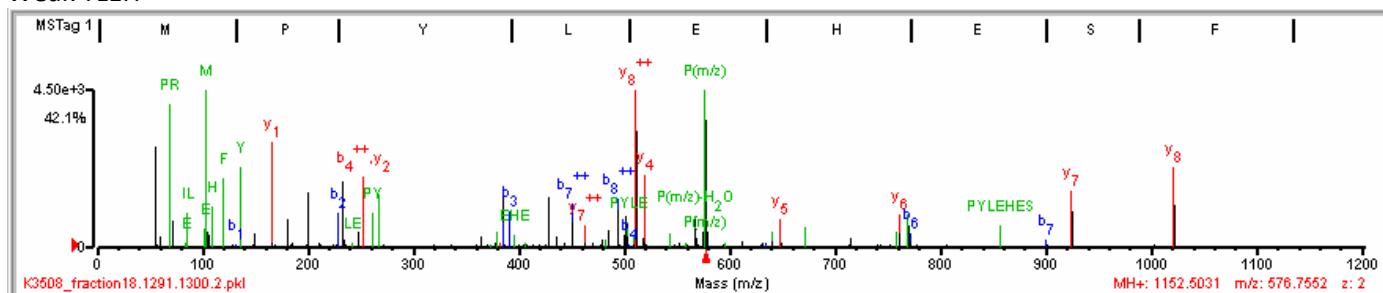


Weak y4,b6



HLA-B3508

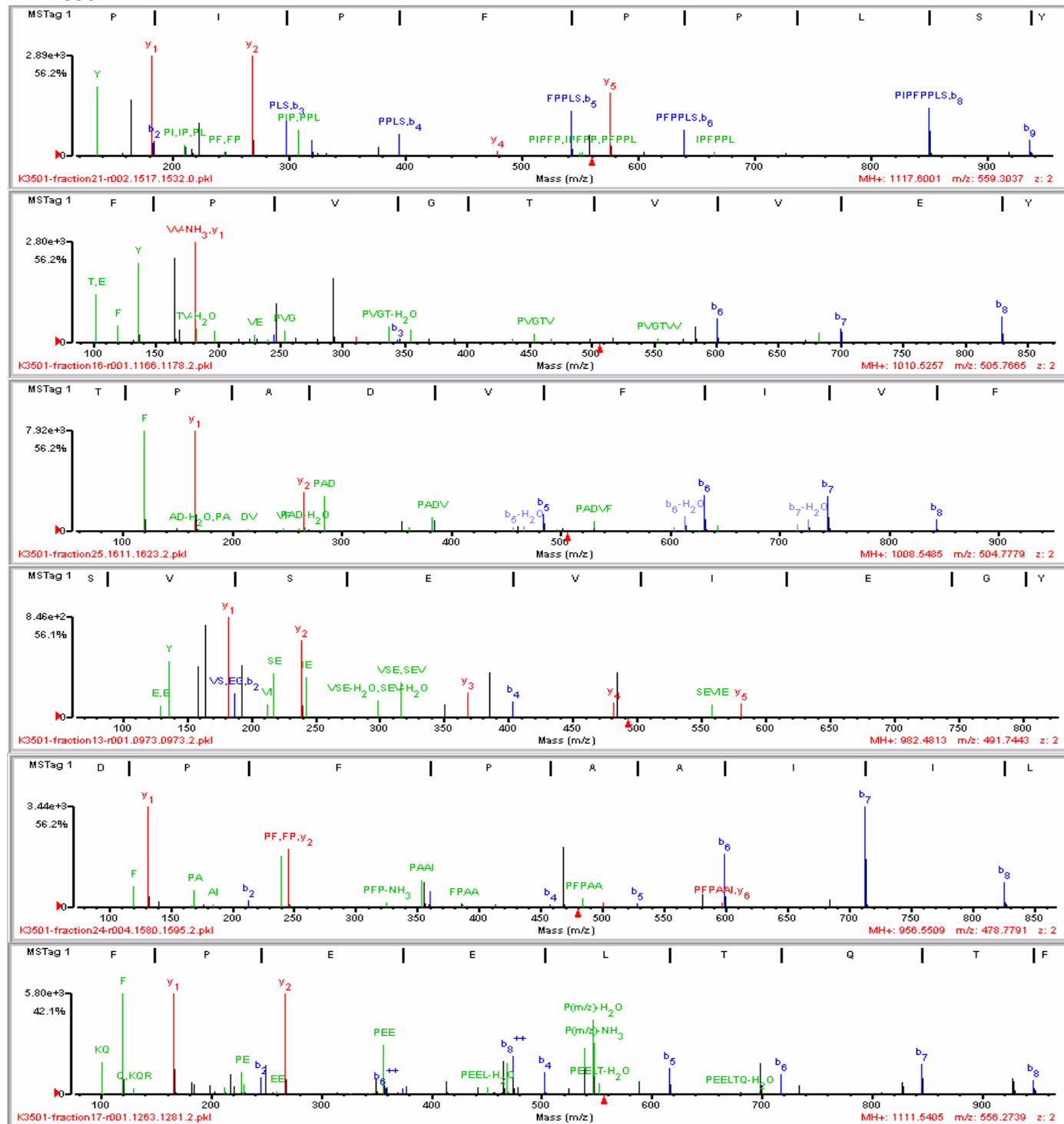
Weak YLEH

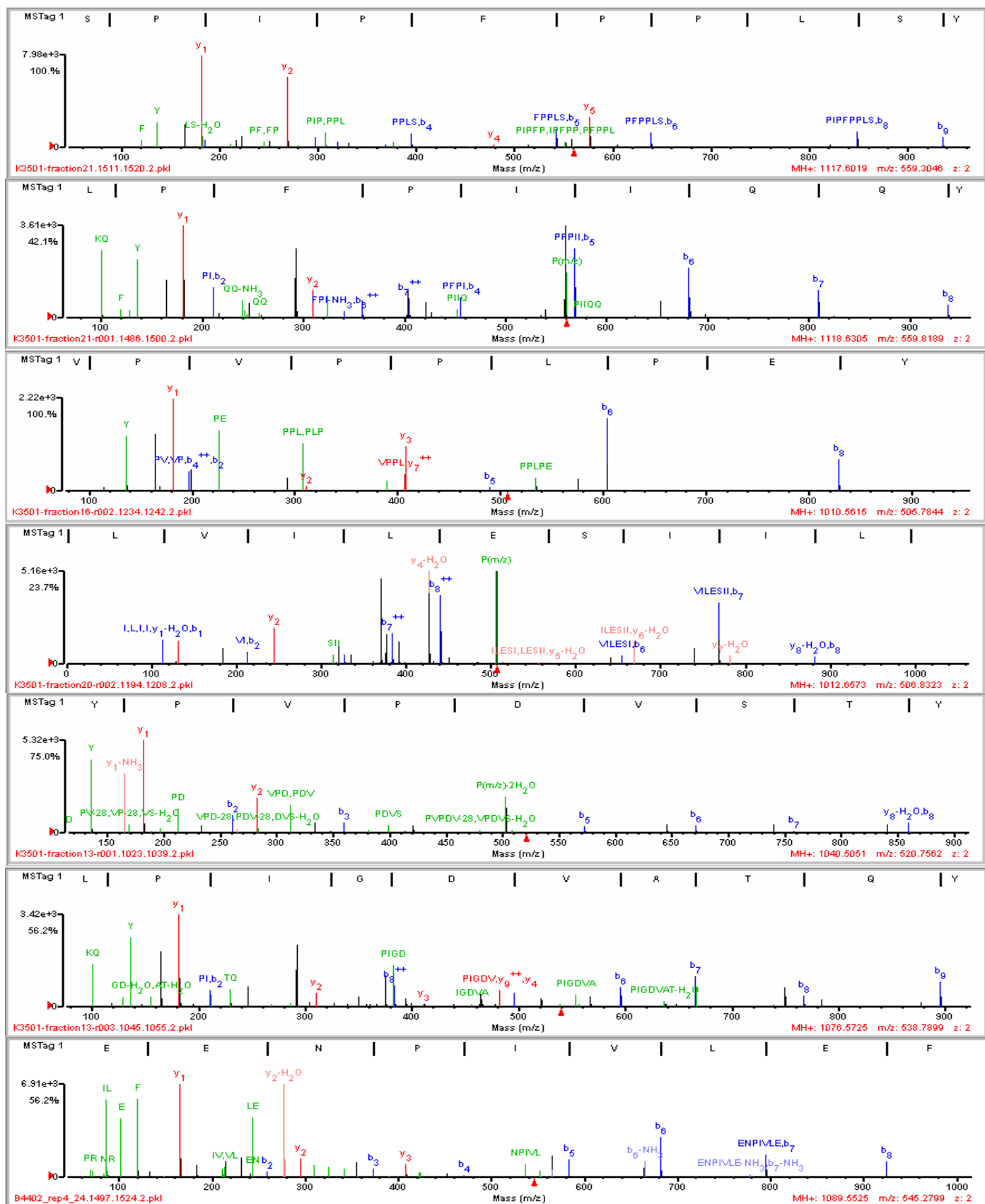


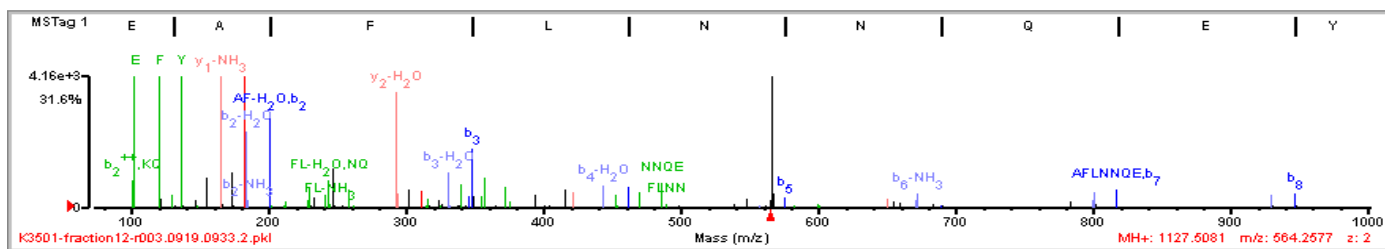
B. QTOF spectra peptides MHC class I without basic residues inside.

(Please note the software design does not default display all ion signals in the follow spectra. Complete or specific count of ions may require additional manipulations. When a spectrum is repeated to show additional signals both are emphasized in one box).

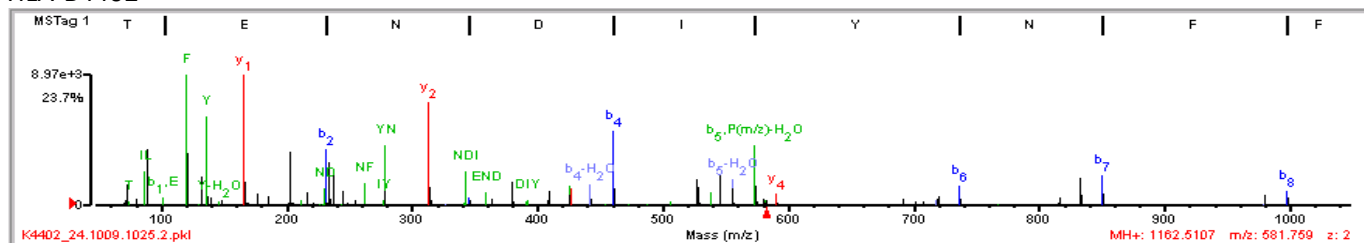
HLA-B3501



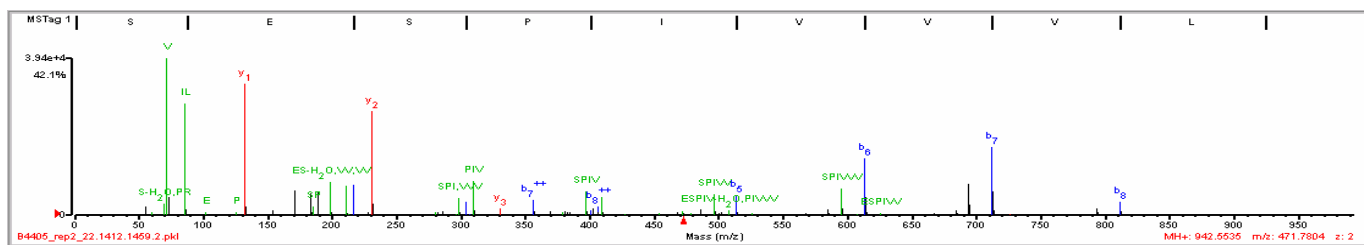
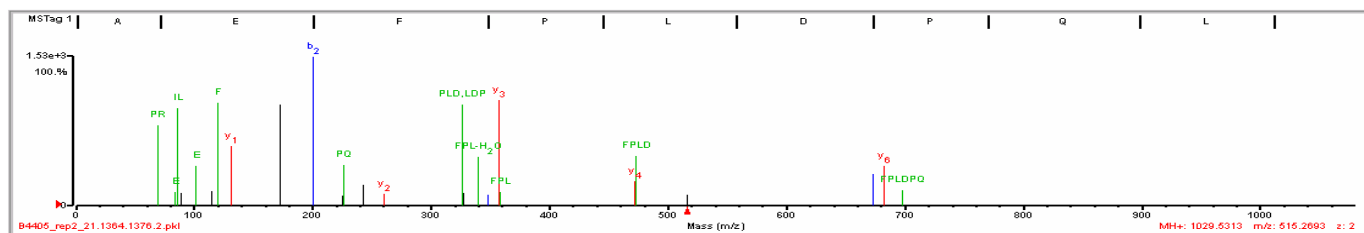
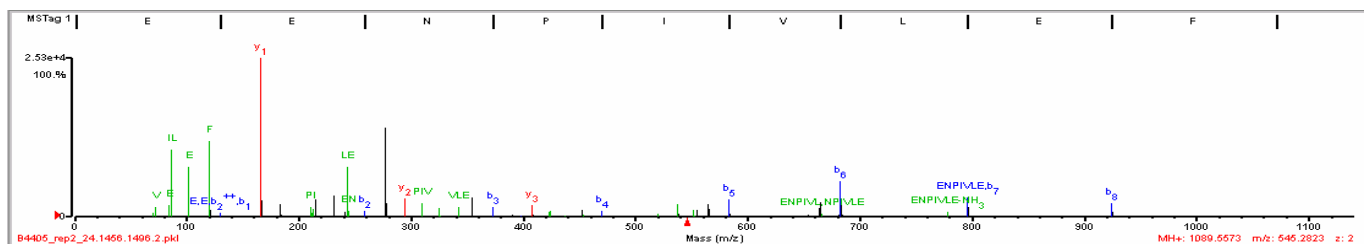
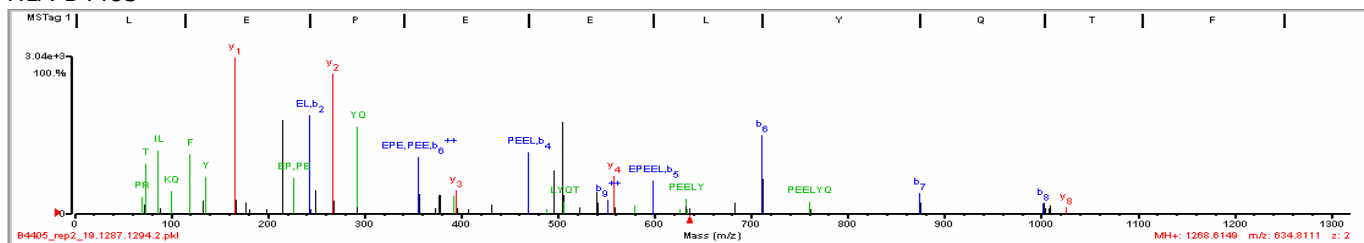


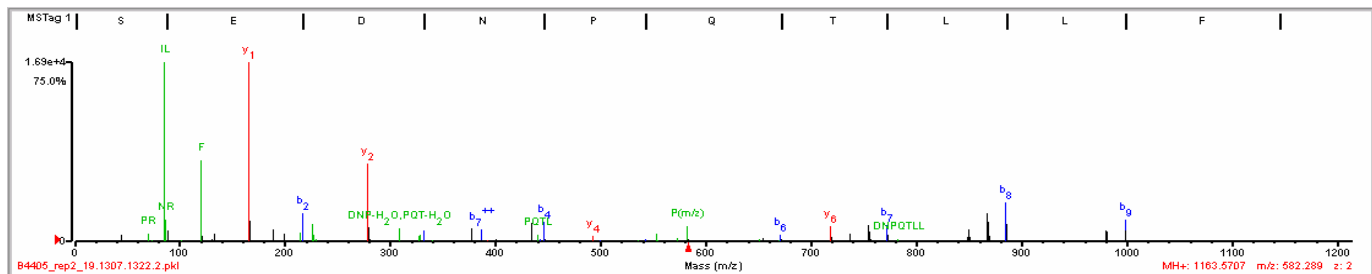
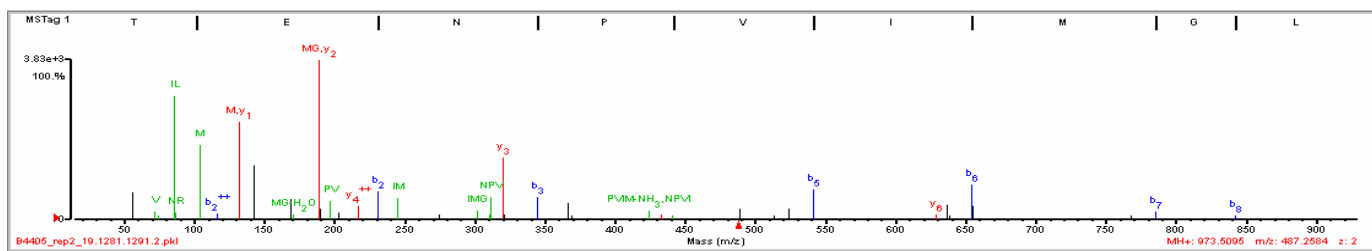
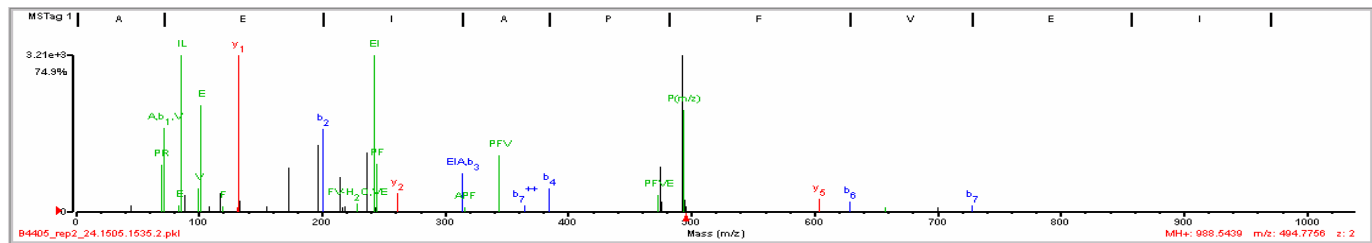
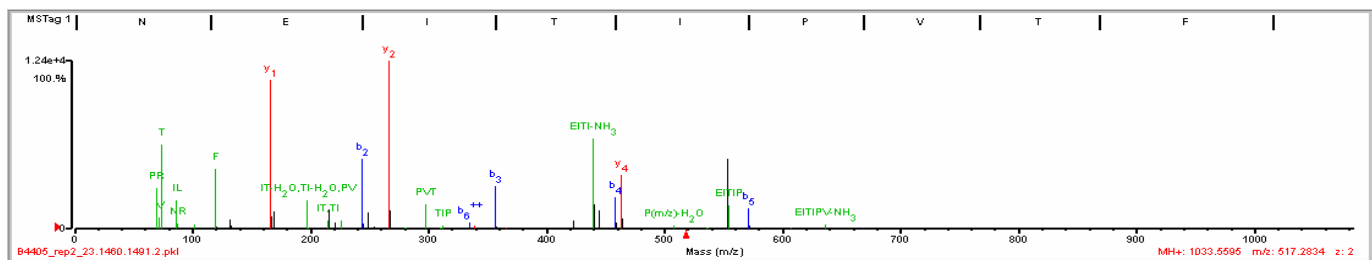


HLA-B4402

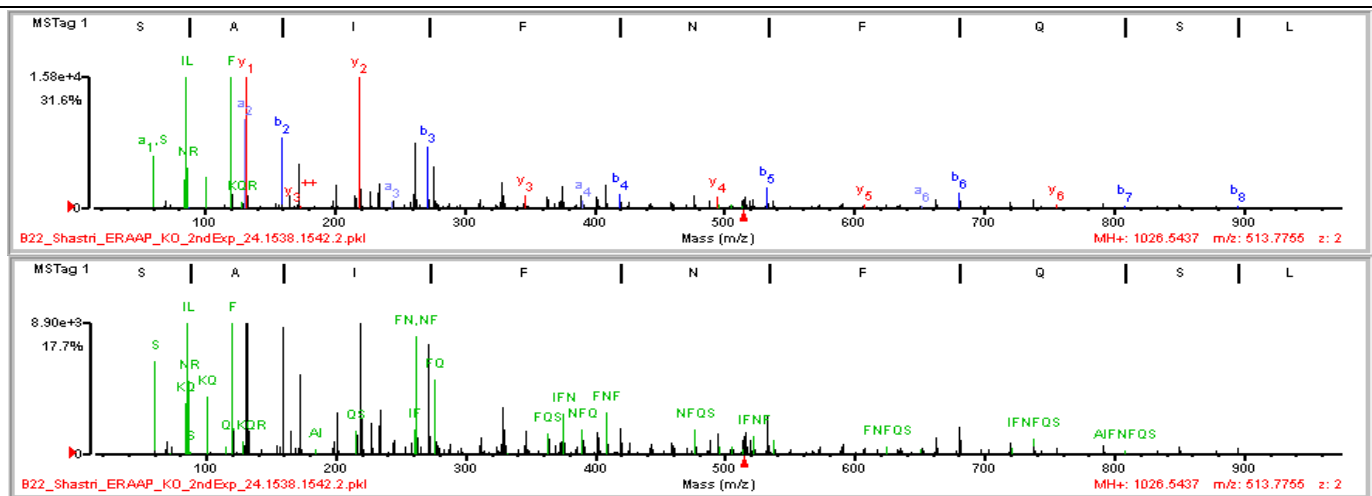


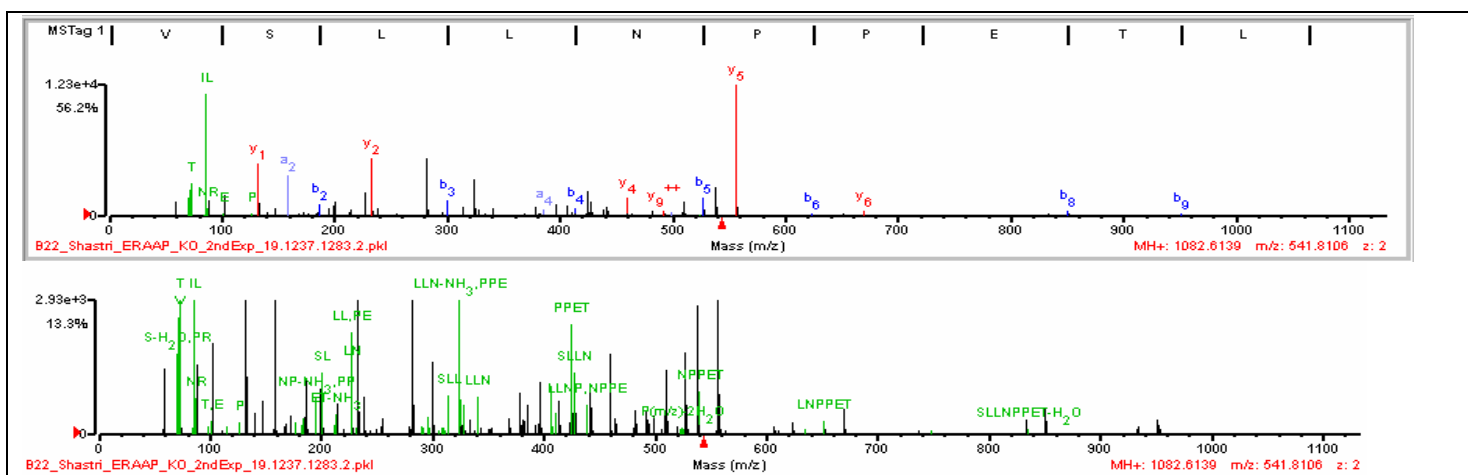
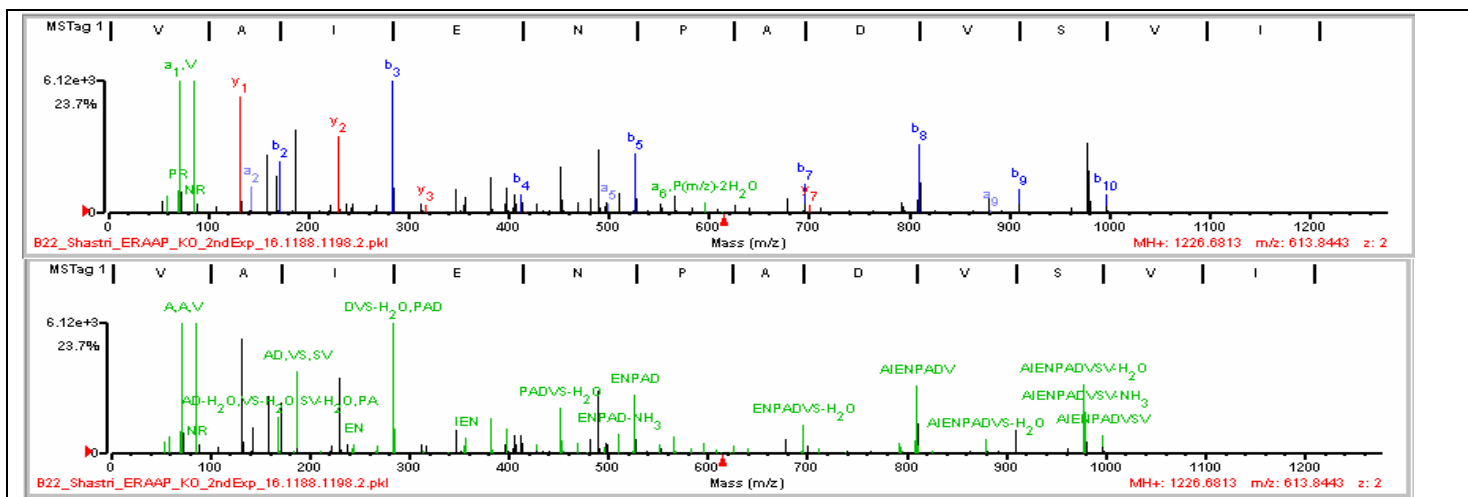
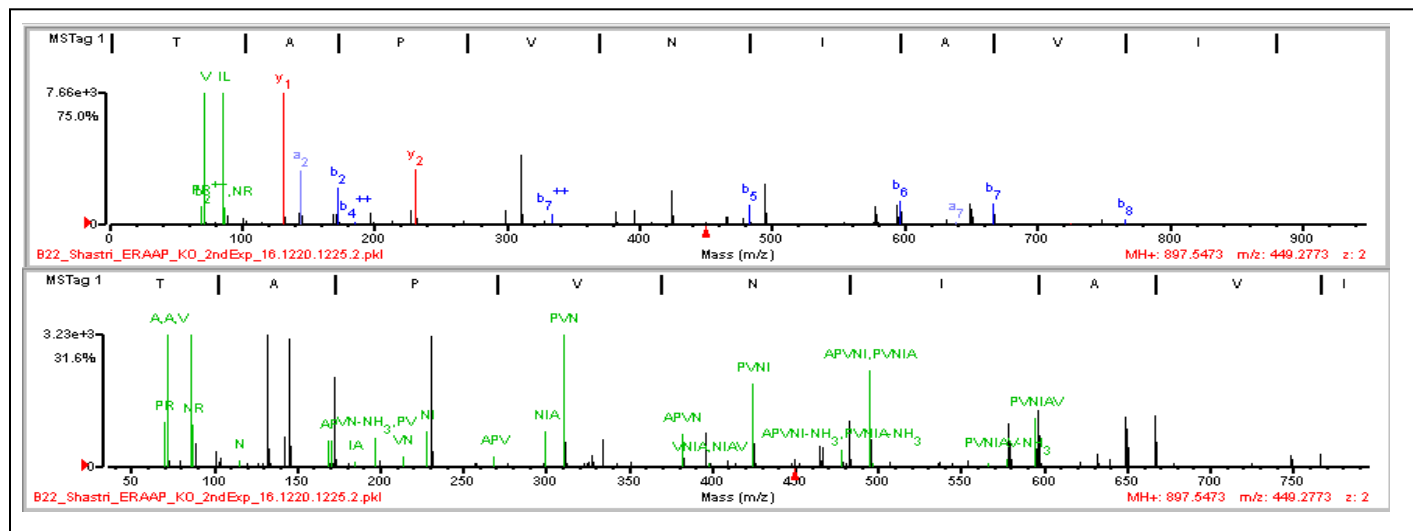
HLA-B4405

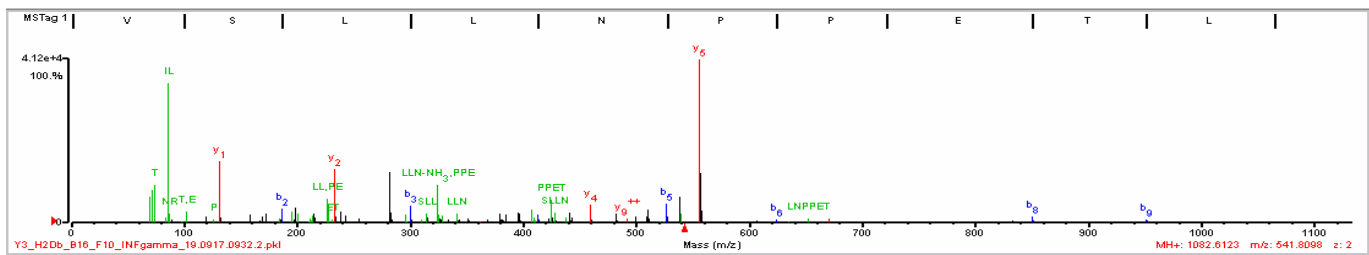
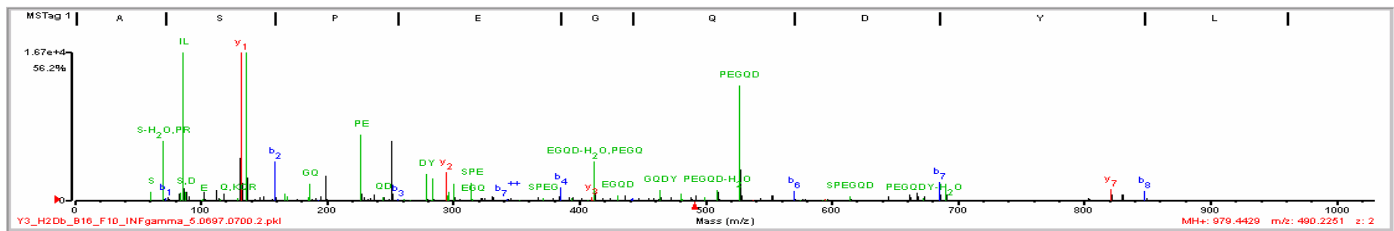
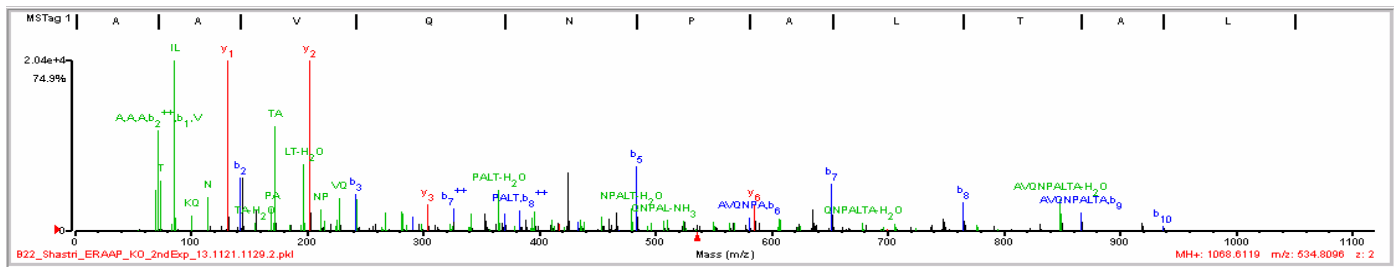
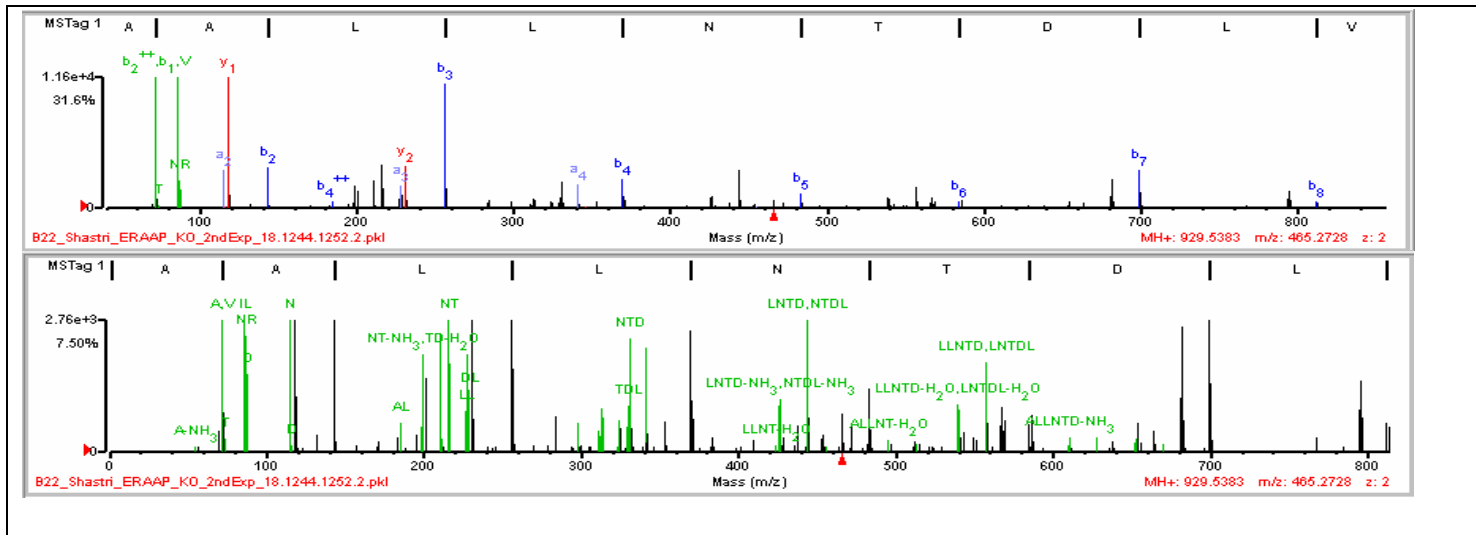
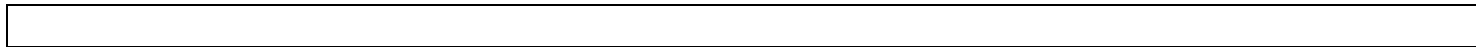




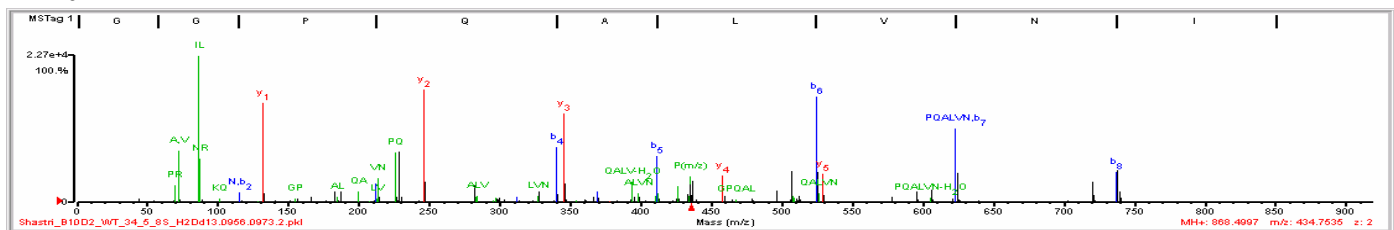
H2-Db

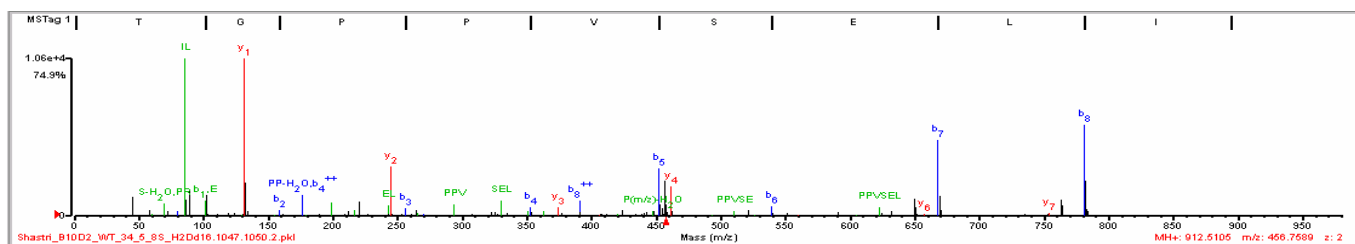
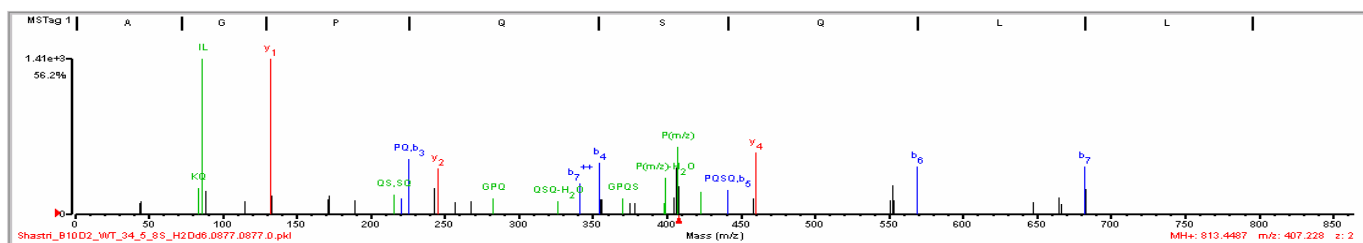
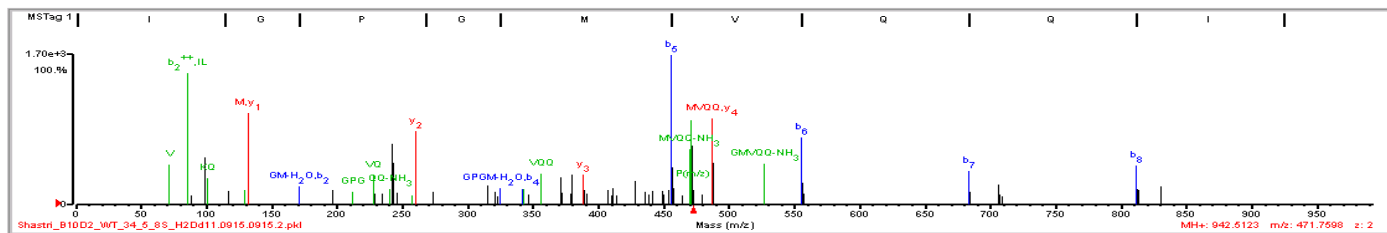
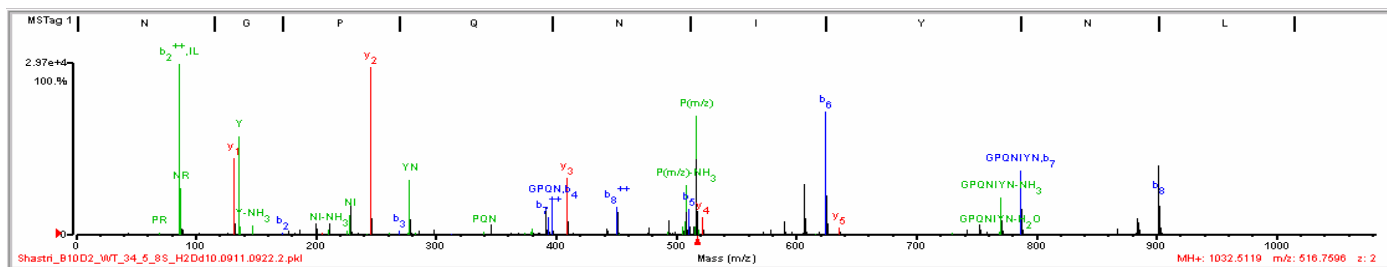




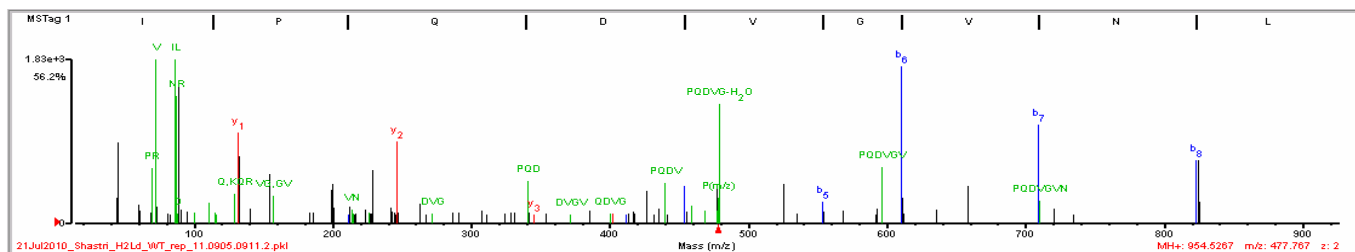
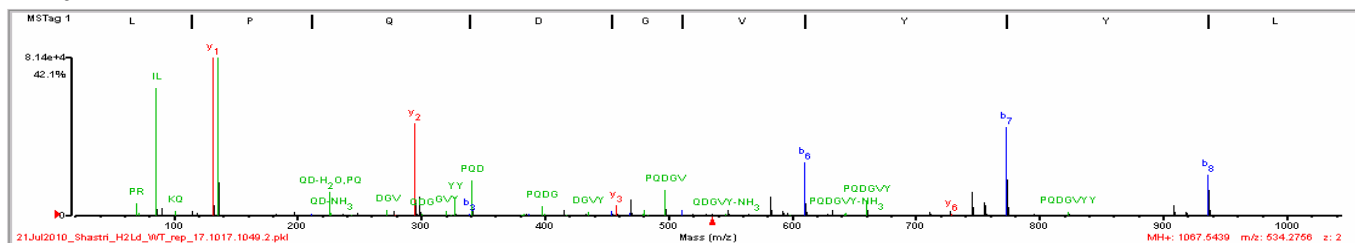


H2-Dd





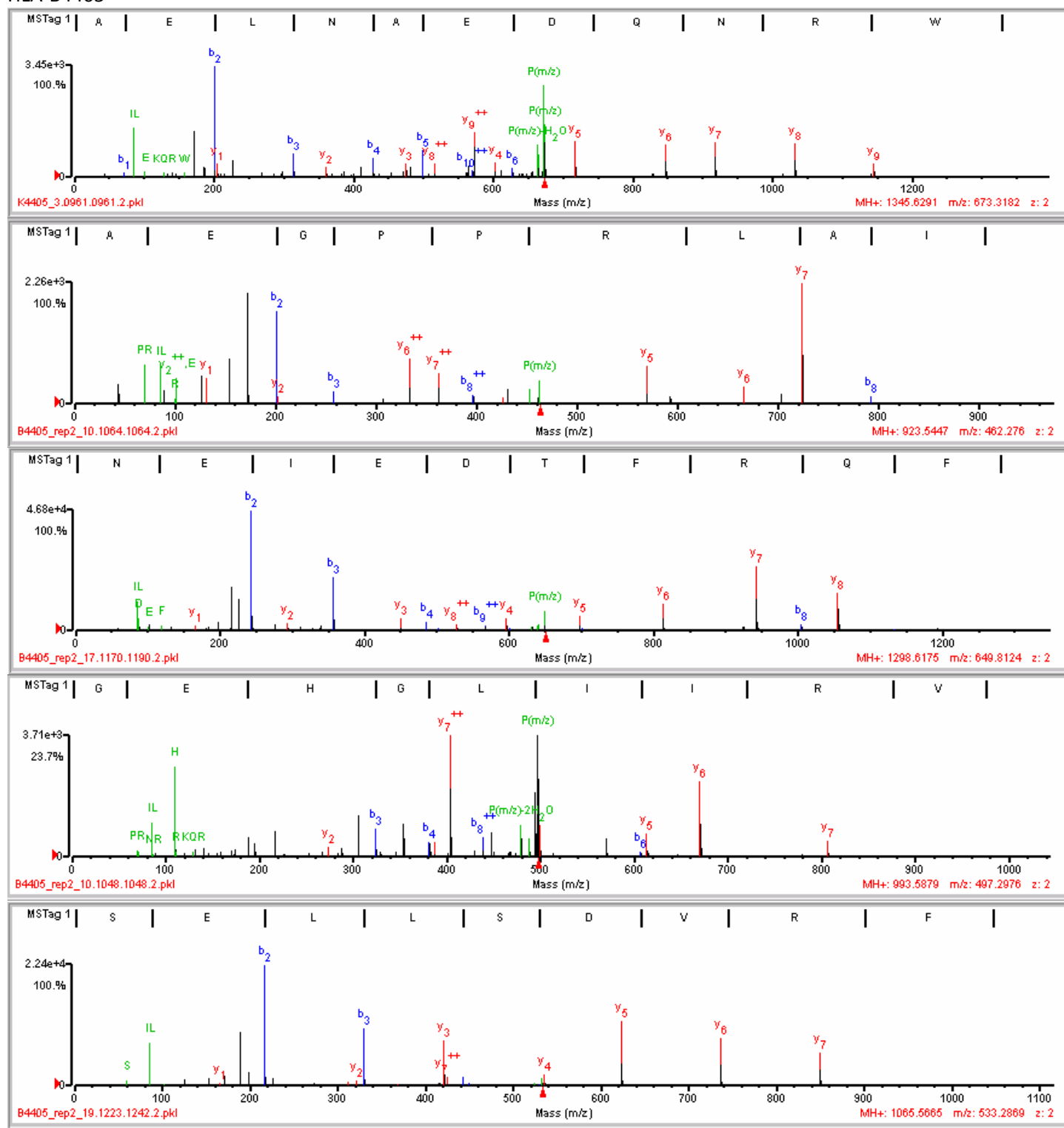
H2-Ld

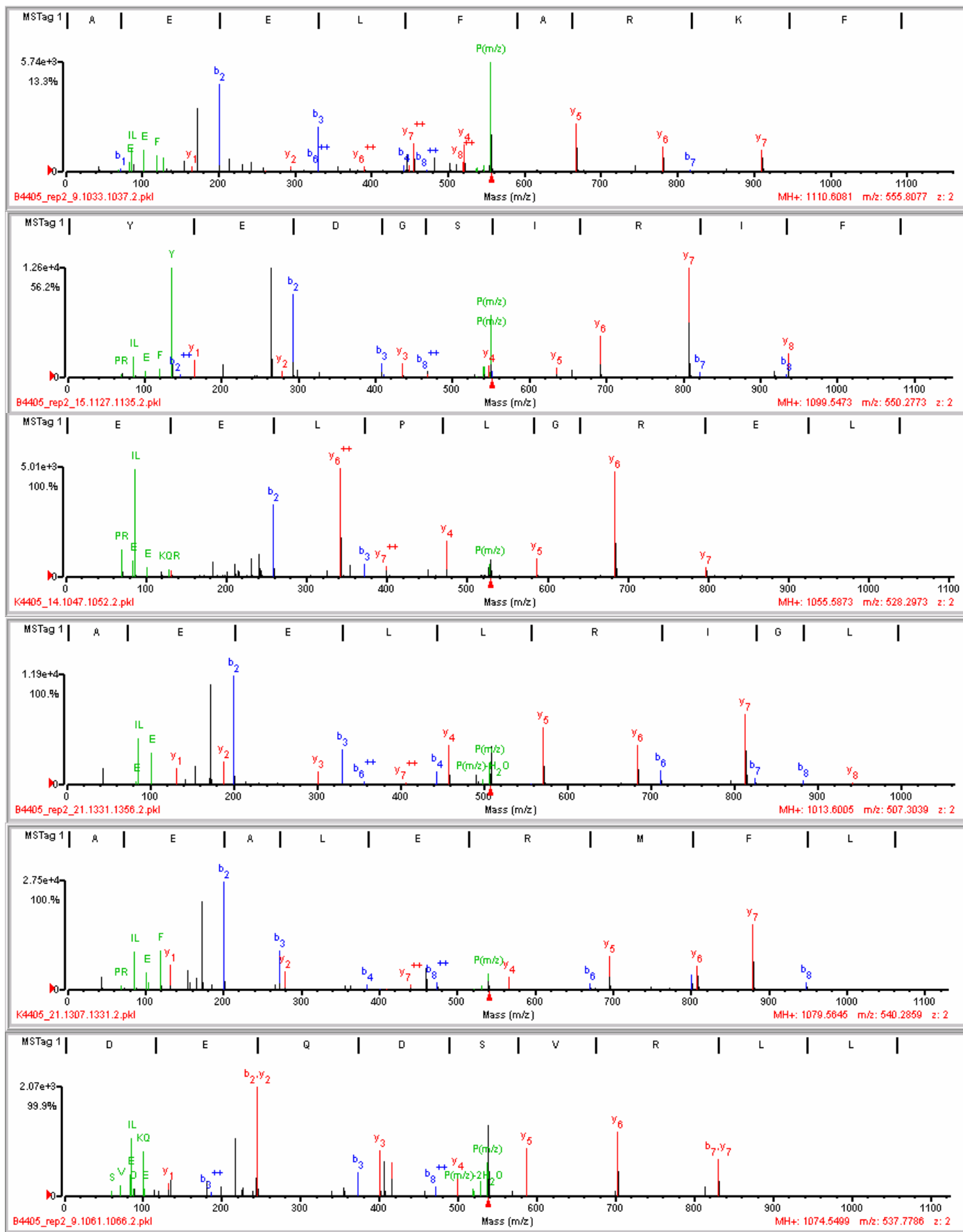


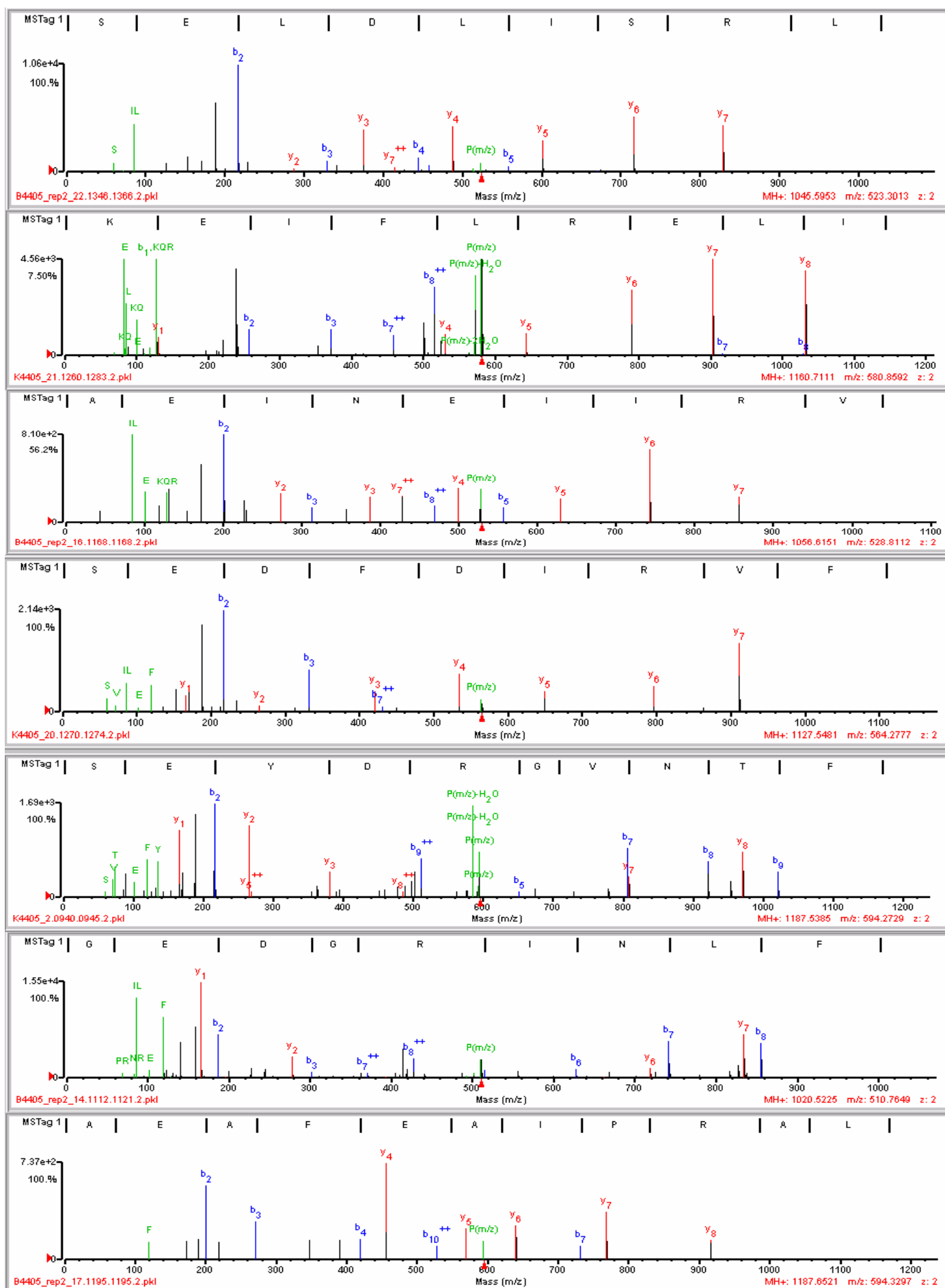
C. QTOF spectra MHC class I peptides containing Arginine.

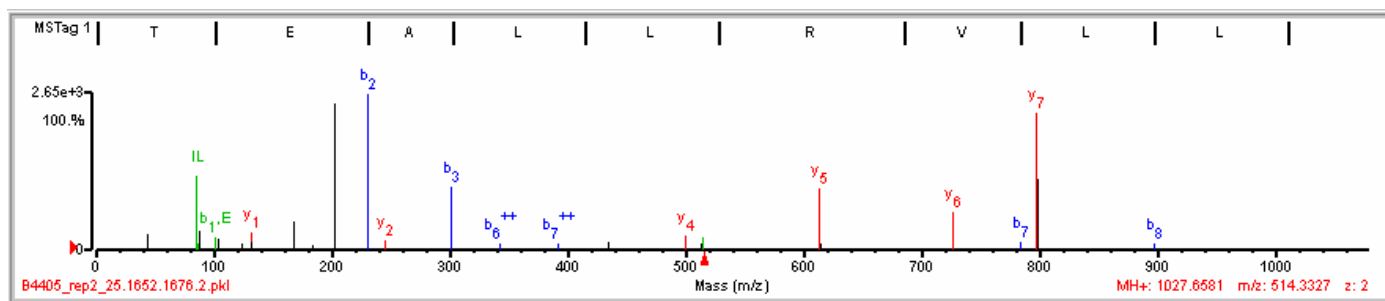
(Please note the software design does not default display all ion signals in the follow spectra. Complete or specific count of ions may require additional manipulations. When a spectrum is repeated to show additional signals both are emphasized in one box).

HLA-B4405

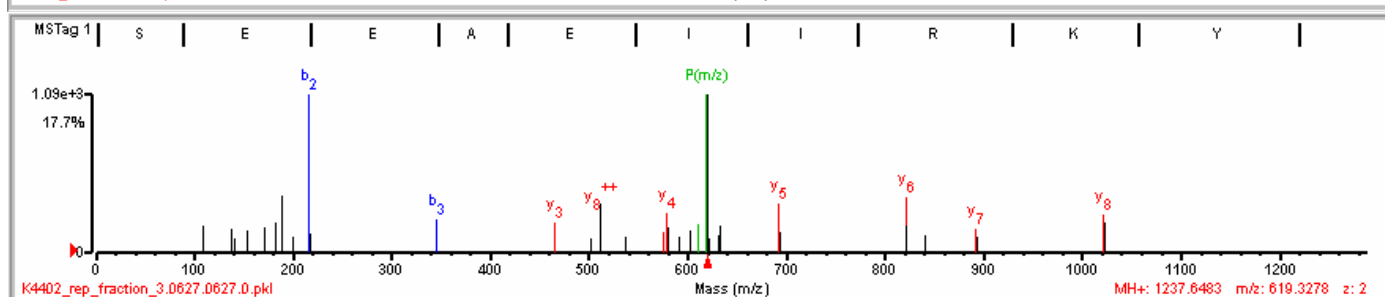
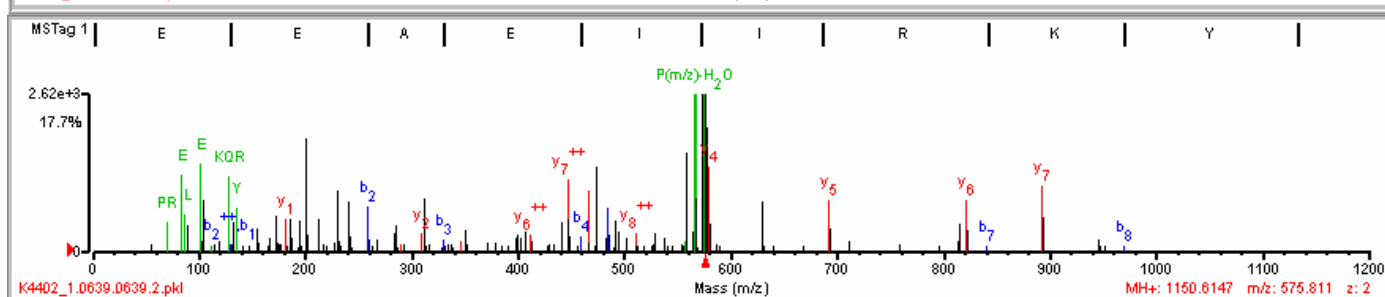
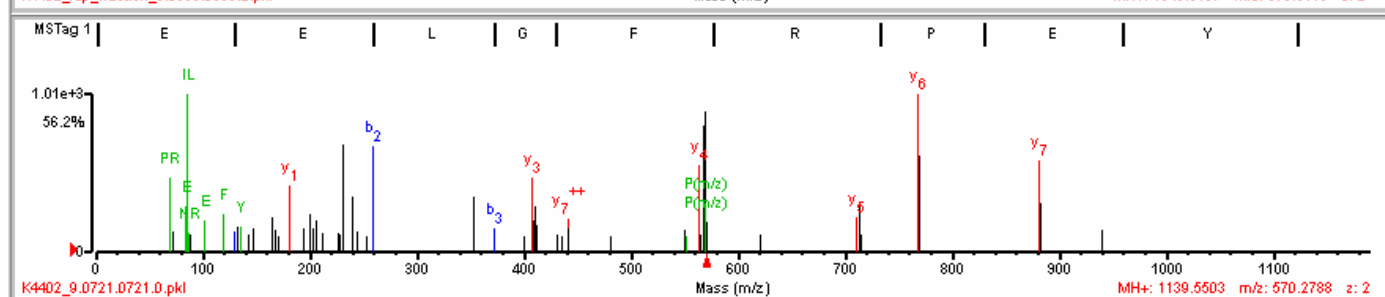
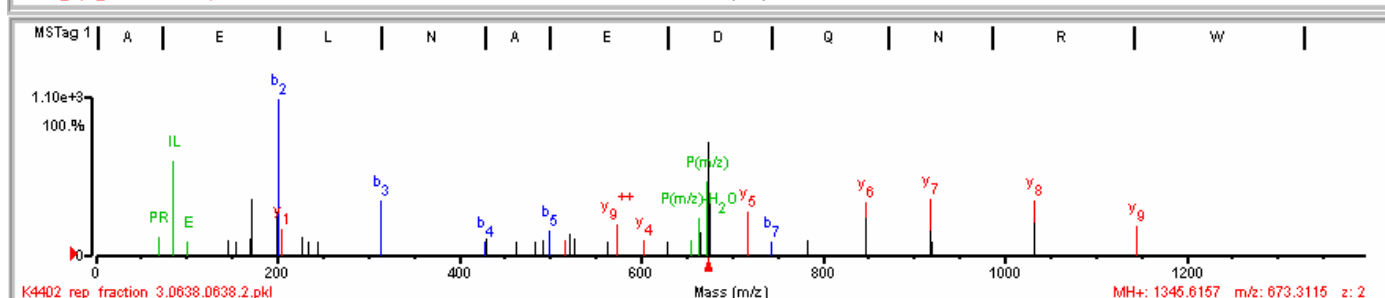
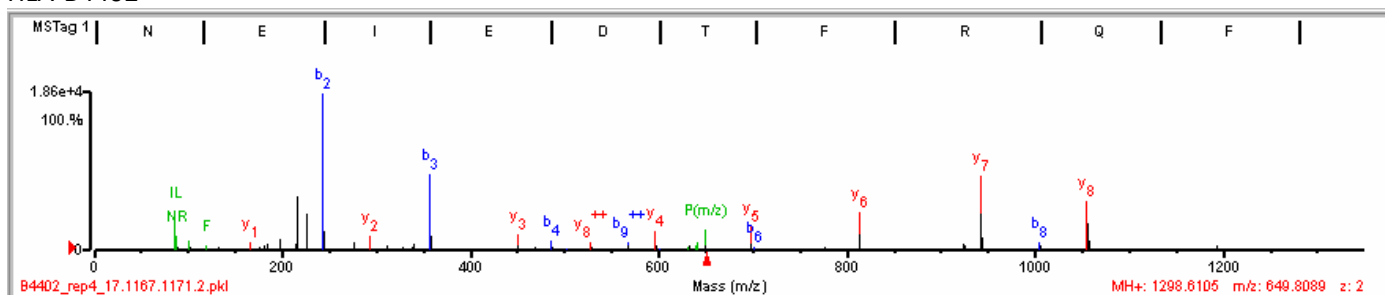


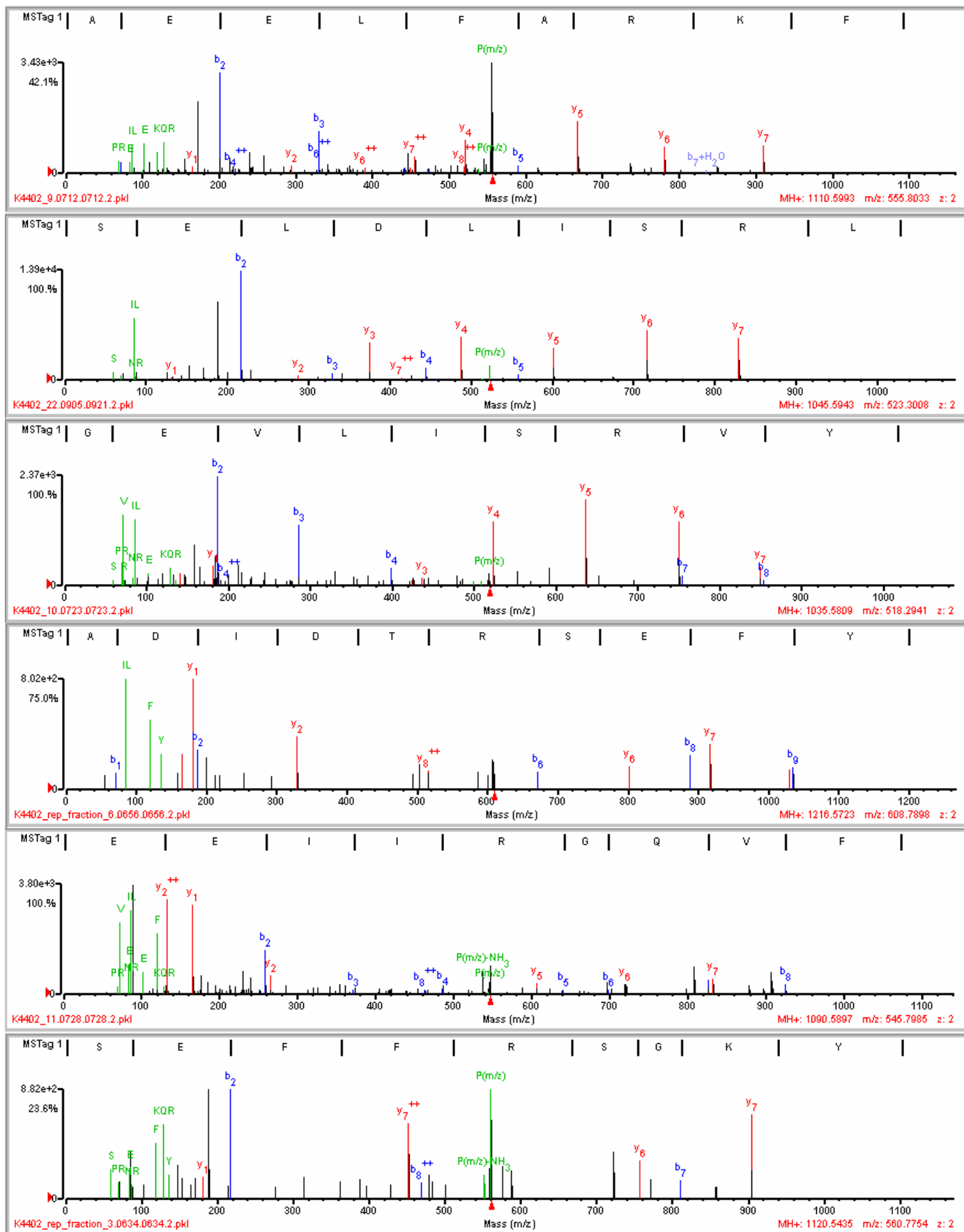


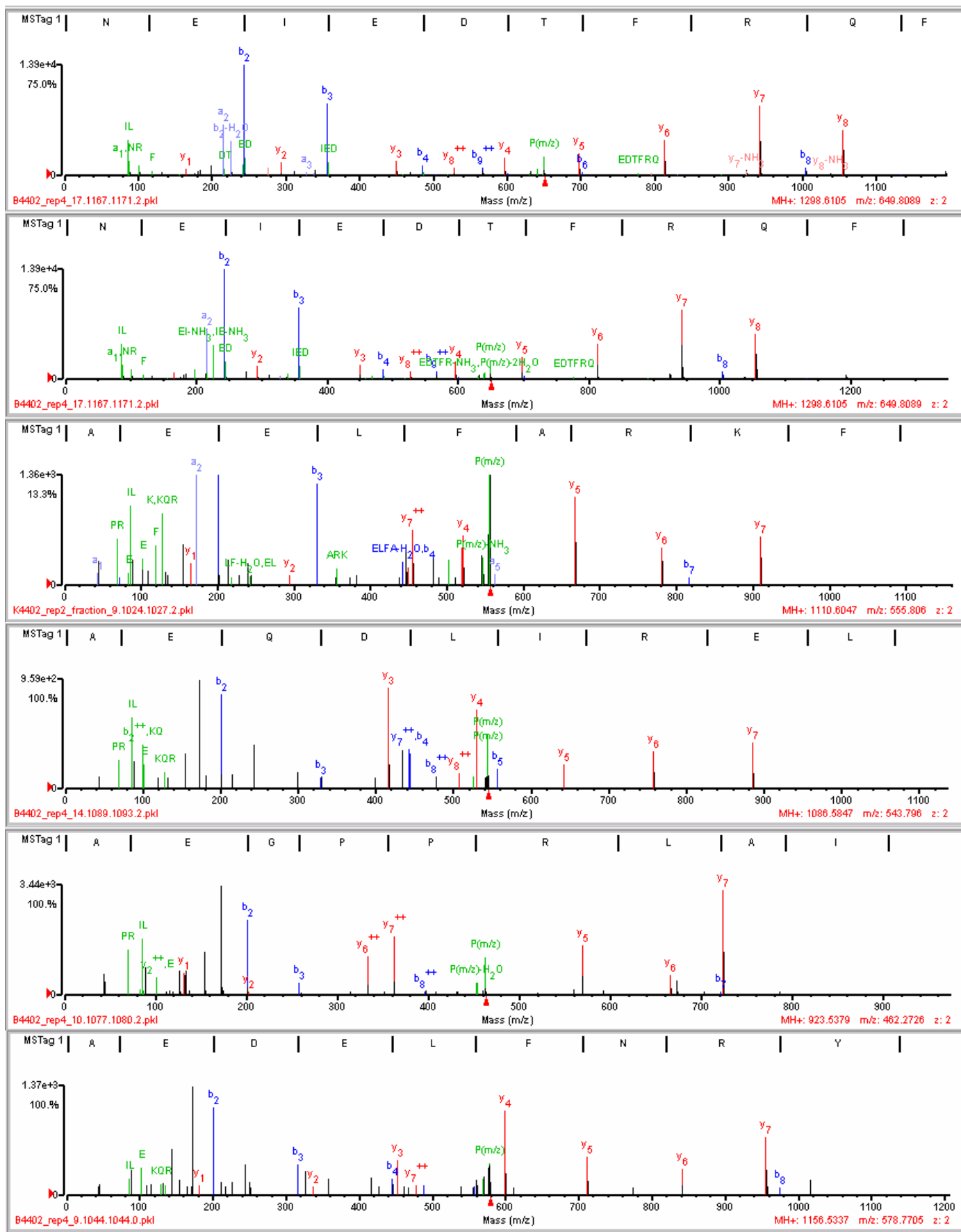


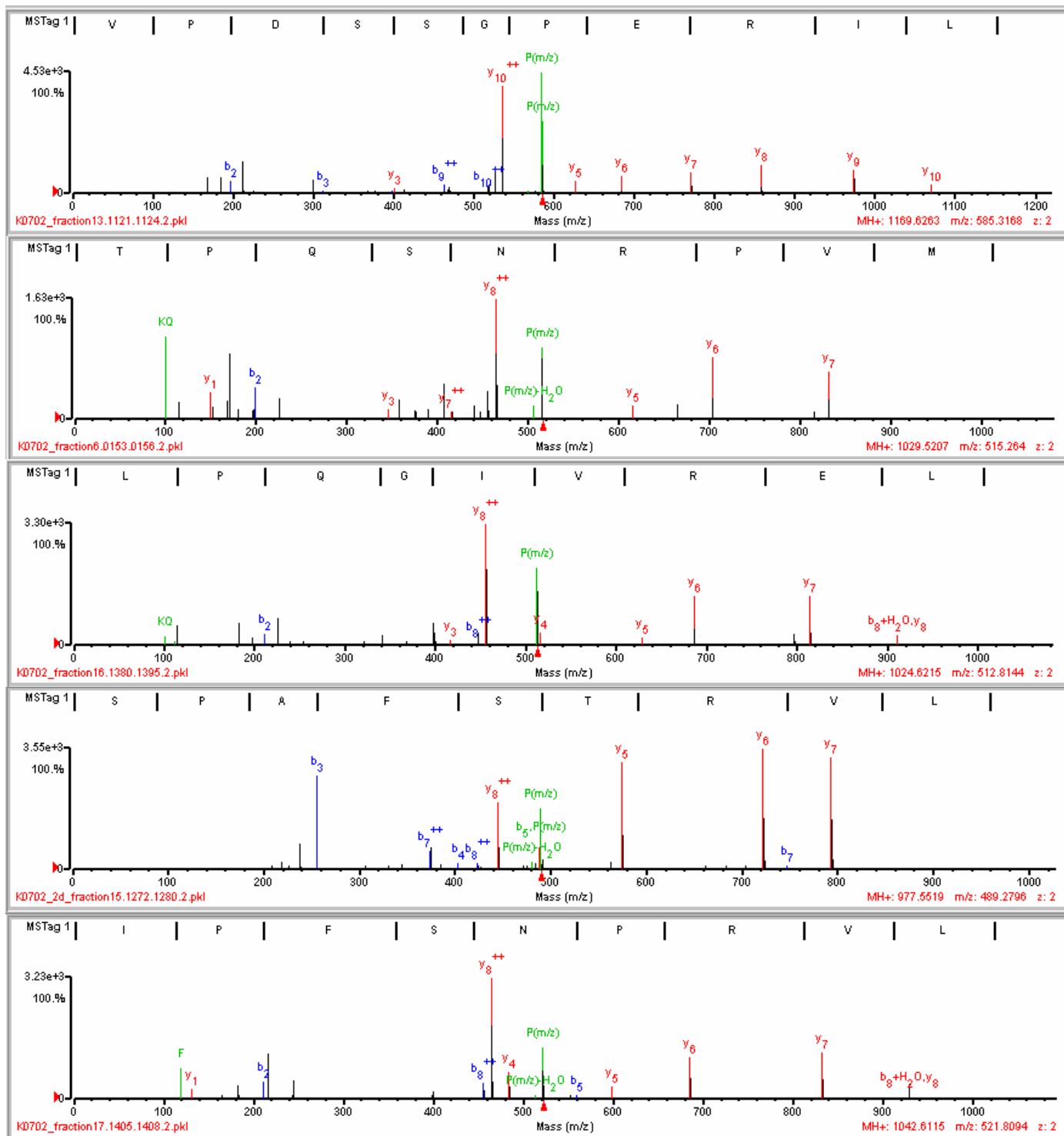


HLA-B4402

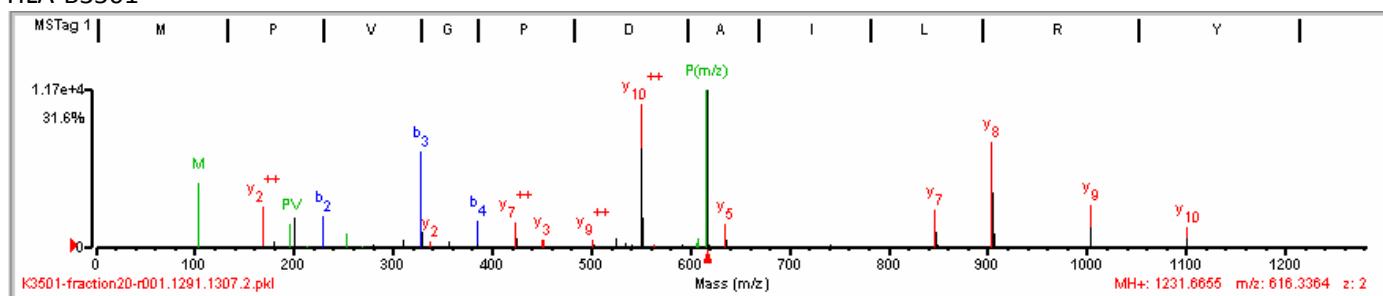


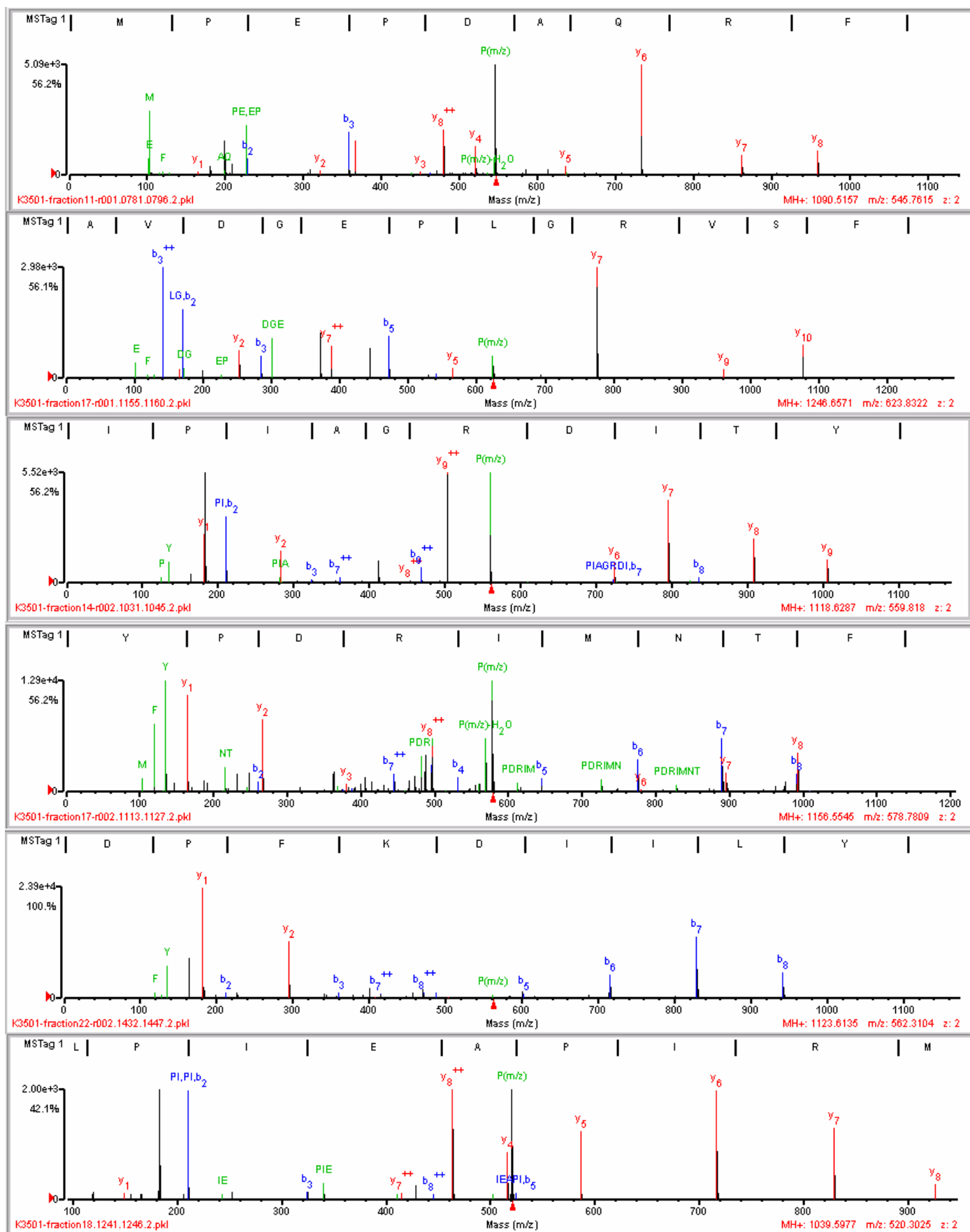


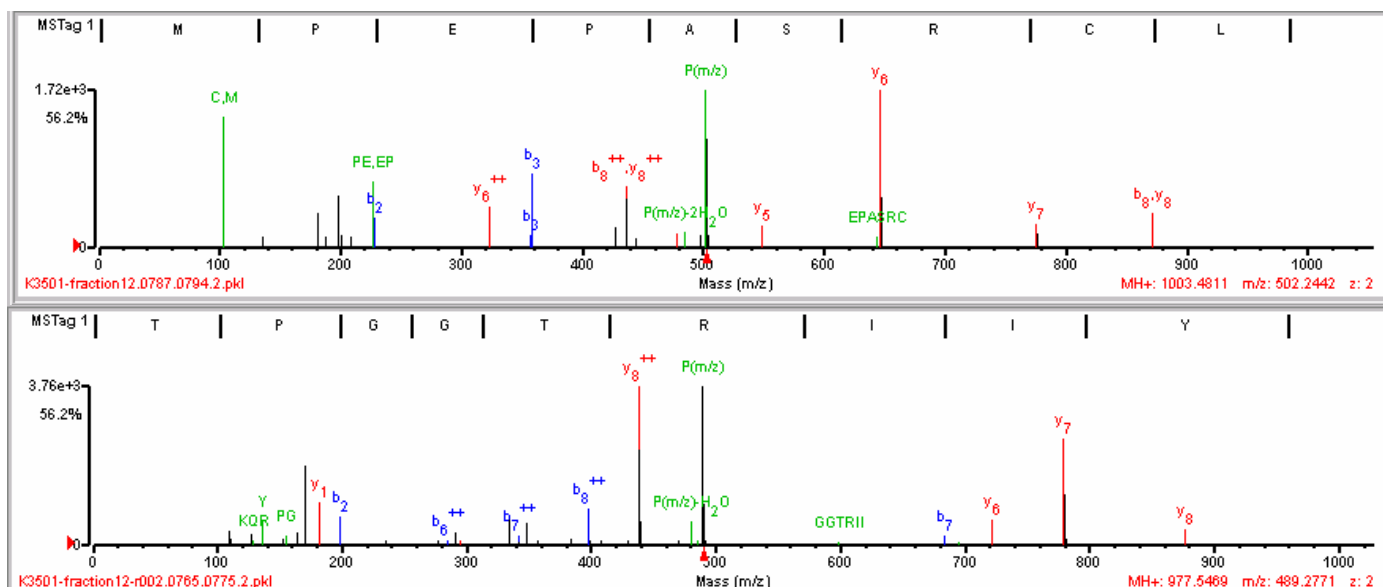




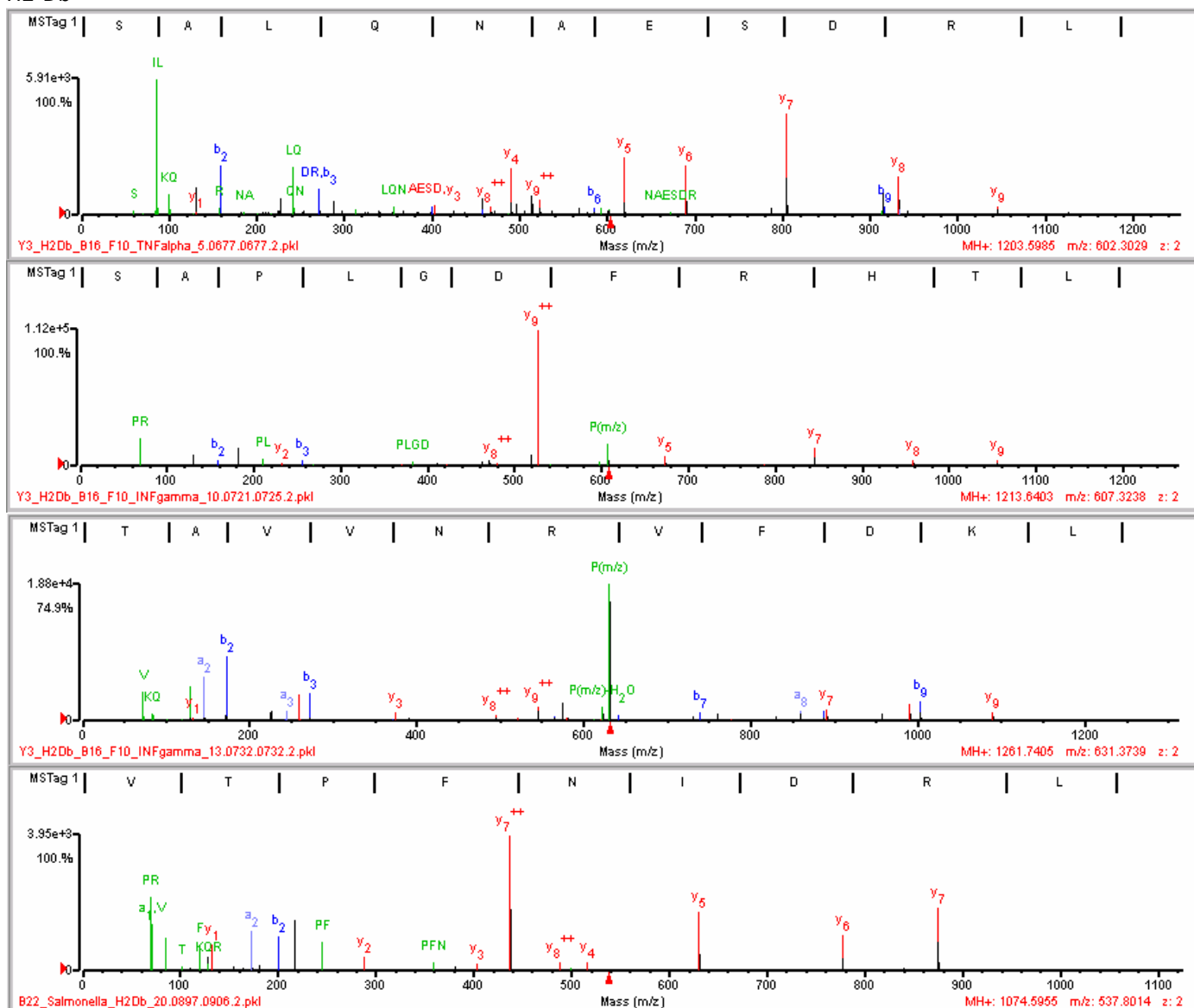
HLA-B3501

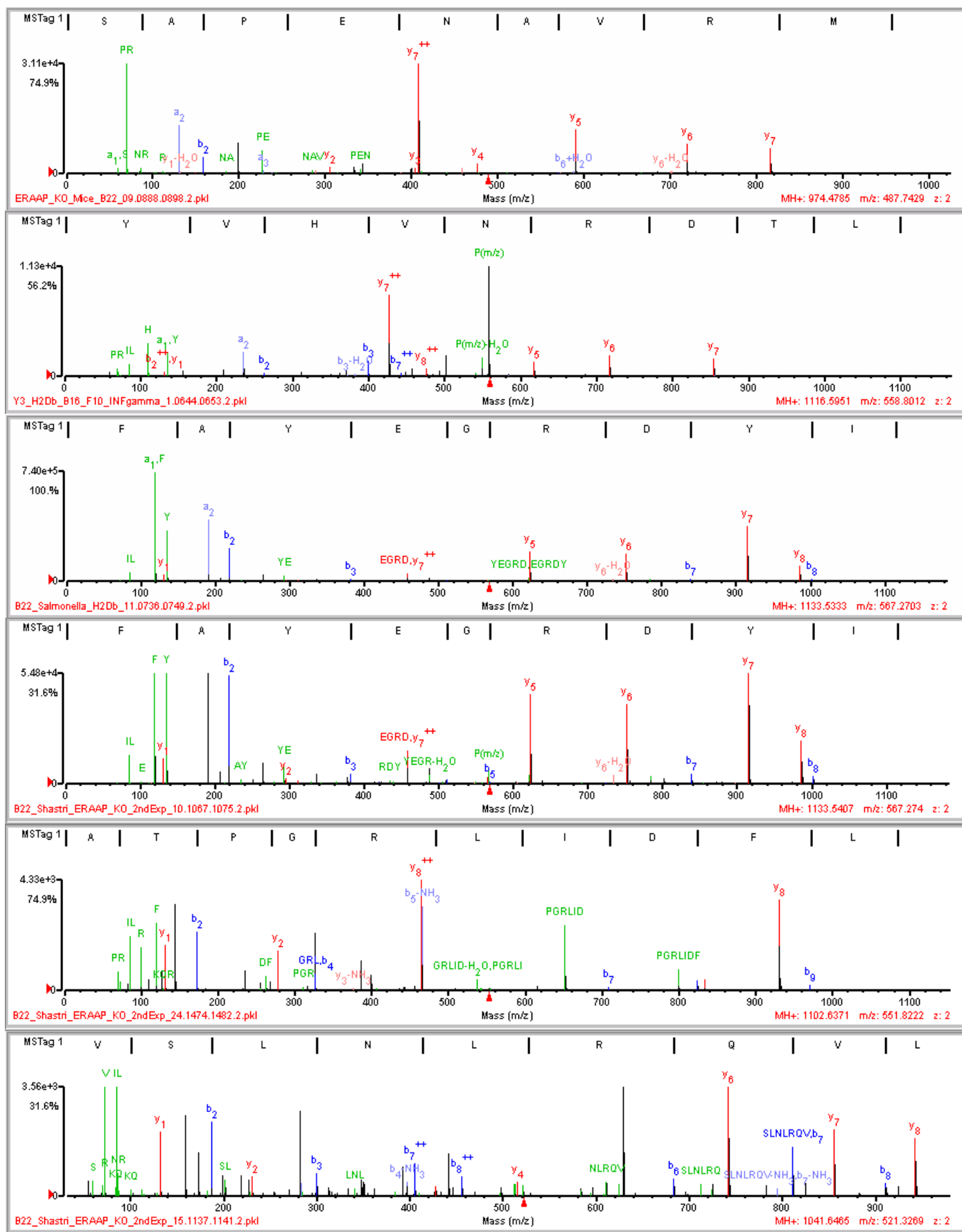


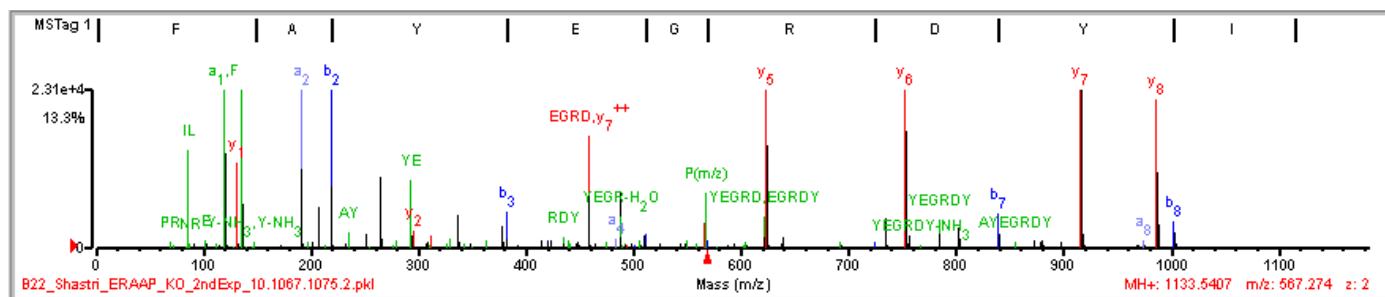




H2-Db



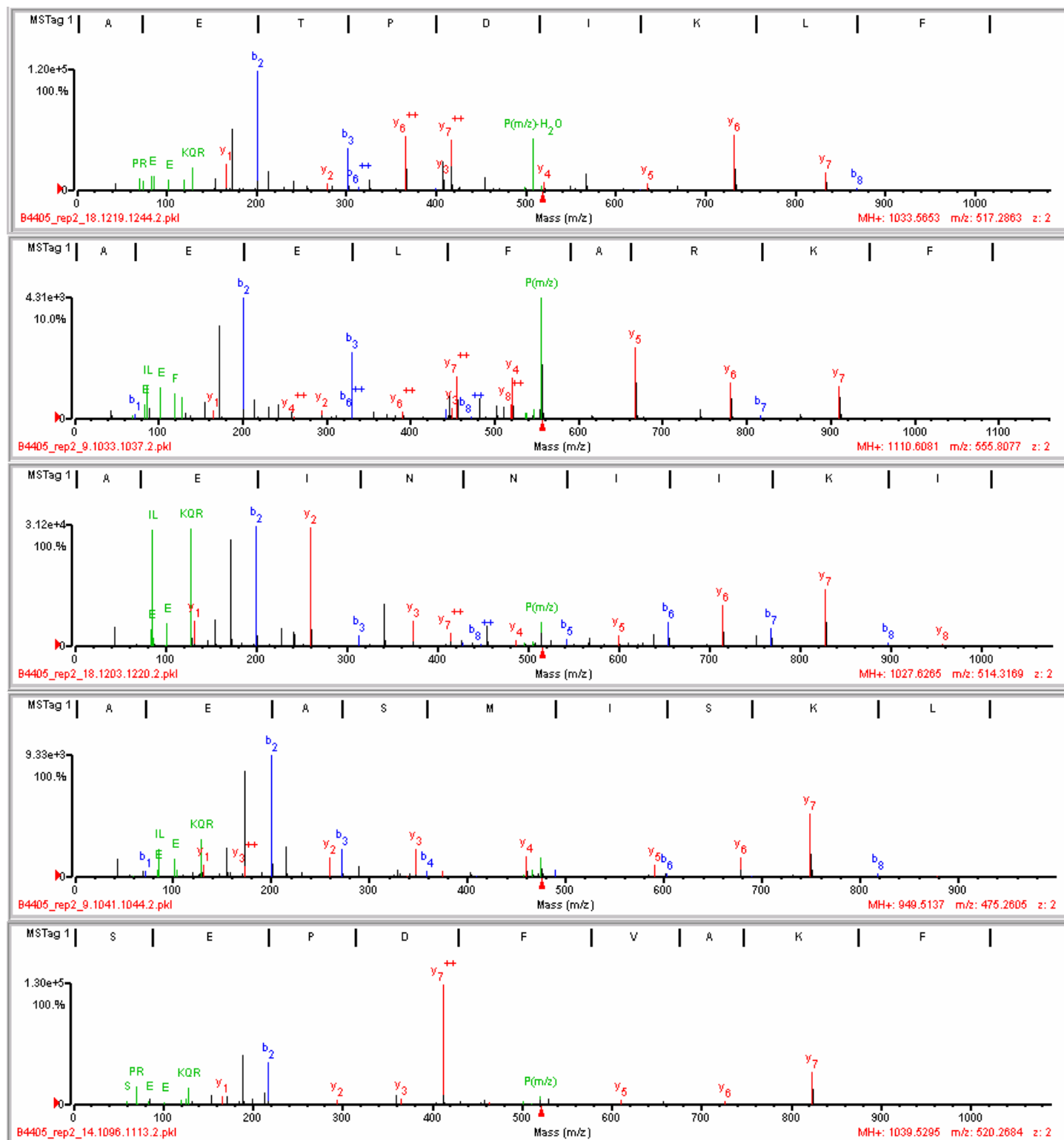


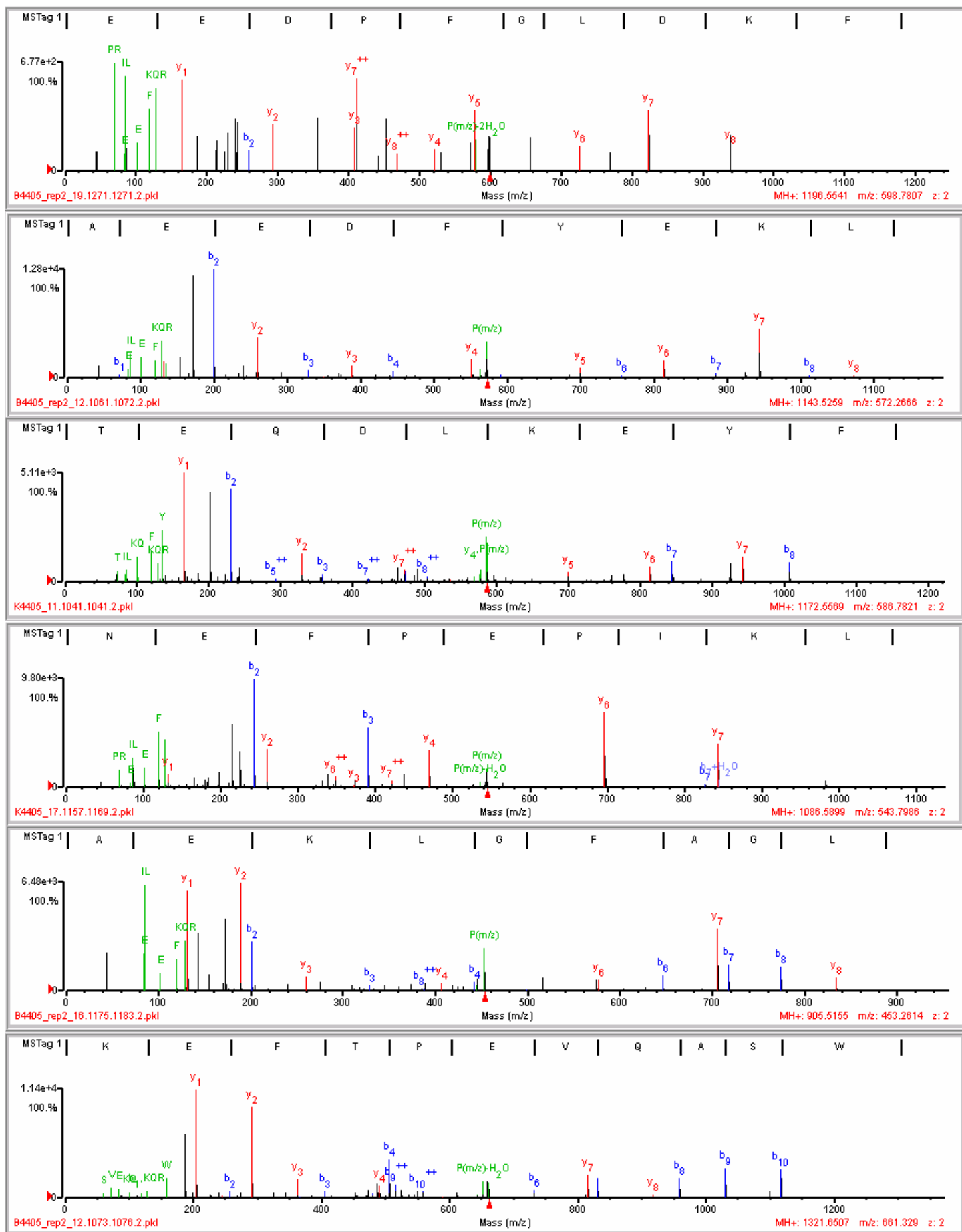


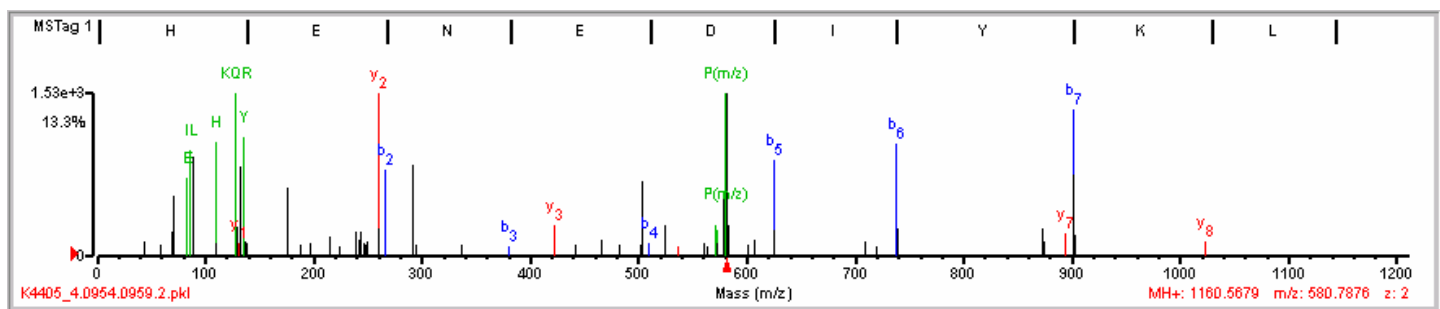
D. QTOF spectra MHC class I peptides containing Lysine.

Not.e. By the software tool design not all the signals would be visible in the following spectra. A complete or specific count of them may require additional manipulations. When a spectrum is repeated to show additional signals both are emphasized in one box.

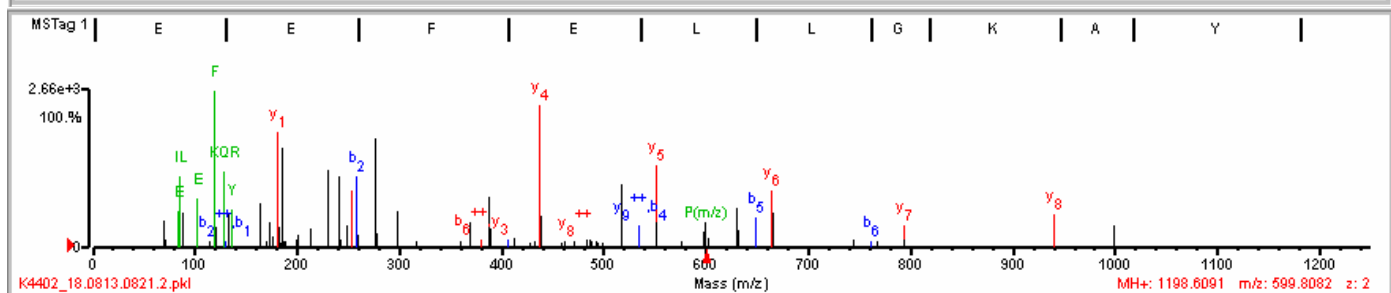
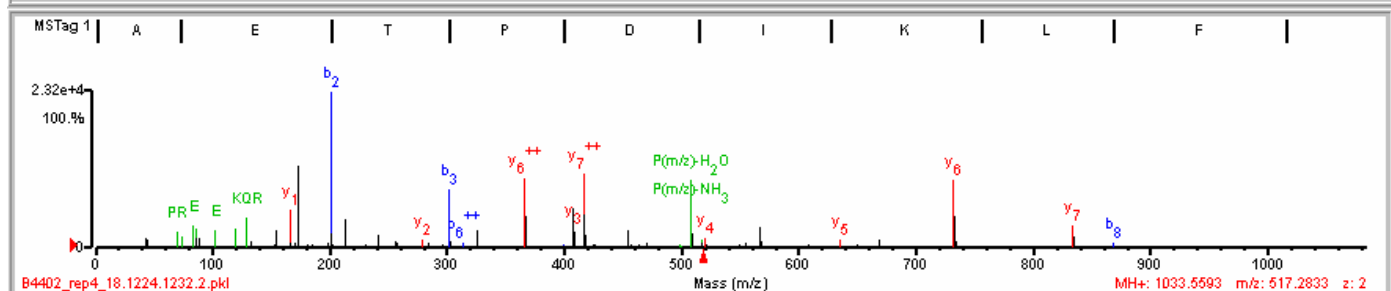
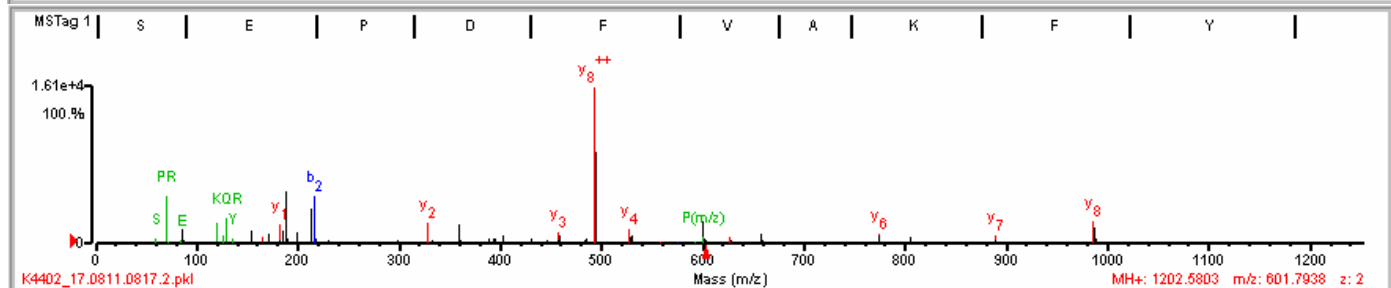
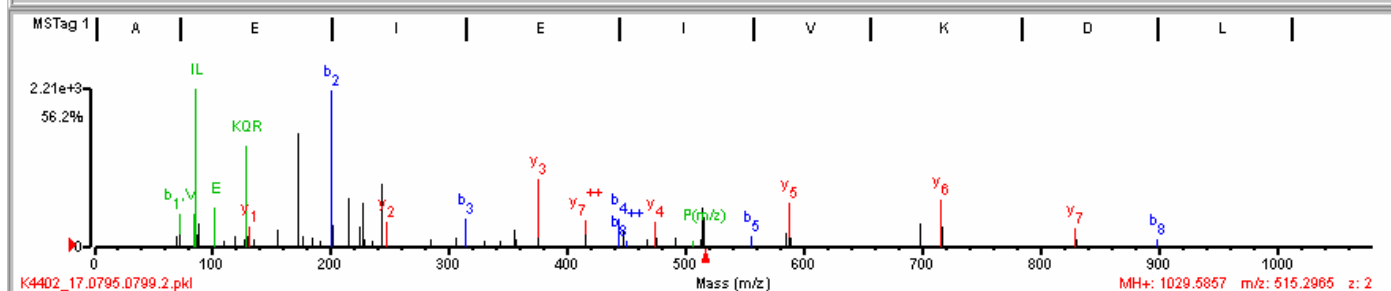
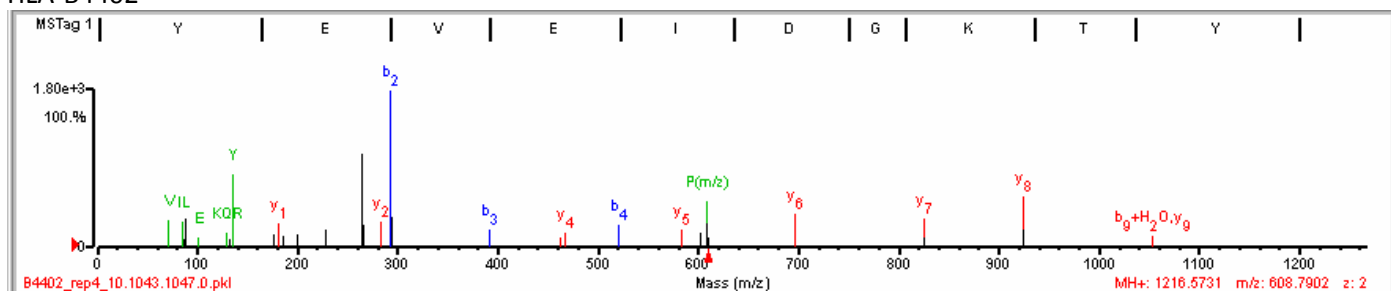
HLA-B4405

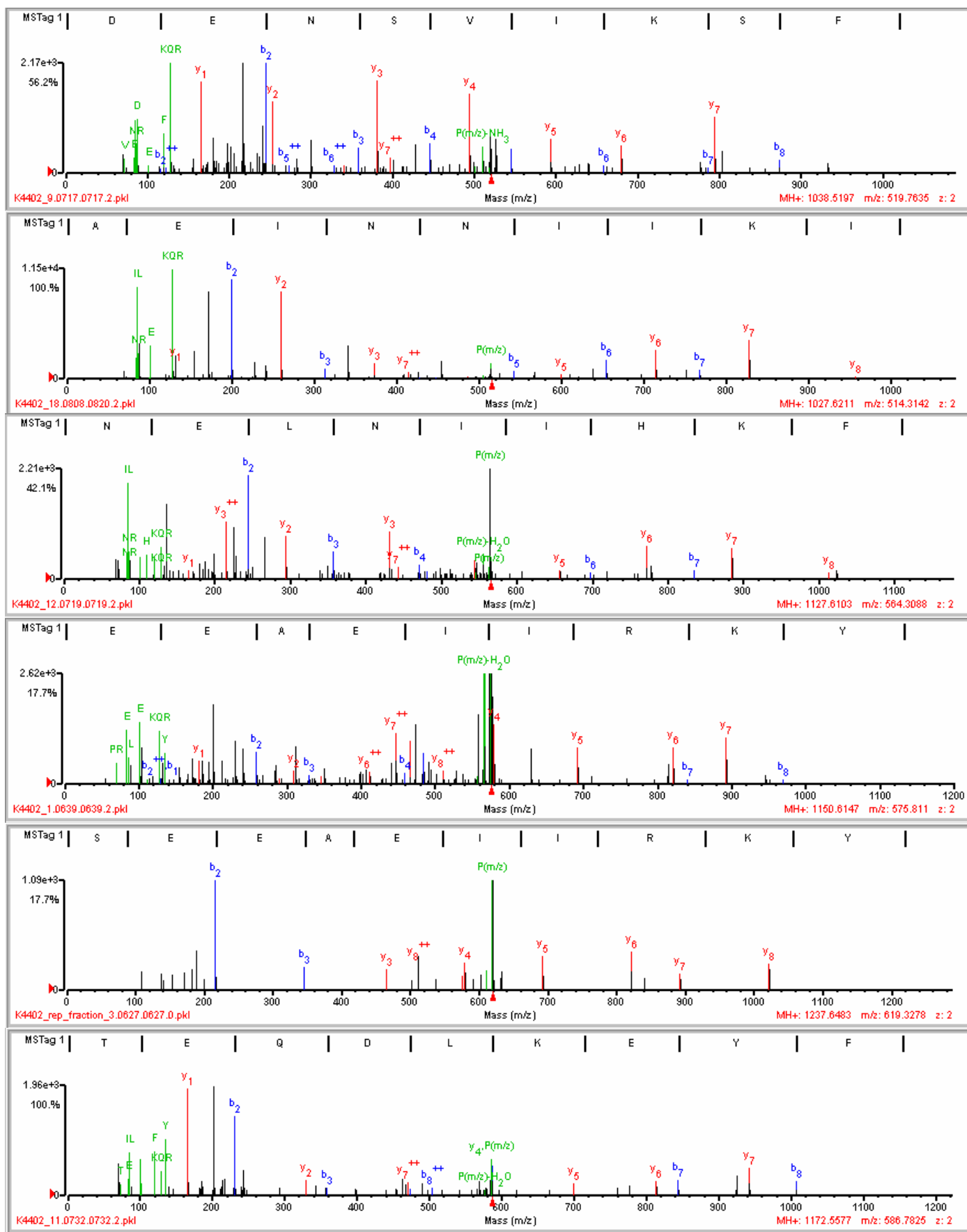


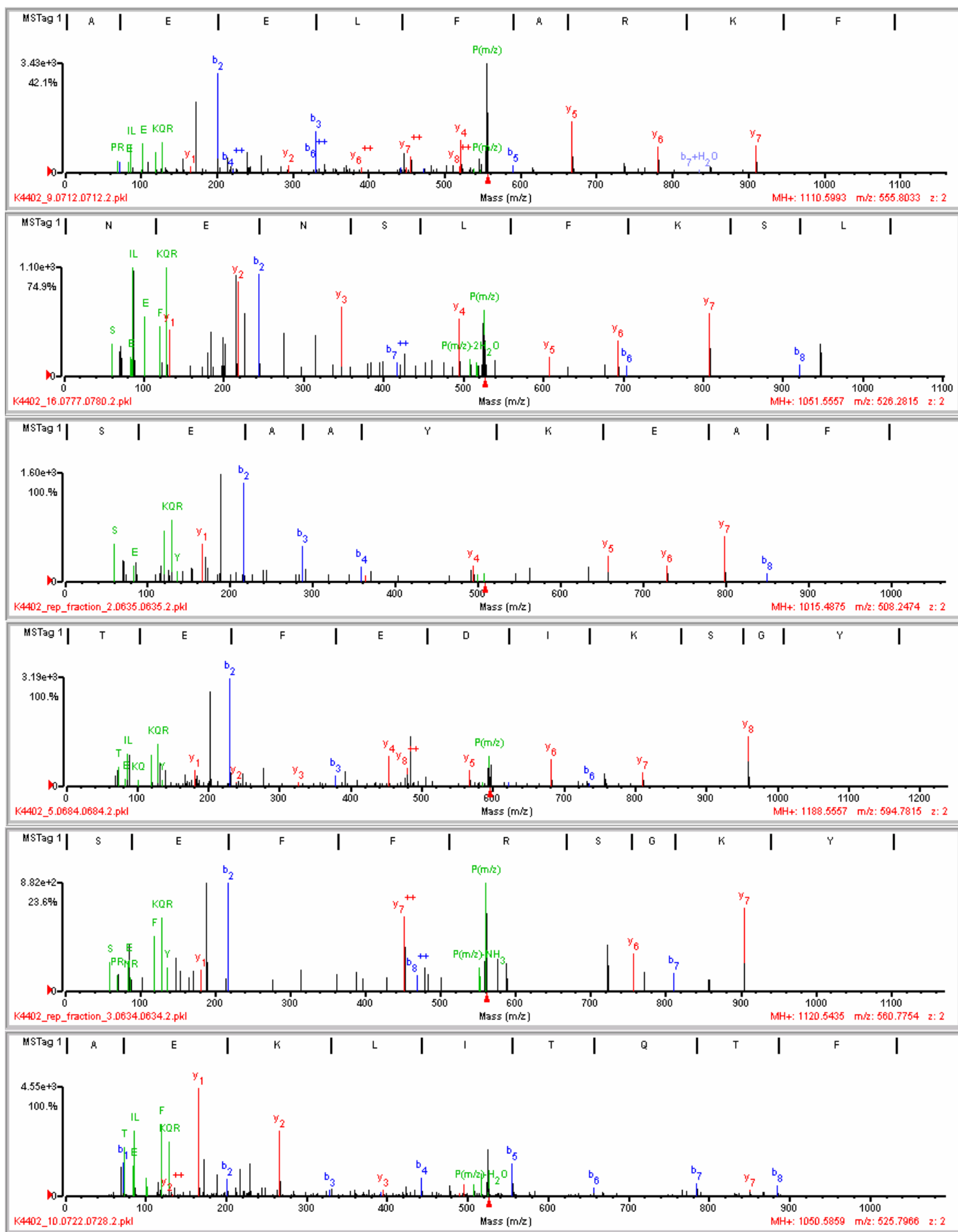


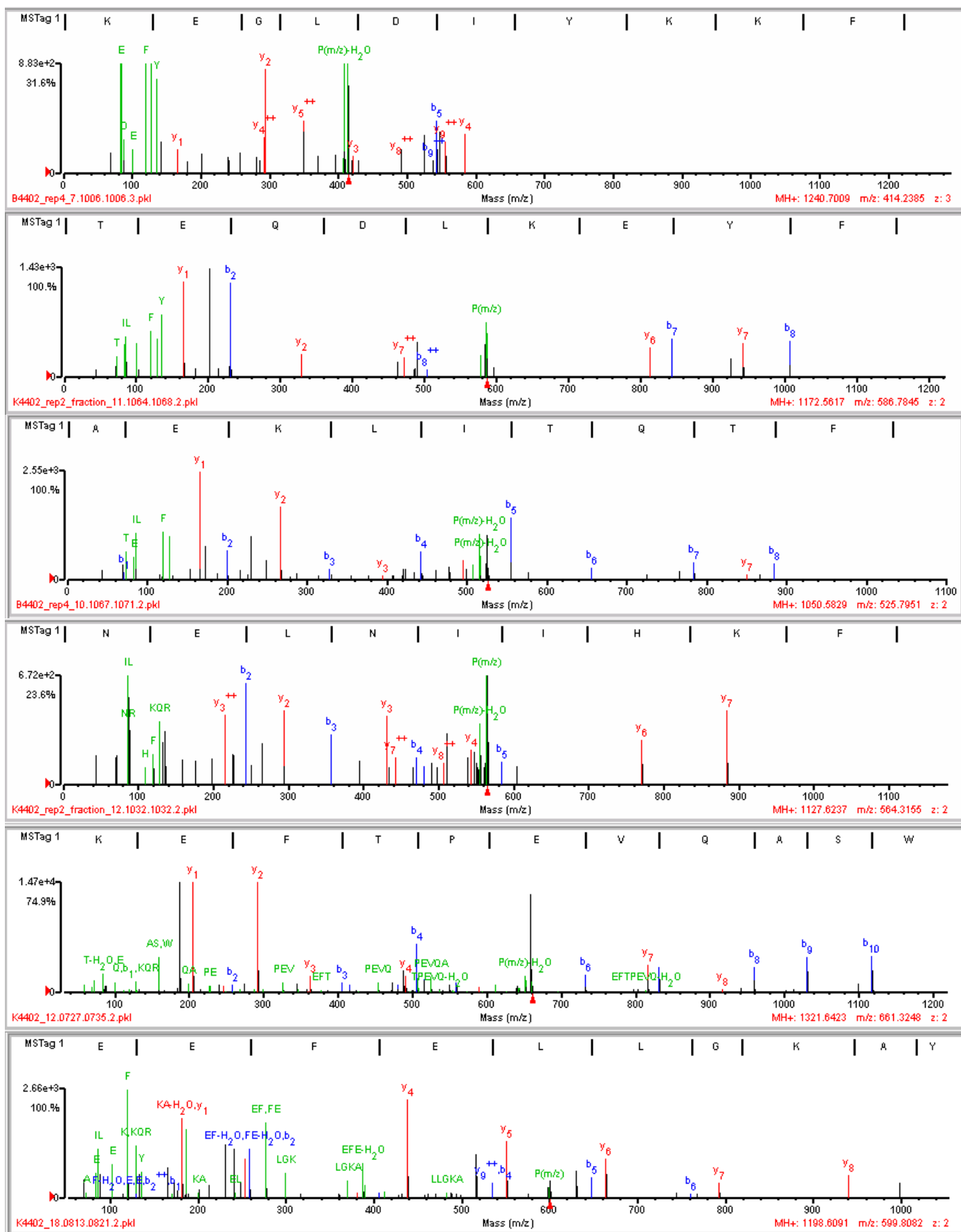


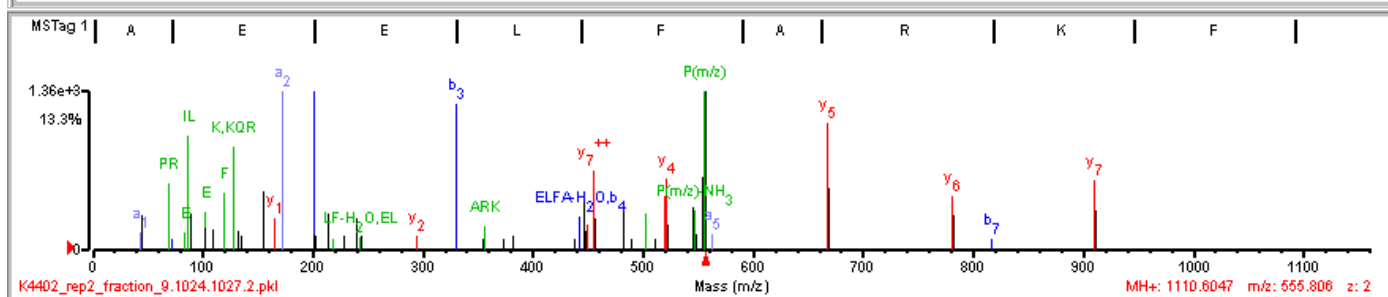
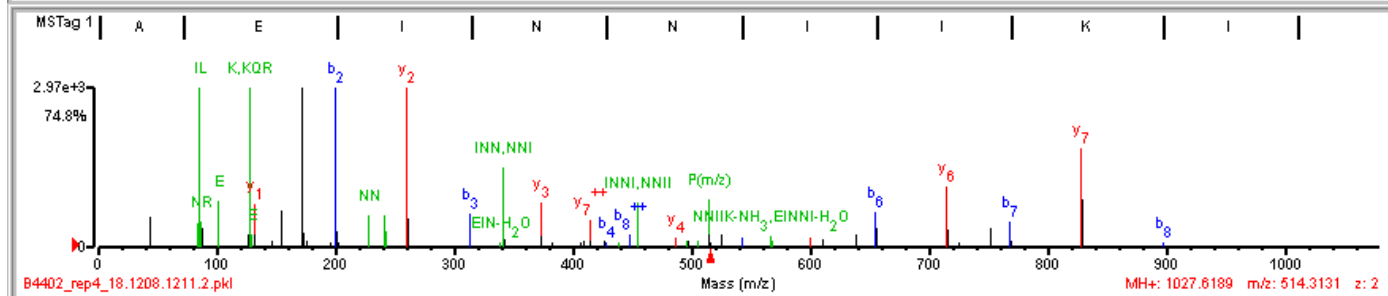
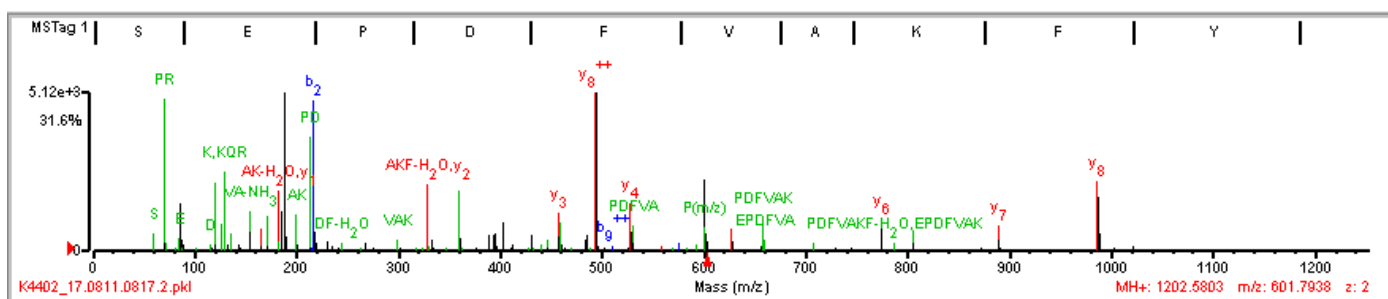
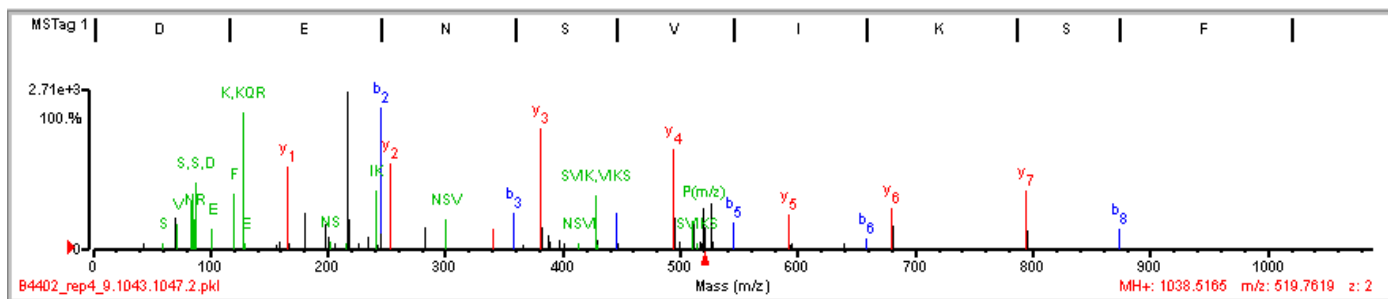
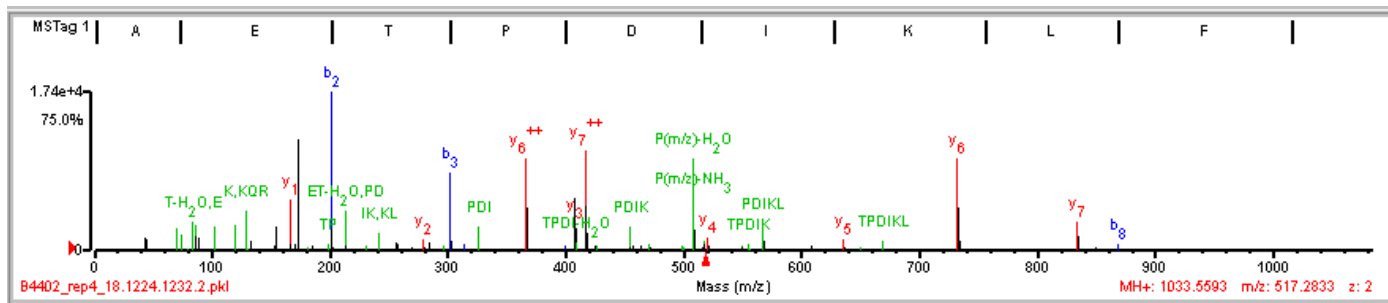
HLA-B4402

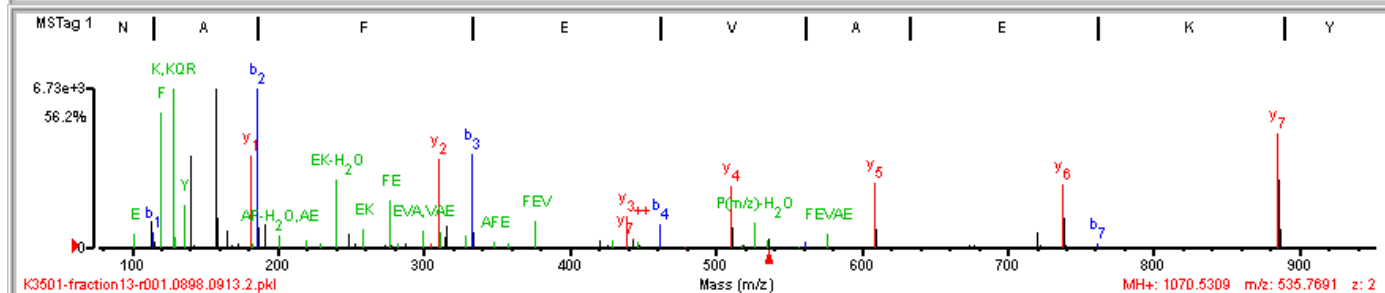
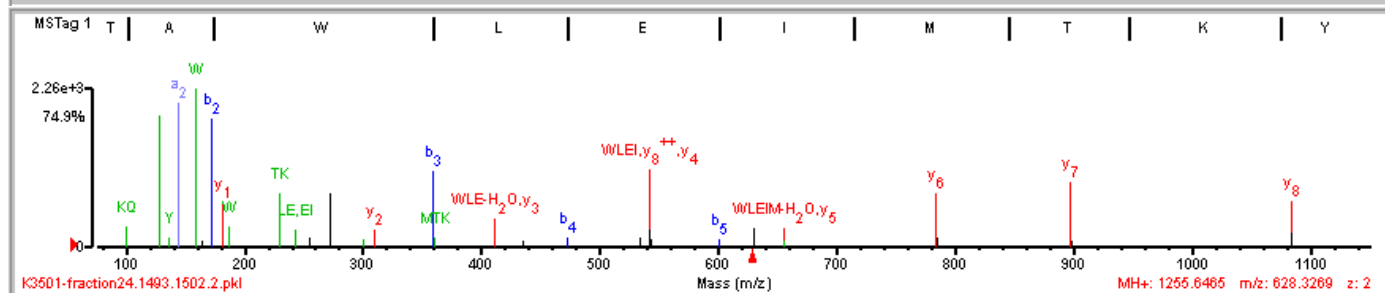
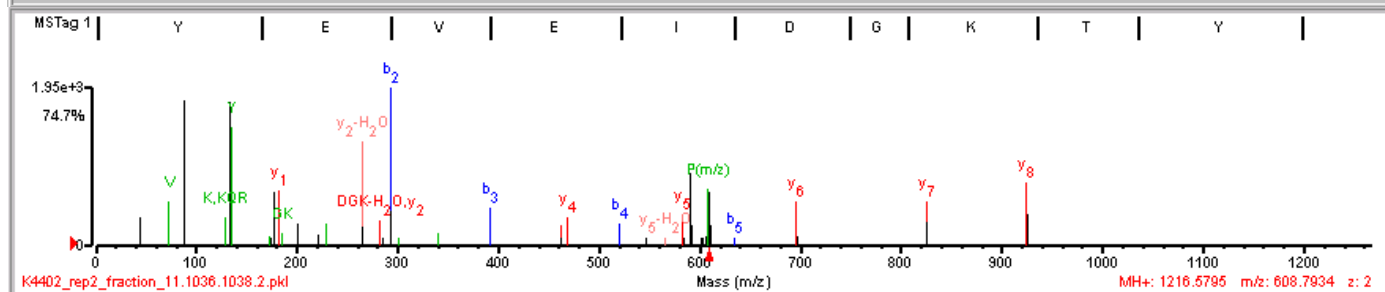
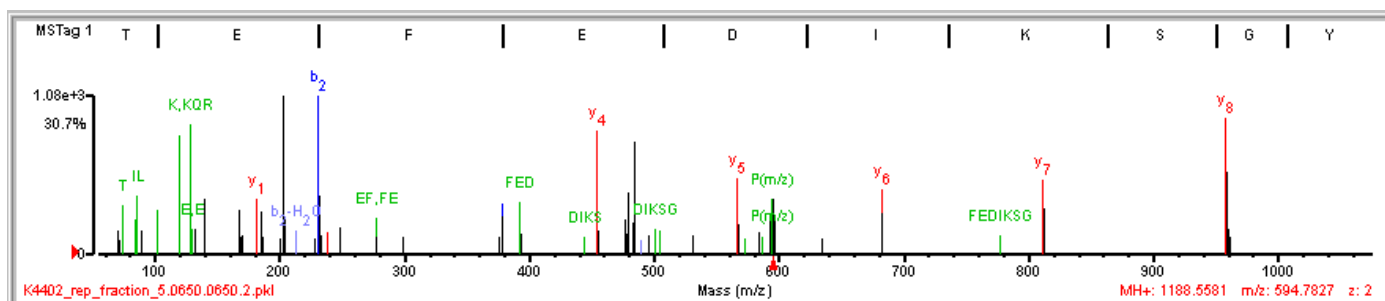
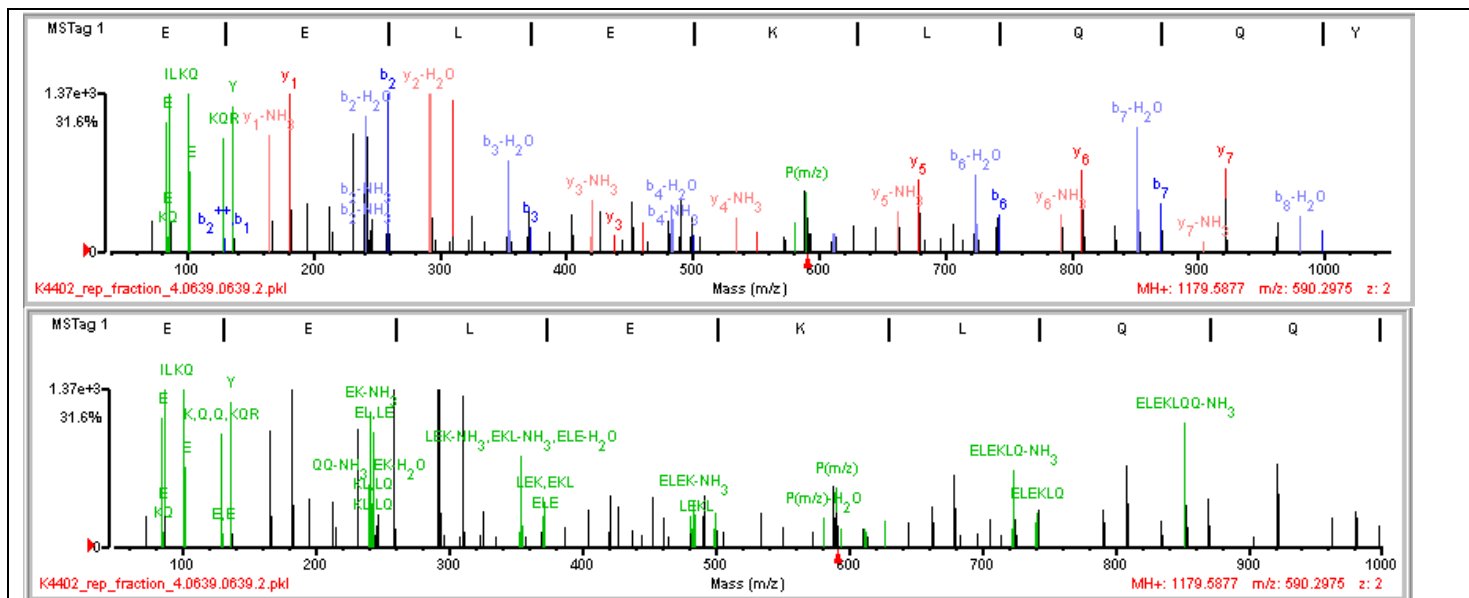




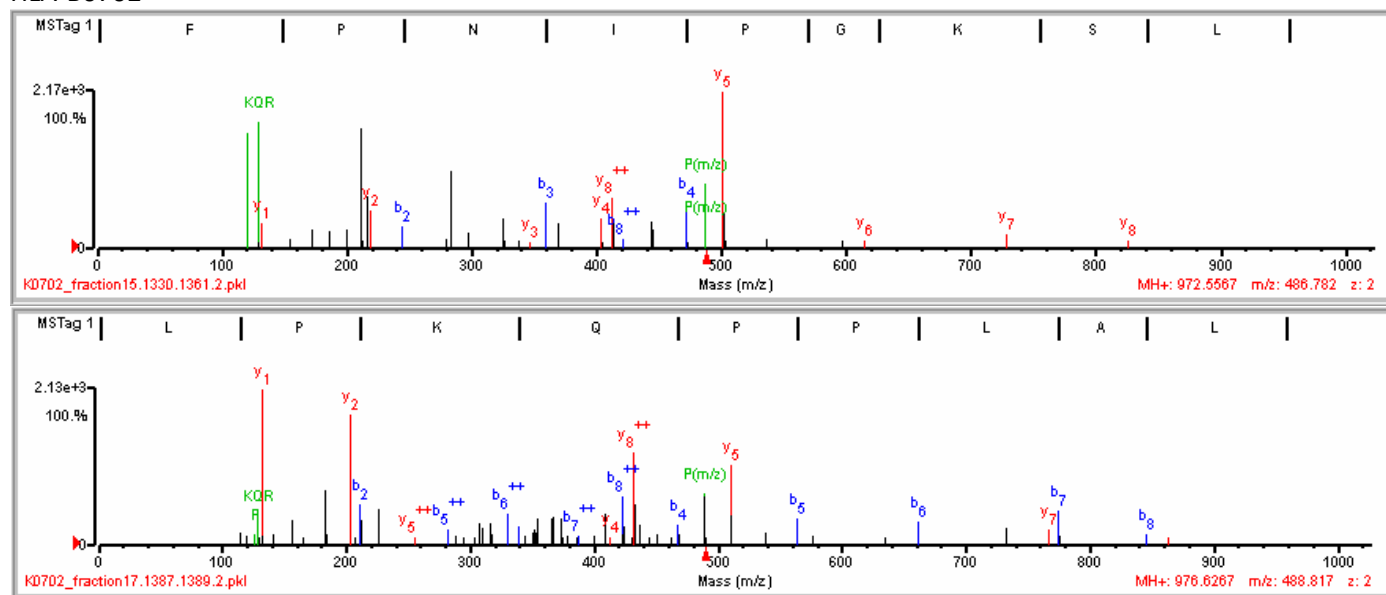




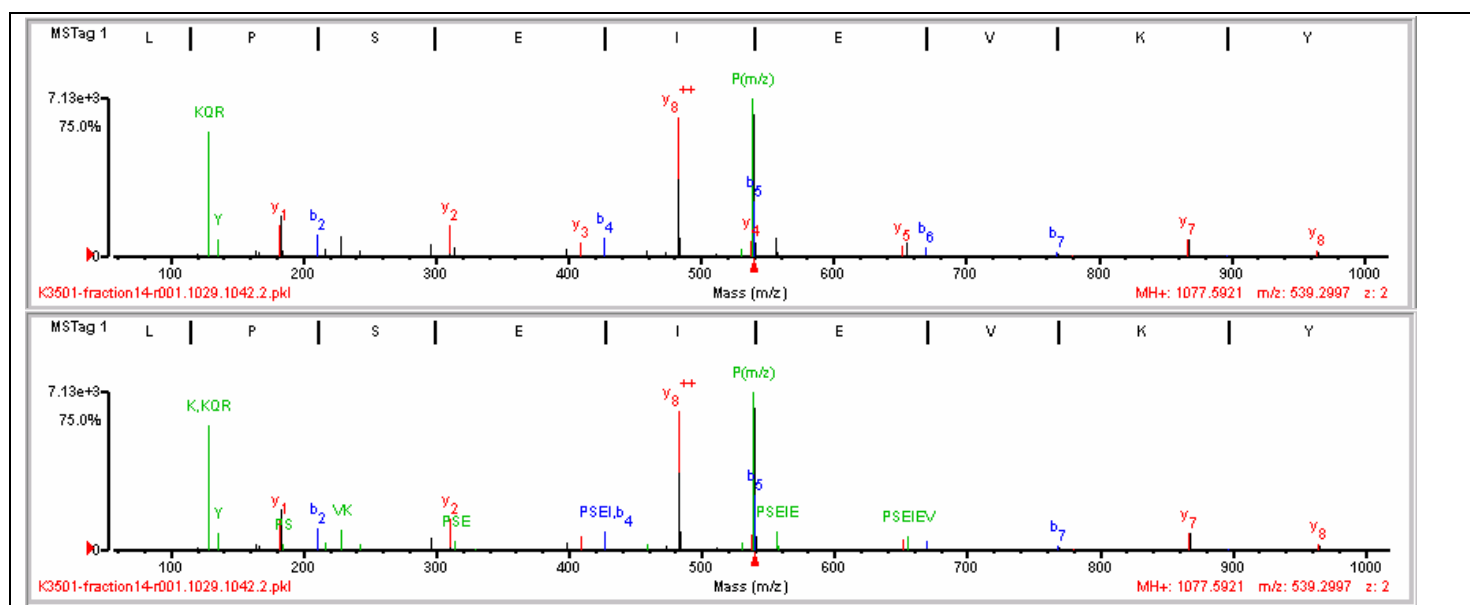
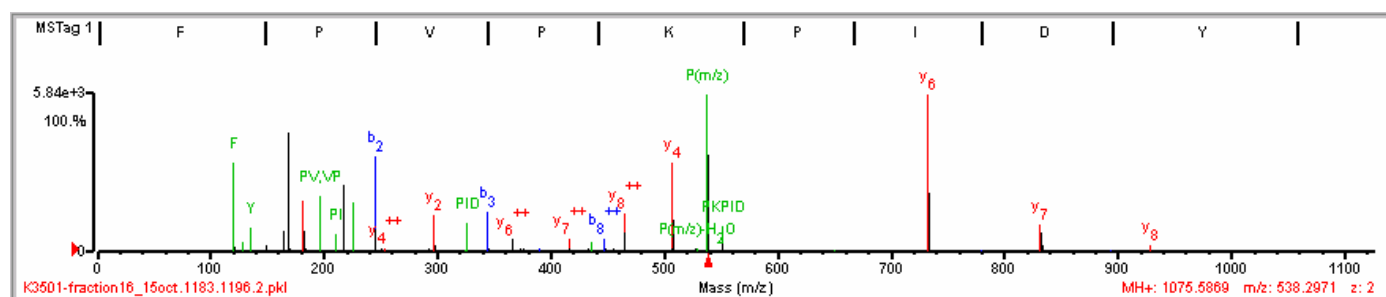


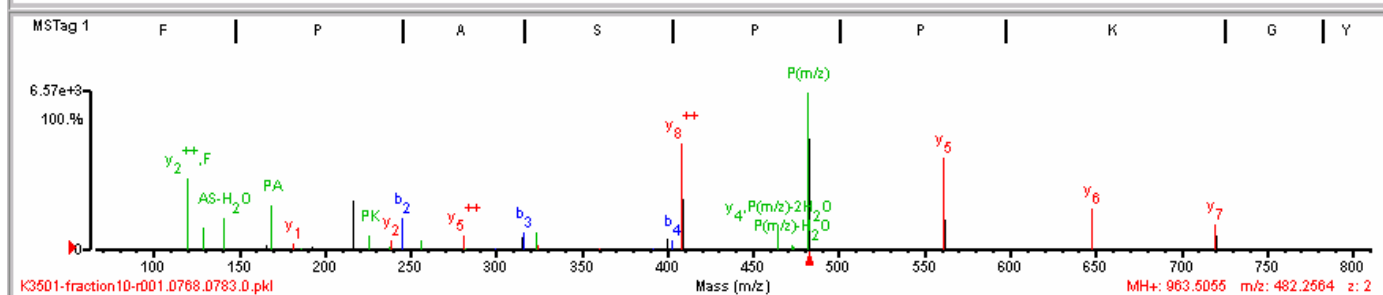
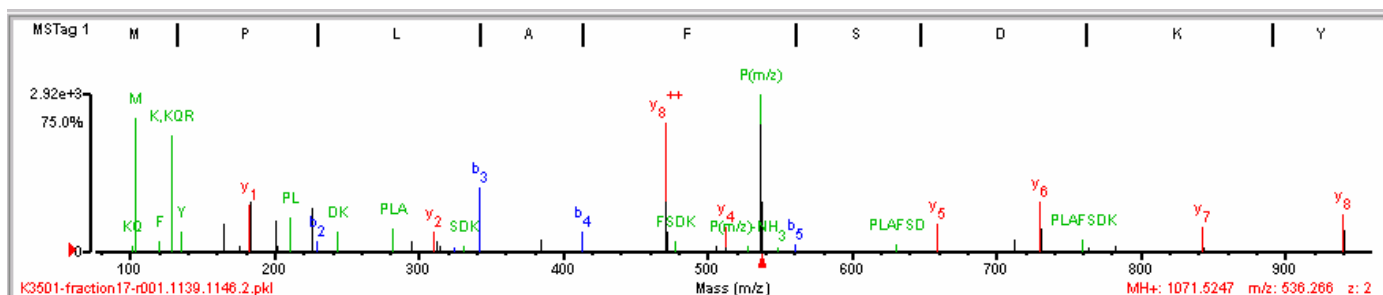
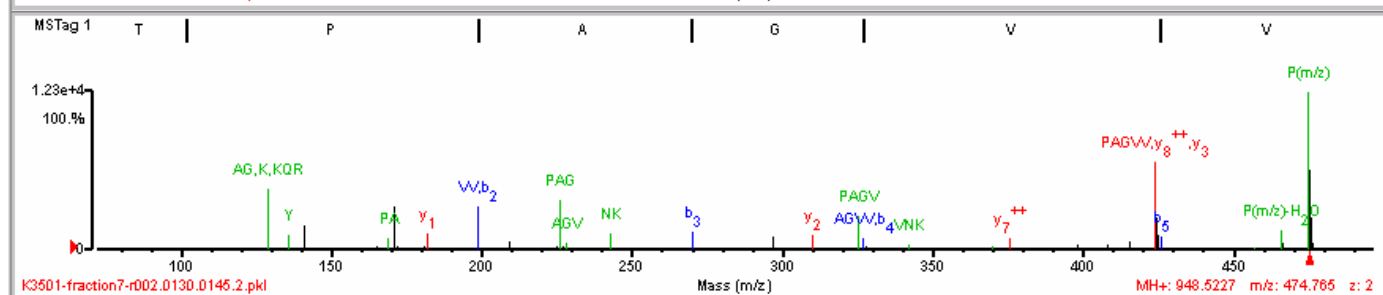
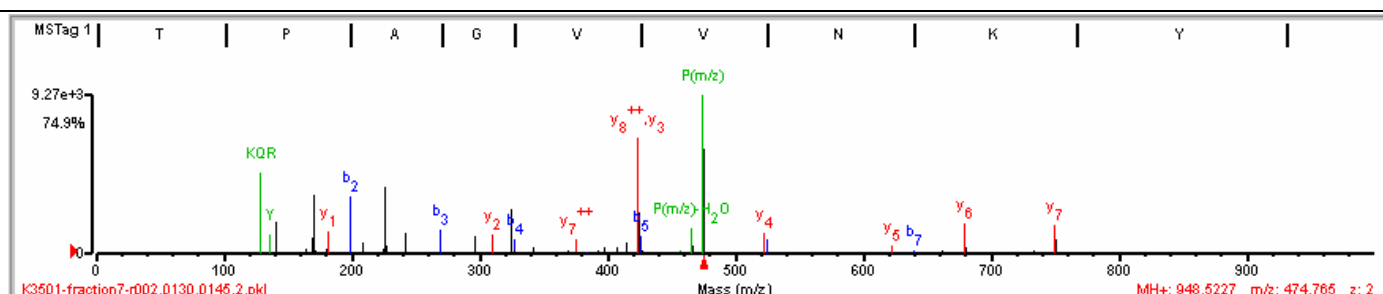
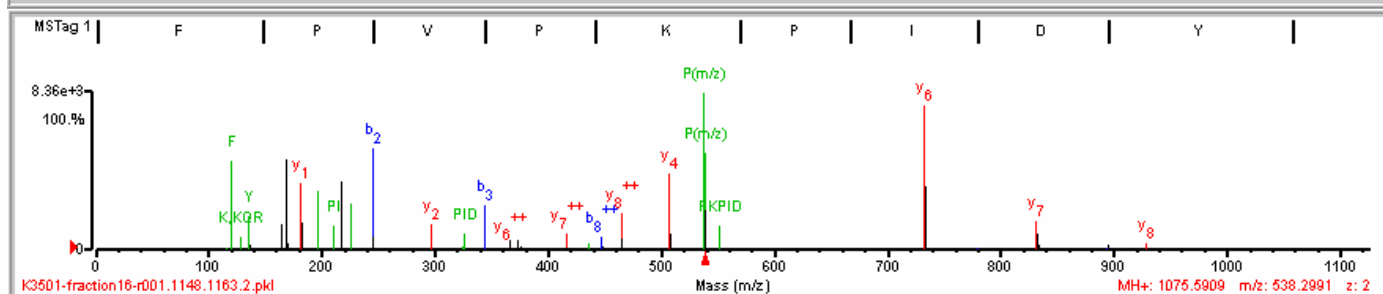
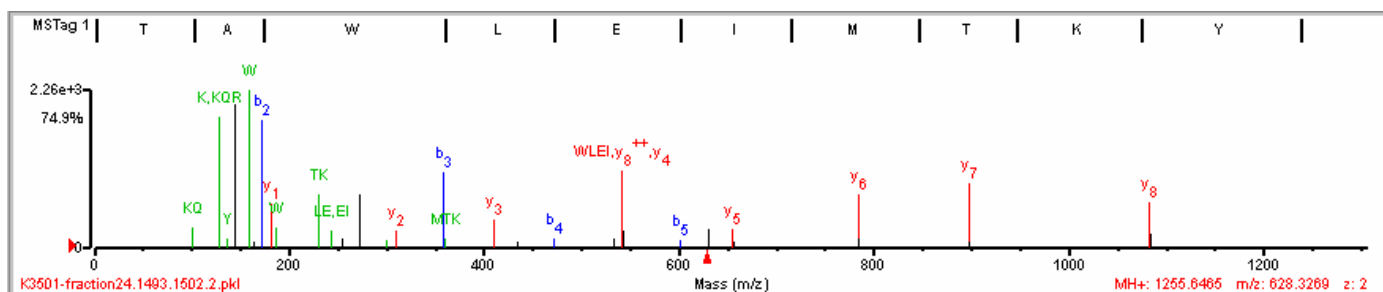


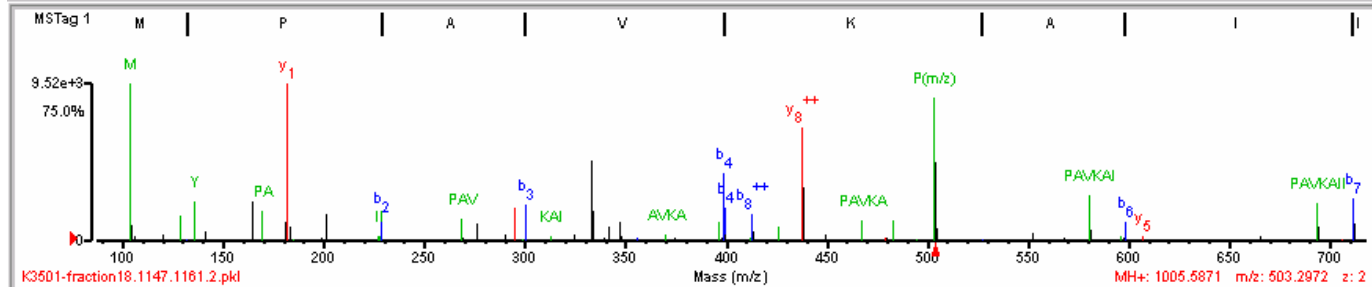
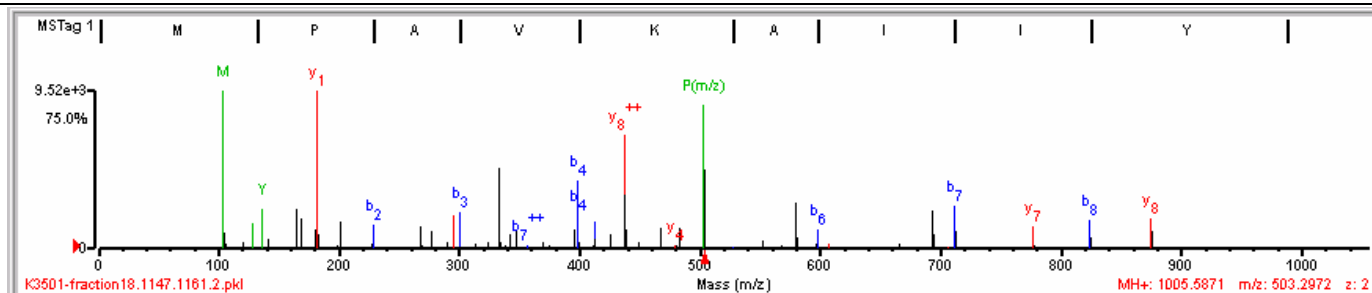
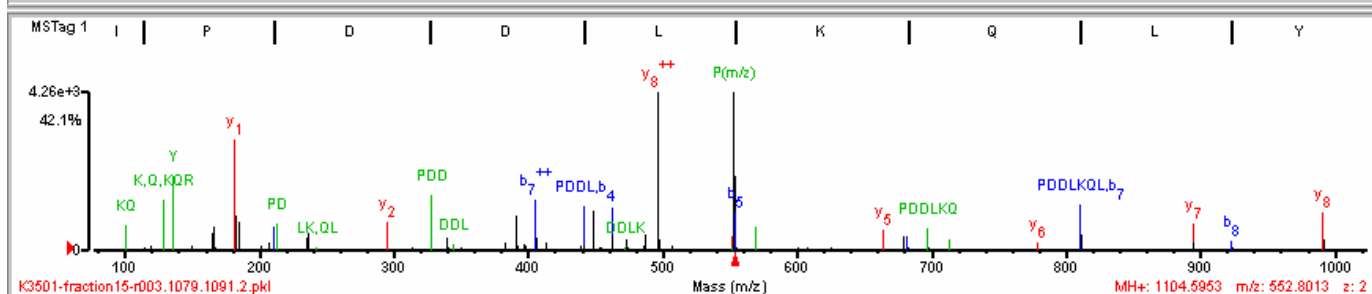
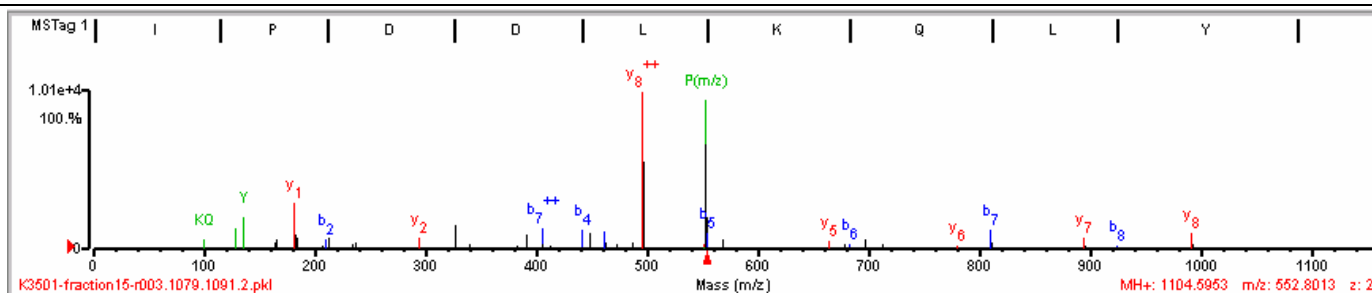
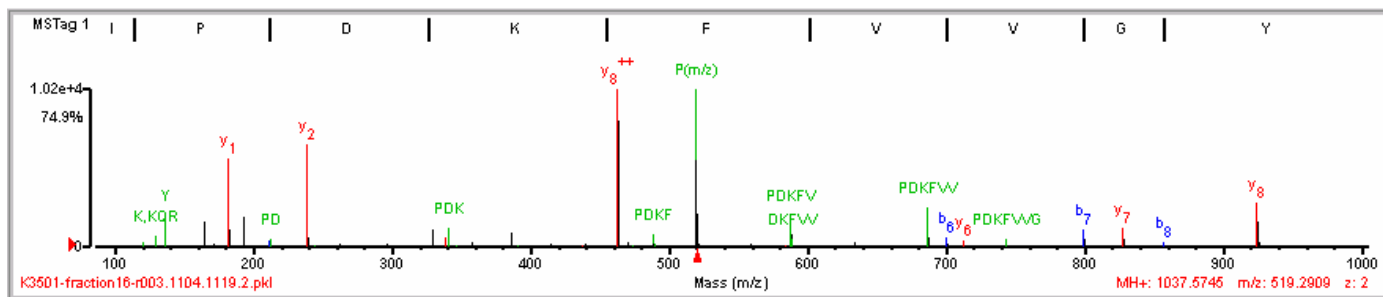
HLA-B0702

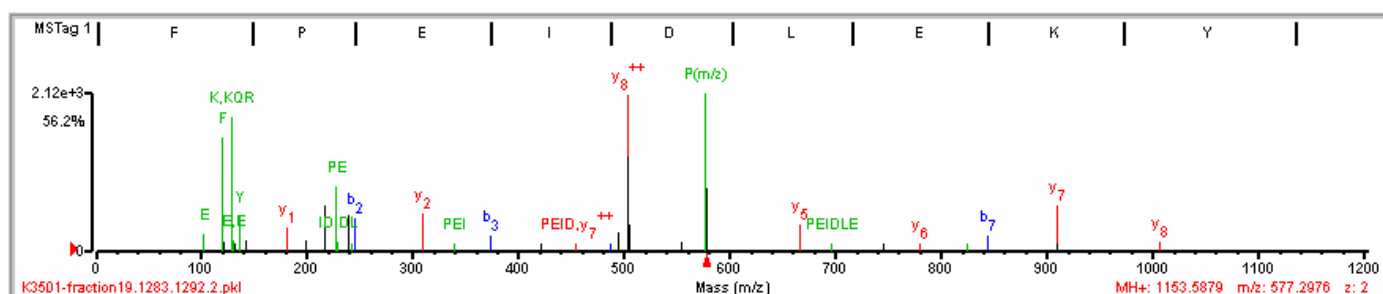
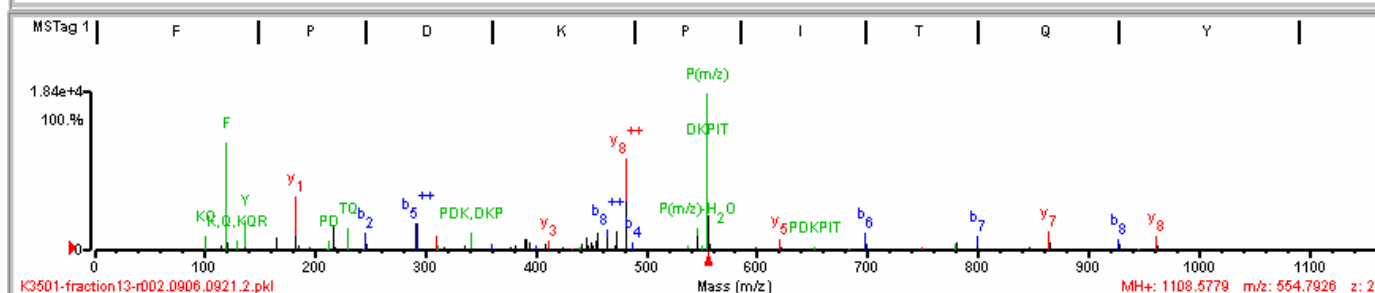
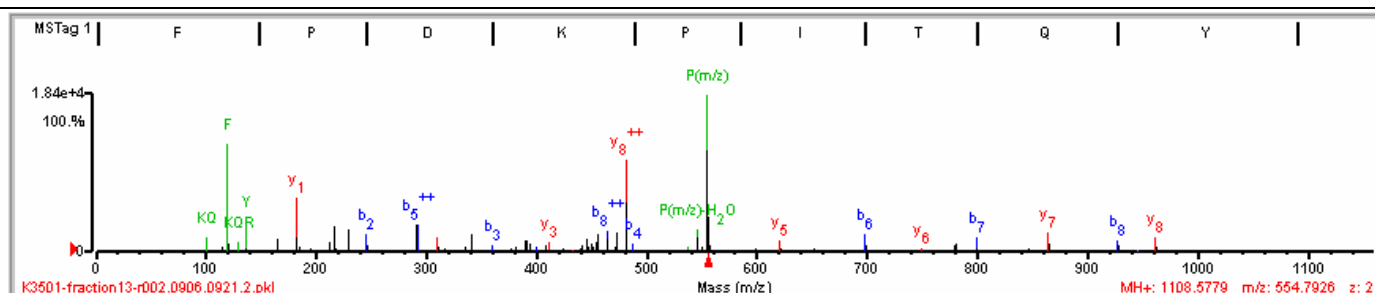
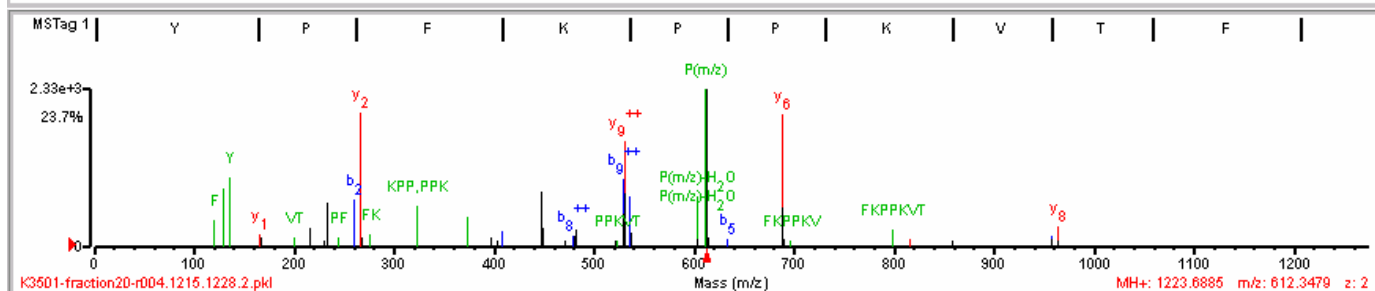
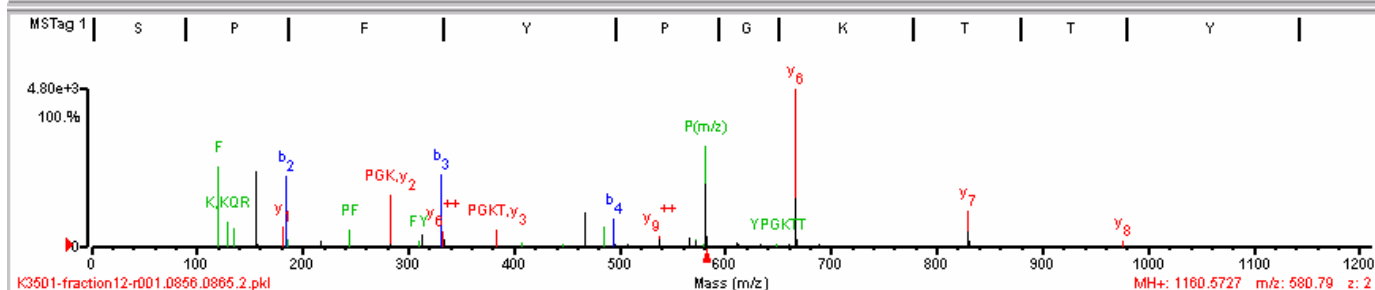
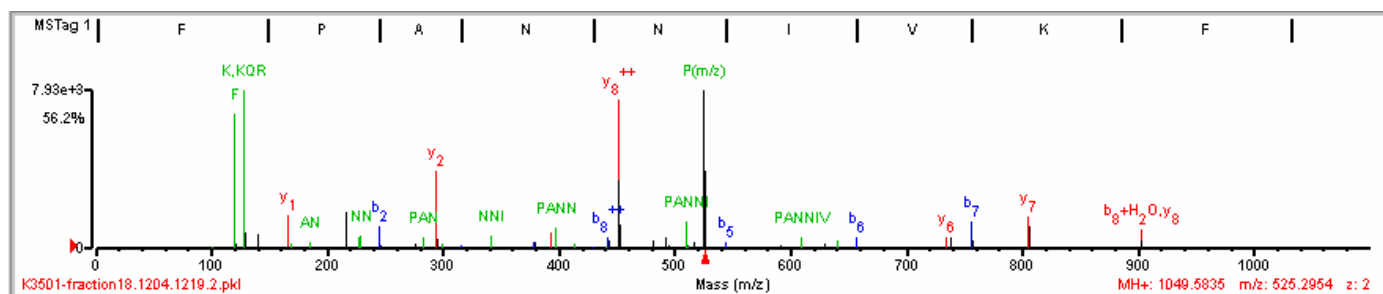


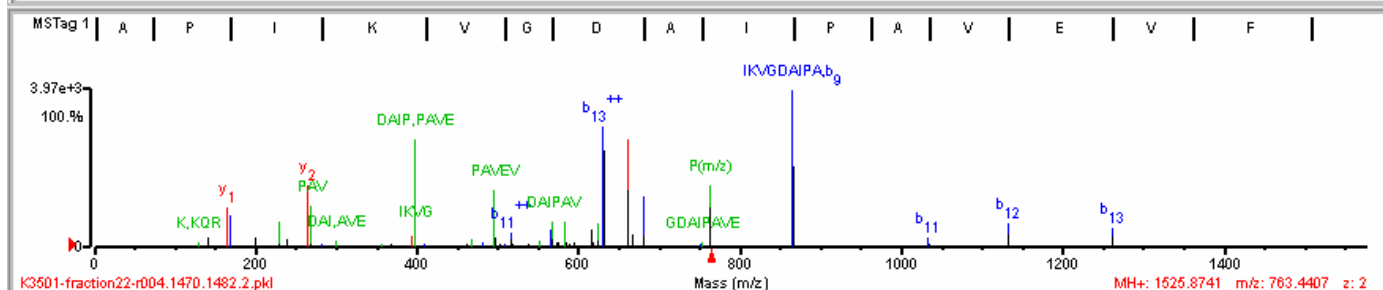
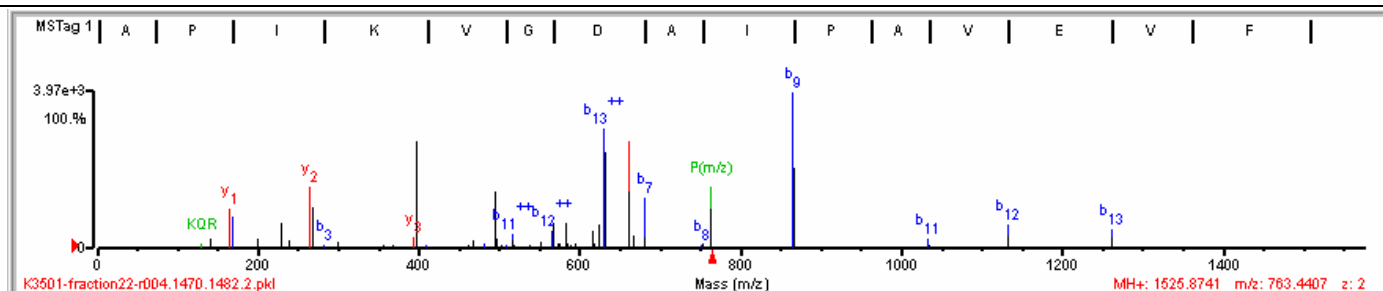
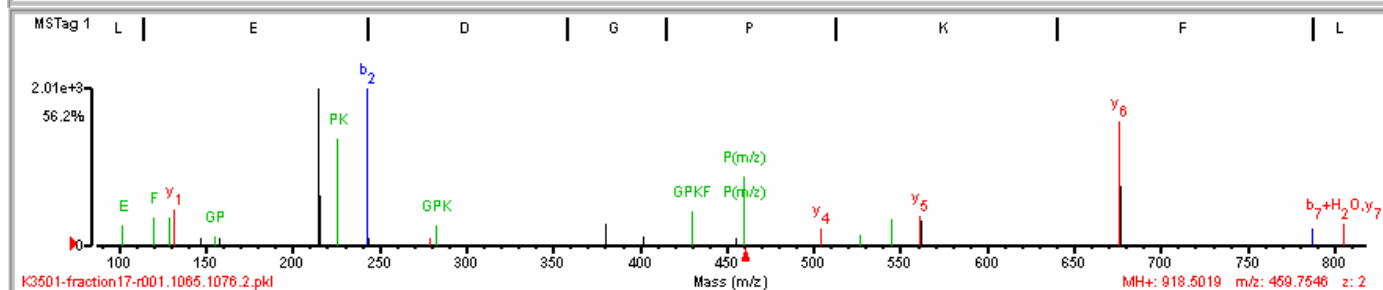
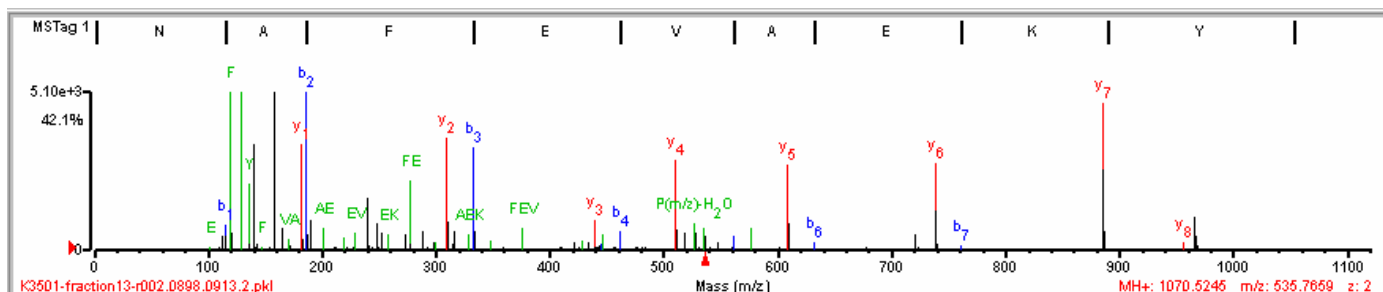
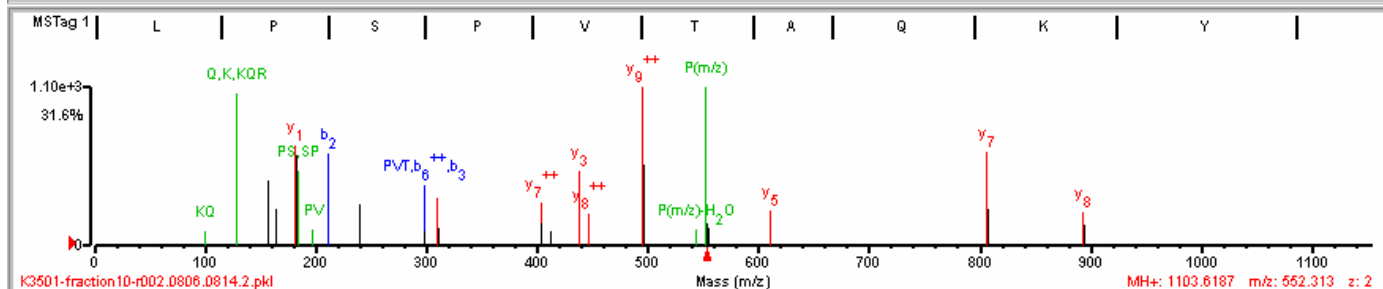
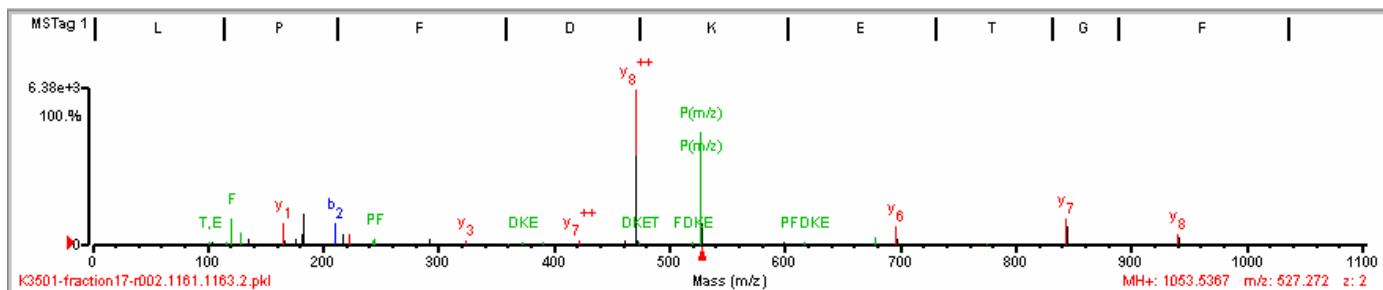
HLA-B3501

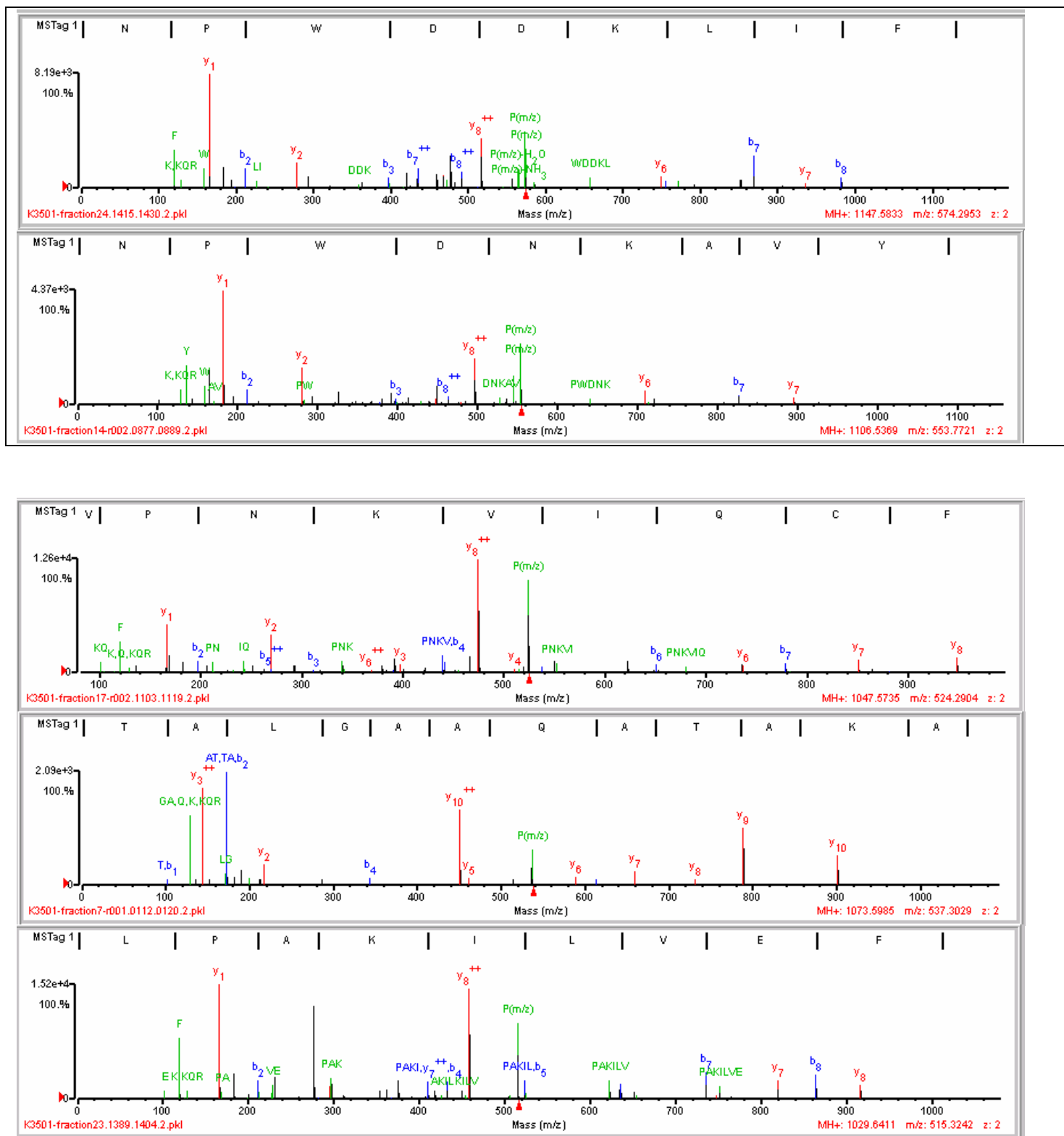


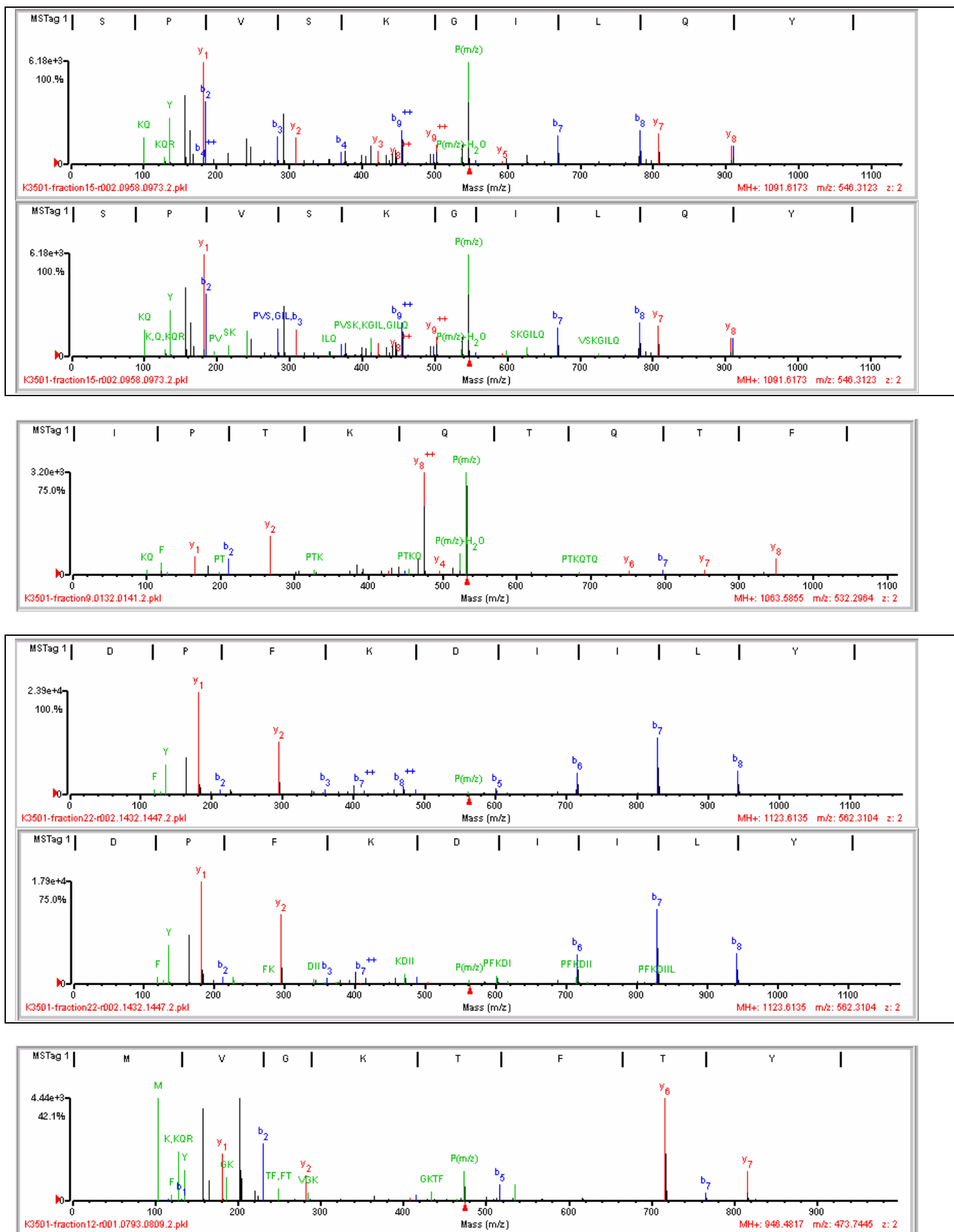


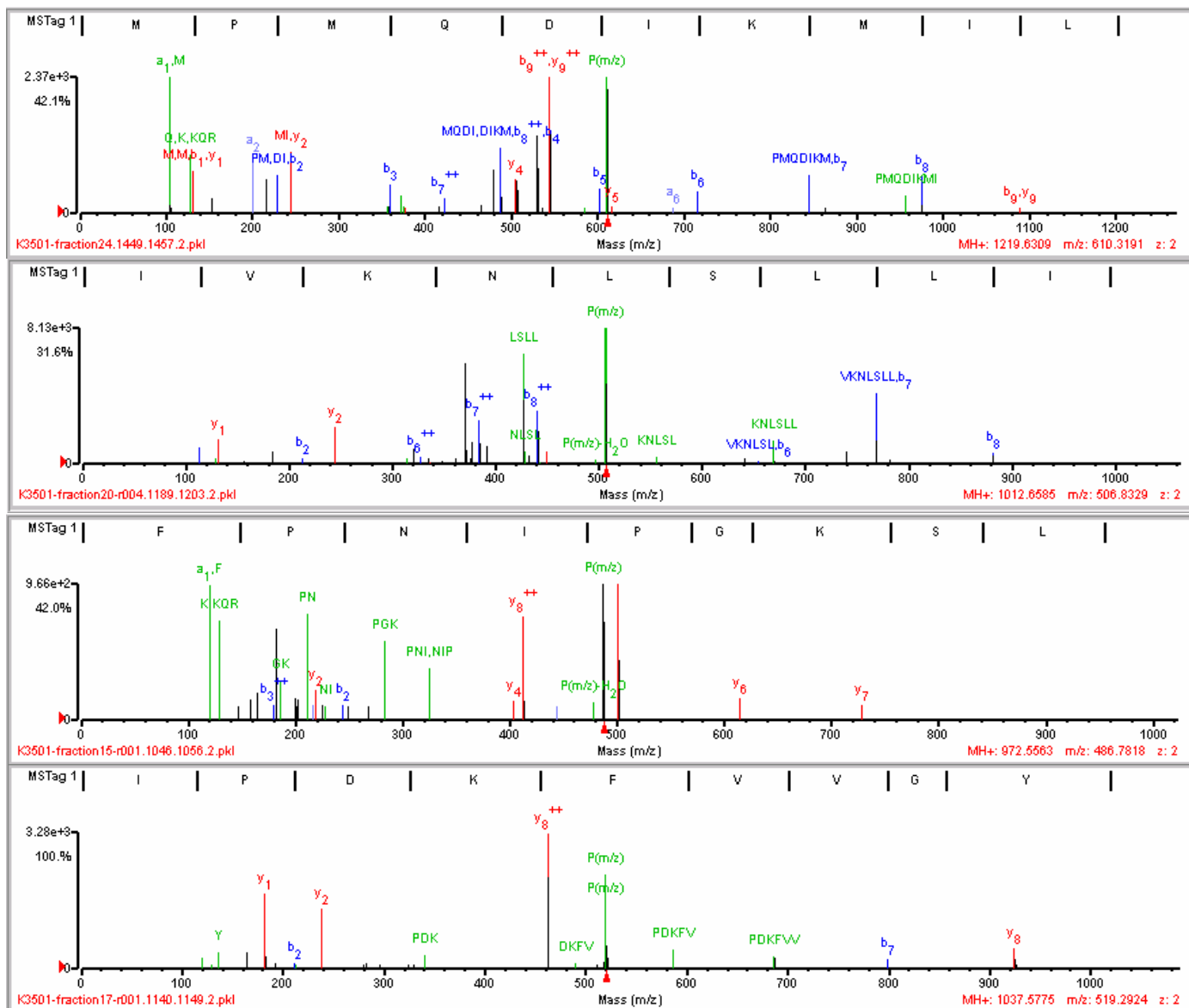












H2-Db

