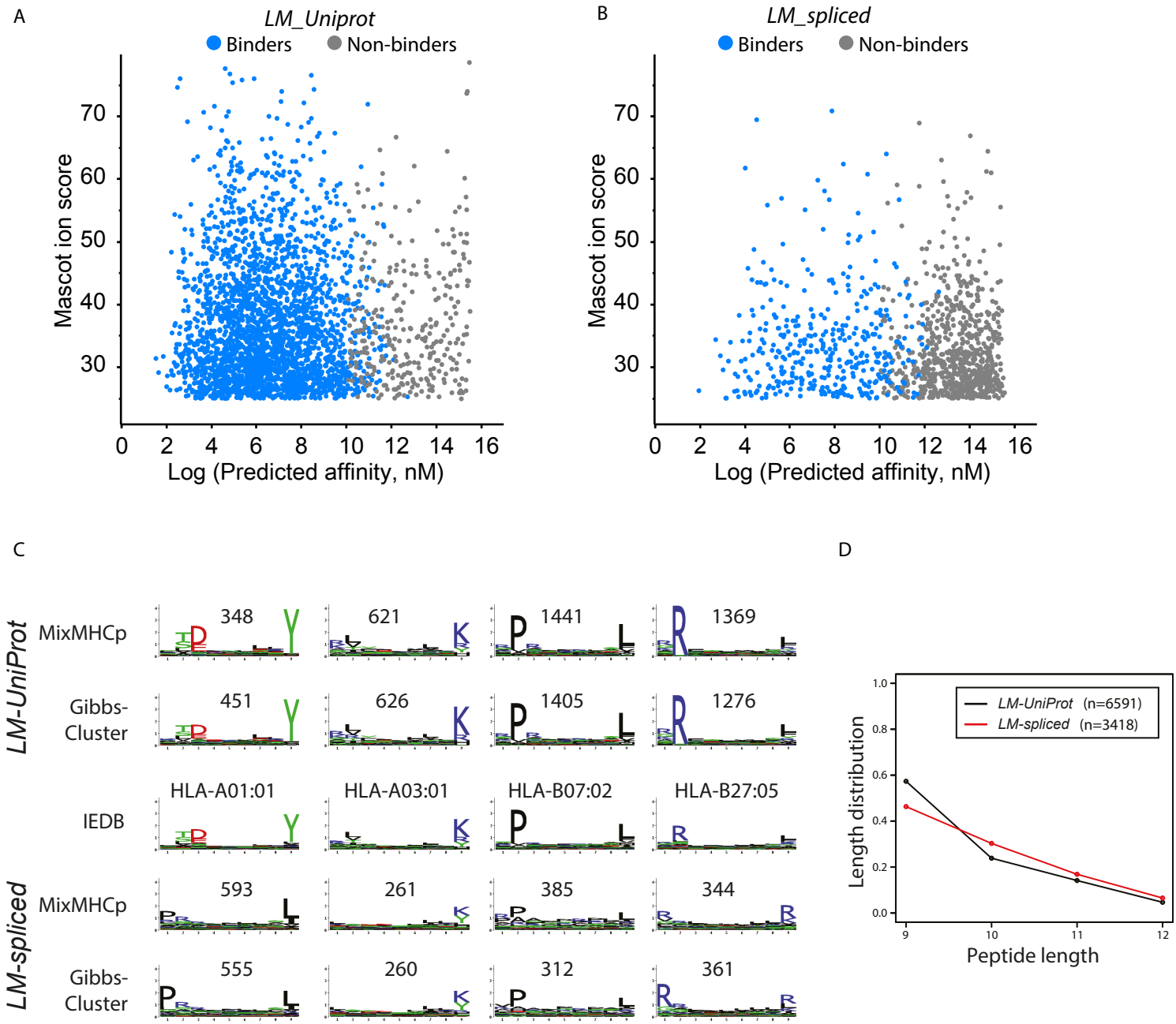
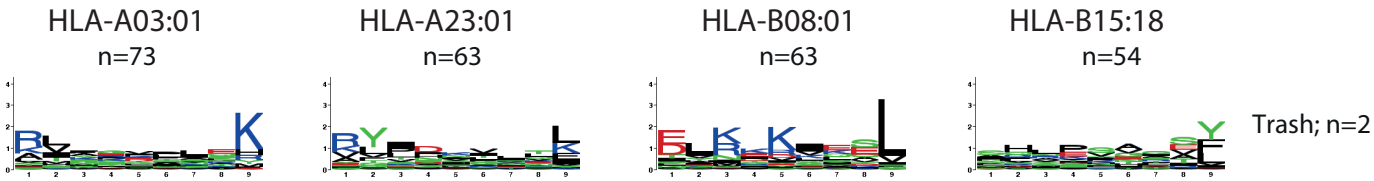
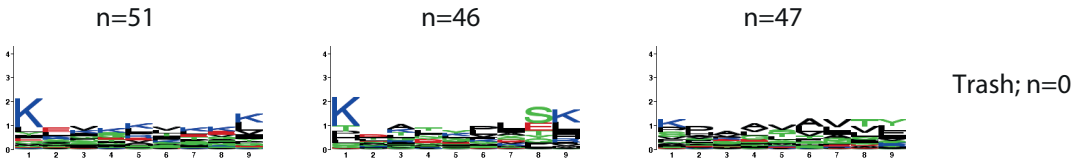


Supplemental Figure 1: Re-analysis of the Liepe et al. data available as Supplemental Data 2. A) Correlation of the predicted binding affinity by NetMHCpan and MS/MS identification scores of peptides identified by Liepe et al. as *LM_UniProt* (A) and *LM_spliced* (B) in the Fib dataset. Binders (rank top 2%) are marked in blue and non-binders in gray C) Motif deconvolution analysis with MixMHCp and GibbsCluster of the 9-mer *LM-UniProt* and *LM-spliced* peptides identified by Liepe et al. in the GR-LCL 2D dataset and comparison to known logos from IEDB for HLA-A01:01, HLA-A03:01 and HLA-B07:02 and HLA-B27:05. Motifs found in *LM-UniProt* peptides are highly reproducible and comparable to the known motifs from IEDB, while this is not the case for motifs found in *LM-spliced* peptides. D) Length distribution of the *LM-UniProt* and *LM-spliced* peptide identified by Liepe et al. in the GR-LCL 2D dataset.

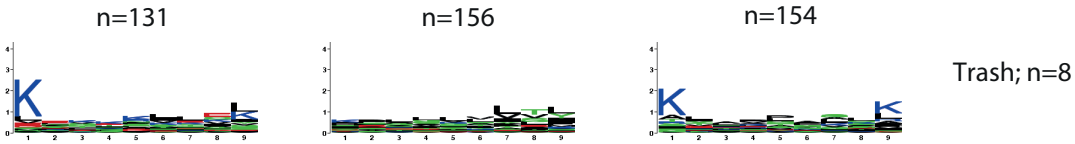
A 9 mer peptides assigend by MaxQuant as UniProt while originally identified as *LM_spliced* by Mascot



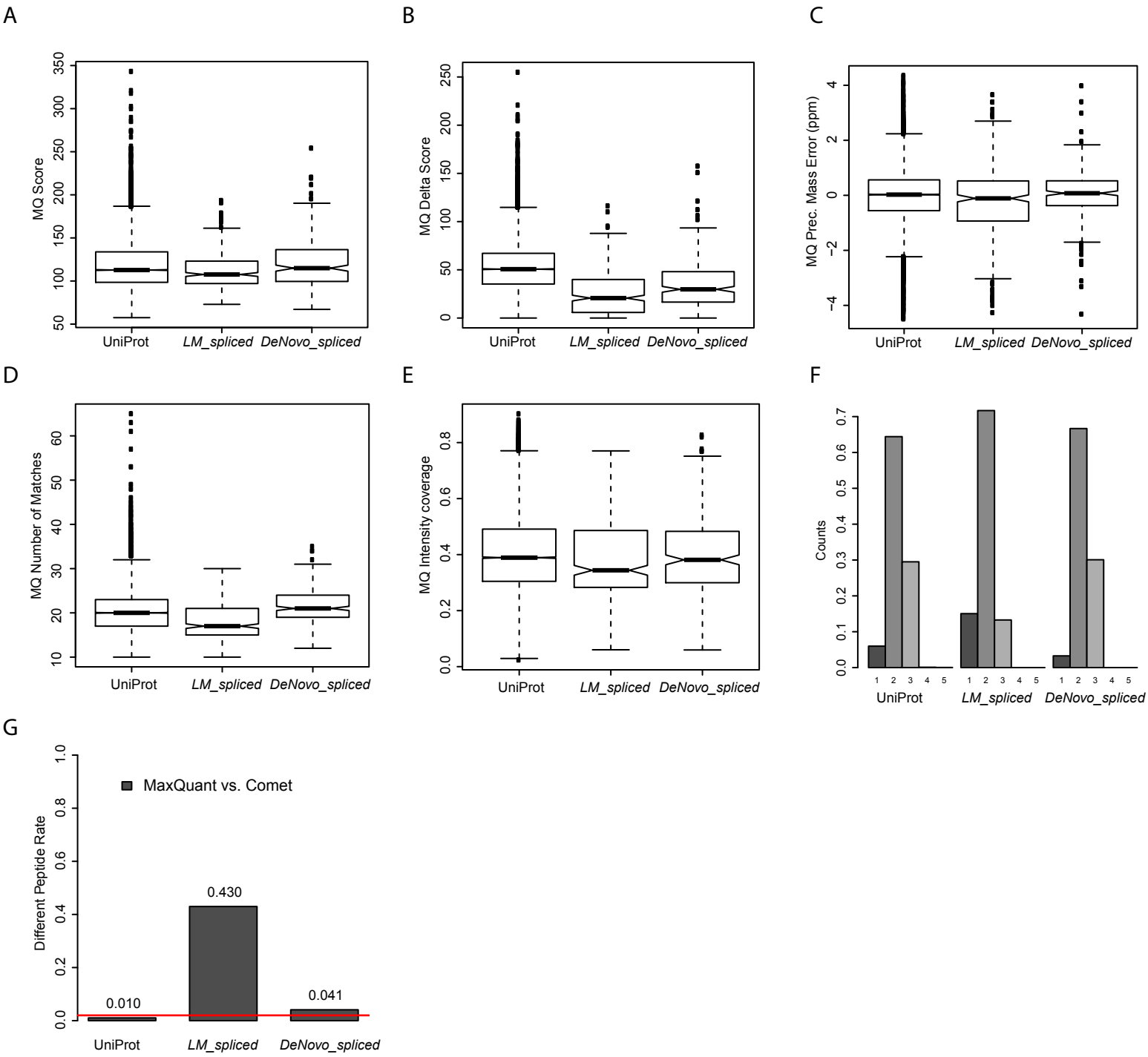
B 9 mer peptides Mascot assigned as *LM_spliced* while MaxQuant as Uniprot



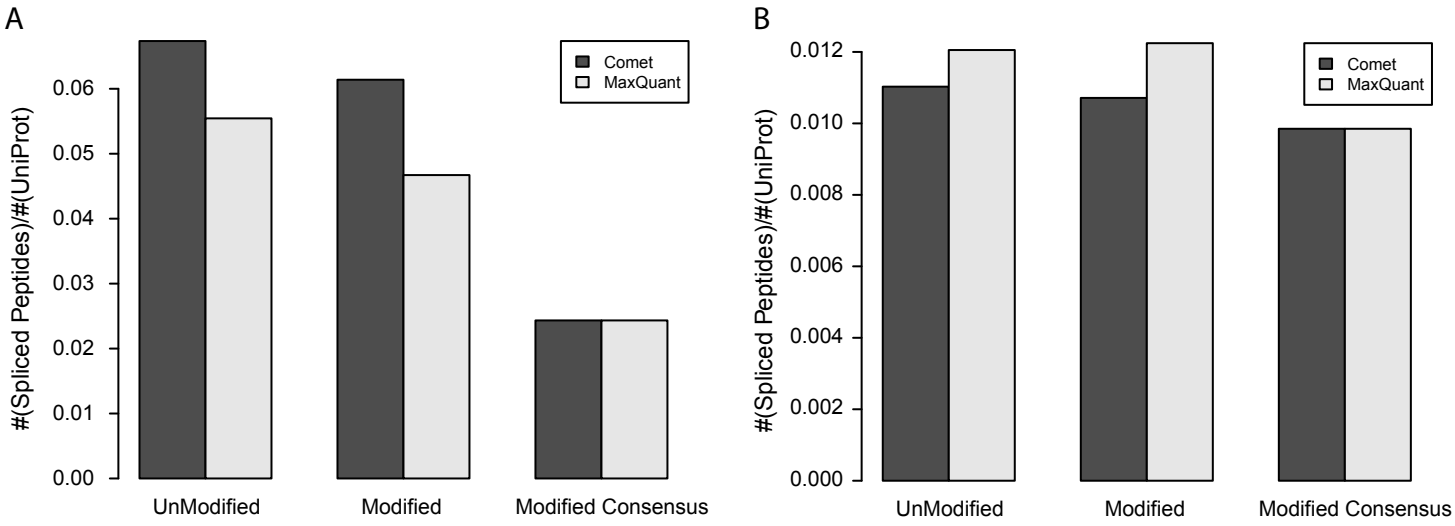
C All remianing 9 mer *LM_spliced* after removal of PSMs assigend by MaxQuant as Uniprot



Supplemental Figure 2: Motif deconvolution analysis with MixMHCp of the 9-mer peptide fraction of A) sequences matching UniProt proteins by MaxQuant that the same MS/MS scans have been originally identified as *LM_spliced* by Liepe et al., B) Mascot assigened *LM_spliced* peptides that MaxQuant assigned the same MS/MS scans as UniProt matches, and C) remaining *LM_spliced* peptides after removal of conflicting PSMs that have been assigned as UniProt in MaxQuant search. The list of peptides has been extracted from Supplemental Data 6.

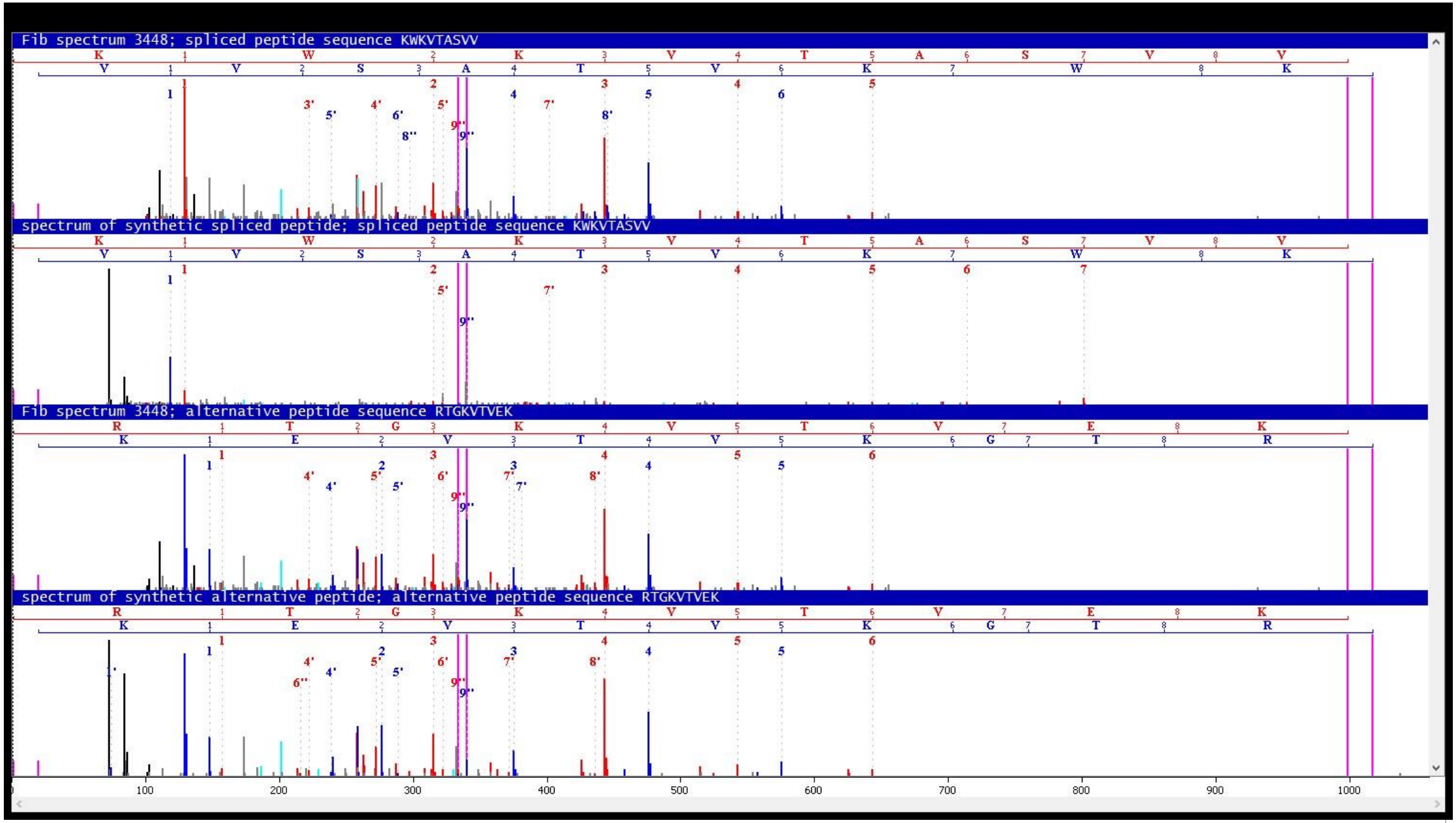


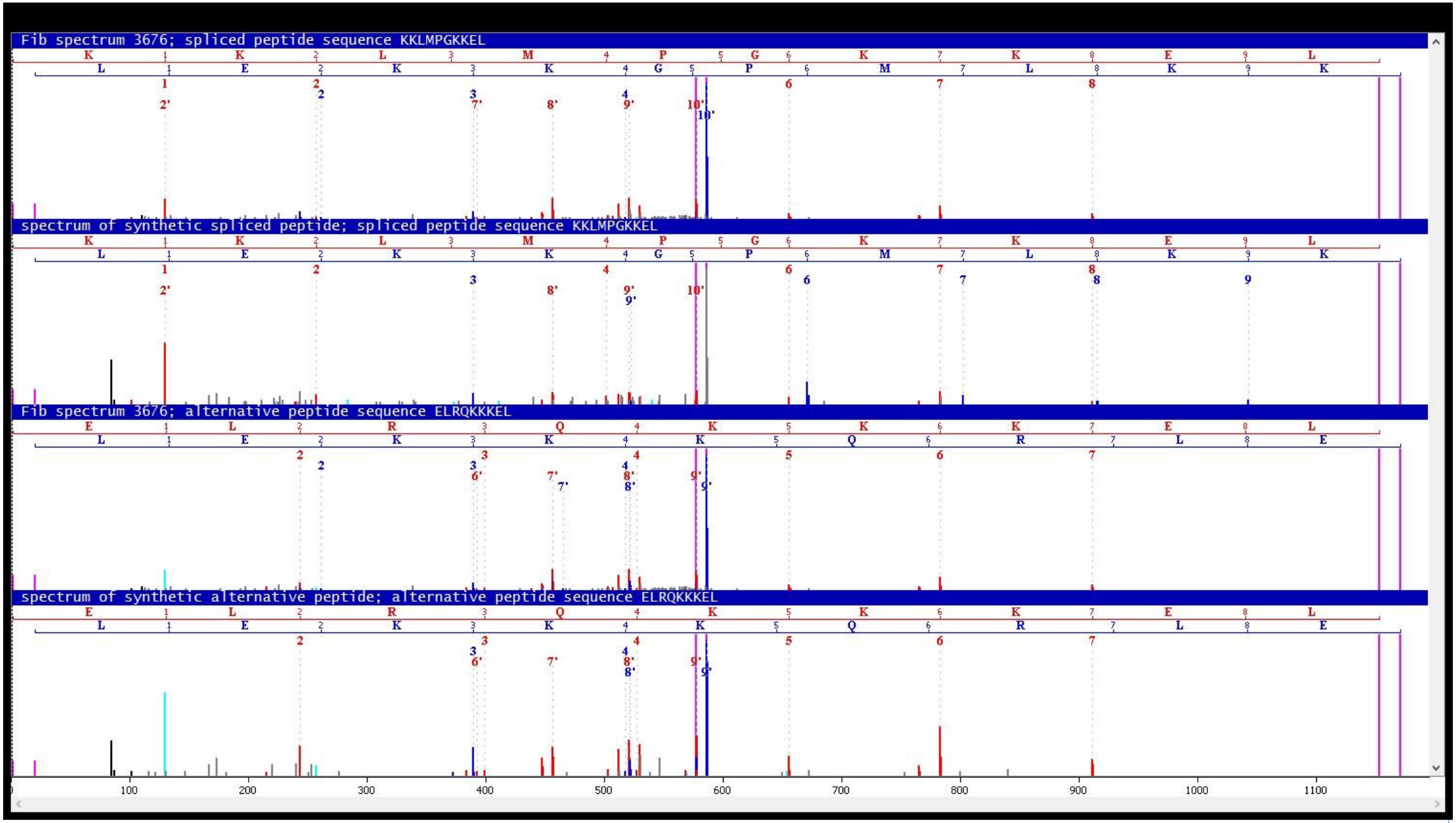
Supplemental Figure 3: Compared to UniProt and *DeNovo_spliced* peptides, *LM_spliced* peptides were characterized with A) lower Andromeda score for the best MS/MS spectrum, B) lower score difference to the second best identified peptide, C) higher absolute precursor mass deviation, D) fewer peaks matching to the predicted fragmentation spectrum, E) lower fraction of total MS/MS peak intensity matched, and F) a larger fraction of singly charged MS/MS spectra matched. G) Rate of disagreement in peptide identification between MaxQuant and Comet for MS/MS scans matched to Uniprot, *LM_spliced* and *DeNovo_spliced* peptides.

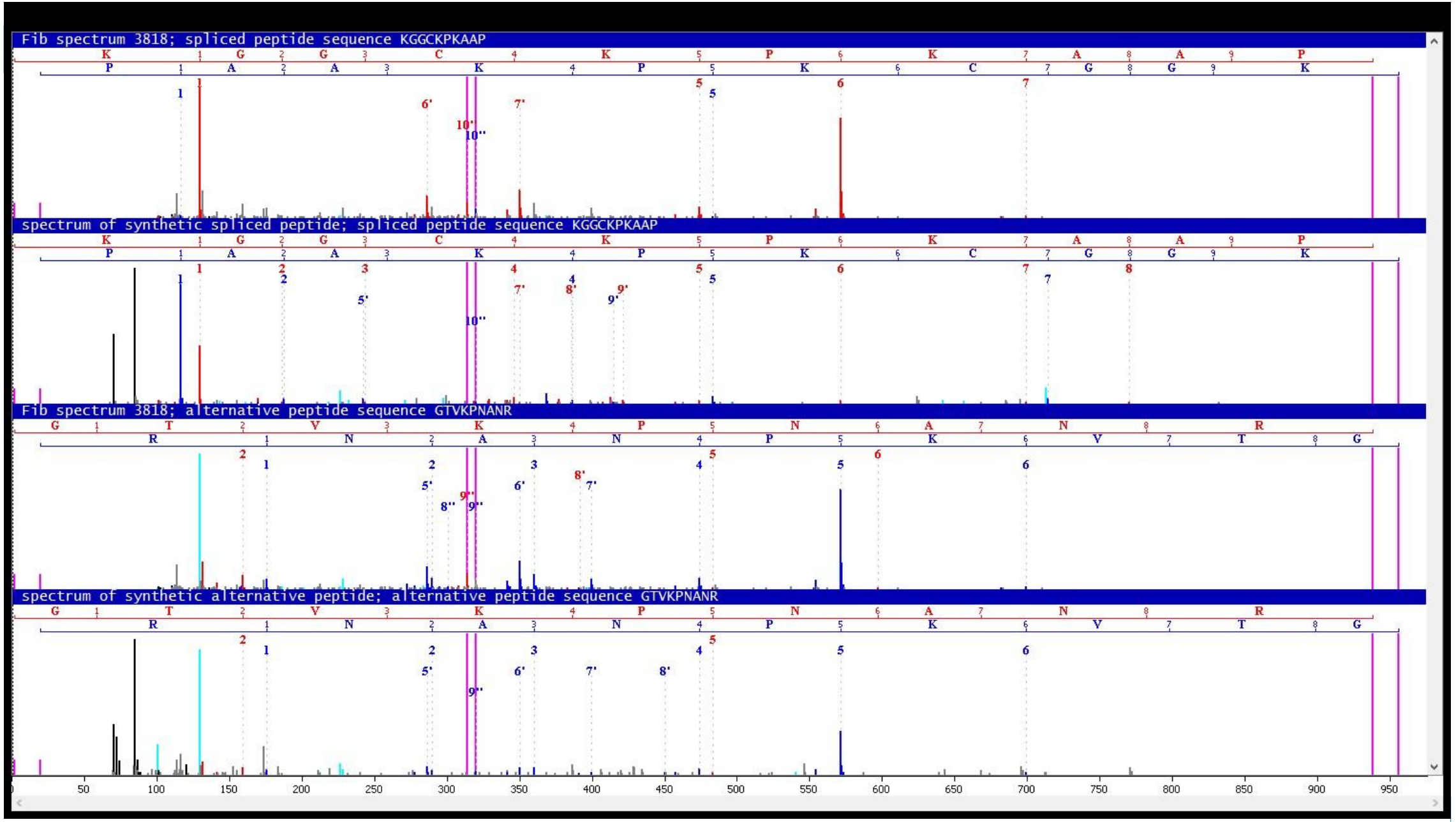


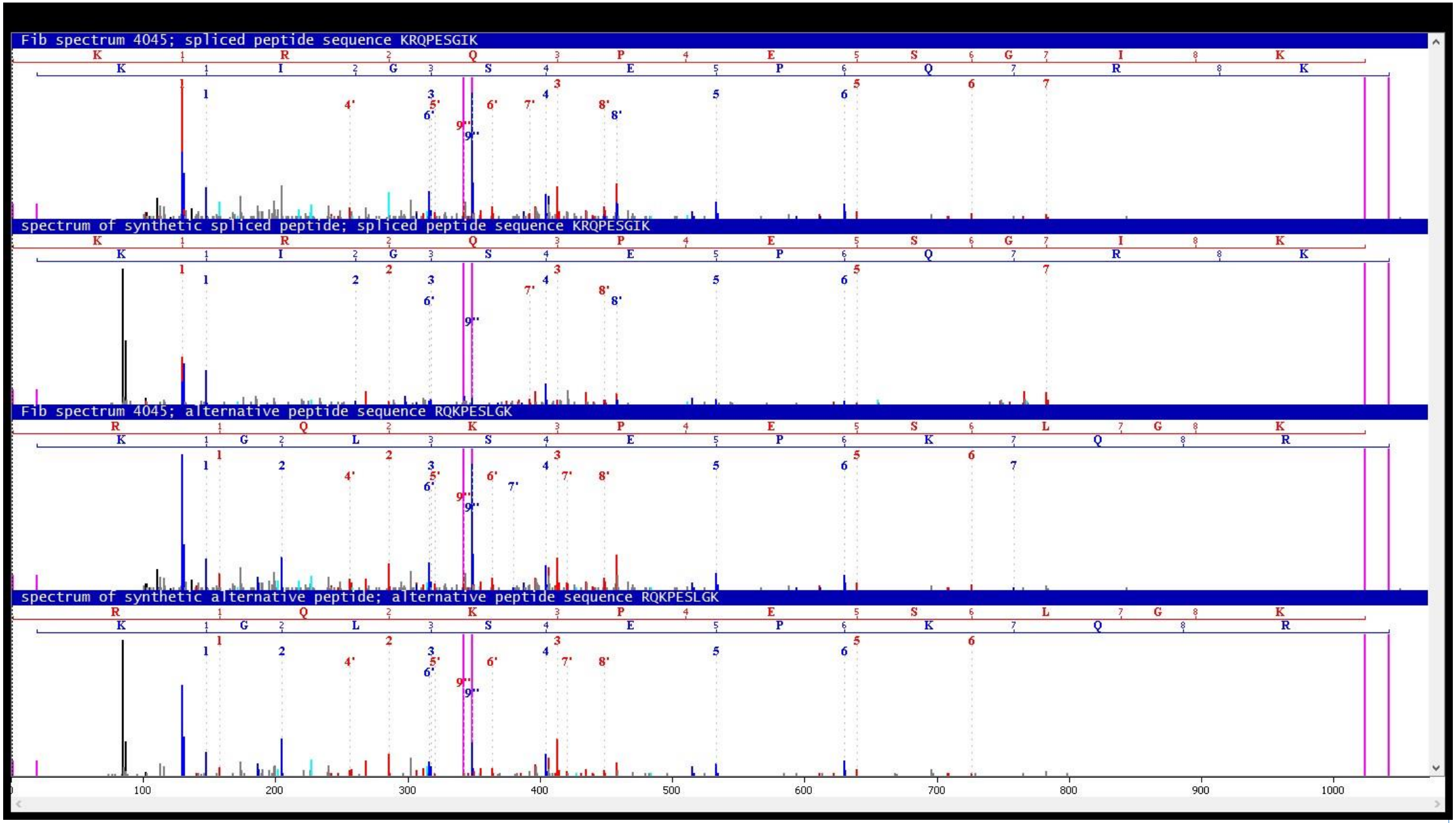
Supplemental Figure 4: Ratio of spliced peptide count and UniProt peptide count for the *LM_spliced* (A) and *DeNovo_spliced* (B) peptides for all Fib samples. Variable modifications are as specified in the text. For 'Modified consensus' the peptide counts are calculated on the subset of PSMs where Comet and MaxQuant agreed, therefore, the values for Comet and MaxQuant are the same.

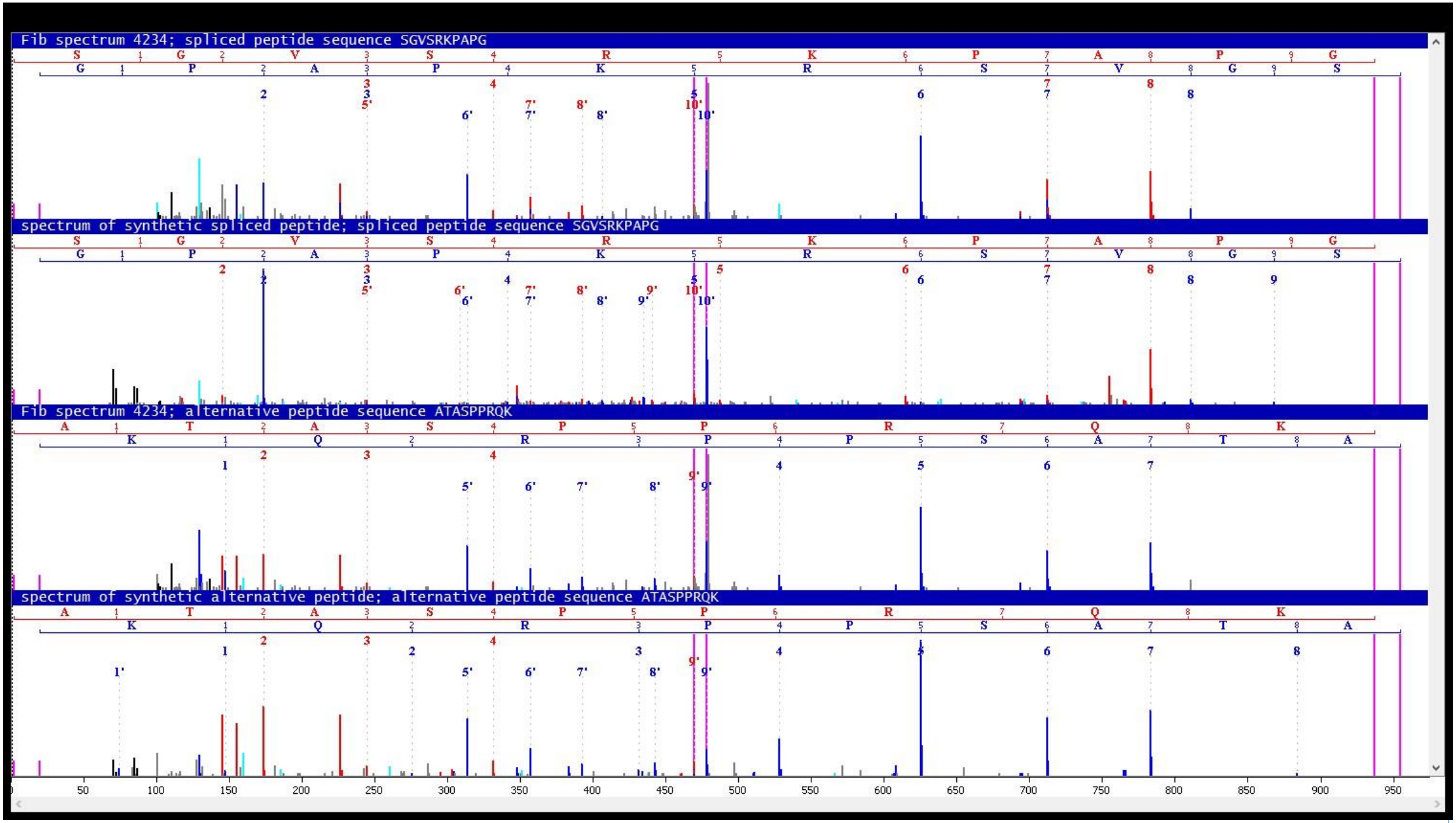
Supplemental Figure 5: Comparison of MS/MS annotations of endogenous and synthetic counterparts of 21 pairs of *LM_spliced* and their UniProt alternatives, and of the three *LM_spliced* peptides. For the 21 pairs we provide: MS/MS annotation of endogenous HLA-Ip as *LM_spliced* peptides, MS/MS annotation of the synthetic counterpart of the *LM_spliced*, MS/MS annotation of the same endogenous HLA-Ip as an alternative UniPro*t* peptide, and MS/MS annotation of the synthetic counterpart of the alternative UniProt peptide. For the three *LM_spliced* peptides we provide MS/MS annotations of endogenous HLA-Ip and of the synthetic counterparts.

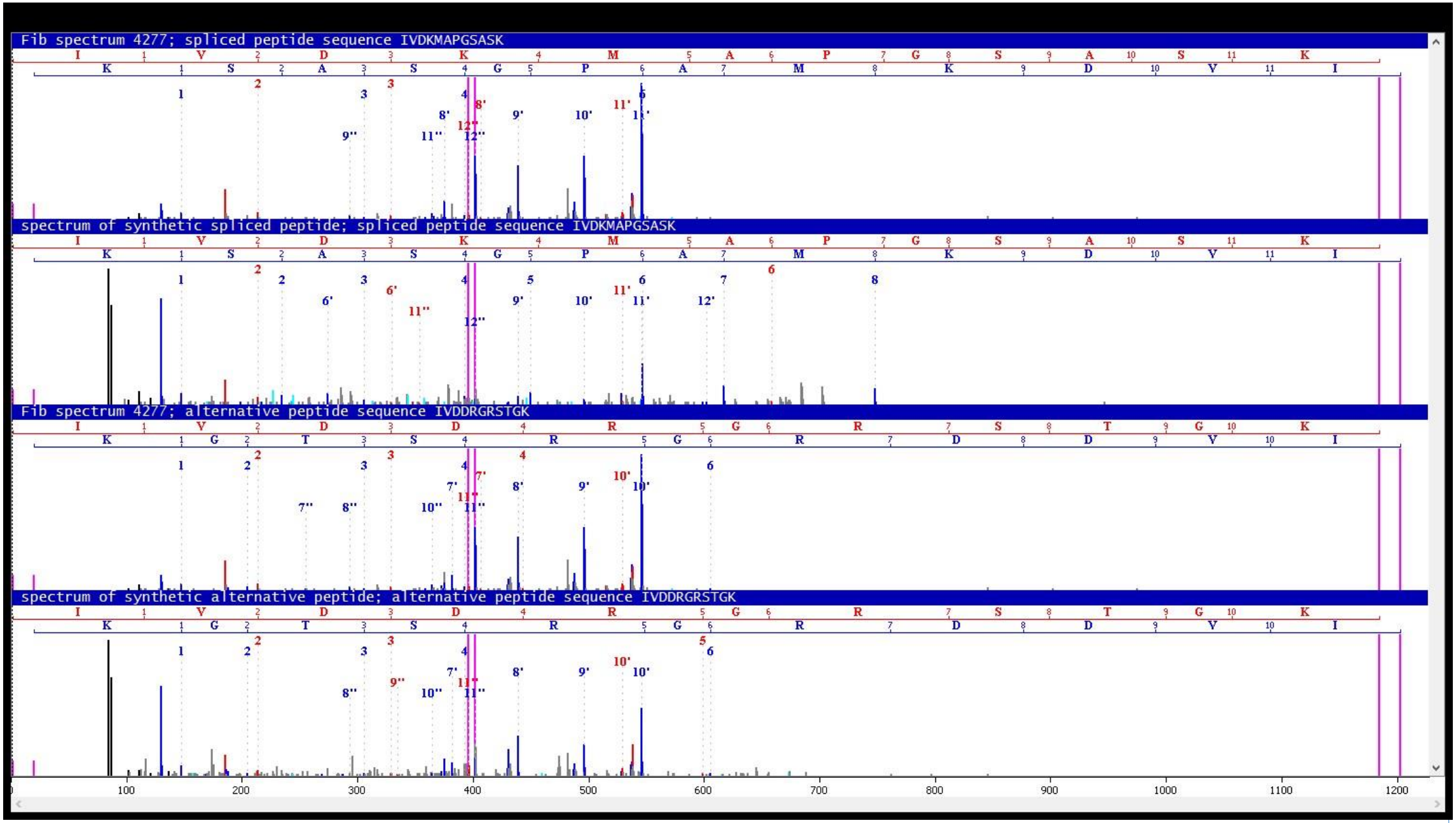

Fib spectrum 3448; spliced peptide sequence KWKVTASVV
spectrum of synthetic spliced peptide; spliced peptide sequence KWKVTASVV
Fib spectrum 3448; alternative peptide sequence RTGKVTVEK
spectrum of synthetic alternative peptide; alternative peptide sequence RTGKVTVEK
100
200
300
400
500
600
700
800
900
1000

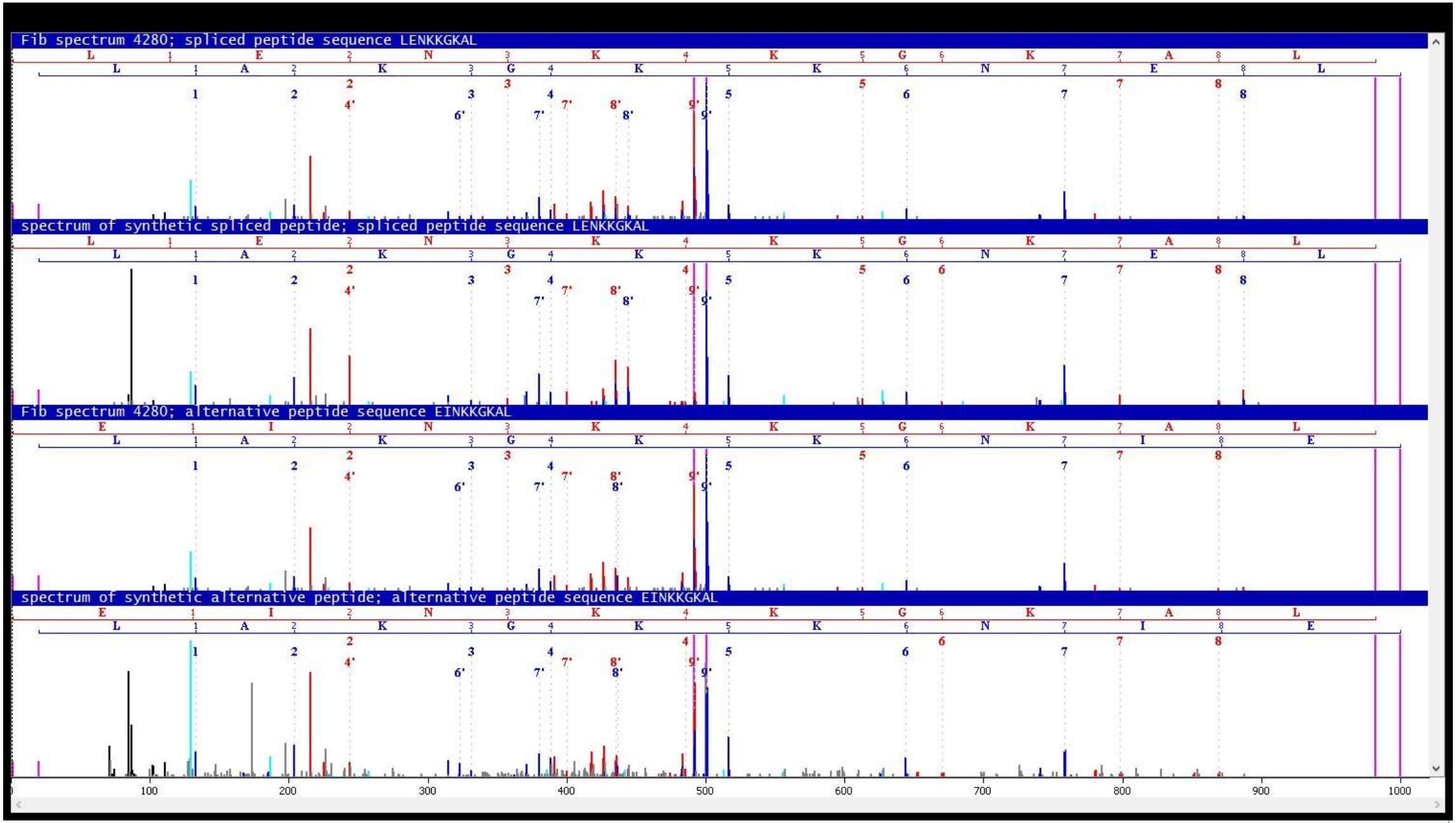

Fib spectrum 3676; spliced peptide sequence KKLMPGKKEL
spectrum of synthetic spliced peptide; spliced peptide sequence KKLMPGKKEL
Fib spectrum 3676; alternative peptide sequence ELRQKKKEL
spectrum of synthetic alternative peptide; alternative peptide sequence ELRQKKKEL
100
200
300
400
500
600
700
800
900
1000
1100

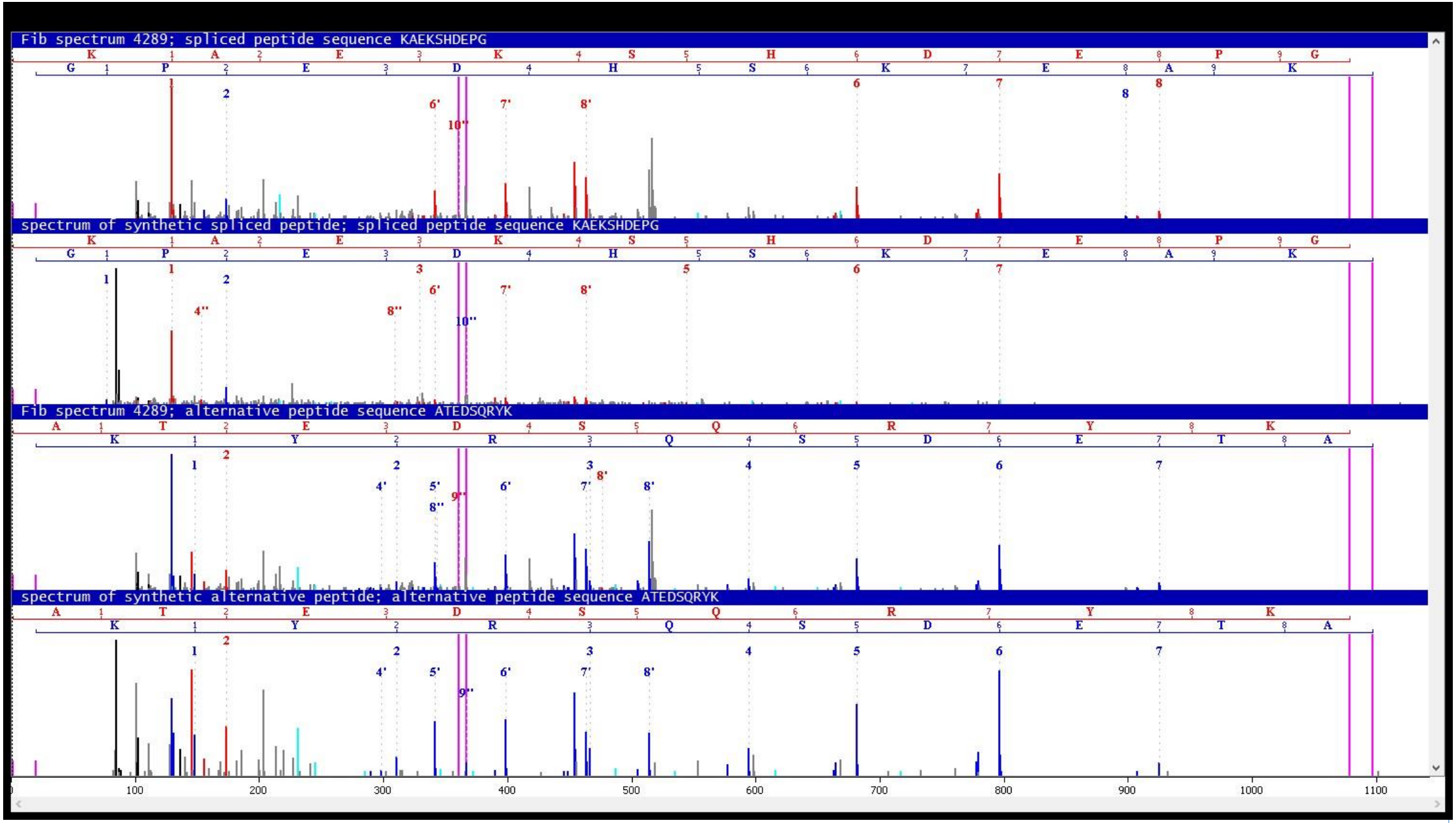

Fib spectrum 3818; spliced peptide sequence KGGCKPKAAP
spectrum of synthetic spliced peptide; spliced peptide sequence KGGCKPKAAP
Fib spectrum 3818; alternative peptide sequence GTVKPNANR
spectrum of synthetic alternative peptide; alternative peptide sequence GTVKPNANR
0
50
100
150
200
250
300
350
400
450
500
550
600
650
700
750
800
850
900
950

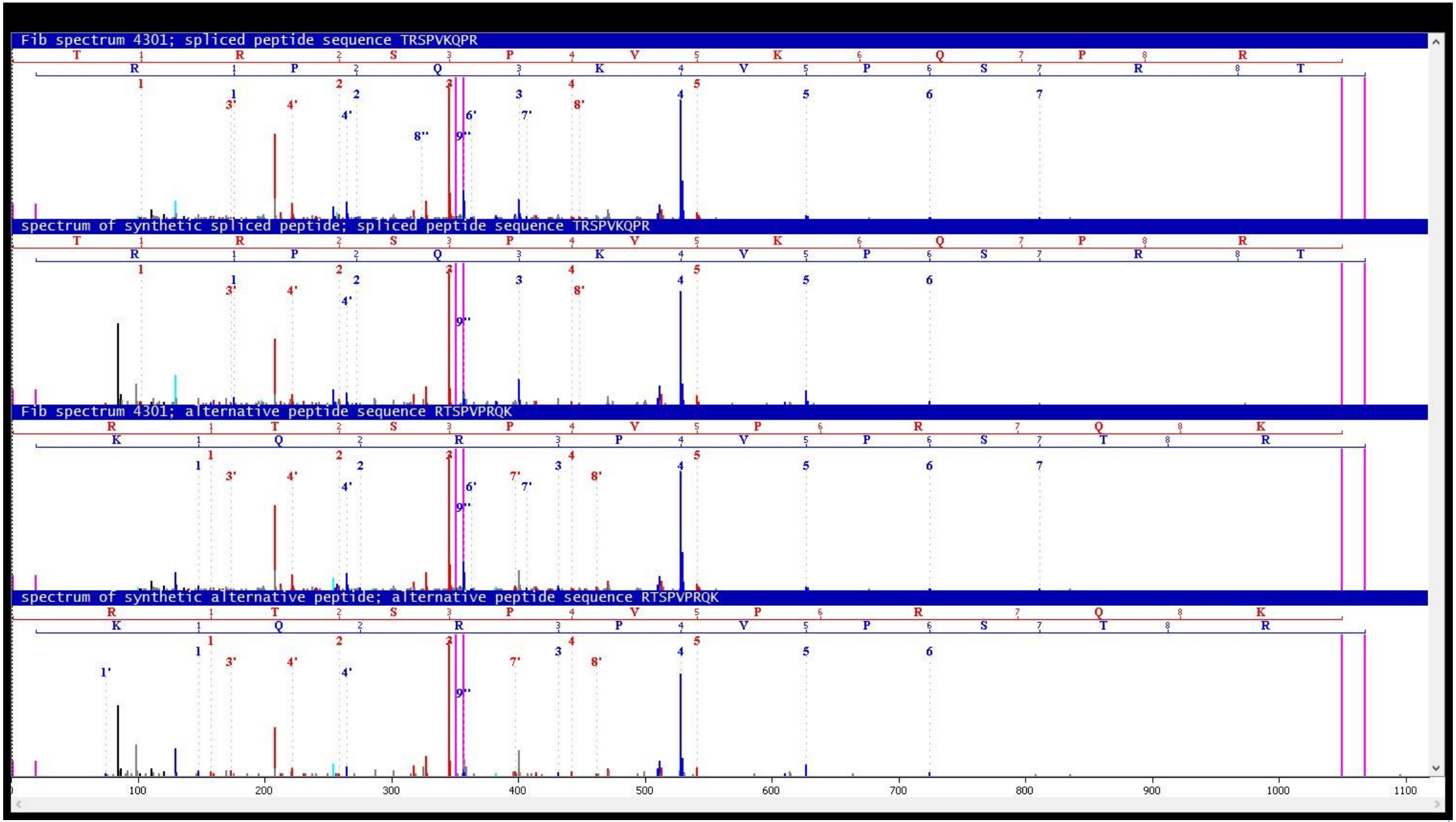

Fib spectrum 4045; spliced peptide sequence KRQPESGIK
spectrum of synthetic spliced peptide; spliced peptide sequence KRQPESGIK
Fib spectrum 4045; alternative peptide sequence RQKPESLGK
spectrum of synthetic alternative peptide; alternative peptide sequence RQKPESLGK
100
200
300
400
500
600
700
800
900
1000

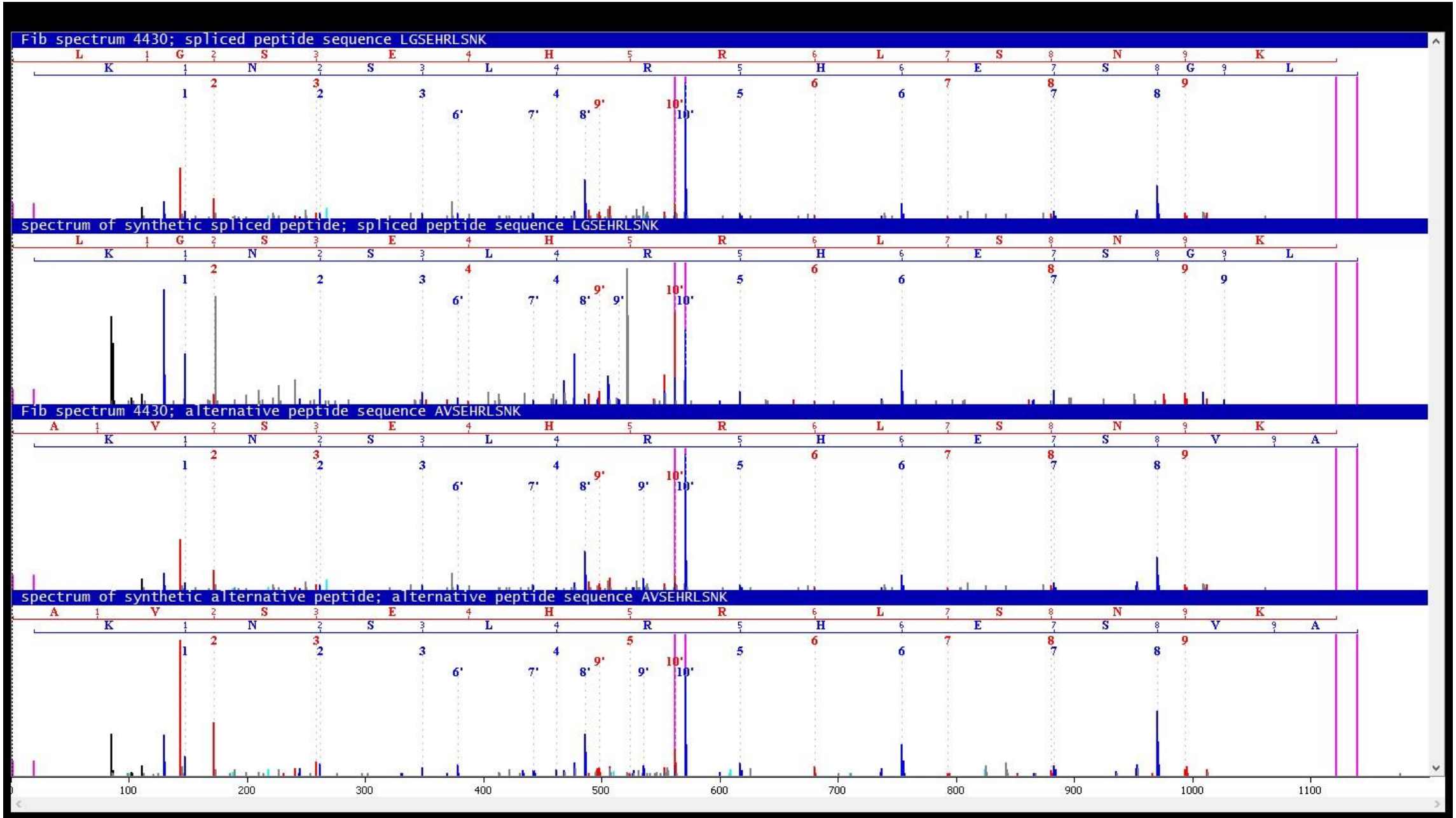

Fib spectrum 4234; spliced peptide sequence SGVSRKPAPG
spectrum of synthetic spliced peptide; spliced peptide sequence SGVSRKPAPG
Fib spectrum 4234; alternative peptide sequence ATASPPRQK
spectrum of synthetic alternative peptide; alternative peptide sequence ATASPPRQK
0
50
100
150
200
250
300
350
400
450
500
550
600
650
700
750
800
850
900
950

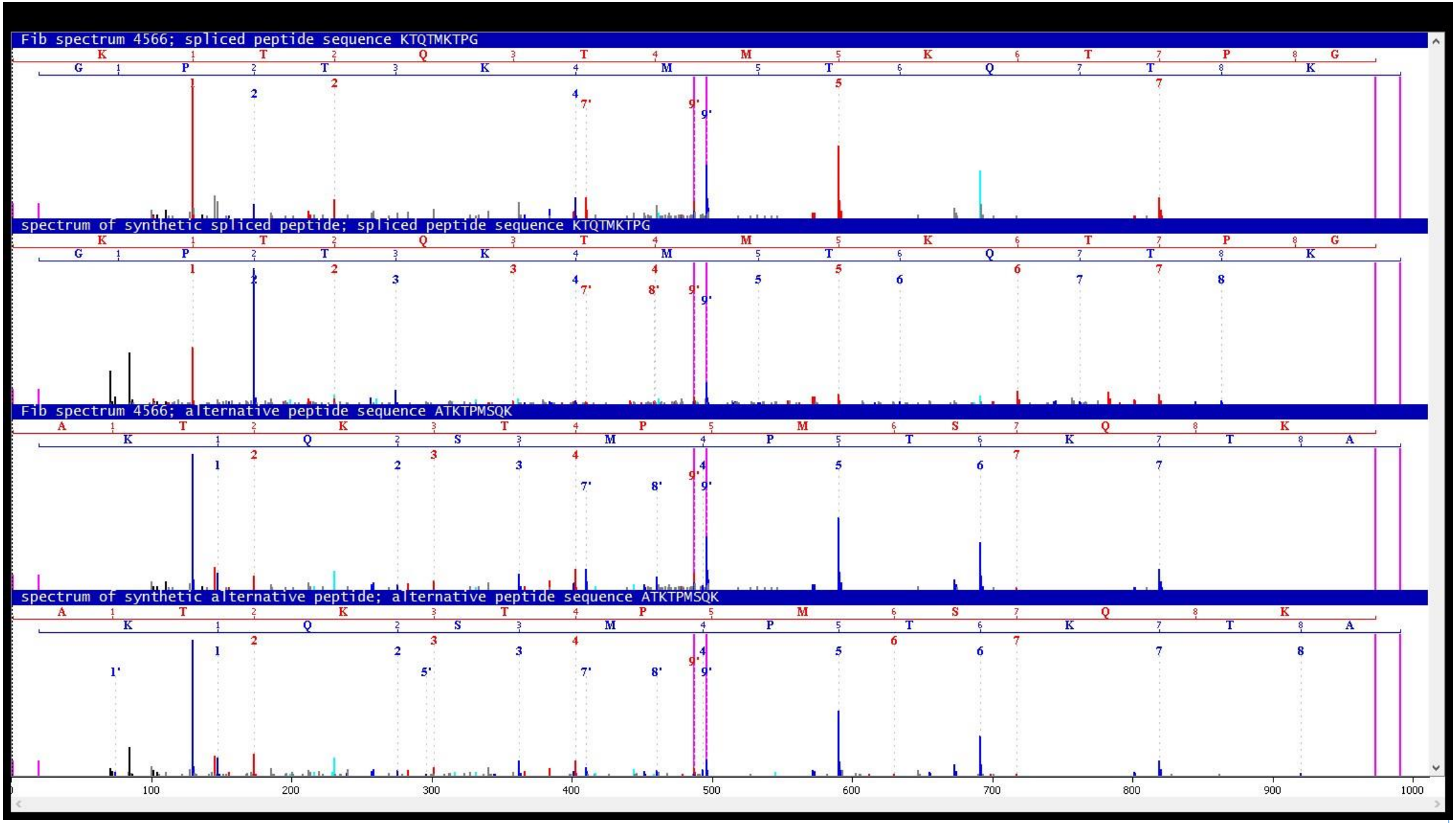

Fib spectrum 4277; spliced peptide sequence IVDKMAPGSASK
spectrum of synthetic spliced peptide; spliced peptide sequence IVDKMAPGSASK
Fib spectrum 4277; alternative peptide sequence IVDDRGRSTGK
spectrum of synthetic alternative peptide; alternative peptide sequence IVDDRGRSTGK
100
200
300
400
500
600
700
800
900
1000
1100
1200

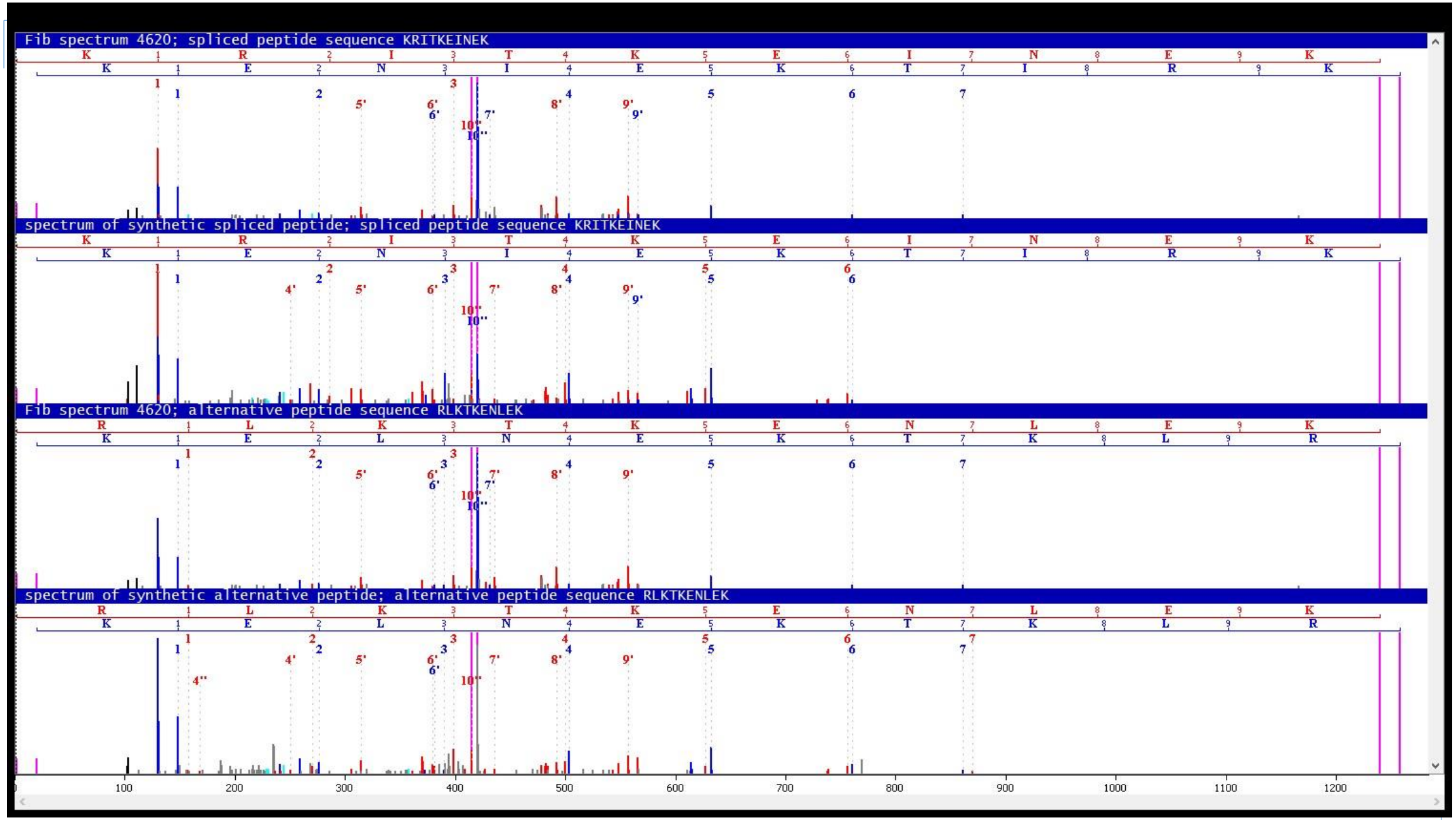

Fib spectrum 4280; spliced peptide sequence LENKKGKAL
spectrum of synthetic spliced peptide; spliced peptide sequence LENKKGKAL
Fib spectrum 4280; alternative peptide sequence EINKKGKAL
spectrum of synthetic alternative peptide; alternative peptide sequence EINKKGKAL
100
200
300
400
500
600
700
800
900
1000

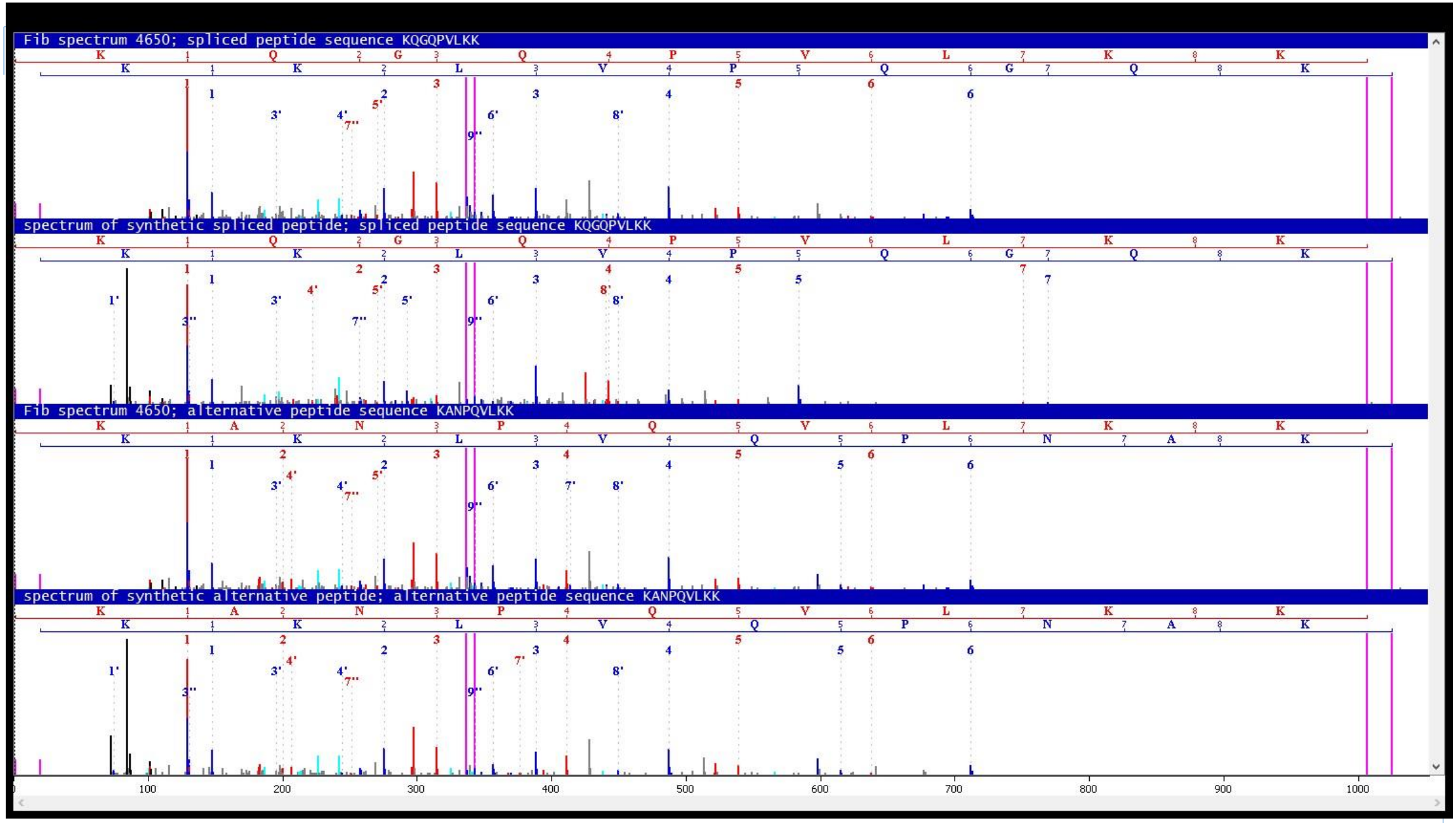

Fib spectrum 4289; spliced peptide sequence KAEKSHDEPG
spectrum of synthetic spliced peptide; spliced peptide sequence KAEKSHDEPG
Fib spectrum 4289; alternative peptide sequence ATEDSQRYK
spectrum of synthetic alternative peptide; alternative peptide sequence ATEDSQRYK
100
200
300
400
500
600
700
800
900
1000
1100

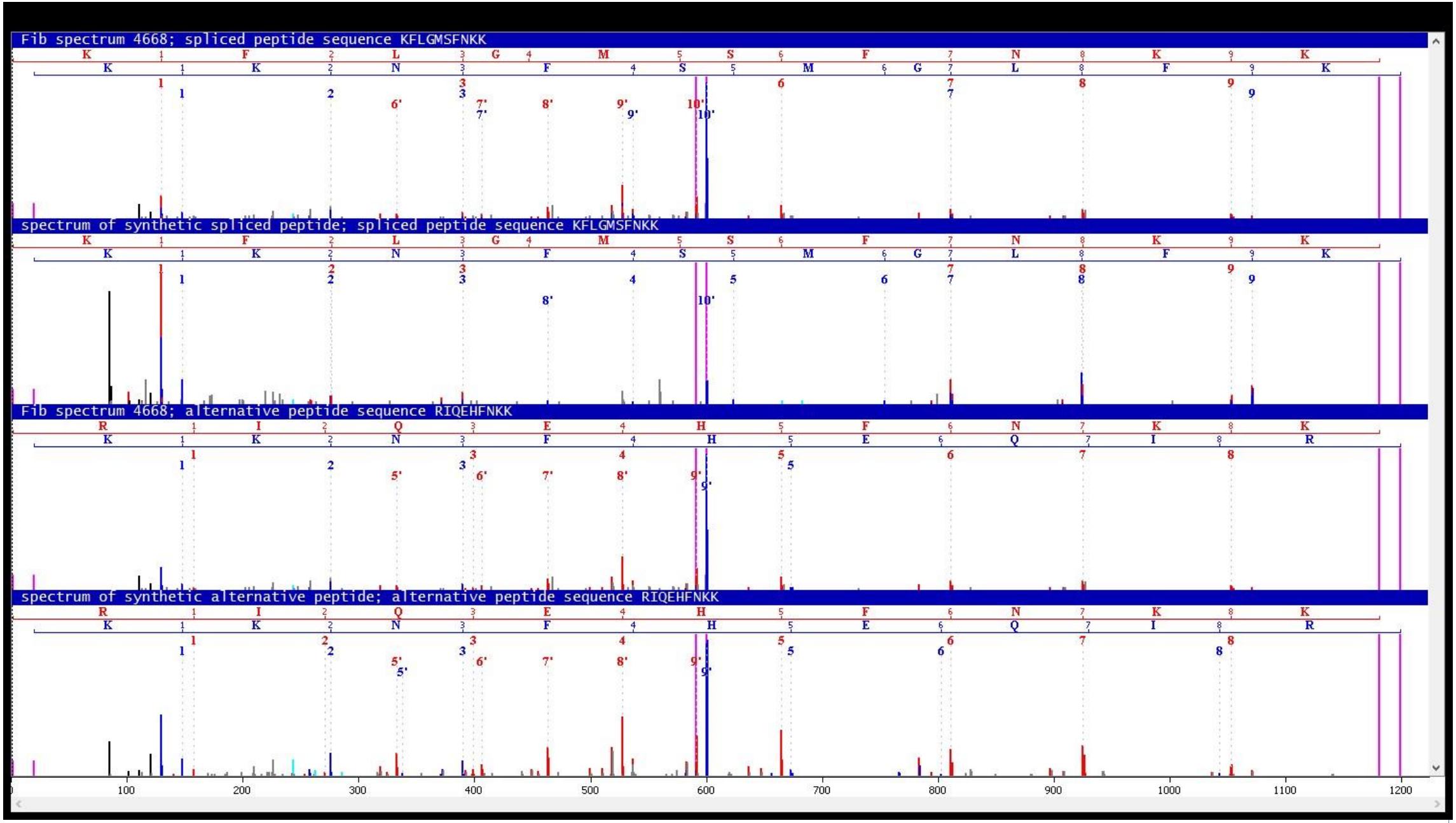

Fib spectrum 4301; spliced peptide sequence TRSPVKQPR
spectrum of synthetic spliced peptide; spliced peptide sequence TRSPVKQPR
Fib spectrum 4301; alternative peptide sequence RTSPVPRQK
spectrum of synthetic alternative peptide; alternative peptide sequence RTSPVPRQK
100
200
300
400
500
600
700
800
900
1000
1100

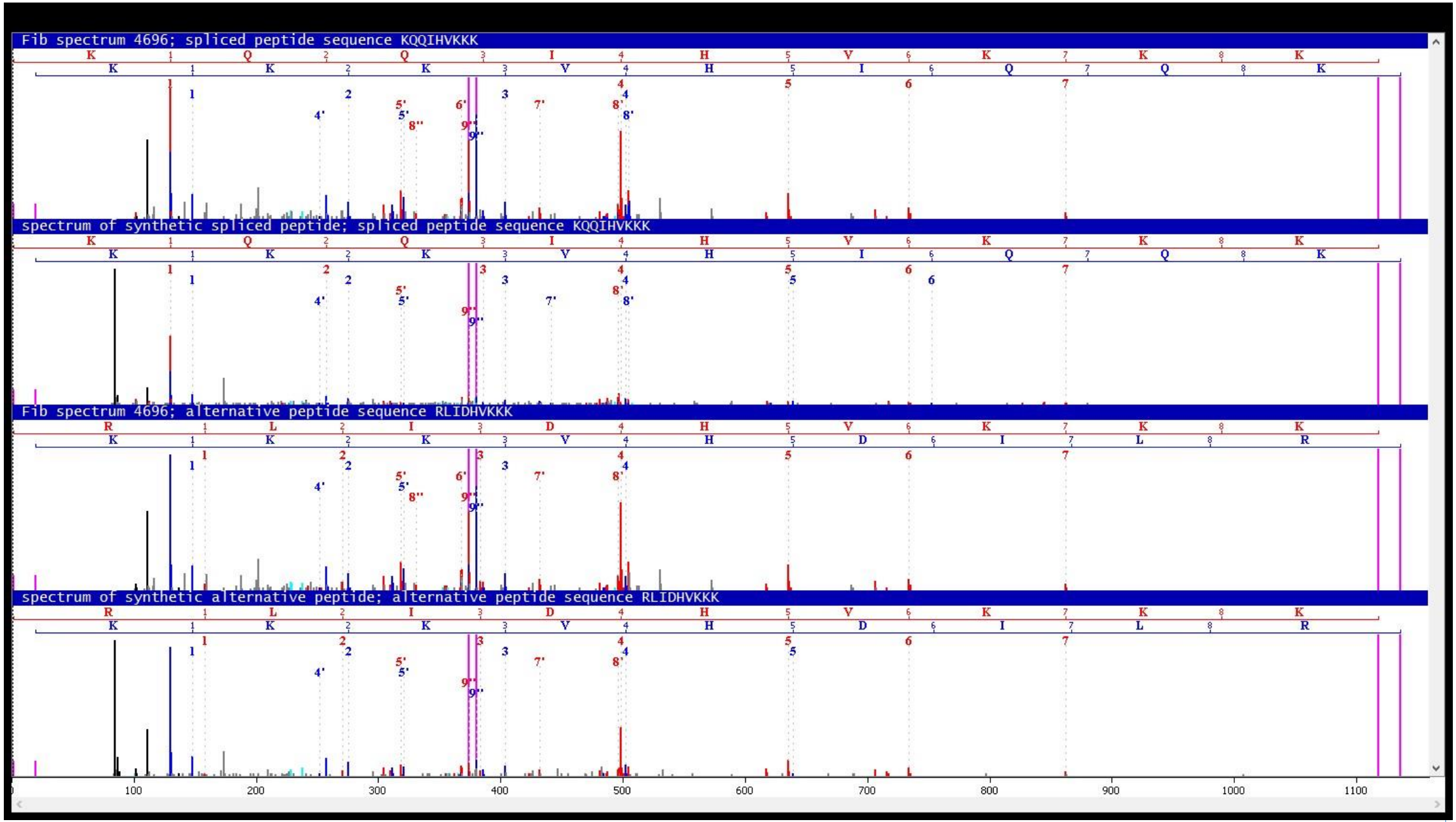

Fib spectrum 4430; spliced peptide sequence LGSEHRLSNK
spectrum of synthetic spliced peptide; spliced peptide sequence LGSEHRLSNK
Fib spectrum 4430; alternative peptide sequence AVSEHRLSNK
spectrum of synthetic alternative peptide; alternative peptide sequence AVSEHRLSNK
100
200
300
400
500
600
700
800
900
1000
1100

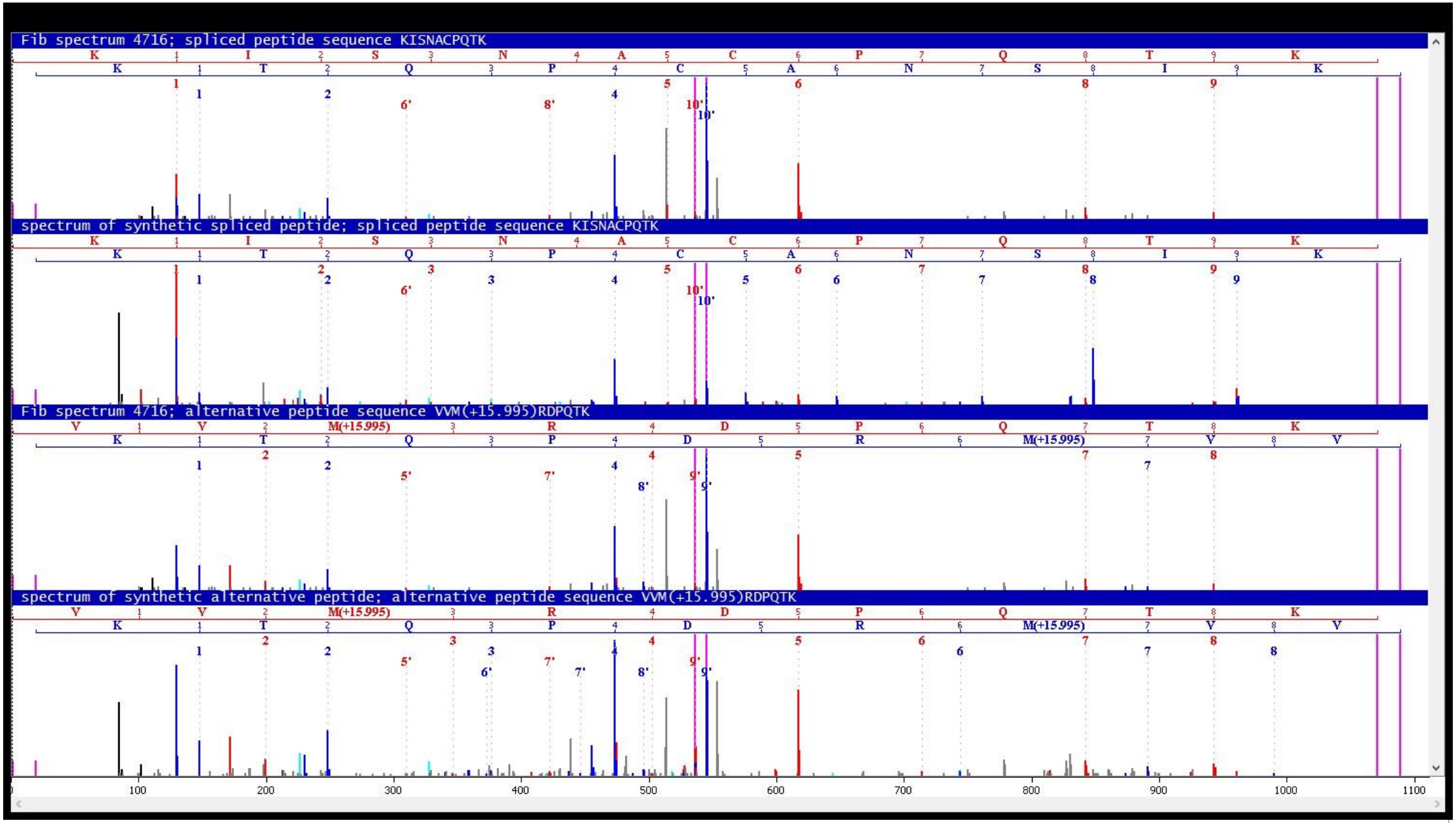

Fib spectrum 4566; spliced peptide sequence KTQTMKTPG
spectrum of synthetic spliced peptide; spliced peptide sequence KTQTMKTPG
Fib spectrum 4566; alternative peptide sequence ATKTPMSQK
spectrum of synthetic alternative peptide; alternative peptide sequence ATKTPMSQK
100
200
300
400
500
600
700
800
900
1000

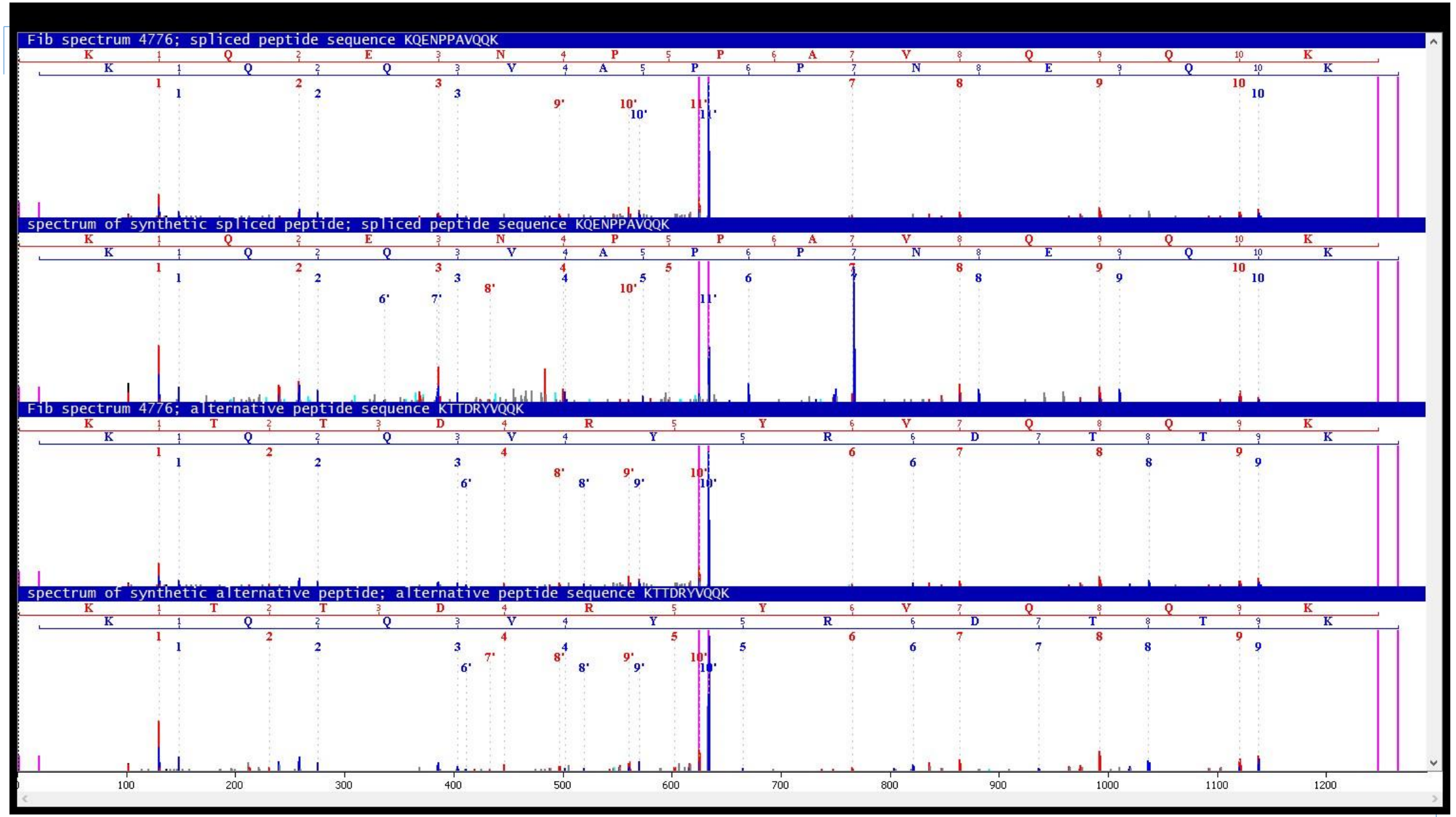

Fib spectrum 4620; spliced peptide sequence KRITKEINEK
spectrum of synthetic spliced peptide; spliced peptide sequence KRITKEINEK
Fib spectrum 4620; alternative peptide sequence RLKTKENLEK
spectrum of synthetic alternative peptide; alternative peptide sequence RLKTKENLEK
100
200
300
400
500
600
700
800
900
1000
1100
1200

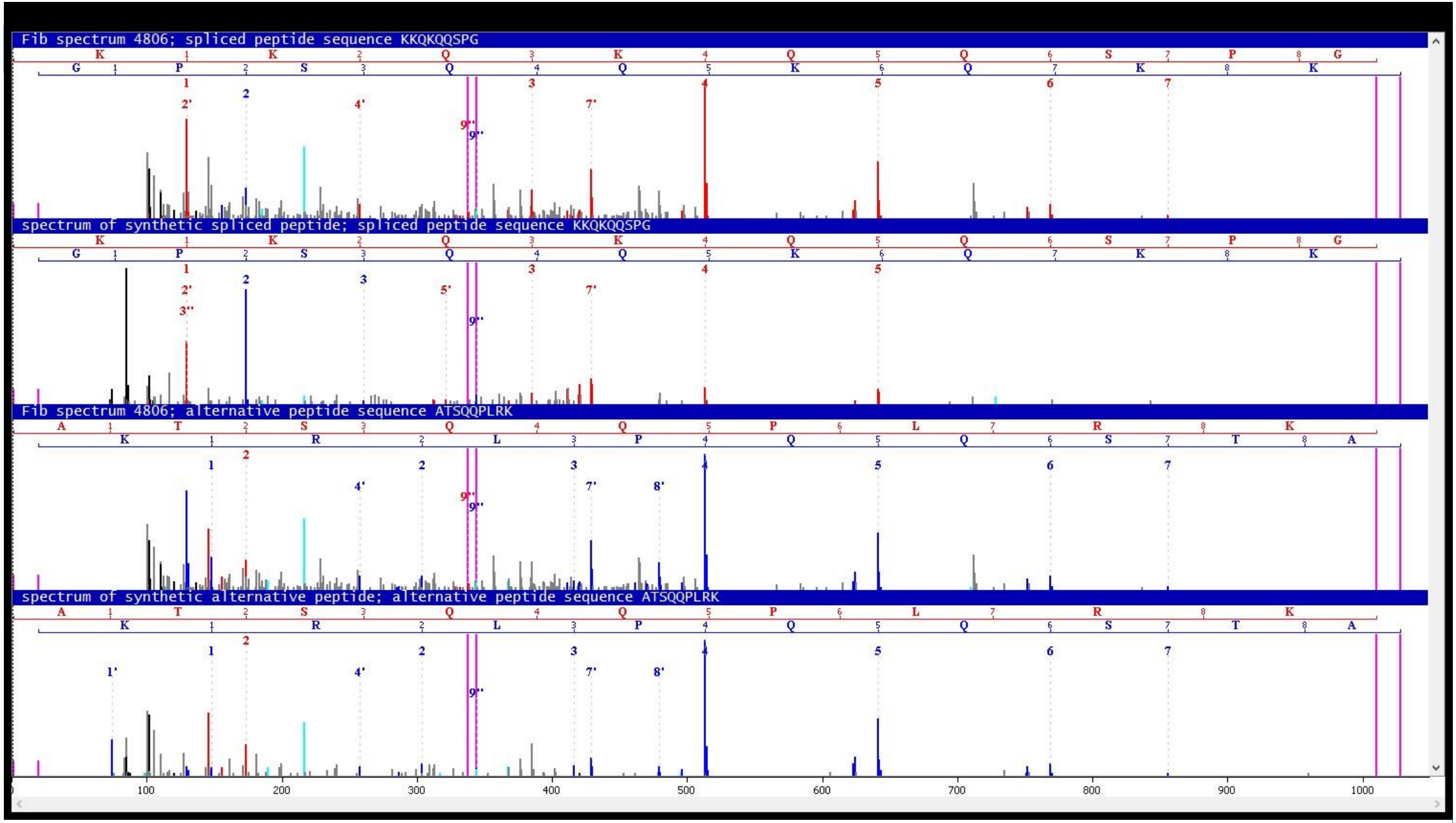

Fib spectrum 4650; spliced peptide sequence KQGQPVLKK
spectrum of synthetic spliced peptide; spliced peptide sequence KQGQPVLKK
Fib spectrum 4650; alternative peptide sequence KANPQVLKK
spectrum of synthetic alternative peptide; alternative peptide sequence KANPQVLKK
100
200
300
400
500
600
700
800
900
1000

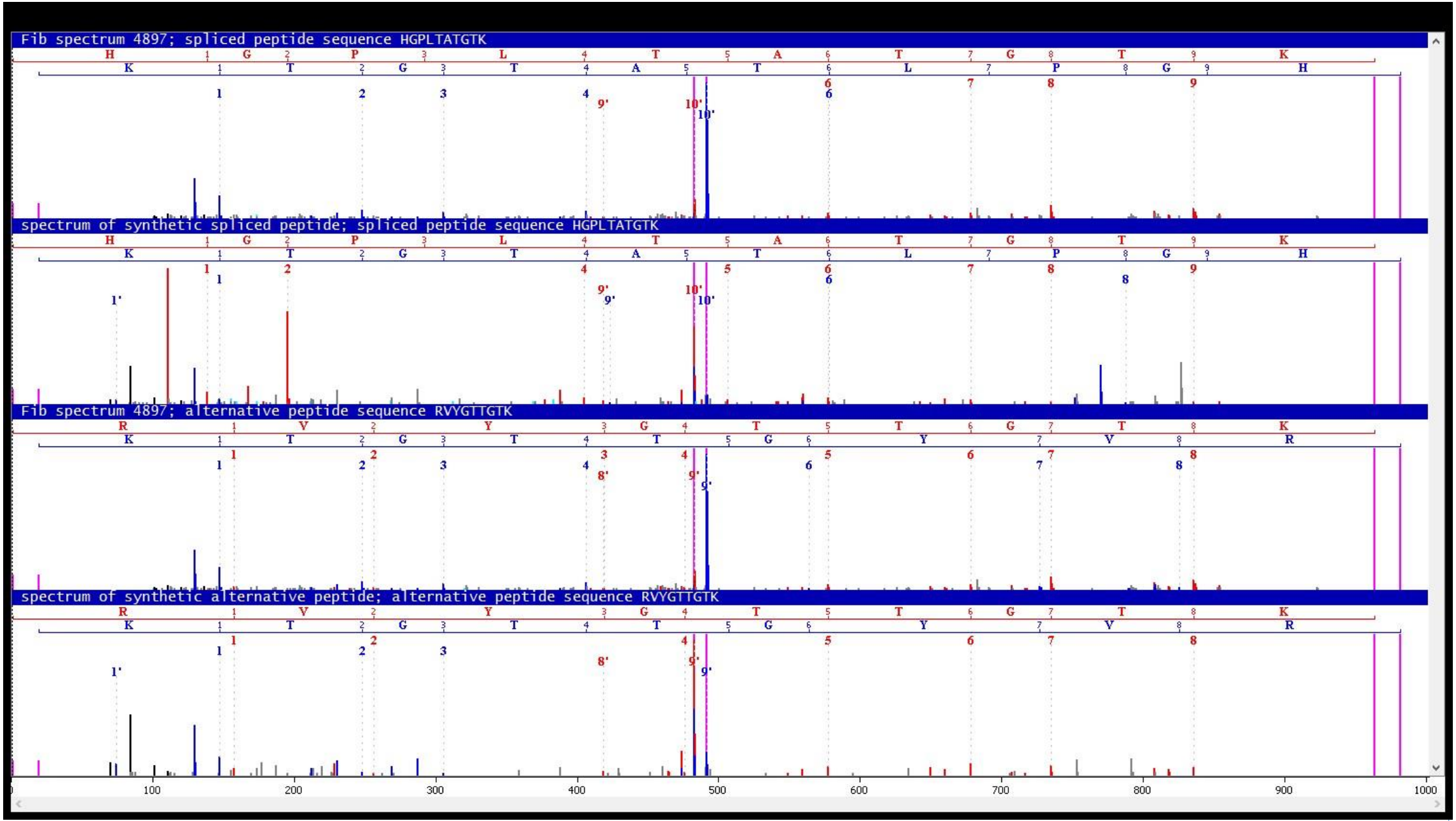

Fib spectrum 4668; spliced peptide sequence KFLGMSFNKK
spectrum of synthetic spliced peptide; spliced peptide sequence KFLGMSFNKK
Fib spectrum 4668; alternative peptide sequence RIQEHFNKK
spectrum of synthetic alternative peptide; alternative peptide sequence RIQEHFNKK
100
200
300
400
500
600
700
800
900
1000
1100
1200

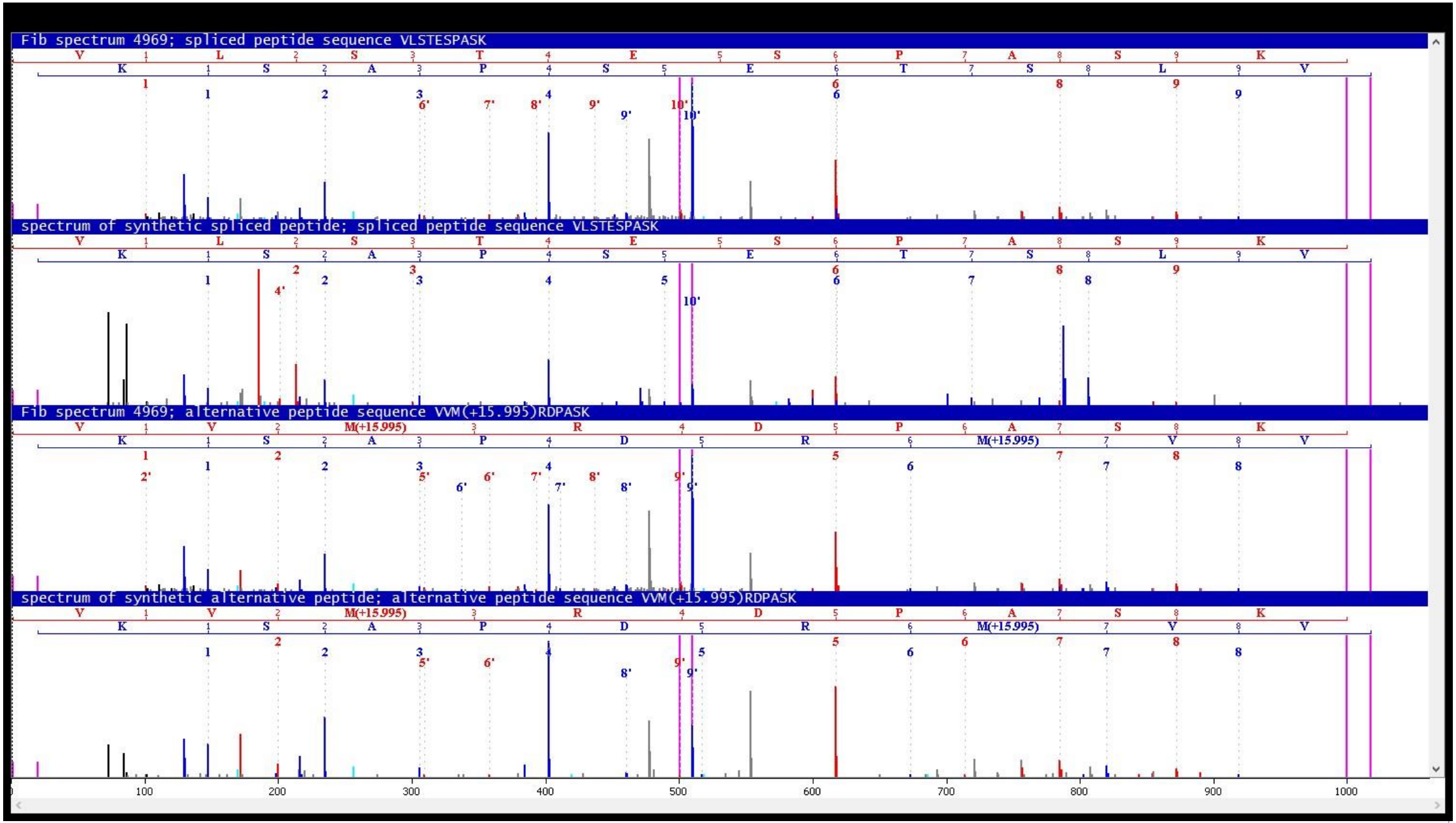

Fib spectrum 4696; spliced peptide sequence KQQIHVKKK
spectrum of synthetic spliced peptide; spliced peptide sequence KQQIHVKKK
Fib spectrum 4696; alternative peptide sequence RLIDHVKKK
spectrum of synthetic alternative peptide; alternative peptide sequence RLIDHVKKK
100
200
300
400
500
600
700
800
900
1000
1100

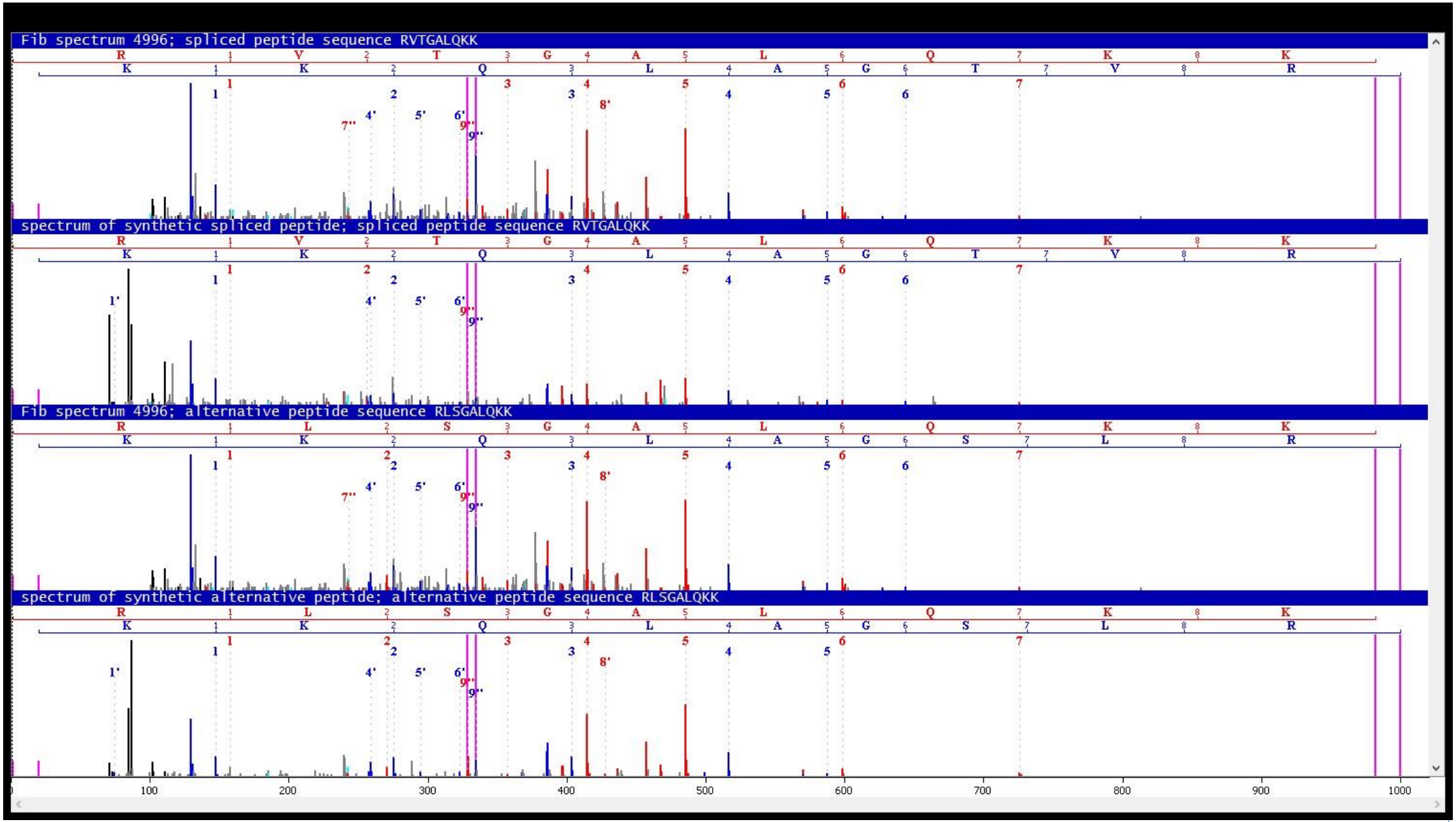

Fib spectrum 4716; spliced peptide sequence KISNACPQTK
spectrum of synthetic spliced peptide; spliced peptide sequence KISNACPQTK
Fib spectrum 4716; alternative peptide sequence VVM(+15.995)RDPQTK
spectrum of synthetic alternative peptide; alternative peptide sequence VVM(+15.995)RDPQTK
100
200
300
400
500
600
700
800
900
1000
1100

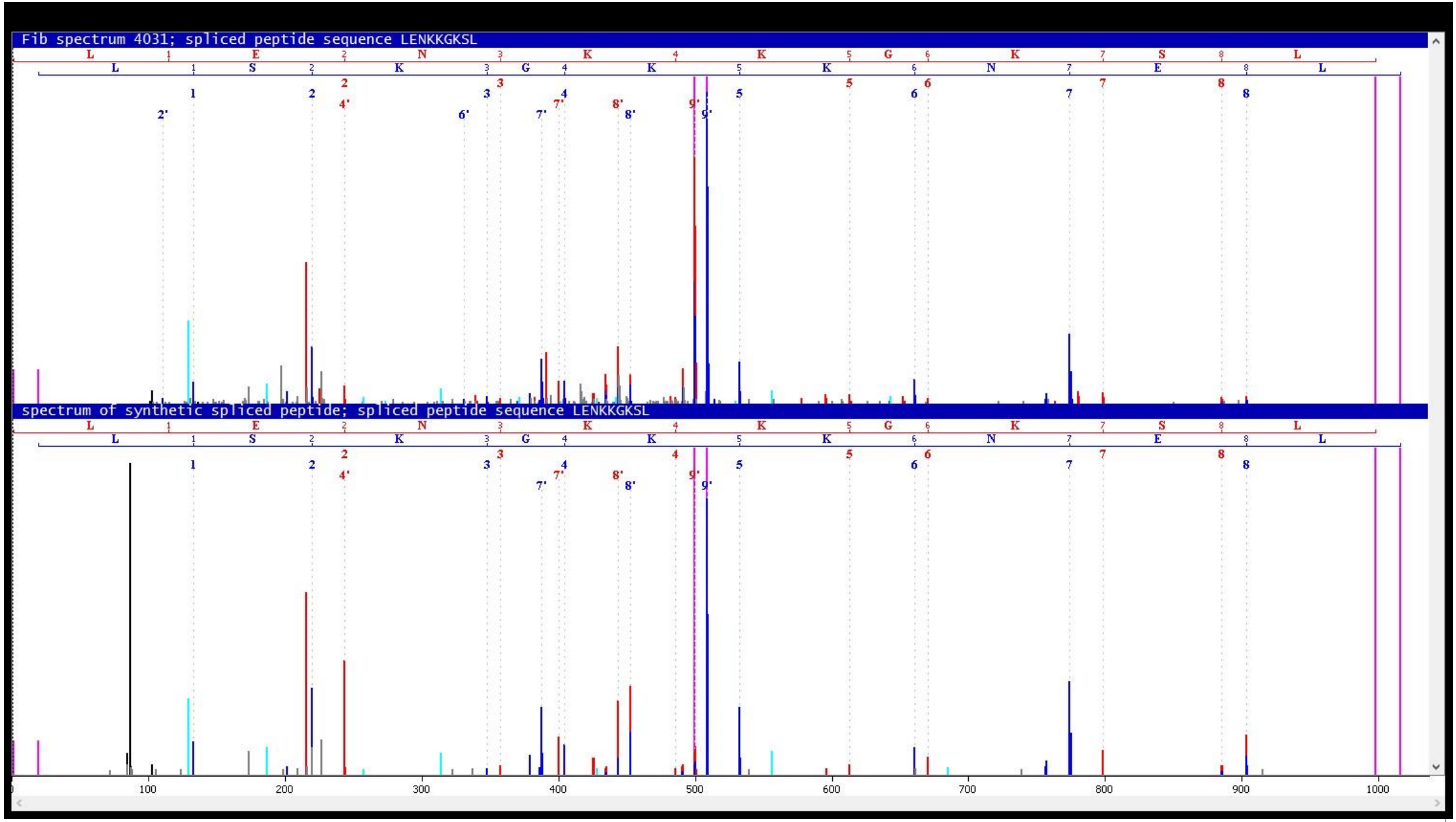

Fib spectrum 4776; spliced peptide sequence KQENPPAVQQK
spectrum of synthetic spliced peptide; spliced peptide sequence KQENPPAVQQK
Fib spectrum 4776; alternative peptide sequence KTTDRYVQQK
spectrum of synthetic alternative peptide; alternative peptide sequence KTTDRYVQQK
100
200
300
400
500
600
700
800
900
1000
1100
1200

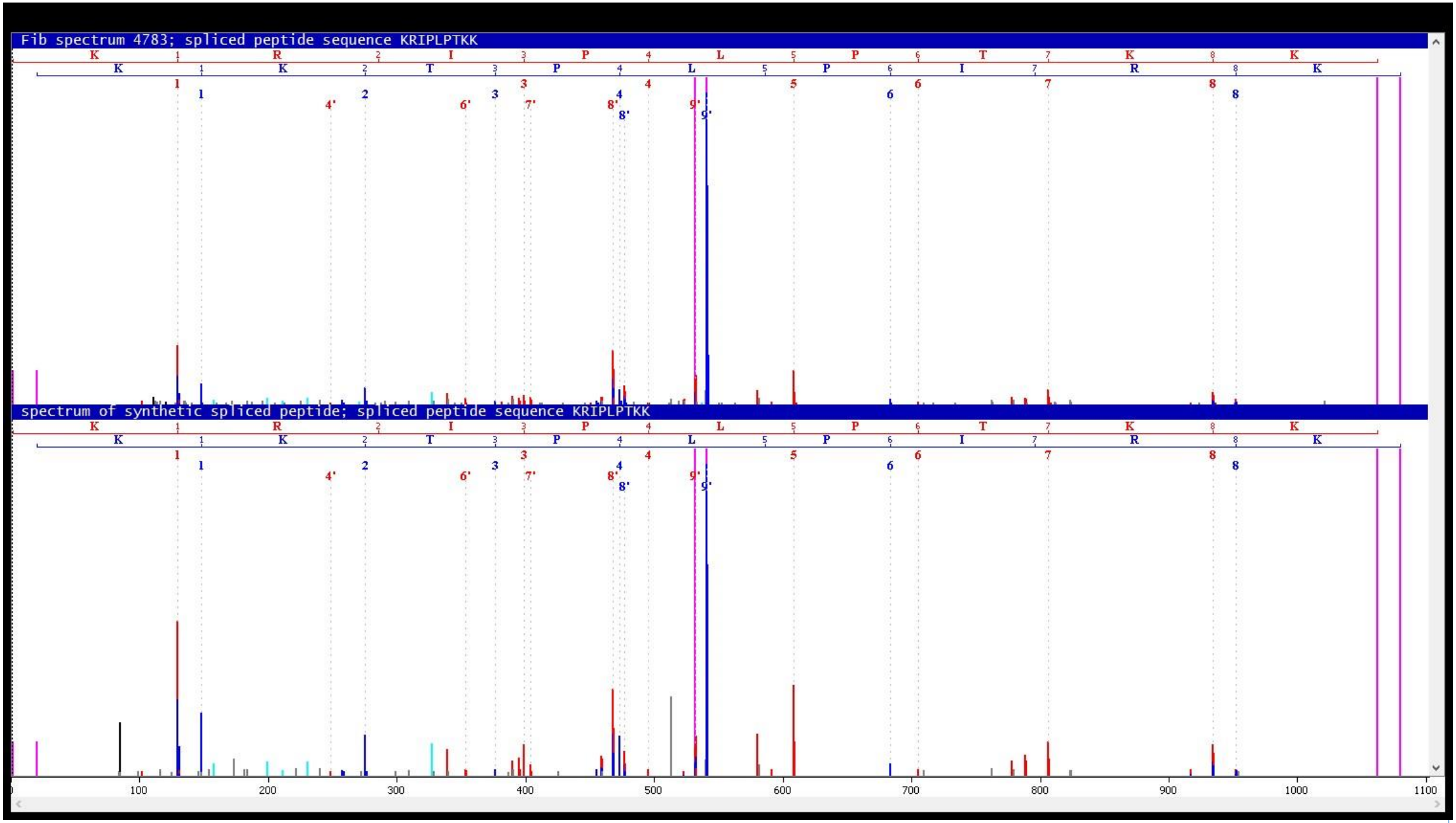

Fib spectrum 4806; spliced peptide sequence KKQKQQSPG
spectrum of synthetic spliced peptide; spliced peptide sequence KKQKQQSPG
Fib spectrum 4806; alternative peptide sequence ATSQQPLRK
spectrum of synthetic alternative peptide; alternative peptide sequence ATSQQPLRK
100
200
300
400
500
600
700
800
900
1000

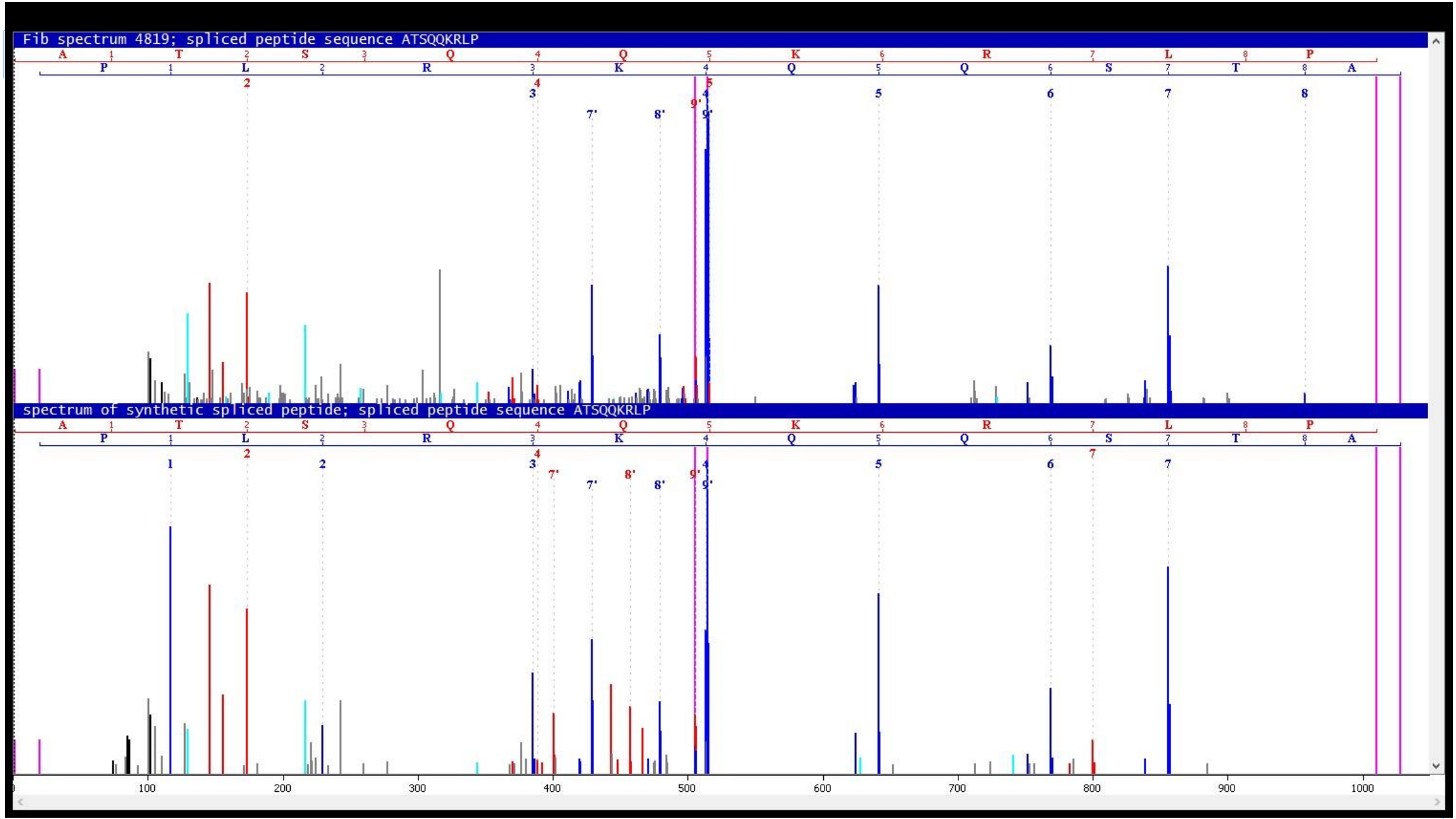

Fib spectrum 4897; spliced peptide sequence HGPLTATGTK
spectrum of synthetic spliced peptide; spliced peptide sequence HGPLTATGTK
Fib spectrum 4897; alternative peptide sequence RVYGTTGTK
spectrum of synthetic alternative peptide; alternative peptide sequence RVYGTTGTK
0
100
200
300
400
500
600
700
800
900
1000

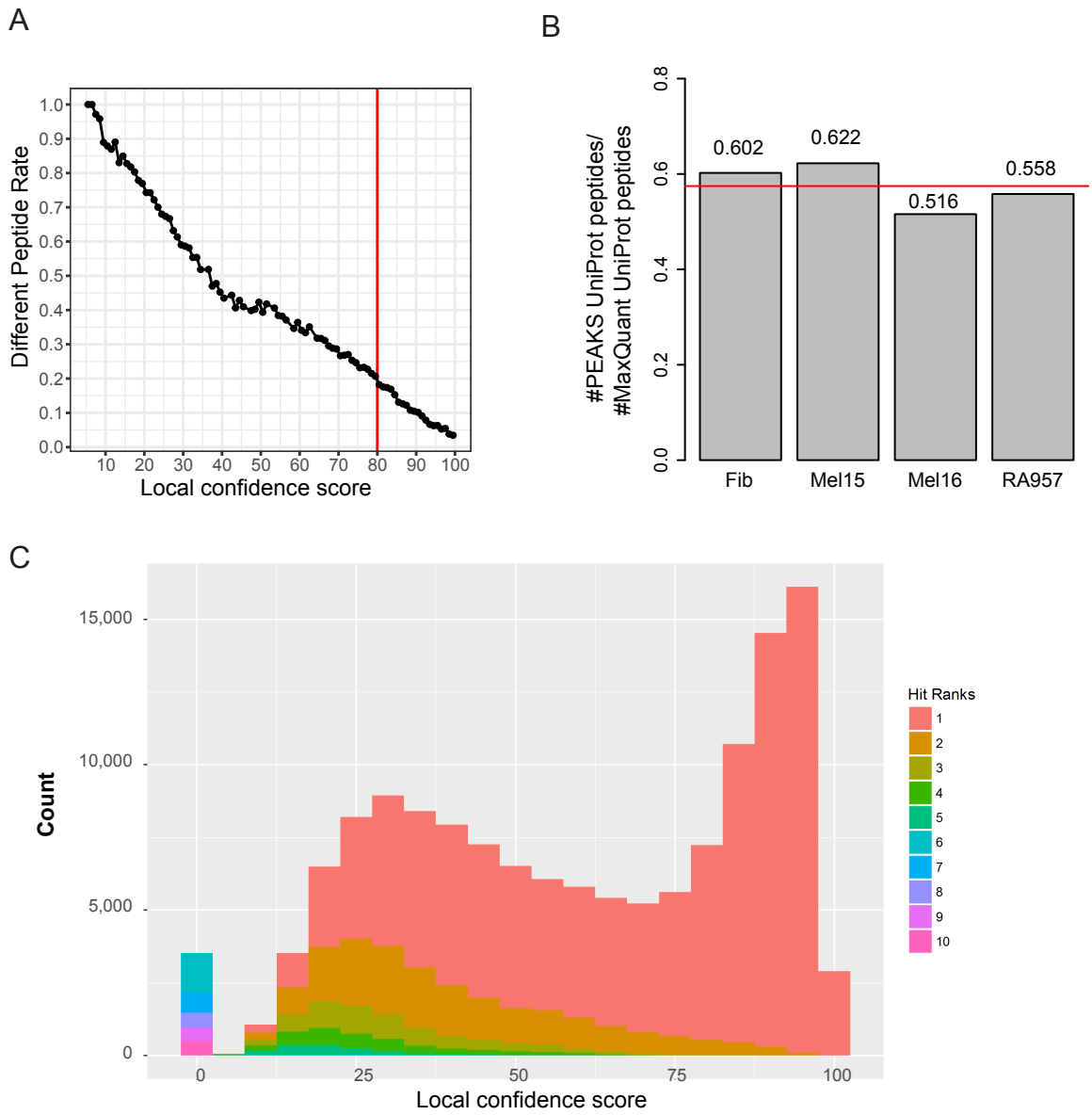

Fib spectrum 4969; spliced peptide sequence VLSTESPASK
spectrum of synthetic spliced peptide; spliced peptide sequence VLSTESPASK
Fib spectrum 4969; alternative peptide sequence VVM(+15.995)RDPASK
spectrum of synthetic alternative peptide; alternative peptide sequence VVM(+15.995)RDPASK
M(+15.995)
0
100
200
300
400
500
600
700
800
900
1000


Fib spectrum 4996; spliced peptide sequence RVTGALQKK
spectrum of synthetic spliced peptide; spliced peptide sequence RVTGALQKK
Fib spectrum 4996; alternative peptide sequence RLSGALQKK
spectrum of synthetic alternative peptide; alternative peptide sequence RLSGALQKK
0
100
200
300
400
500
600
700
800
900
1000

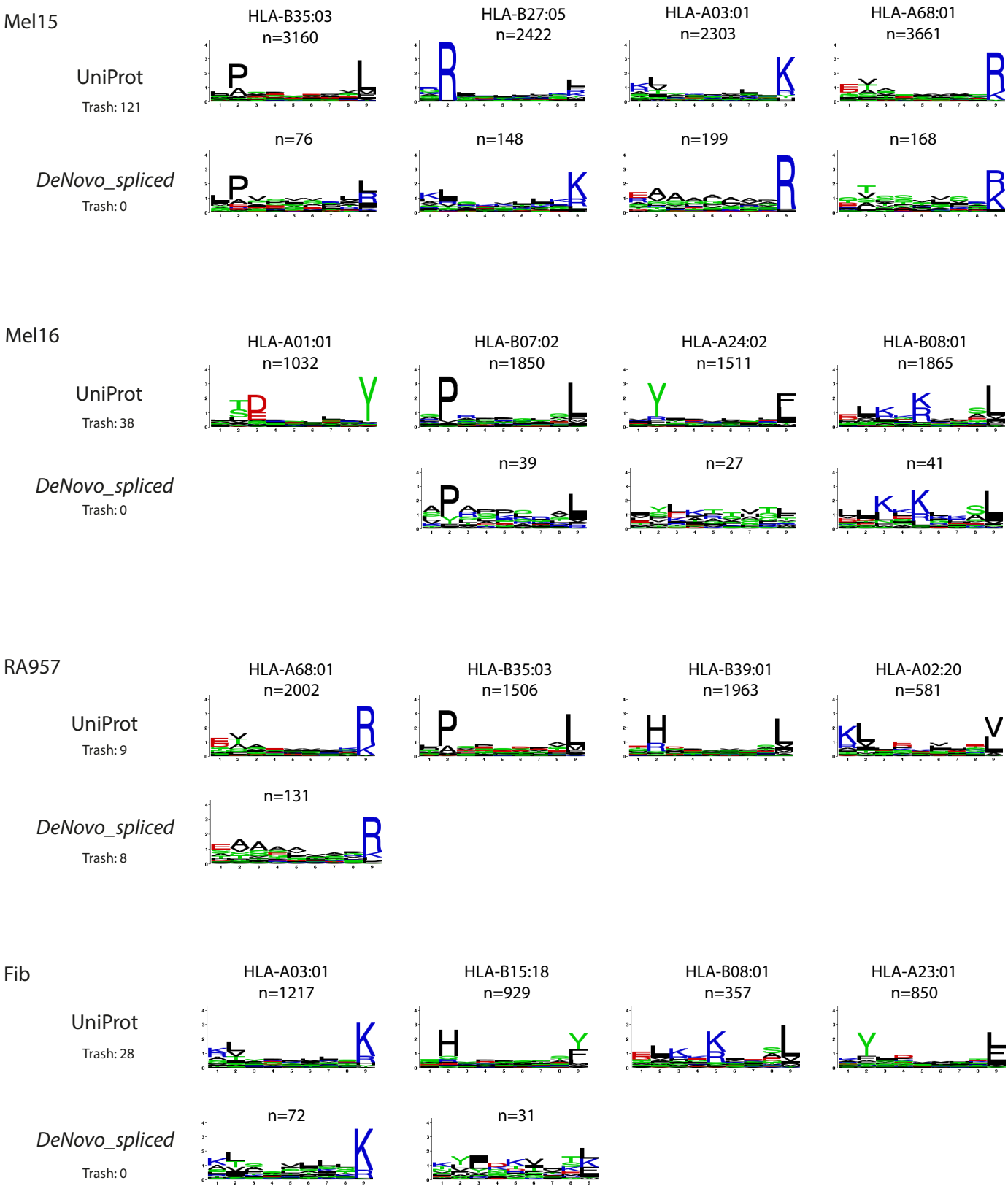

Fib spectrum 4031; spliced peptide sequence LENKKGKSL
spectrum of synthetic spliced peptide; spliced peptide sequence LENKKGKSL
100
200
300
400
500
600
700
800
900
1000


Fib spectrum 4783; spliced peptide sequence KRIPLPTKK
spectrum of synthetic spliced peptide; spliced peptide sequence KRIPLPTKK
100
200
300
400
500
600
700
800
900
1000
1100


Fib spectrum 4819; spliced peptide sequence ATSQQKRLP
A 1 T 2 S 3 Q 4 Q 5 K 6 R 7 L 8 P
P 1 L 2 R 3 K 4 Q 5 Q 6 S 7 T 8 A
spectrum of synthetic spliced peptide; spliced peptide sequence ATSQQKRLP
A 1 T 2 S 3 Q 4 Q 5 K 6 R 7 L 8 P
P 1 L 2 R 3 K 4 Q 5 Q 6 S 7 T 8 A
0
100
200
300
400
500
600
700
800
900
1000

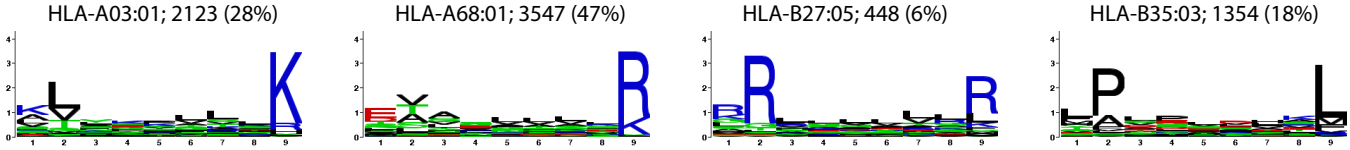


Supplemental Figure 6: A) Estimation of the error rate in PEAKS results by the fraction of Mel15 UniProt PSMs (y-axis) with a local confidence score higher than a threshold (x-axis), for which PEAKS assigned a different peptide compared to MaxQuant. Since the MaxQuant FDR is in agreement with the Comet FDR and Mascot FDR for the UniProt peptides (Figure 1 C), we assume that the error by MaxQuant is indeed about 1% and most of the sequence differences in this case are caused by erroneous PEAKS matches. PEAKS shows a FDR of about 20% at a local confidence score of 80 (red line). B) Fraction of UniProt HLA peptides identified by PEAKS at a local confidence score of 80 compared to MaxQuant UniProt peptides at spectrum level of FDR 1% for the different samples. C) A histogram presenting for each PSM the local confidence scores of the ten best matches identified by PEAKS in Mel15 data as an example.



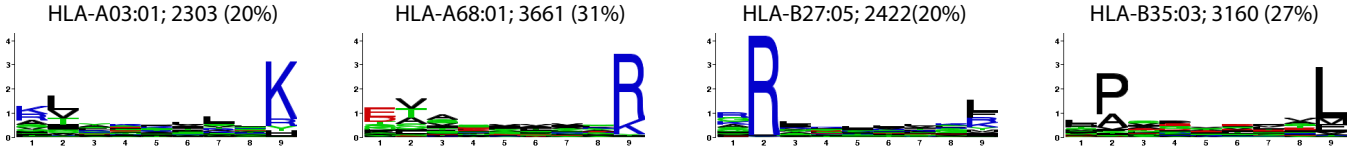
Supplamental Figure 7: Sequence motifs for the 9-mer *DeNovo_spliced* peptides and UniProt peptides identified by MaxQuant in Mel15, Mel16, RA957 and Fib samples.

A. PEAKS



Trash cluster: 4 (0.1%)

B. MaxQuant



Trash cluster: 121 (1%)

Supplemental Figure 8: A) Clustering analysis with MixMHCp of HLA-Ip identified as UniProt hits by PEAKS in Mel15 sample revealed three dominant clusters referring to HLA-B35:03, HLA-A68:01 and HLA-A03:01 and only 6% of the peptides clustered into the motif of the HLA-B27:05 allele. B) Ssimilar clustering analysis of HLA-Ip identified as UniProt hits by MaxQuant from the same MS data revealed four dominant clusters and 20% of the peptides clustered into the HLA-B27:05 motif.