

EXTENDED DATA TABLES

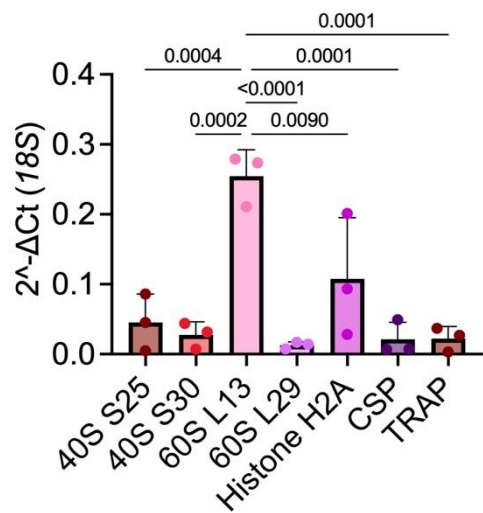
Extended data table 1 | *Plasmodium falciparum* homology

Protein name	Number of patients	<i>P. falciparum</i> homology (%)
60S L4/L1	7	83
ETRAMP A5KBH5	7	63
40S S2	6	97
Don Juan	6	53
ETRAMP A5K676	6	43
40S S11	5	99
Histone H2A A5K214	5	98
EF 1-alpha	5	98
40S S23	5	98
60S L32	5	94
60S L29	5	91
40S S8	5	89
40S S6	5	89
40S S30	5	89
40S S25	5	86
60S L13	5	80
Histone H2A A5K7L8	4	97
60S L37	4	97
60S L30	4	92
Alba domain	4	91
Ubiq/Rib	4	91
60S L36	4	90
60S L35	4	81
60S L23a	4	64
Unch. A5K8G9	4	33
Histone H3	3	100
60S L37a	3	100
50S S28e	3	97
60S L44	3	97
60S L10	3	96
60S L8	3	95
60S L9	3	94
60S L33-B	3	92
60S L18a	3	92
60S L13a	3	91
60S L34-A	3	91
60S L31	3	90
60S L17	3	89
60S L7a	3	84
PPR_long domain	3	59
Unch. A5K9Y3	3	43
PI3K/PI4K	3	36
Unch. A5K4L2	3	32
Unch. A5K8G8	3	0
Unch. A5KC18	3	23

Extended data table 2 | Primers used for *Plasmodium falciparum* liver stage expression

Primer	Sequence
NF135 40S S25-1F	5'-TCTCCCAGAGCTGTTTACCC-3'
NF135 40S S25-1R	5'-TCATCTTCCAAGTTGTCACACA-3'
NF135 40S S30-1F	5'-ATGGGAAAGGTACATGGATCATTAG-3'
NF135 40S S30-1R	5'-ACTTTGGGACTTTAGGAGTCTGG-3'
NF135 60S L13-1F	5'-AAGAGTTAAAGGGAGCCGGA-3'
NF135 60S L13-1R	5'-ACCTCCAATACCTTTCTTTGCT-3'
NF135 60S L29-1F	5'-ACCGTAAAGCCCATAAGAATGGA-3'
NF135 60S L29-1R	5'-GGATCCAACCCTTTACGCGA-3'
NF135 H2A-1F	5'-TAGGTTCCACAGCAGCAGTT-3'
NF135 H2A-1R	5'-AACACCACCACCAGCAATTG-3'
NF135 CSP-1F	5'-GCCAAATGACCCAAACCGAA-3'
NF135 CSP-1R	5'-GGTTTATTAGCAGAGCCAGGC-3'
NF135 TRAP-1F	5'-AGAGAAACACGTCCACATGG-3'
NF135 TRAP-1R	5'-TCCAGCTATTCCACCTGCAA-3'
18S rRNA (3D7) -1R	5'-GGGTTCTGGGGCGAGTATT-3'
18S rRNA (3D7) -1R	5'-GCCATGCATCACCATCCAAG-3'

EXTENDED DATA FIGURE



Extended data figure 1 | Antigen gene expression in humanized mouse livers infected with *Plasmodium falciparum* (Pf). The expression of genes related to antigens found in *P. vivax* immunopeptidomics was analysed in the liver stage of *Pf* infection. Pf 18S rRNA was used as endogenous gene. CSP and TRAP genes were used as positive control. Statistical analyse was performed by one-way ANOVA, with Turkey’s multiple comparison test. Actual P values are shown.

Extended data figure 2 | Spectral match validation

Peptide			Biological peptide				Synthetic peptide				Validation
Sequence	ID	n	[m/z]	z	RT	Peaks Score	[m/z]	z	RT	Peaks Score	(y/n)
AIEPSLAQLAQK	Ubqt/ribos- 1	12	635.3643	2	41.42	37.49	634.8646	2	41.64	48.35	Y
AKAVKKQKKTLKPV	60S L31	14	523.0190	3	21.67	45.30	NA	NA	NA	NA	NA
ANKALLPTAGDD	60S L4/L1- 1	12	593.3108	2	34.53	50.03	593.3093	2	34.31	52.84	Y
APISAGIKKPHR	Hist. H3	12	425.5951	3	25.01	40.59	425.5946	3	23.20	23.34	Y
AQIKKKKKLTPA	Don Juan	13	494.6678	3	24.09	36.08	494.6680	3	20.52	41.55	Y
ASGGVLPNIHNV	Hist. H2A- 2	12	589.3206	2	39.78	25.89	589.3203	2	39.89	37.05	Y
AVKKVGKIVK	60S L32-2	10	357.2537	3	22.97	22.03	357.2534	3	20.62	27.81	Y
DNNEHVVEKETV	Unch A5JZN7.2	12	706.3463	2	27.81	53.74	706.3451	2	26.94	49.45	Y
DNNEHVVEKETVSF	Unch A5JZN7.1	14	549.2657	3	35.20	61.62	549.2657	3	34.69	47.01	Y
EDYSPRKV	Unch A5K8G9	8	497.2546	2	26.23	39.03	497.2538	2	24.87	29.67	Y
GAFGKPNGV	60S L10	9	423.7279	2	30.37	19.14	423.7272	2	29.41	29.60	Y
GKGKNKEKL	40S S25- 1	9	501.3087	2	20.20	33.76	501.3089	2	19.19	25.14	Y
GKGKNKEKLNHAVF	40S S25- 2	14	523.9681	3	25.17	48.81	523.9667	3	24.09	25.39	Y
GLNQKQPTKGSNIQ	ETRAMP- 2	14	757.4133	2	24.53	37.29	756.9106	2	23.51	45.90	Y
GPKRATKIRK	40S S6- 3	10	385.5873	3	20.09	38.30	385.5876	3	18.95	33.14	Y
GRIGRYLKKGYA	Hist. H2A- 4	13	378.2308	4	32.42	33.32	378.2307	4	26.84	33.94	Y
GRIGRYLKKGYAK	Hist. H2A- 1	14	410.2536	4	30.86	33.56	410.2541	4	25.27	17.95	Y
GSKQVHVV	40S S8- 2	8	427.2482	2	24.78	41.76	427.2479	2	23.31	37.69	Y
GTGRKKGPNSKL	40S S30- 2	12	414.9137	3	20.78	46.04	414.9140	3	19.77	36.32	Y
GTPIEKLHPI	60S L13- 2	10	552.8251	2	34.59	27.43	552.8255	2	34.19	35.45	Y
GVKKDVAK	40S S6- 2	8	422.7661	2	19.95	34.31	422.7660	2	18.91	26.58	Y
KKVAAGYKKLTD	ETRAMP- 1	12	441.2663	3	24.33	41.58	441.2662	3	22.62	39.46	Y
KKYKNKKFKPY	60S L35	11	491.2969	3	27.30	25.41	491.6320	3	23.68	19.14	Y
KNIKSKNGIGGIPAD	60S L13- 3	15	504.6231	3	27.71	31.77	504.6229	3	26.84	16.62	Y
KSAGADSKSLKKLD	ETRAMP- 5	14	483.6093	3	23.95	34.26	483.2758	3	22.52	44.23	Y
KVPVPPTQAKPKKN	Hist. H2A- 5	15	415.7615	4	23.19	35.10	415.7620	4	21.87	40.76	Y
LETYQNMKIQQTP	40S S2	14	861.4510	2	32.68	40.56	861.9502	2	31.64	48.98	Y
LGGLNQKQPT	ETRAMP- 4	10	528.2960	2	28.65	37.38	528.2964	2	27.59	39.32	Y
LILRAAIKTK	Unch A5K8J8	10	1126.7635	1	54.45	17.89	376.2605	3	29.11	20.42	N
LNQKQPTKGSNIQ	ETRAMP- 3	13	728.4005	2	23.52	25.98	728.4000	2	22.49	41.12	Y
MYKKVYVID	60S L3a	9	579.8156	2	32.90	39.80	579.8154	2	31.92	31.61	Y
NKALLPTAGDD	60S L4/L1- 2	11	557.7906	2	34.65	43.43	557.7903	2	35.38	50.46	Y
NQLRPKKLKL	Ubqt/ribos- 2	10	418.2789	3	21.63	26.12	418.2785	3	19.94	19.81	Y
PKFFKNQRY	60S L29	9	409.8929	3	28.72	38.08	409.8927	3	27.66	31.37	Y
RLTGGKKKIHKKK	40S S8- 1	13	381.5072	4	20.20	43.53	NA	NA	NA	NA	NA
SDGTGRKKGPNSKL	40S S30- 1	14	482.2673	3	21.44	48.01	482.2670	3	20.18	37.03	Y
SFFNSKKIKKGSKS	40S S11	14	529.3069	3	25.90	41.15	397.2319	4	23.01	37.69	Y
SGGVLPNIHNV	Hist. H2A- 3	11	553.8022	2	39.63	29.28	533.8016	2	39.69	34.16	Y
SKNGKNRFKPKIQ	40S S6- 1	14	415.2520	4	26.02	36.88	415.2518	4	24.19	30.73	Y
SSHAKGIVVEKV	40S S23	12	418.5784	3	27.41	44.55	418.5784	3	25.94	43.89	Y
TGRKKGPNSKL	40S S30- 3	11	395.9069	3	20.74	37.56	297.1819	4	19.67	25.63	N
TPIITNKPFPG	ETRAMP- 6	10	544.3114	2	38.69	36.66	544.3113	2	38.87	46.35	Y
VITDVGDSDIIKTNE	60S L30	15	809.9135	2	41.55	55.90	809.9135	2	41.11	55.13	Y
VSEVTTVEKDE	60S L9	11	618.3045	2	30.31	35.50	618.3038	2	29.67	41.63	Y
YESIEVSKID	60S L13- 1	10	591.7995	2	38.03	40.35	591.7988	2	37.62	41.05	Y
YGRIFKKKITKK	60S L4/L1- 3	12	378.4963	4	28.58	35.21	378.2462	4	23.62	15.05	N

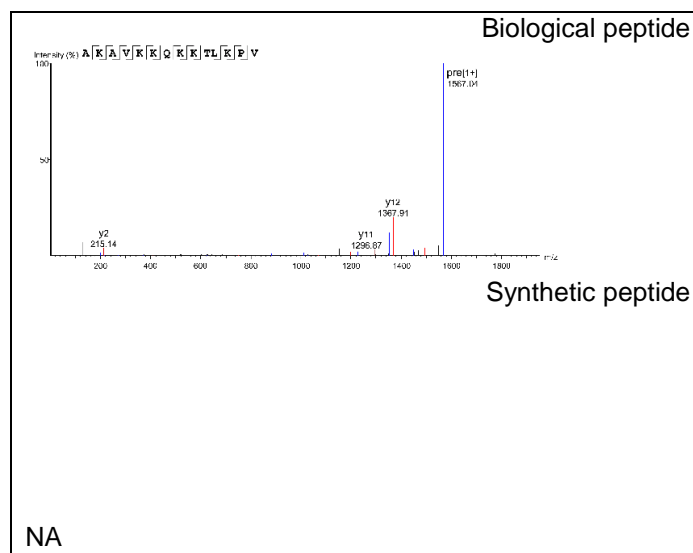
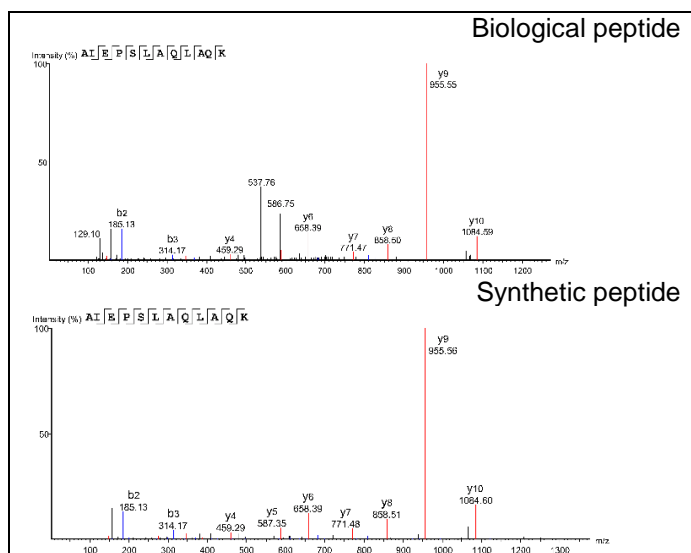
n amino acid length

[m/z] Mass over Charge

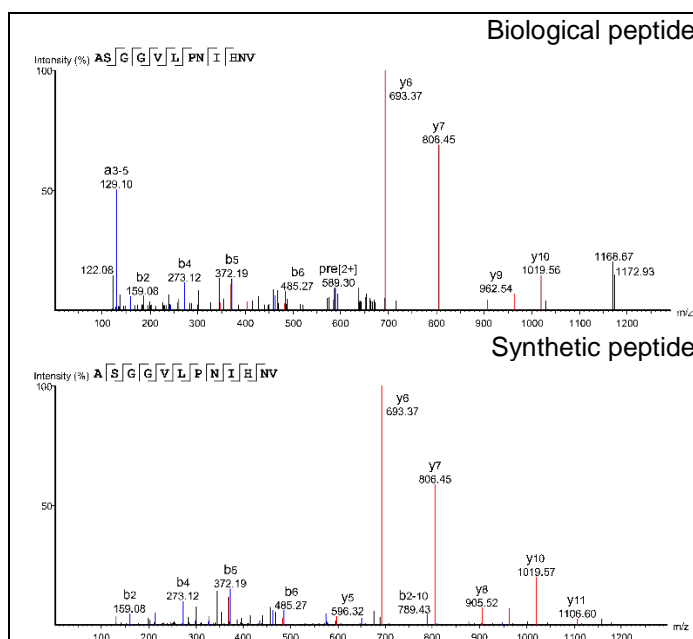
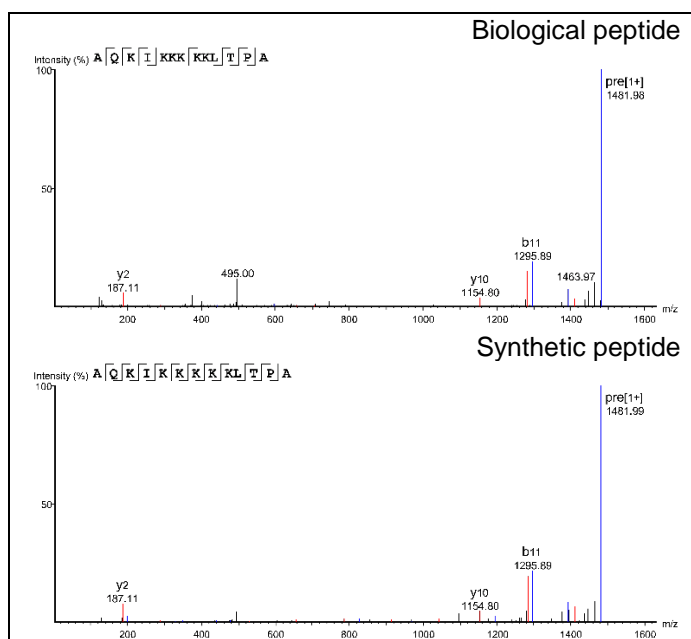
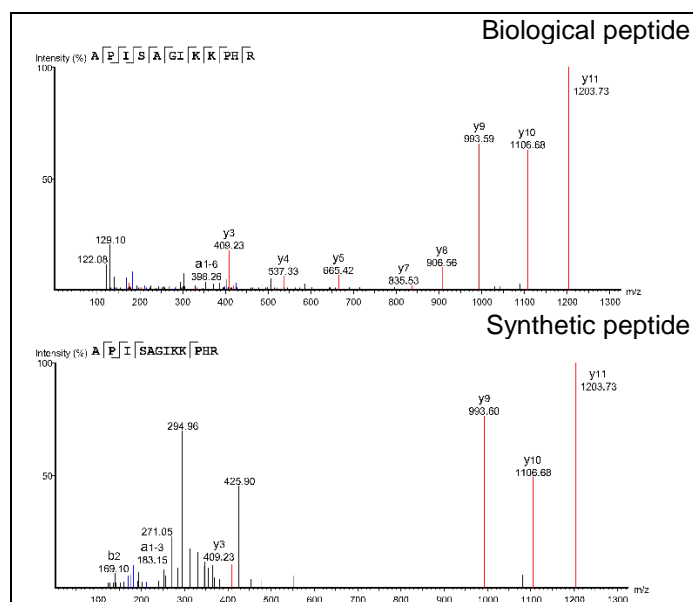
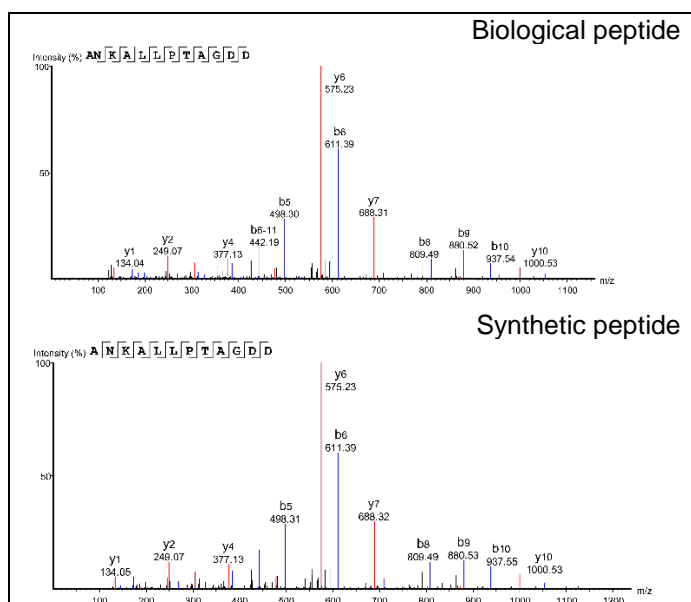
z charge state

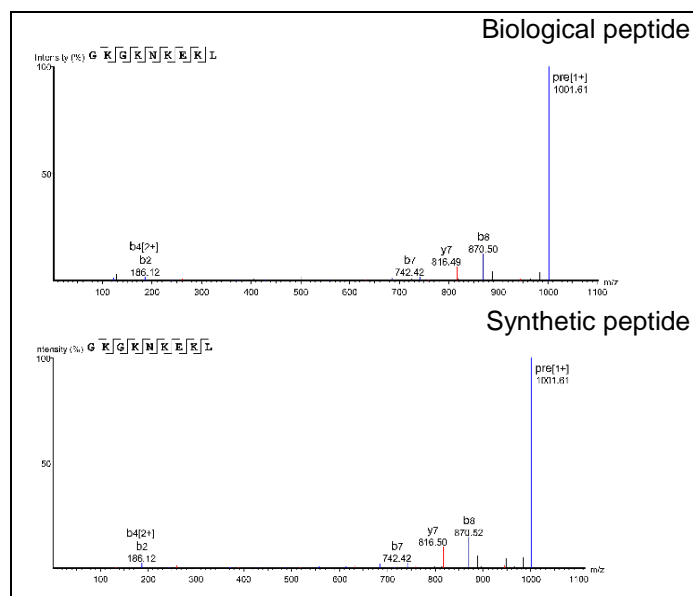
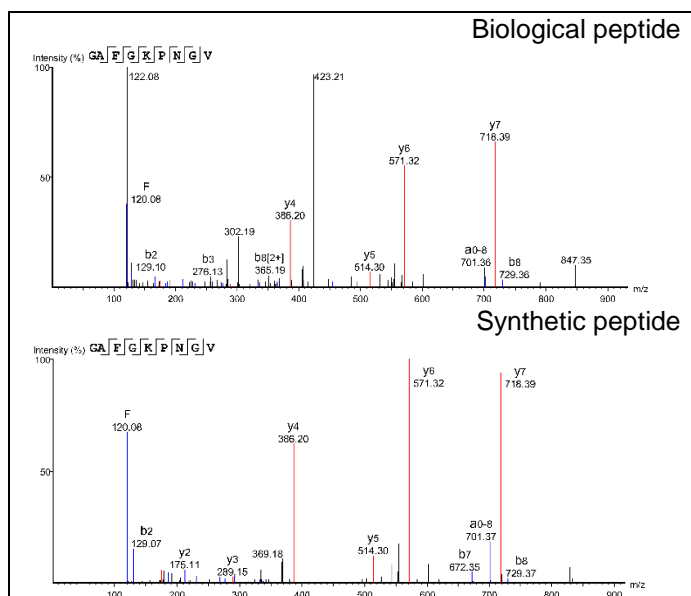
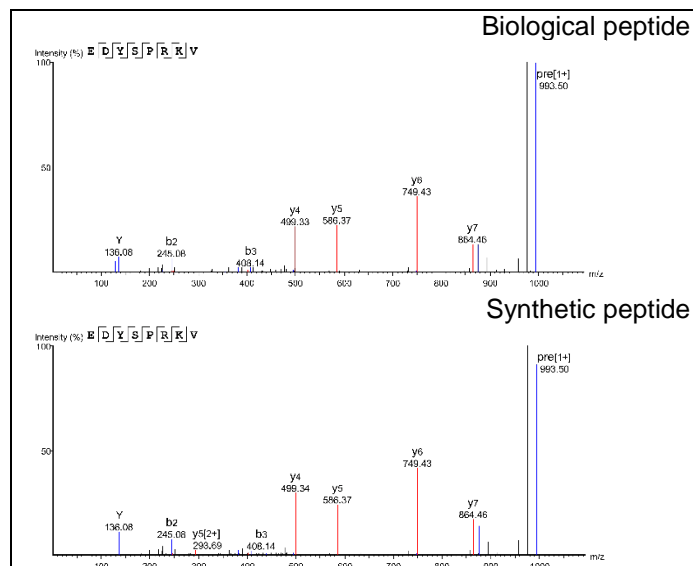
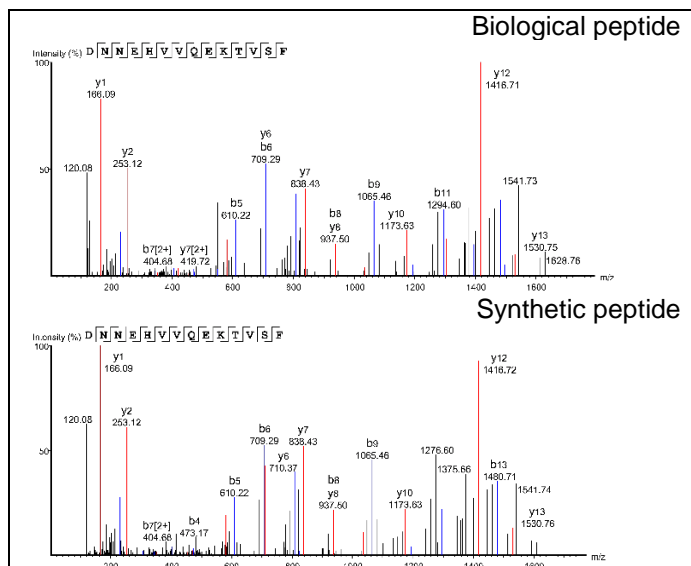
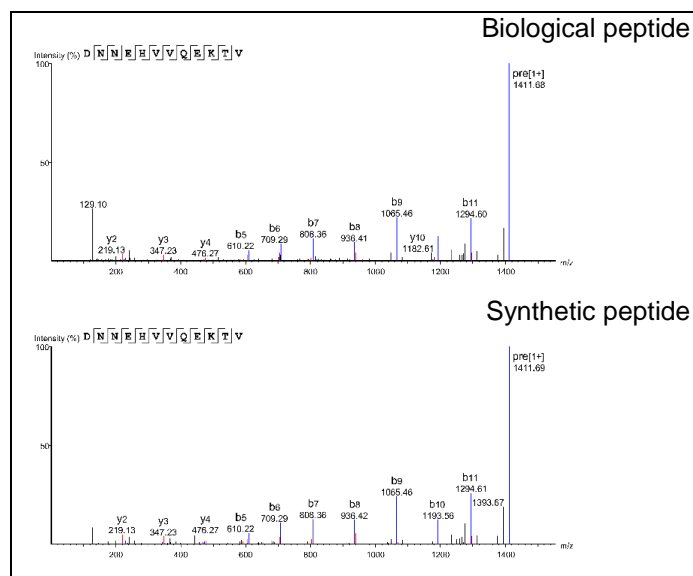
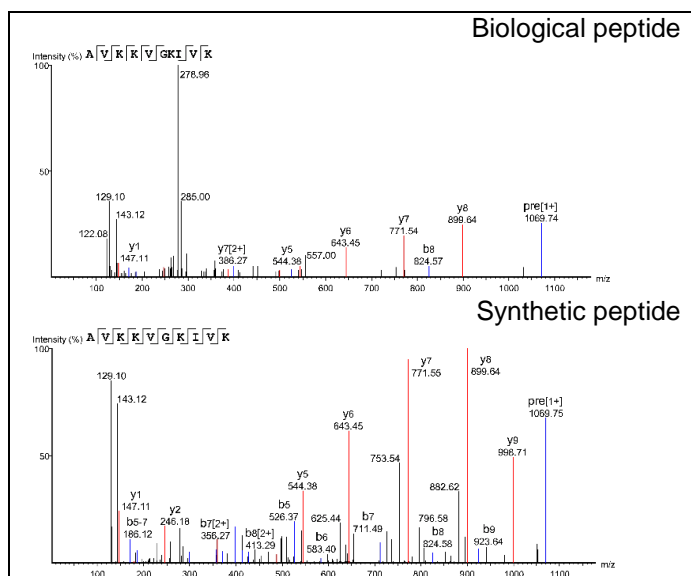
RT Retention Time

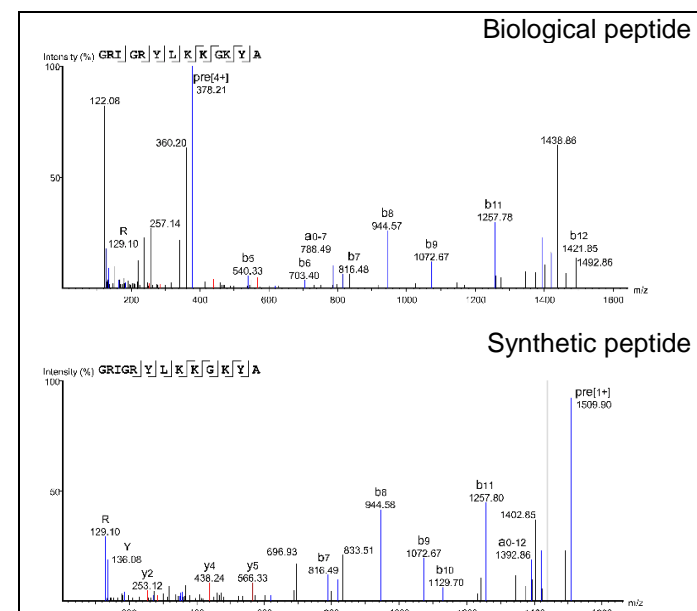
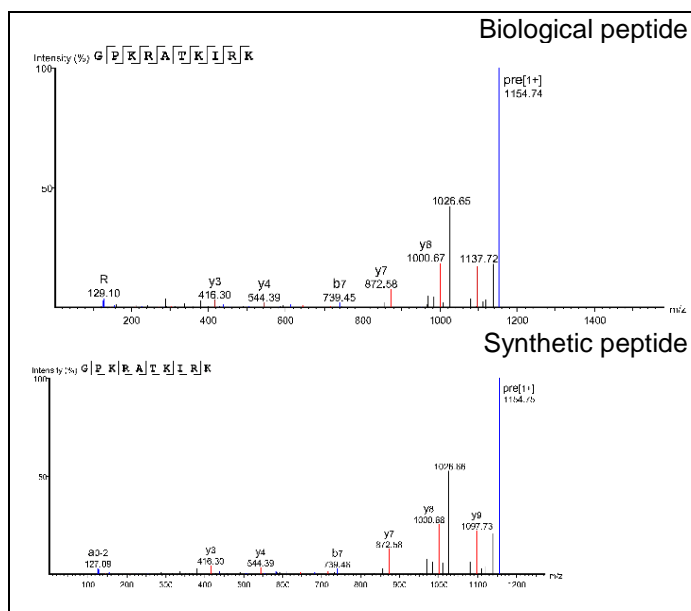
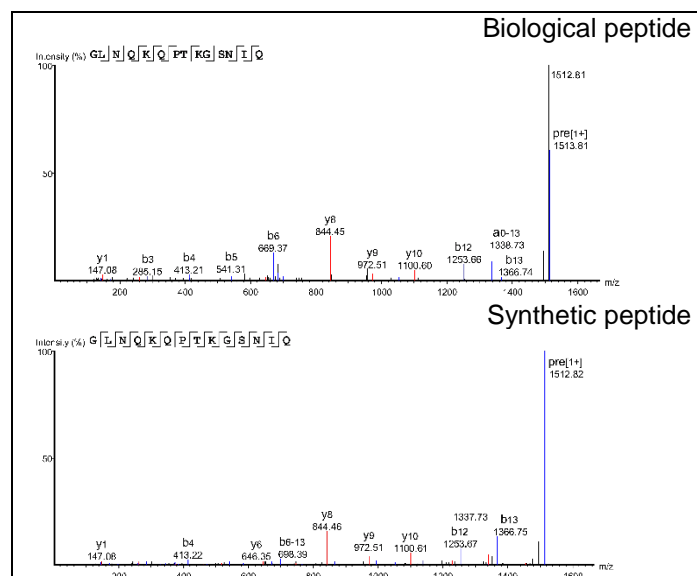
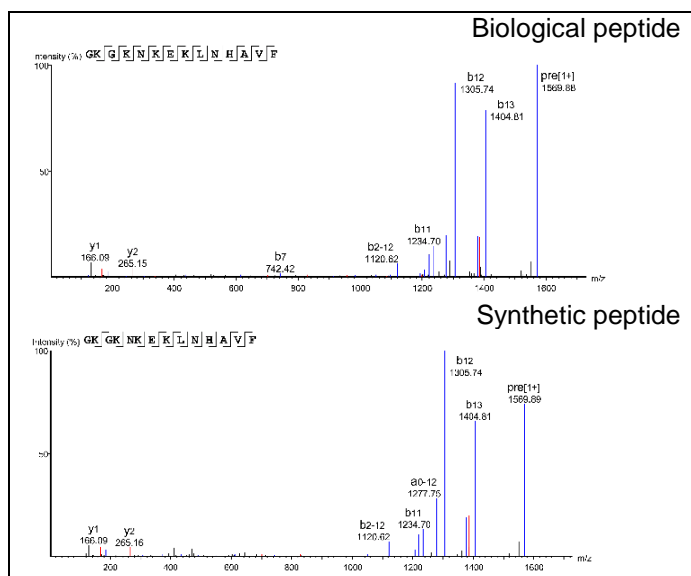
NA Not Available

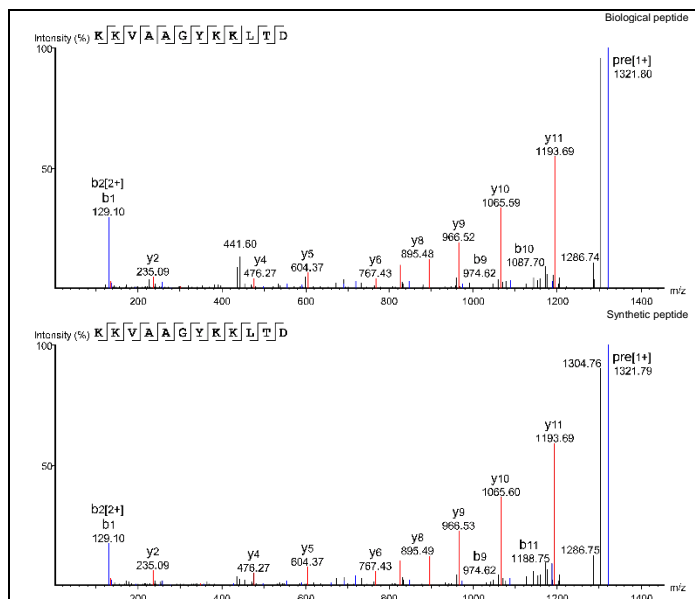
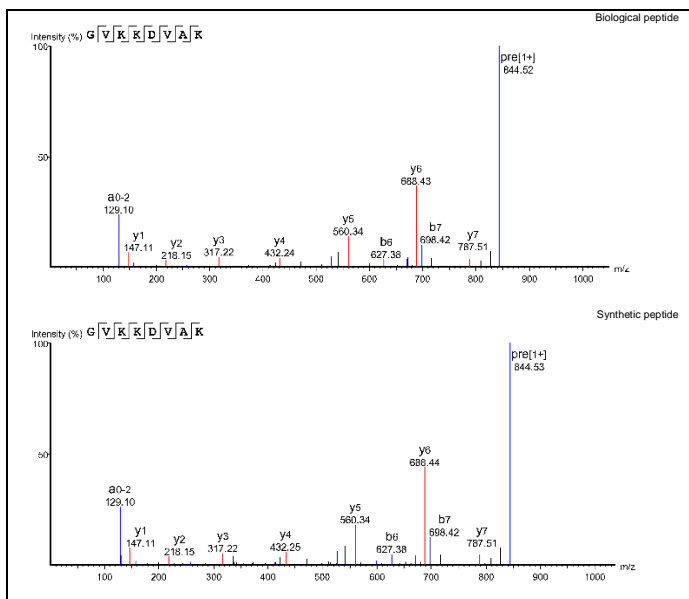
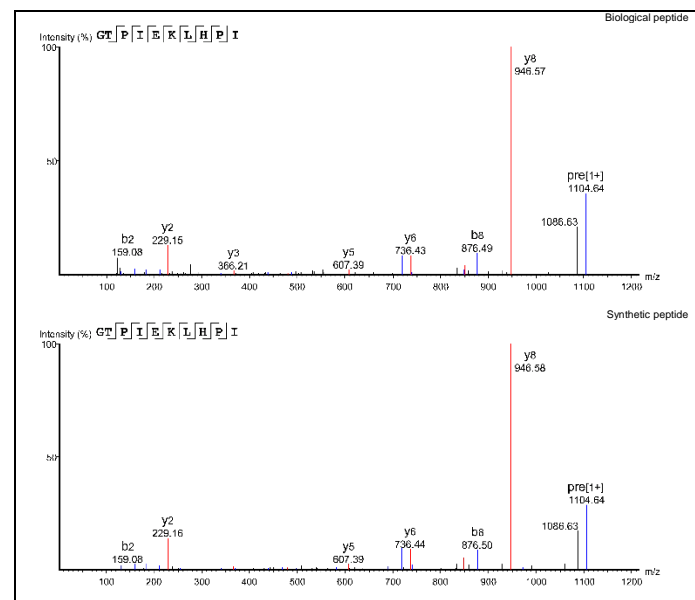
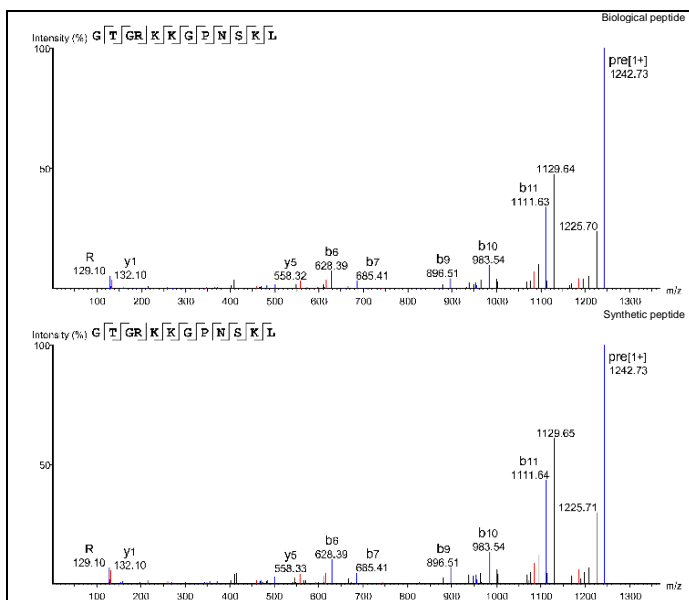
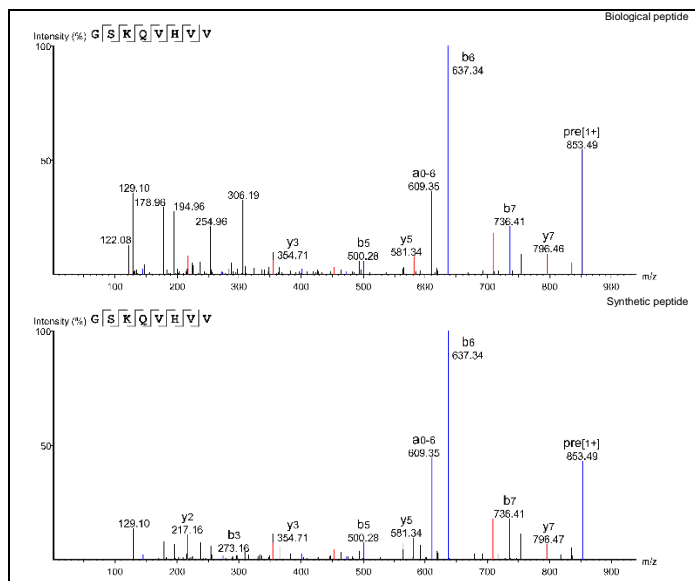
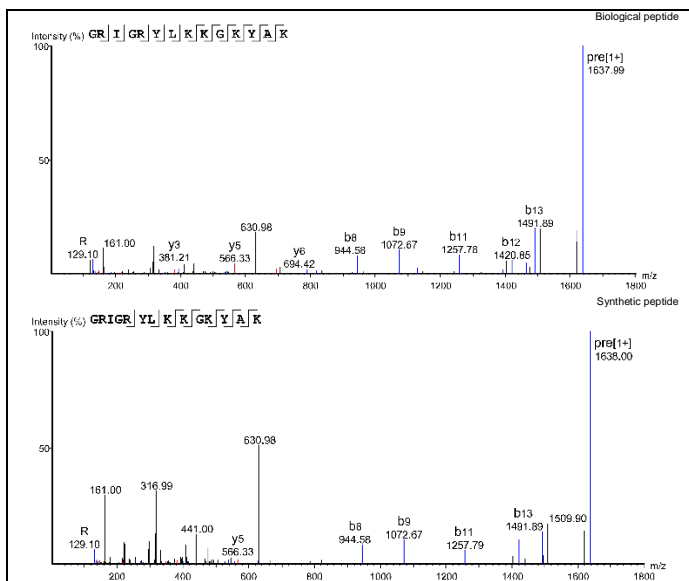


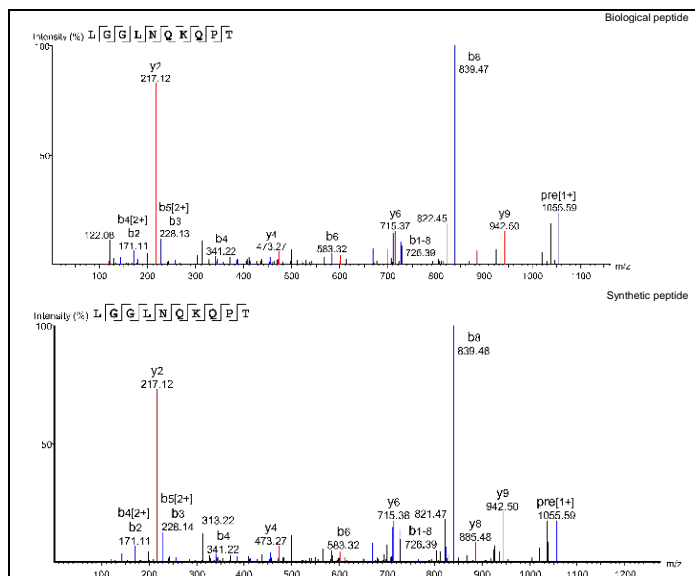
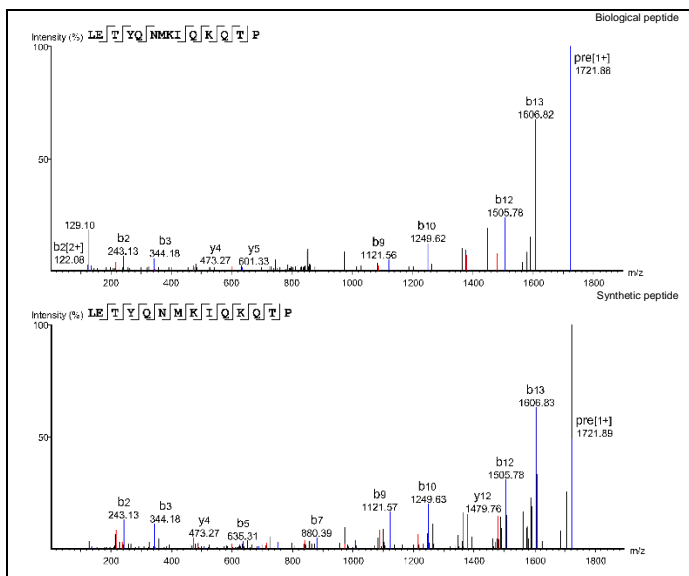
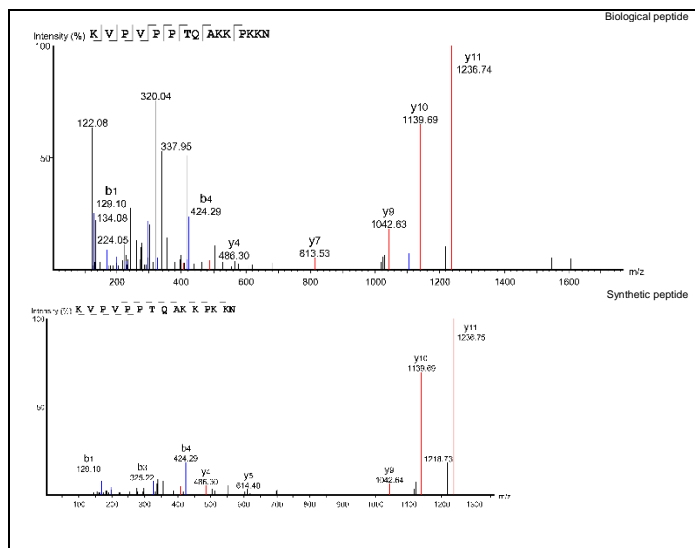
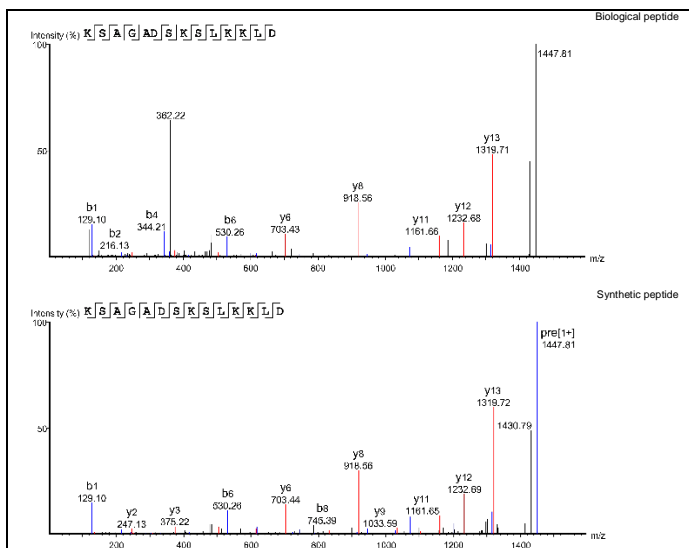
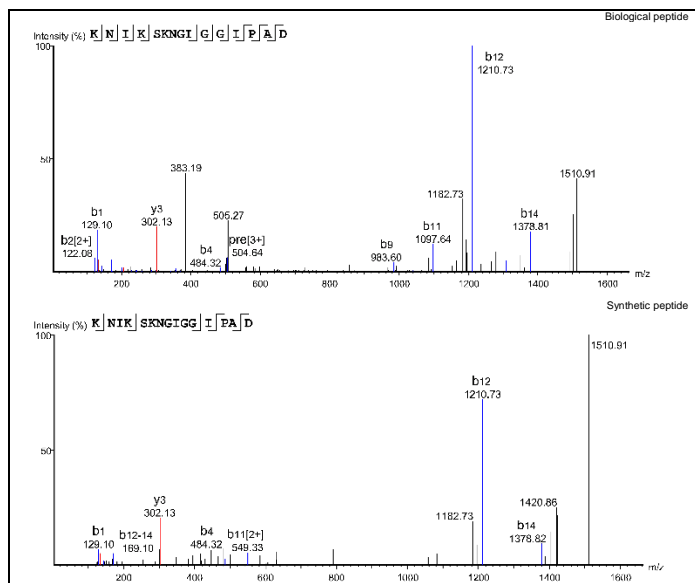
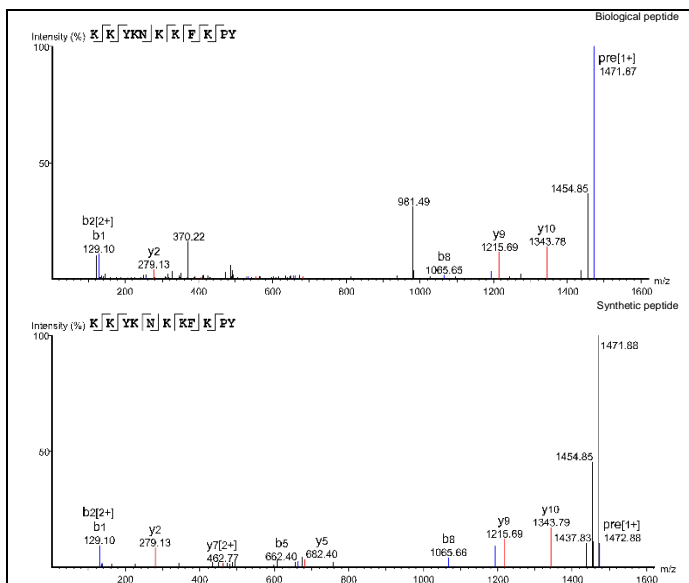
NA

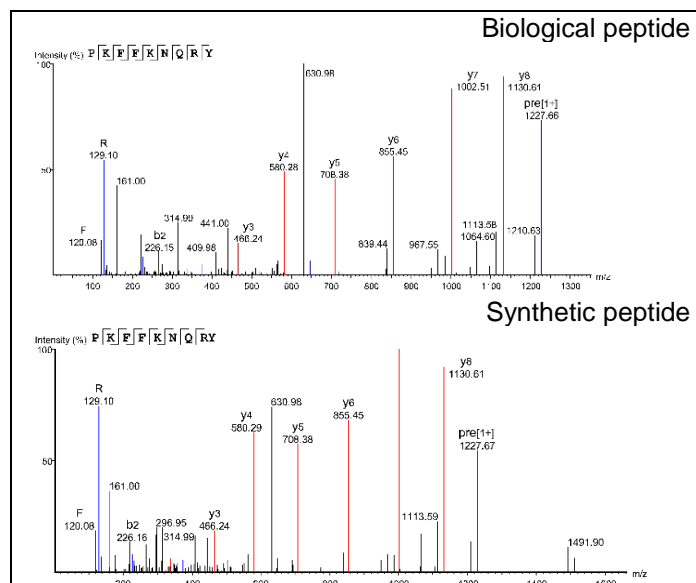
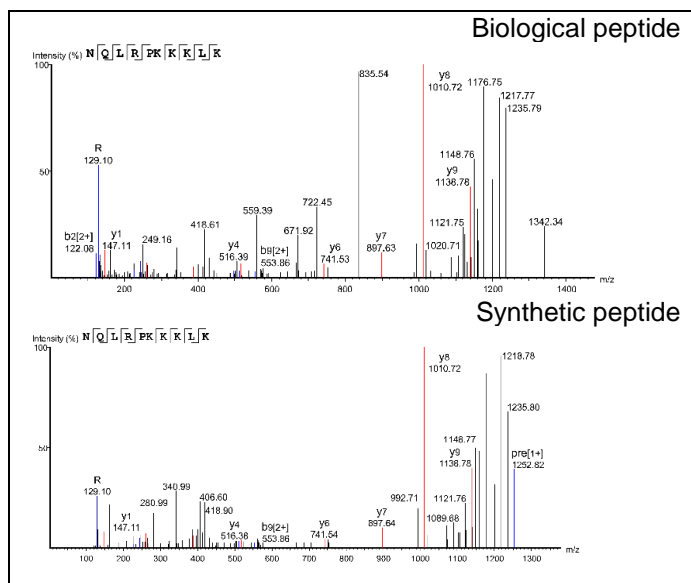
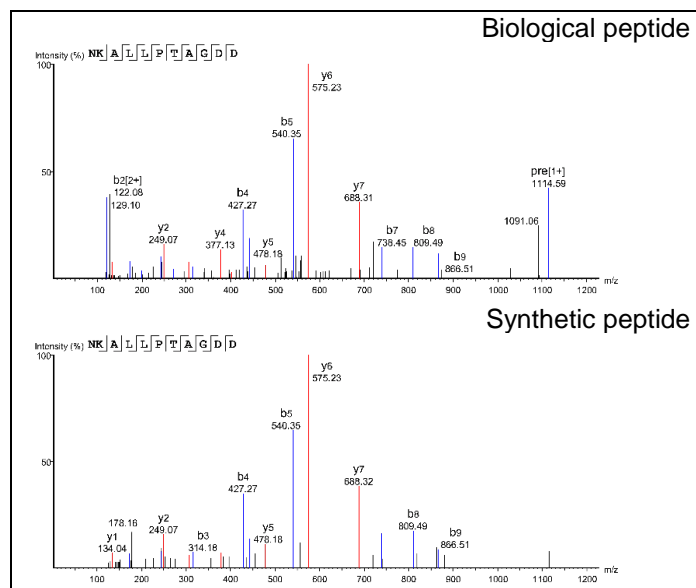
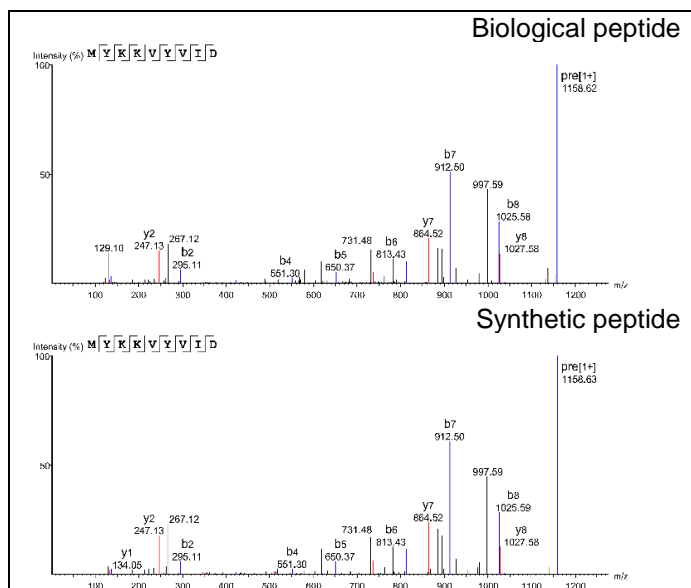
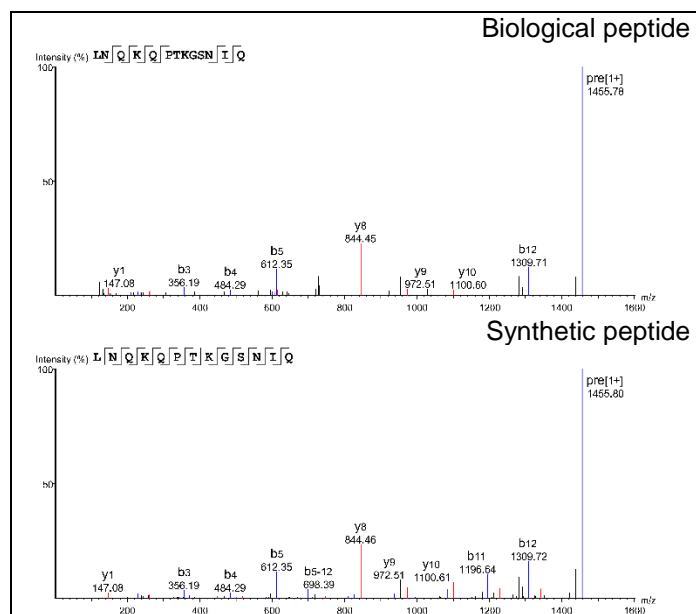
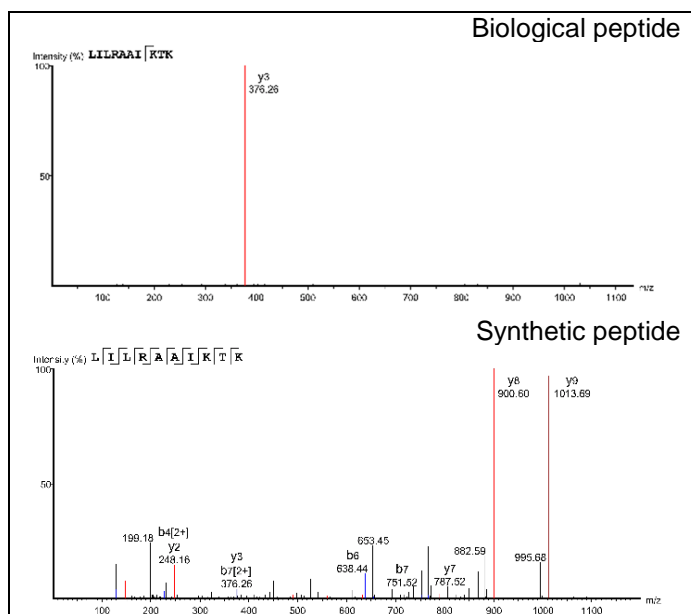


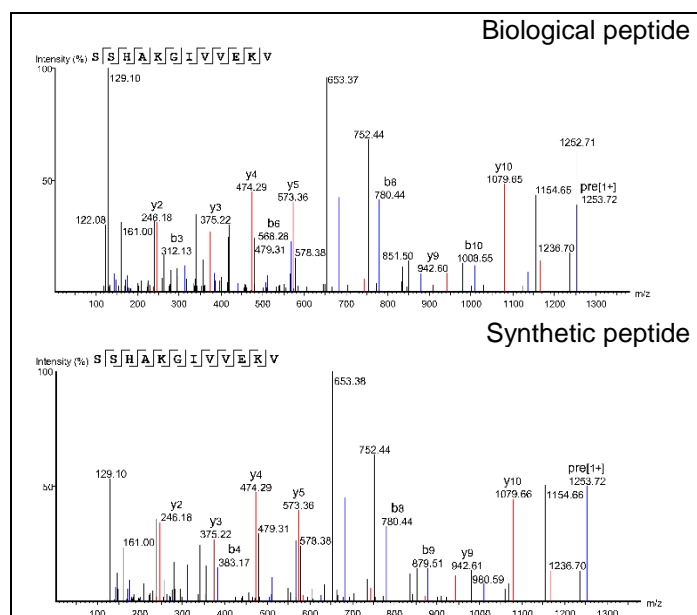
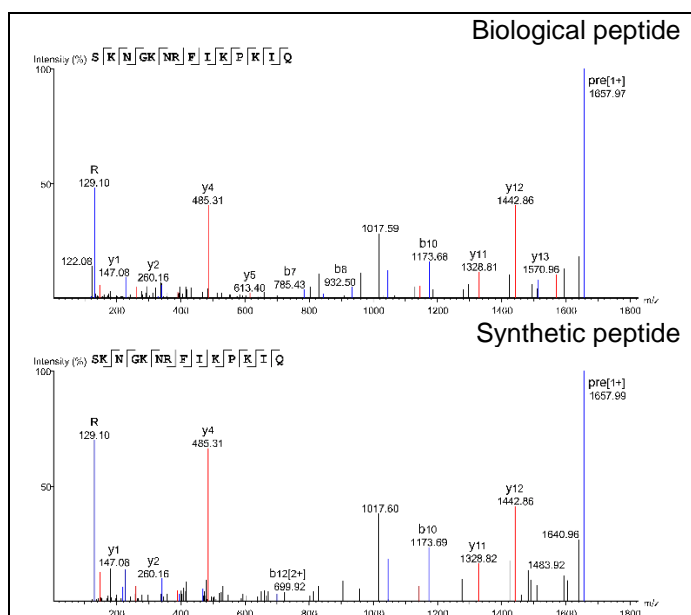
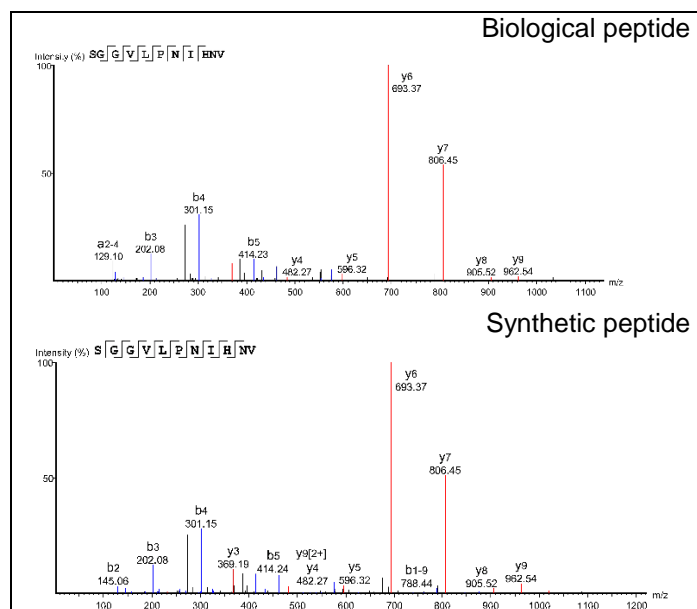
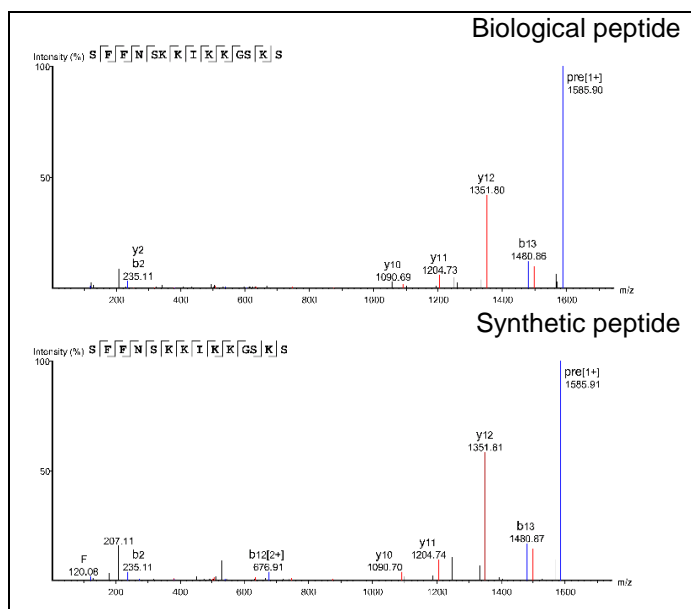
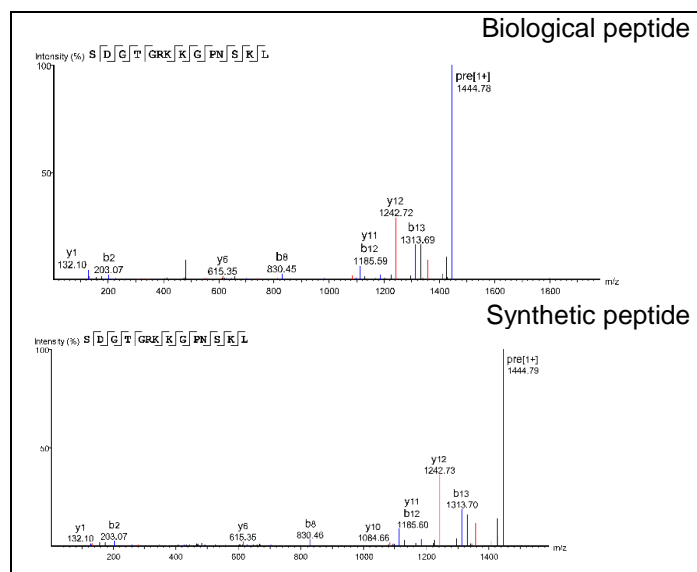
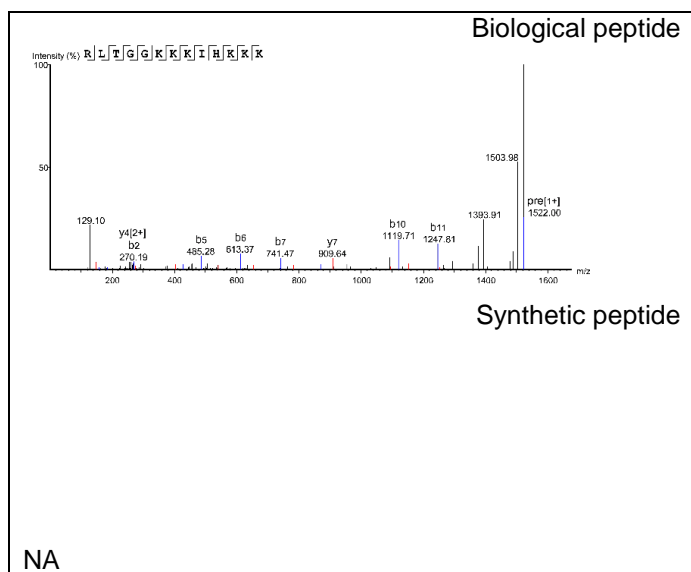


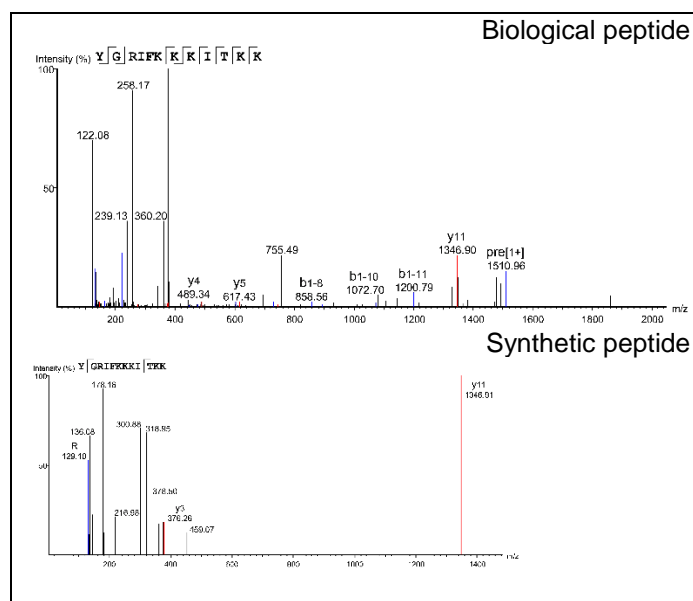
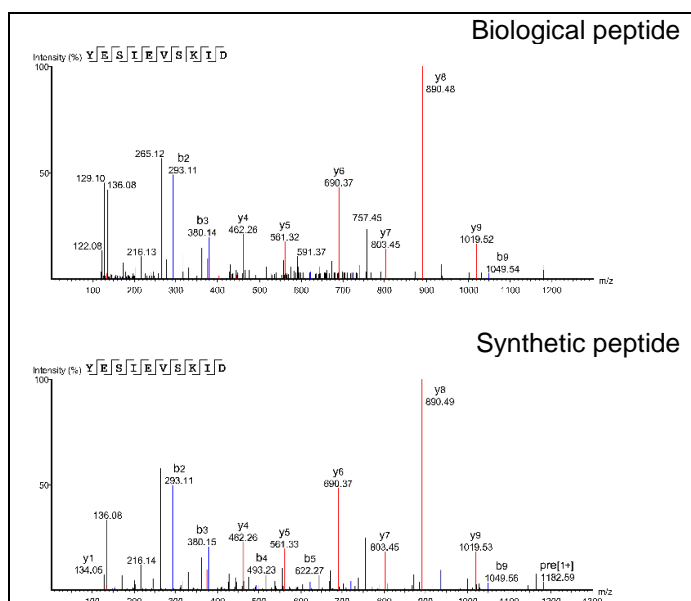
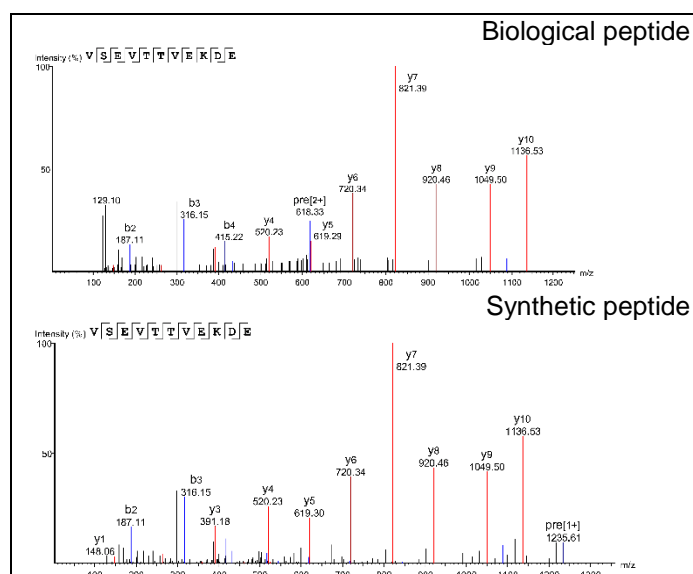
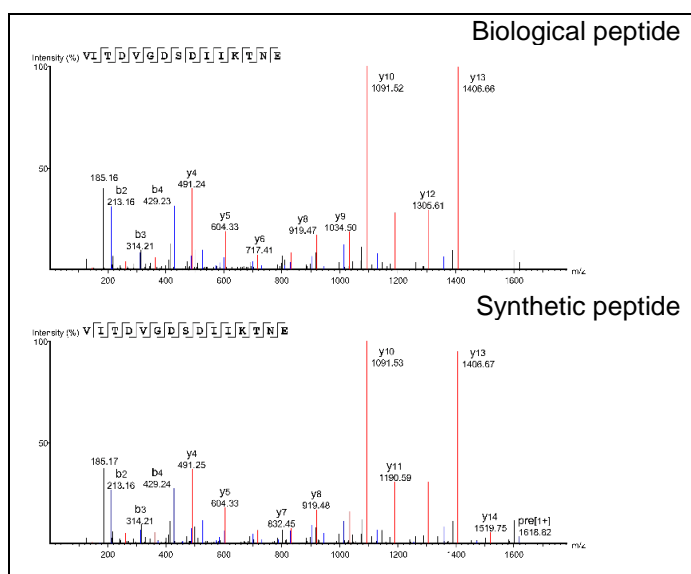
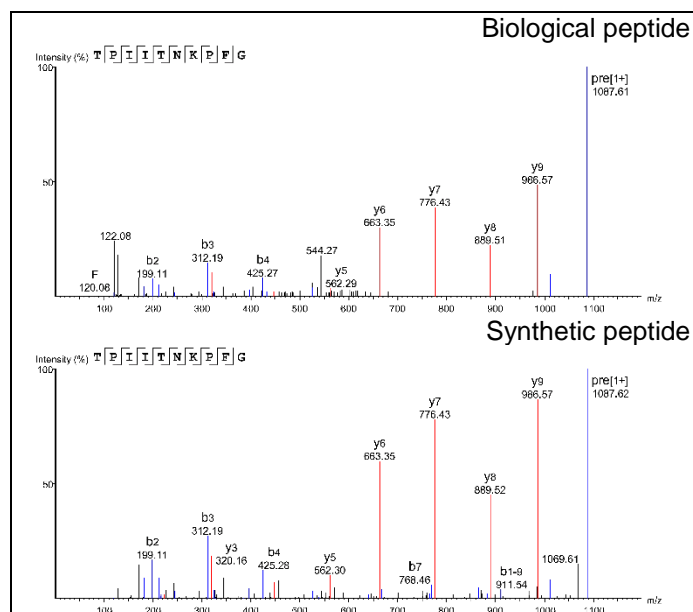
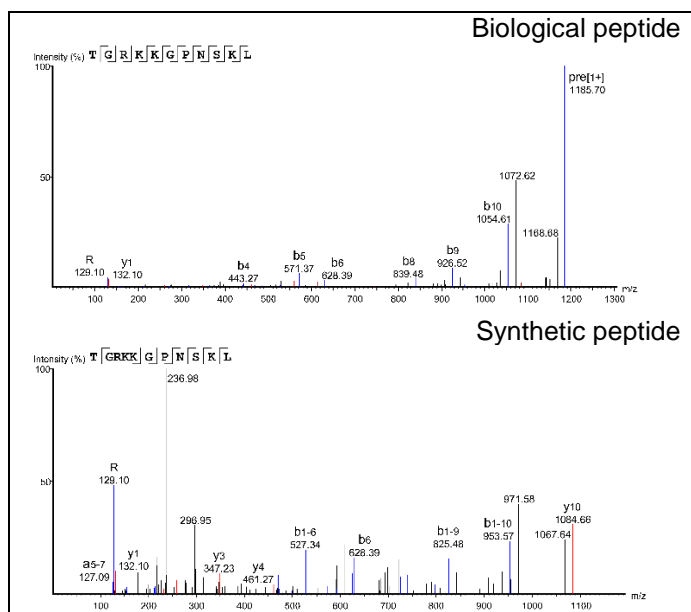


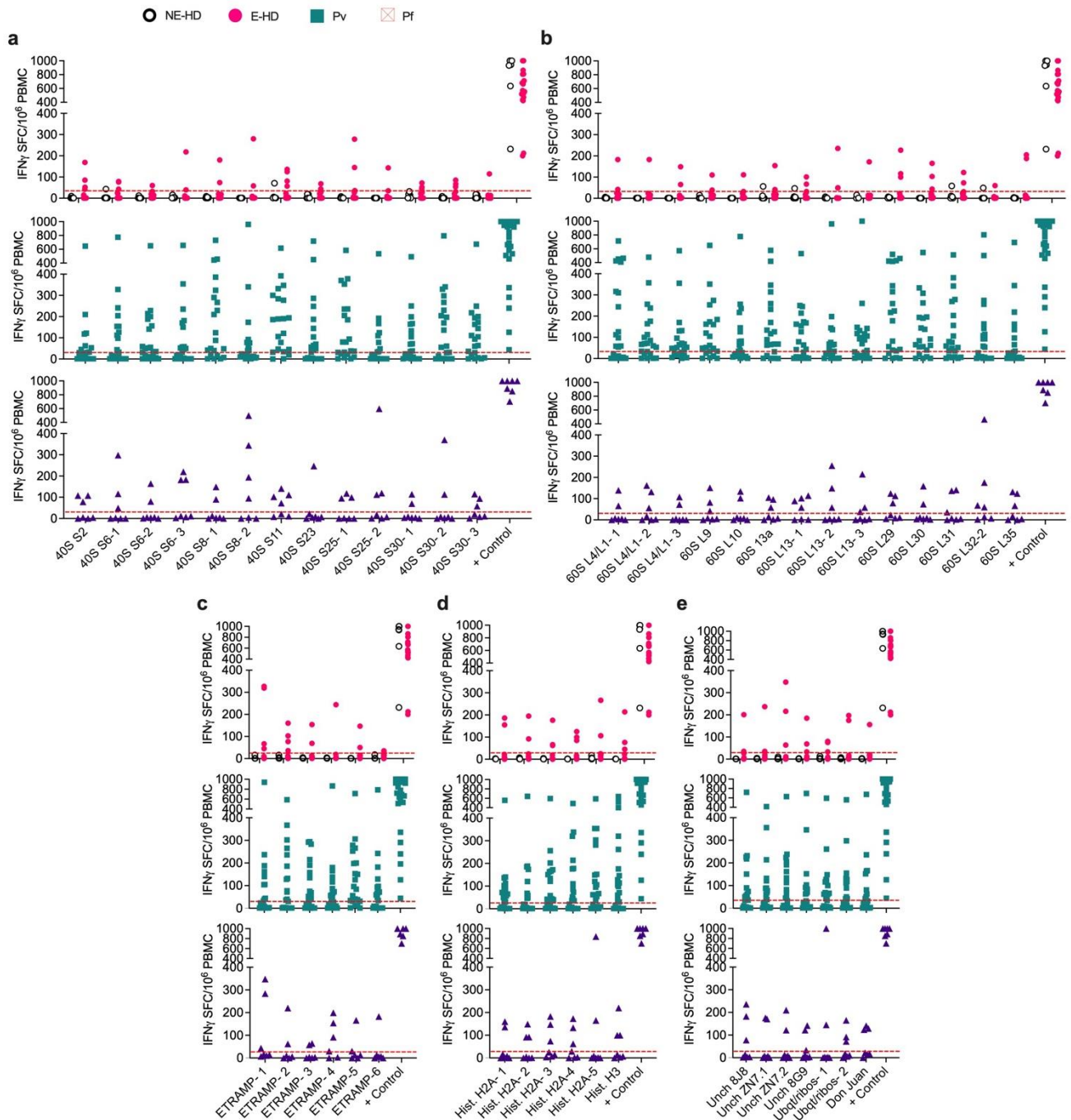




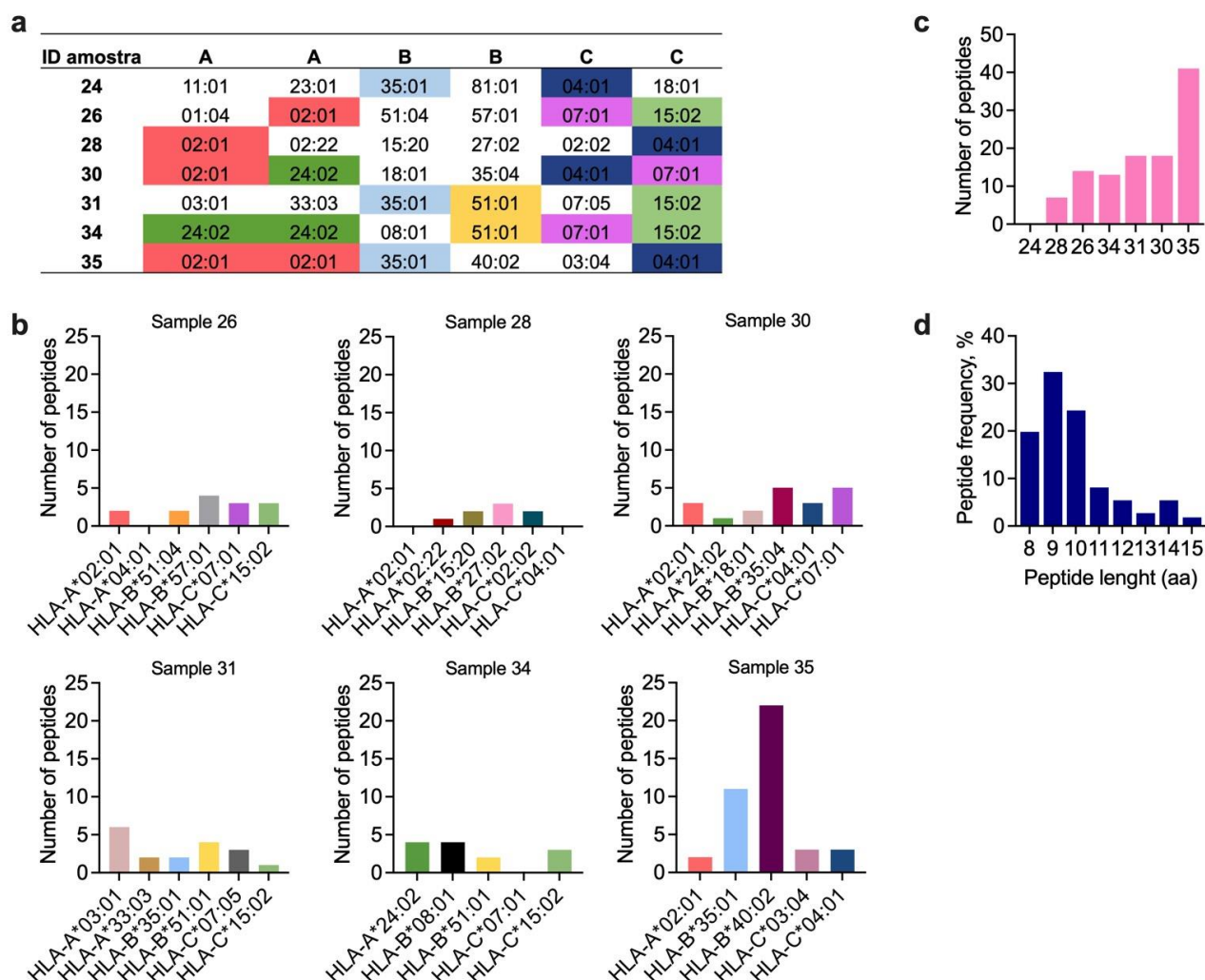




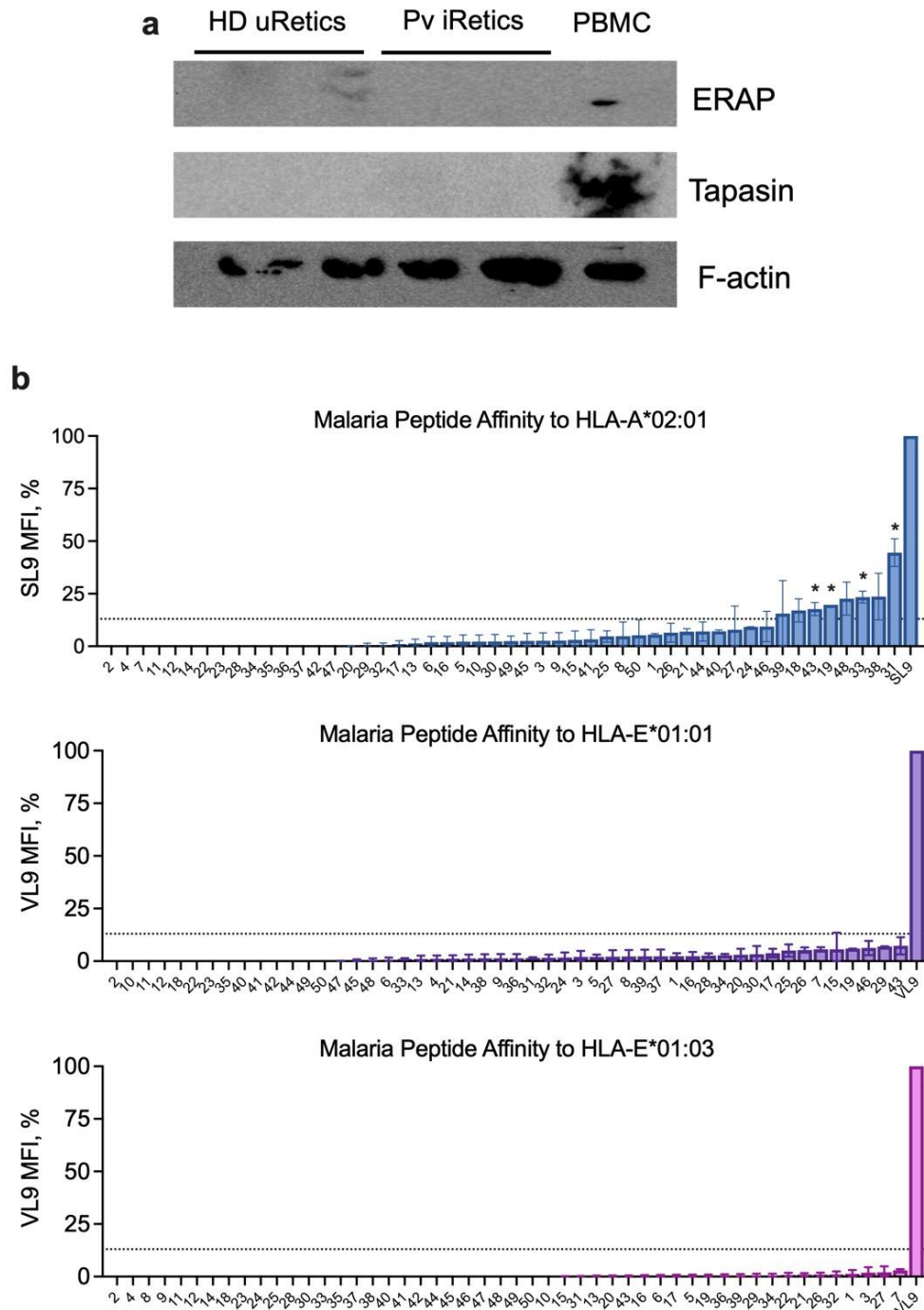




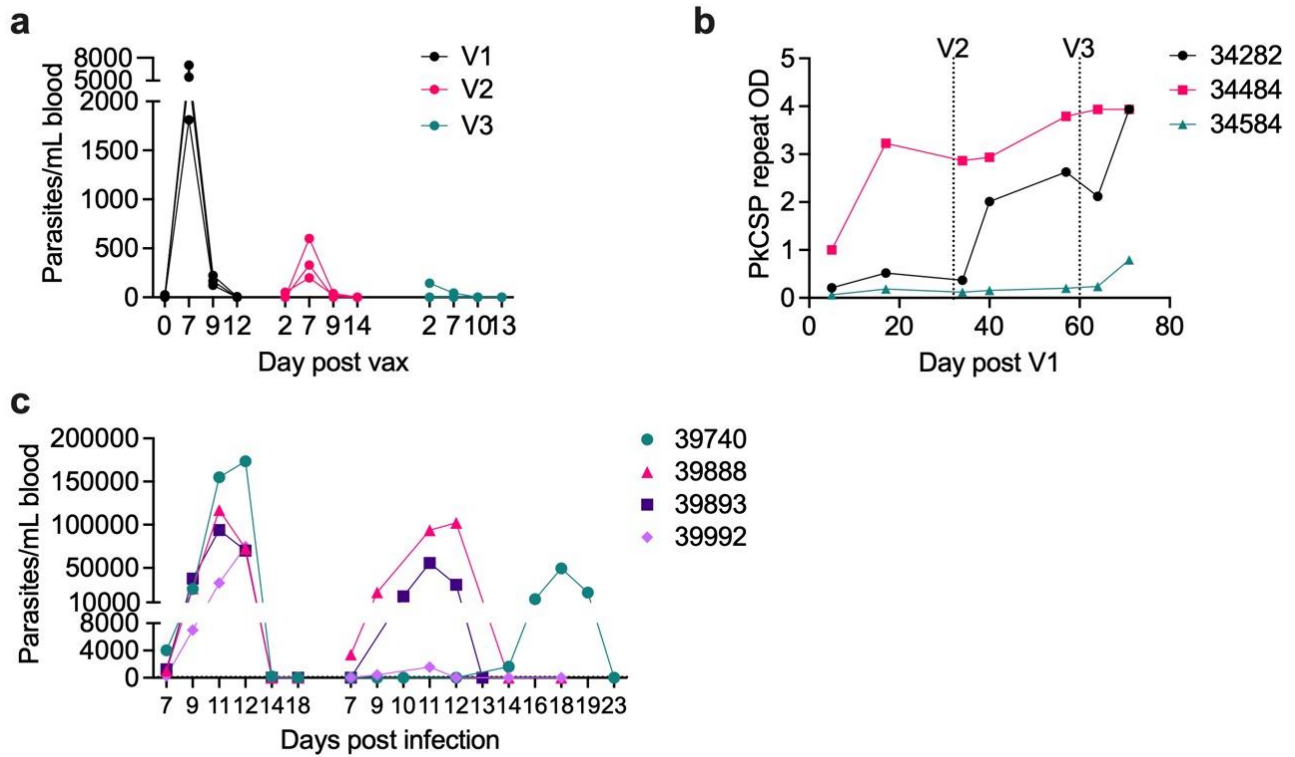
Extended data figure 3 | Peptide validation by ex vivo ELISpot assay. Selected peptides were tested using PBMC isolated from patients infected with *Plasmodium vivax* (*Pv*) (n = 24), *P. falciparum* (*Pf*) (n=7) and healthy donors from endemic (n = 15) and non-endemic (n = 6) regions for malaria. Cells were stimulated with **a**, 40S ribosomal protein peptides, **b**, 60S ribosomal protein peptides, **c**, ETRAMP peptides, **d**, Histone peptides, and **e**, other peptides. Each symbol represents one individual. Green squares are *Pv* patients, purple triangles are *Pf* patients, pink circles are endemic healthy donors, and white circles are non-endemic healthy donor. IFN γ production was measured by spot counting and the results are expressed as spot-forming cells (SFC) by 1×10^6 PBMC. Positive peptides were considered as responses that induced ≥ 30 spots in each patient- above the red dashed line.



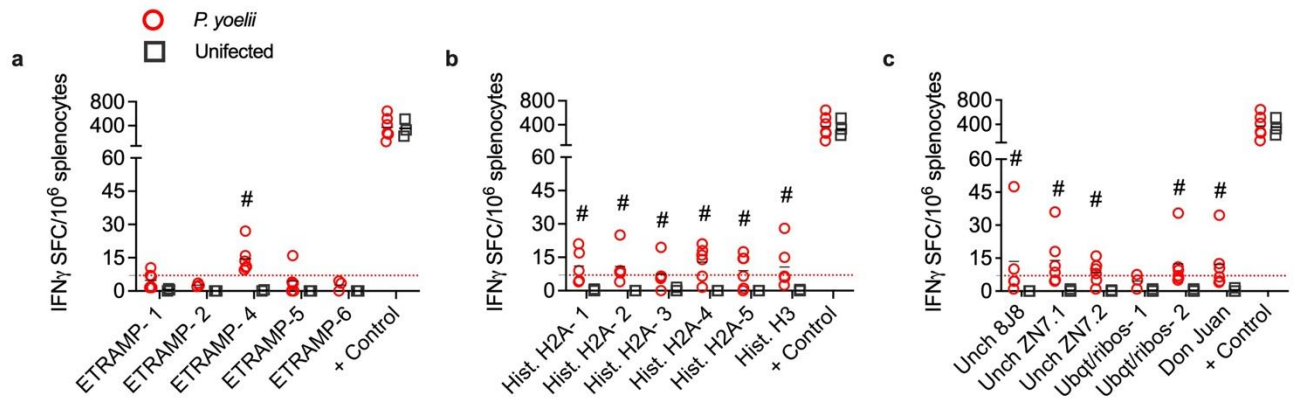
Extended data figure 4 | *Plasmodium vivax* peptides are presented by different HLAs. a, HLA-ABC genotyping typing of the 7 patients used in the immunopeptidome analysis. **b-d**, NetMHC bind analysis of peptides identified in each sample to their donor HLA type. **b**, Number of peptides identified as HLA-ABC binders for each sample. HLA-ABC genotyping of the donor of samples used in the immunopeptidomics assays. Alleles represented with same color mean that they were shared across different samples.



Extended data figure 5 | Tapasin and ERAP1 expression on reticulocytes. Western blot of cell lysates from two healthy donors (HD): uninfected Retics (uRetics) and PBMC; and three *Plasmodium vivax*-infected reticulocytes (Pv iRetics) from three patients were probed for the antigen processing protein Tapasin and Endoplasmic reticulum aminopeptidase 1 (ERAP1). PBMC was used as positive control. F-actin was loaded as endogen protein expression control. *In vitro* binding assay of all selected peptides pulsed in an HLA-A*02:01, HLA-E*01:01, and HLA-E*01:03 cell line, including control peptides. Those peptides marked with asterisk were considered HLA-A: 0201 binders by *in silico* affinity predictions.



Extended data figure 6 | Parasitemia during *Plasmodium knowlesi* (*Pk*) and *P. cynomolgi* (*Pcy*) infection of non-human primates. a, Parasitemia of *Pk* as measured by qRT-PCR following each immunization. b, Antibodies to *Pk* CSP repeat region throughout immunization as measured by single dilution ELISA OD value. c, Parasitemia of *Pcy* as measured by Giemsa-stained thin smears.



Extended data figure 7 | Peptide validation by *ex vivo* IFN γ ELISpot in mouse model. Immunogenicity in *Plasmodium yoelii* (*Py*) infected mice. Mice were intraperitoneally injected with 10^5 *Py* iRBC. At 12 days post-infection (dpi), mouse splenocytes were isolated and incubated with selected peptides. Cells were stimulated with **a**, ETRAMP peptides, **b**, Histone peptides and **c**, Miscellaneous peptides. Each symbol represents one individual. Red circles are *Py* infected mice (n=6), and black ones are non-infected mice (n=3). IFN γ production was measured by spot counting, and the results are expressed as SFC by 1×10^6 splenocytes. Positive peptides (#) were considered as responses that induced ≥ 10 spots in each mouse – above the red dashed line – and $P \geq 0.05$ by multiple unpaired student *t* tests.

MAIN FIGURES

- Figure 1 | Peptide identification and immunogenicity validation.
Figure 2 | Classical and non-classical HLA peptide binders
Figure 3 | Novel *Plasmodium* antigens are recognized by T cells elicited during *P. knowlesi* and *P. cynomolgi* infection of non-human primates.
Figure 4 | Antigen validation in the *P. yoelii* malaria experimental model.

EXTENDED DATA TABLES

- Extended data table 1 | *Plasmodium falciparum* homology
Extended data table 2 | Primers used for *Plasmodium falciparum* liver stage expression

EXTENDED DATA FIGURES

- Extended data figure 1 | Antigen gene expression in humanized mouse livers infected with *P. falciparum*.
Extended data figure 2 | Spectral match validation
Extended data figure 3 | Peptide validation by *ex vivo* ELISpot assay.
Extended data figure 4 | *Pv* peptides are presented by different HLAs
Extended data figure 5 | Tapasin and ERAP1 expression on reticulocytes
Extended data figure 6 | Parasitemia during *P. knowlesi* and *P. cynomolgi* infection of non-human primates.
Extended data figure 7 | Peptide validation by *ex vivo* IFN γ ELISpot in mouse model.

SUPPLEMENTARY INFORMATION (SI)

- Supplementary information 1 | Clinical information
Supplementary information 2 | Identified peptides
Supplementary information 3 | *Plasmodium vivax* orthologous gene expression in parasite life cycle.
Supplementary information 4 | Selected peptides used for *in vitro* stimulation
Supplementary information 5 | HLA binding