Supplementary Material

The effect of thermal processing on the behaviour of peanut allergen peptide targets used in multiple reaction monitoring mass spectrometry experiments

R. L. Sayers,[a] P.E. Johnson,[a] J.T. Marsh,[a] P. Barran,[b] H. Brown[c] and E. N. C. Mills[a]

Supplementary Section I: Materials and methods

Table S-1 MS tune parameters for untargeted analysis on Orbitrap Elite (Thermo Fisher Scientific, Waltham, MA) mass spectrometer

Tune File Values
Source Type: NSI
Capillary Temp (C): 350.00
APCI Vaporizer Temp (C): 0.00
Sheath Gas Flow (): 0.00
Aux Gas Flow (): 0.00
Sweep Gas Flow (): 0.00
Injection Waveforms: On
Ion Trap MSn AGC Target: 10000.00
FTMS Injection Waveforms: On
FTMS Full AGC Target: 1000000.00

POSITIVE POLARITY
Source Voltage (kV): 1.70
Source Current (uA): 100.00
S-Lens RF Level (%): 60.00
Skimmer Offset (V): 0.00
Multipole RF Amplifier (Vp-p): 800.00
Multipole 00 Offset (V): -2.00
Lens 0 Voltage (V): -3.00
Multipole 0 Offset (V): -9.00
Lens 1 Voltage (V): -15.00
Gate Lens Offset (V): -90.00
Multipole 1 Offset (V): -20.00
Front Lens (V): -9.50
Ion Trap MSn Micro Scans: 1
Ion Trap MSn Max Ion Time (ms): 100.00
FTMS Full Micro Scans: 1
FTMS Full Max Ion Time (ms): 10.00

Scan Event Details:
1: FTMS + p norm o(350.0-1500.0)
CV = 0.0V
2: ITMS + c norm Dep Rapid MS/MS Most intense ion from (1)
Activation Type: CID
Min. Signal Required: 500.0
Isolation Width: 2.00
Normalized Coll. Energy: 35.0
Default Charge State: 2
Activation Q: 0.250
Activation Time: 10.000

Table S-2 MRM transitions and collision energies for heavy and light target peptides generated by Skyline. Parent and daughter ion m/z are monoisotopic.

Target peptide ID			y ion	Collision Energy (z)	Un-labelled		Labelled	
					Parent Ion (m/z)	Daughter Ion (m/z)	Parent Ion (m/z)	Daughter Ion (m/z)
Arah1	(P43237)	329-342	y9	26.5	786.88	989.43	791.88	999.44
			y8	26.5	786.88	875.39	791.88	885.39
			y7	26.5	786.88	804.35	791.88	814.36
Arah1	(P43237)	555-577	y9	23.6	688.84	930.45	692.85	938.47
			y8	23.6	688.84	833.40	692.85	841.41
			y7	23.6	688.84	776.38	692.85	784.39
Arah3	(Q647H4)	103-115	y10	25.9	767.84	1278.55	772.84	1288.56
			y7	25.9	767.84	923.42	772.84	933.42
			y6	25.9	767.84	809.37	772.84	819.38
Arah3	(Q647H4)	147-155	y9	23.8	695.35	977.51	699.36	985.52
			y8	23.8	695.35	814.44	699.36	822.46
			y7	23.8	695.35	700.40	699.36	708.41
Arah2	(Q6PSU2)	25-41	y9	28.8	863.85	1163.54	868.85	1173.55
			y8	28.8	863.85	1050.46	868.85	1060.47
			y7	28.8	863.85	936.42	868.85	946.43
Arah2	(Q6PSU2)	372-384	y7	19.2	543.28	858.43	548.28	868.43
			y6	19.2	543.28	761.37	548.28	771.38
			y5	19.2	543.28	633.31	548.28	643.32
Arah6	(Q647G9)	136-144	y8	17.6	489.72	818.40	494.72	828.41
			y7	17.6	489.72	703.37	494.72	713.38
			y6	17.6	489.72	590.29	494.72	600.30
Arah7	(B4X1D4)	143-151	y7	19.5	553.26	878.39	558.27	888.40
			y6	19.5	553.26	781.34	558.27	791.35
			y5	19.5	553.26	653.28	558.27	663.29

Equation S-1 Calibration plot method for calculation of limits of detection (LOD) and quantification (LOQ).

This method uses the standard error of the y-estimate for the least-squares regression line is used in place of the standard deviation.

$$LOD = 3\ x\ sy|x/slope$$

$$LOQ = 3\ x\ LOD$$

The limit of detection is equal to the signal corresponding to the lowest point in the linear regression plus 3 sy|x/slope and the limit of quantification is equal to 3 times the LOD. [1, 2]

1. D. Anderson, *Clinical chemistry*, 1989, **35**, 2152 - 2153.

2. D. R. Mani, S. Abbatiello and S. Carr, *BMC Bioinformatics*, 2012, **13**, S9.

Supplementary Section II: Results and discussion

a) Effects of thermal processing on protein solubility

Table S-3 Effect of thermal processing conditions on extractability of peanut proteins using five different buffer conditions.

Total protein was determined by Kjeldahl analysis and extracted protein determined by 2D Quant-Kit™ assay. The extraction efficiency was calculated as the ratio of extracted to total protein. Results are presented as the mean of triplicate analyses of each extract. SD – standard deviation.

Sample	Total Protein (% w/w)		Buffer 1: 0.01 M PBS (0.002 M NaH_2PO_4, 0.008 M Na_2HPO_4, 0.85% (w/v) NaCl), pH 7.4				Buffer 2: 50 mM Tris-HCl, pH 8.8				Buffer 3: 50 mM Tris-HCl, 50 mM DTT, pH 8.8				Buffer 4: 50 mM Tris-HCl, 50 mM DTT, 0.04% (w/v) RapiGest, pH 8.8				Buffer 5: 7 M Urea, 2 M Thiourea, 2% (w/v) CHAPS, 50 mM DTT, pH 8.8			
			[Protein] mg/mL		Extraction Efficiency		[Protein] mg/mL		Extraction Efficiency		[Protein] mg/mL		Extraction Efficiency		[Protein] mg/mL		Extraction Efficiency		[Protein] mg/mL		Extraction Efficiency	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Raw	44.64	0.66	3.01	0.26	0.34	0.03	9.43	1.02	1.05	0.11	8.38	2.08	0.95	0.24	12.26	0.47	1.22	0.09	10.41	0.31	1.16	0.04
Raw-Boiled 30 min	39.12	0.34	1.06	0.25	0.14	0.03	3.36	0.32	0.43	0.04	2.59	0.31	0.34	0.04	6.19	0.95	0.81	0.12	8.40	0.34	1.09	0.04
Raw-Boiled 1 hr	39.56	0.89	0.73	0.14	0.09	0.02	2.09	0.13	0.27	0.01	2.26	0.33	0.29	0.04	4.94	1.11	0.64	0.16	5.22	1.76	0.67	0.23
Raw-Boiled 2 hr	35.91	3.48	0.68	0.25	0.10	0.04	1.36	0.45	0.19	0.06	1.56	0.20	0.22	0.03	3.80	0.69	0.54	0.11	9.71	1.34	1.37	0.18
Raw-Boiled 4 hr	27.20	0.19	0.36	0.03	0.07	0.01	1.29	0.58	0.24	0.11	1.20	0.15	0.22	0.03	2.31	0.25	0.43	0.04	8.34	0.65	1.56	0.13
Raw-Boiled 6 hr	25.57	0.14	0.50	0.01	0.10	0.00	1.05	0.18	0.21	0.03	0.89	0.07	0.18	0.02	1.97	0.27	0.39	0.06	7.14	1.17	1.42	0.23
Raw-Roasted 150°C	38.07	0.09	1.01	0.44	0.13	0.06	3.75	0.07	0.50	0.01	3.00	0.56	0.40	0.07	7.87	0.42	1.07	0.06	10.08	1.90	1.35	0.25
Raw-Roasted 160°C	49.03	0.11	0.79	0.13	0.08	0.01	3.75	0.07	0.39	0.01	3.21	0.48	0.33	0.05	5.56	0.24	0.58	0.03	8.67	2.27	0.90	0.24
Raw-Roasted 170°C	47.59	0.13	0.34	0.05	0.04	0.01	2.51	0.40	0.27	0.04	2.04	0.96	0.22	0.10	5.29	0.78	0.57	0.09	6.24	1.31	0.66	0.14
Raw-Roasted 180°C	50.75	0.34	0.24	0.04	0.02	0.00	1.27	0.32	0.13	0.03	1.20	0.62	0.12	0.06	3.60	0.33	0.36	0.04	5.07	0.40	0.50	0.04
Blanched	42.84	2.32	2.14	0.15	0.25	0.02	7.01	0.25	0.81	0.02	4.01	0.43	0.47	0.05	9.02	0.66	0.97	0.12	8.39	0.88	0.99	0.04
Fried 155°C 400 sec	50.50	1.61	0.77	0.22	0.08	0.02	5.10	0.55	0.50	0.05	4.43	0.70	0.45	0.07	11.68	1.21	1.10	0.08	9.18	2.24	0.91	0.22
Commercial flour	54.88	0.60	1.43	0.20	0.01	0.01	4.81	0.94	0.43	0.08	7.21	1.87	0.66	0.18	8.64	1.12	0.69	0.08	10.37	0.67	0.96	0.05

b) Peptide target selection and verification

Figure S-1 Informatics pipeline for identification of peptide targets for MRM experiments.

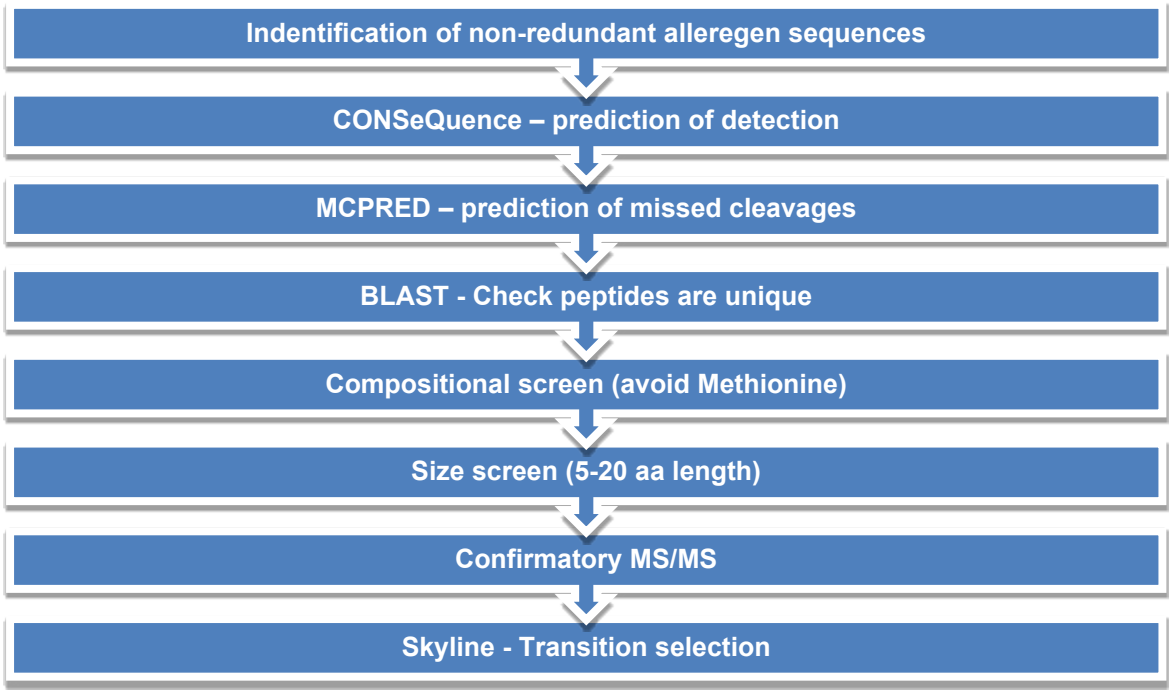
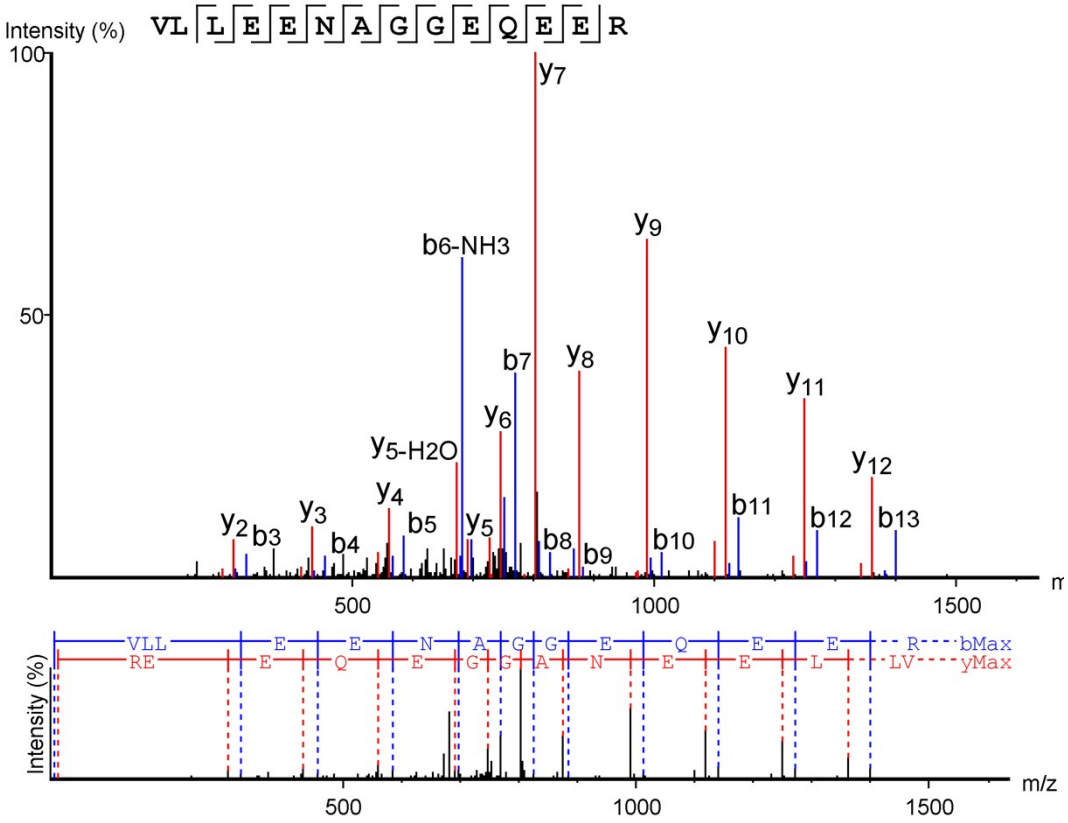


Table S-4 Peanut allergen peptides identified in two samples of raw peanuts using untargeted MS/MS analysis undertaken with data-dependent acquisition. Protein score – PEAKS™ provides a statistical evaluation for each peptide-spectrum match. Isoforms used for conversion of peptide to protein have been marked (*).

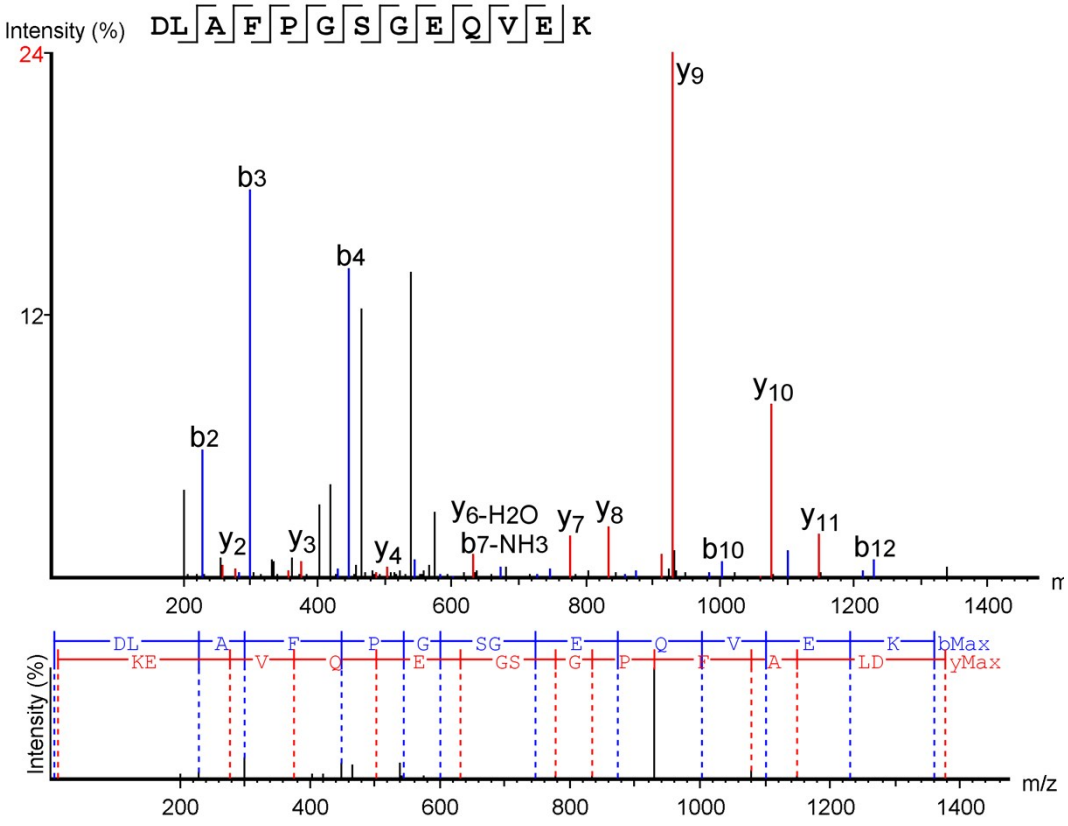
Allergen Name	Accession Number	Calculated Mass	Raw 1				Raw 2			
			Protein Score (-10lgP)	Coverage (%)	No of Peptides	No of unique peptides	Protein Score (-10lgP)	Coverage (%)	No of Peptides	No of unique peptides
Ara h 1	P43237*, B3IXL2, Q6PSU3	70283	277.42	71	94	11	199.53	73	93	10
Ara h 1	P43238, N1NG13	71345	275.42	71	93	6	195.87	74	97	0
Ara h 1	E5G076	70788	263.44	59	76	0	188.92	61	73	0
Ara h 3	Q647H2	54569	185.68	43	21	9	118.5	45	20	9
Ara h 3	A1DZF0	60375	311.81	88	156	32	228.71	87	182	32
Ara h 3	B5TYU1	60624	298.64	85	151	3	221.56	85	181	4
Ara h 3	Q0GM57, E5G077, Q6IWG5	58263	247.13	79	99	68	144.83	53	39	15
Ara h 3	Q5I6T2	60736	300.97	88	155	8	220.31	85	169	4
Ara h 3	Q647H3	61532	296.05	85	146	2	221.31	85	173	2
Ara h 3	Q647H4*	61506	295.13	88	145	0	220.59	85	173	0
Ara h 3	Q6T2T4	61499	290.12	88	137	0	213.69	85	165	0
Ara h 3	Q8LKN1, Q6VT83	61738	94.54	25	4	1	220.25	84	171	0
Ara h 3	Q9FZ11	60449	300.39	84	149	7	221.42	83	170	8
Ara h 3	Q9SQH7	61011	278.98	66	106	0	210	64	126	0
Ara h 3	O82580	58350	262.98	71	102	9	187.53	71	105	4
Ara h 2	Q6PSU2*	20114	227.32	81	62	0	160.06	85	57	0
Ara h 2	Q6PSU2-2	18700	226.86	79	60	0	161.69	84	56	0
Ara h 2	Q6PSU2-3	19795	226.26	82	61	0	159.71	85	56	0
Ara h 2	Q6PSU2-4	18380	225.77	81	59	0	161.33	84	55	0
Ara h 6	Q647G9*, A5Z1R0	16920	190.12	70	27	27	142.92	74	27	27
Ara h 7	B4XID4*	19338	139.1	52	18	7	103.7	52	14	5
Ara h 7	Q647G8, Q9SQH1	18455	133.15	52	15	4	109.2	51	17	8

Figure S-2 Identification of target peptides from peanut allergens present in raw peanut extracts by *de novo* sequencing using untargeted MS analysis.

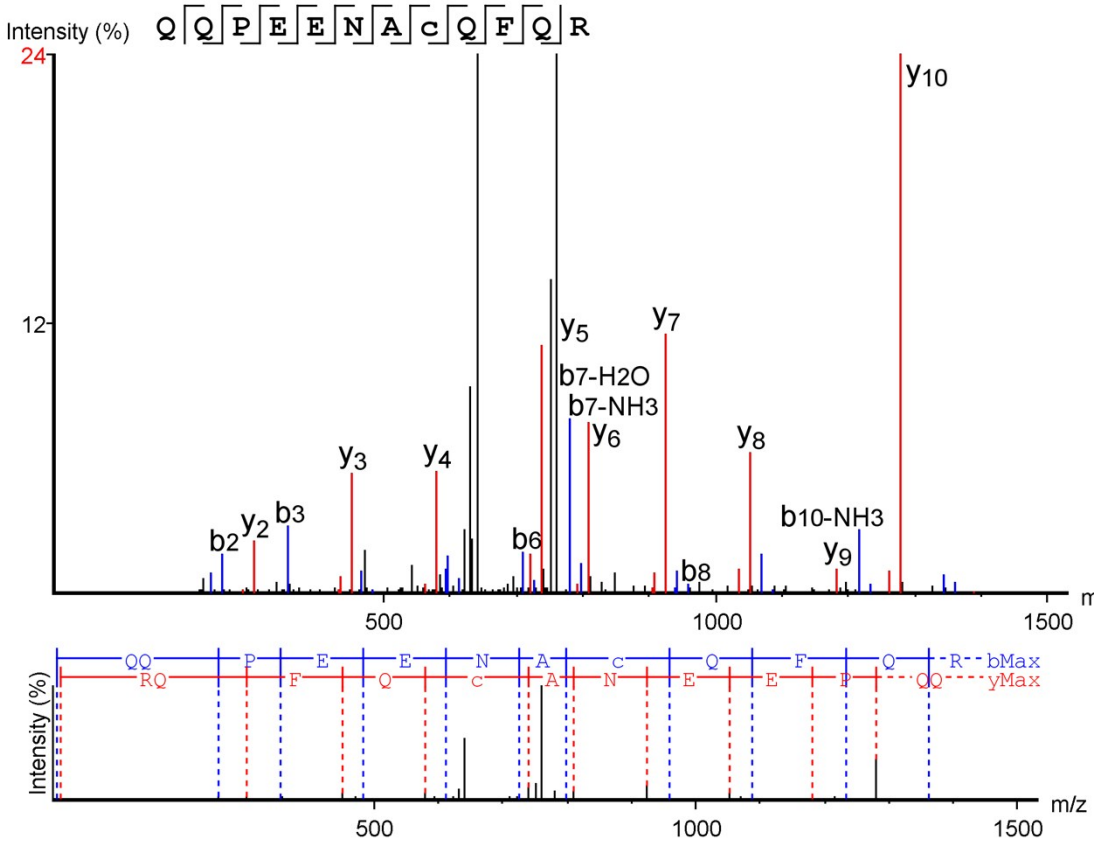
(i) Arah1(P43237)[329-342]



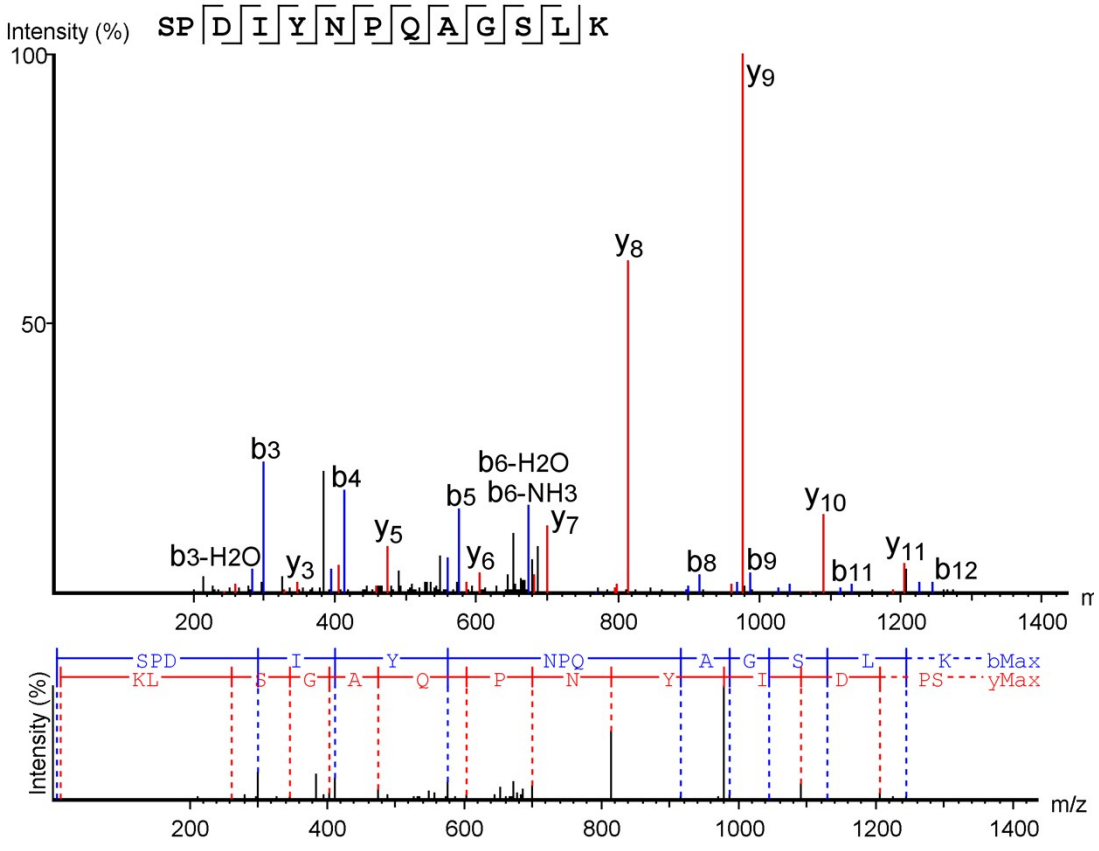
(ii) Arah1(P43237)[555-577]



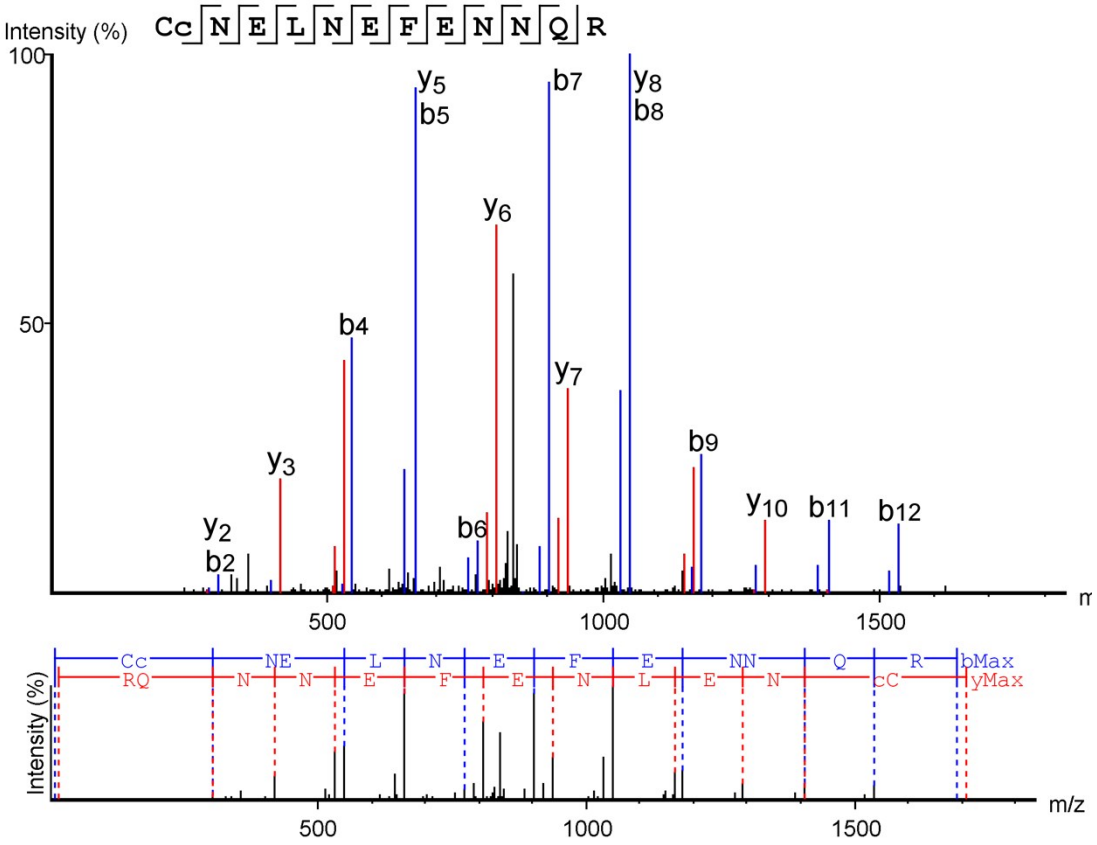
(iii) Arah3(Q647H4)$^{25\text{-}41}$



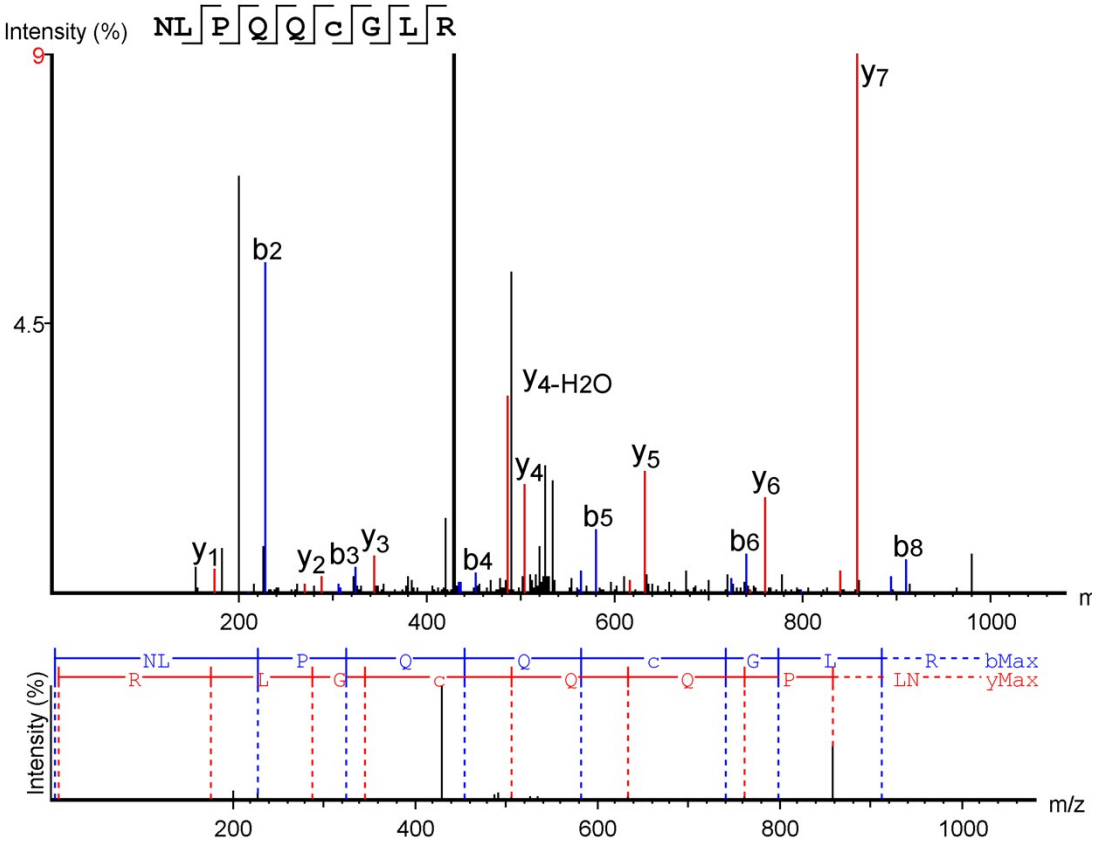
(iv) Arah3(Q647H4)$^{372\text{-}384}$



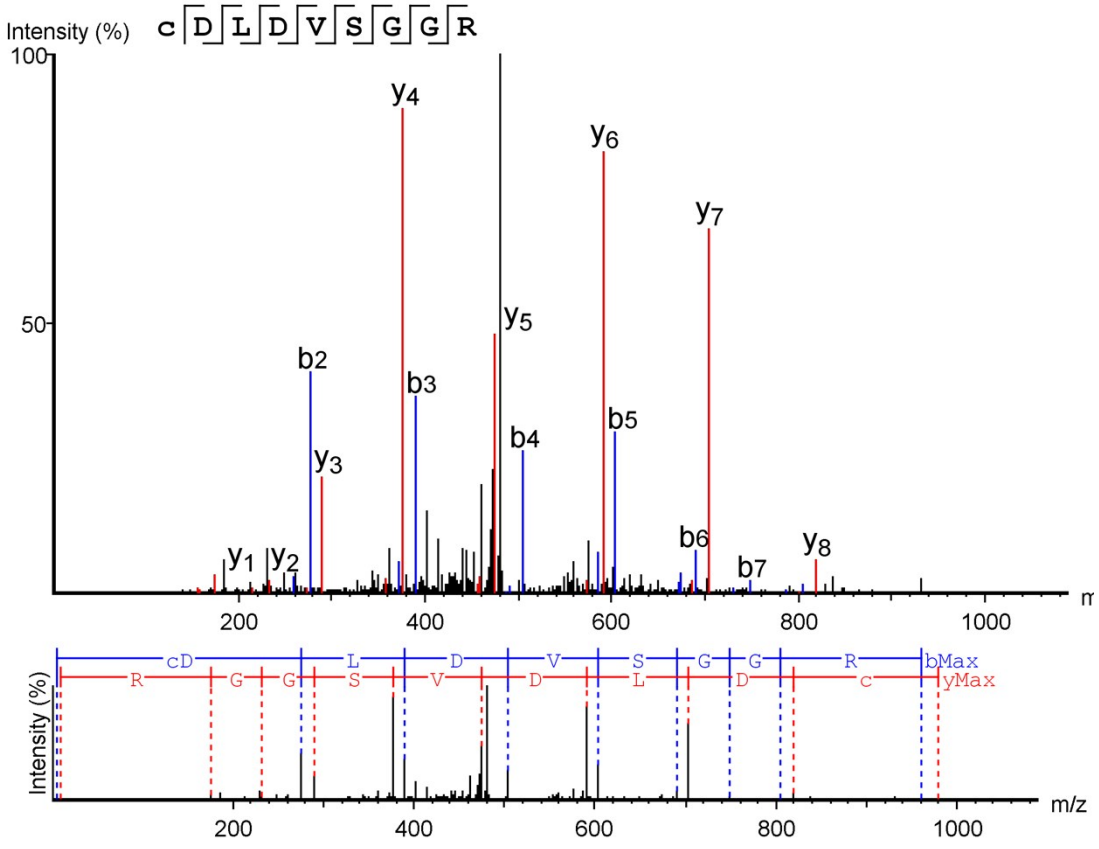
(v) Arah2(Q6PSU2)[103-115]



(vi) Arah2(Q6PSU2)[147-155]



(vii) Arah6(Q647G9)[136-144]



(viii) Arah7(B4XID4)[143-151]

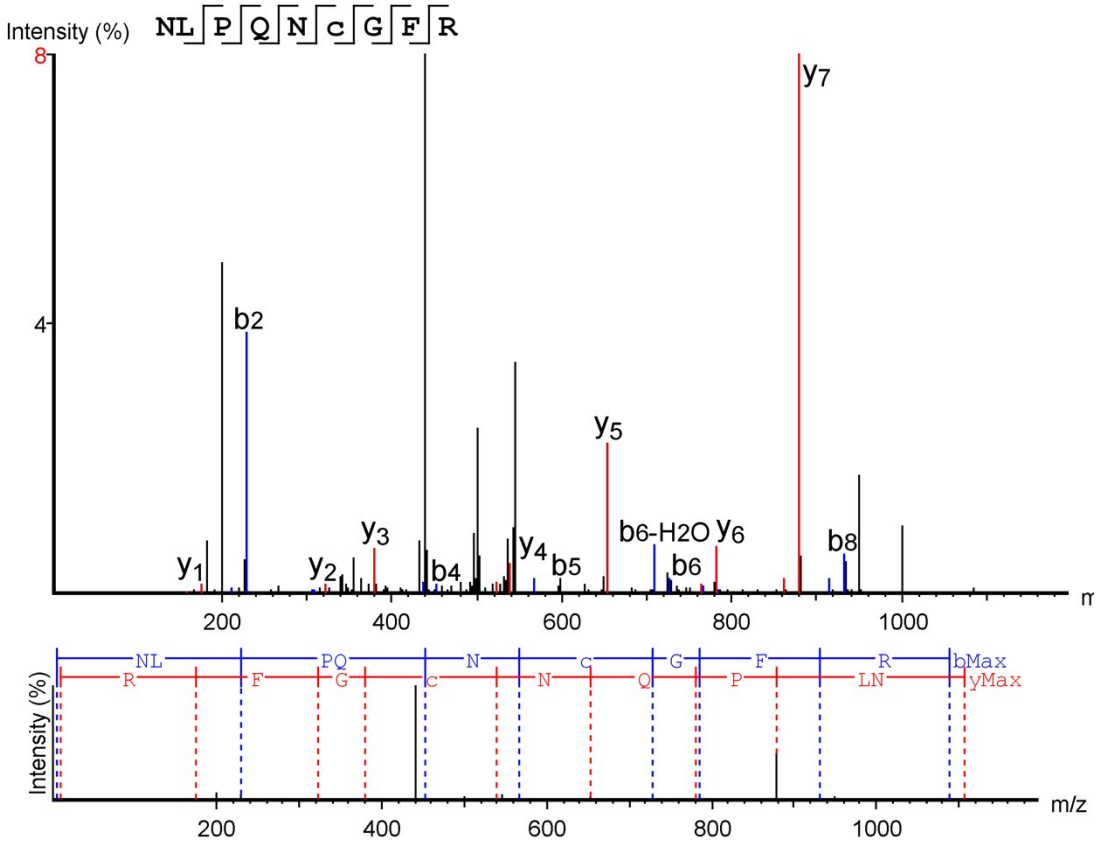


Figure S-3 Peptide targets for Ara h 3 were mapped on to the protein sequence (A) and its 3D structure (B). (A) Ara h 3 protein sequence for isoform Q647H4 (Uniprot) was used and is shown with the signal peptide sequence, N-terminus underlined, target peptides Arah3(Q647H4)$^{25\text{-}41}$ in blue and Arah3(Q647H4)$^{372\text{-}384}$ in red; (B) crystal structure (DOI:10.2210/pdb3c3v/pdb)Ara h 3 monomer, coloured by subunit, acidic in yellow and basic in green and target peptides Arah3(Q647H4)$^{25\text{-}41}$ in blue and Arah3(Q647H4)$^{372\text{-}384}$ in red. Cartoons show van der Waal's space-filling at 80% transparency and cartoon structures modelled using The PyMOL Molecular Graphics System, Version 1.5.0.4 Schrödinger, LLC.

(A) Ara h 3 UniProtKB - Q647H4 (Q647H4_ARAHY)

```
        10         20         30         40         50
MGKLLALSVC FCFLVLGASS ISFRQQPEEN ACQFQRLNAQ RPDNRIESEG
        60         70         80         90        100
GYIETWNPNN QEFECAGVAL SRLVLRRNAL RRPFYSNAPQ EIFIQQGRGY
       110        120        130        140        150
FGLIFPGCPS TYEEPAQQGR RHQSQRPPRR FQGQDQSQQQ QDSHQKVHRF
       160        170        180        190        200
DEGDLIAVPT GVAFWMYNDH DTDVVAVSLT DTNNNDNQLD QFPRRFNLAG
       210        220        230        240        250
NHEQEFLRYQ QQSRRRSLPY SPYSPQTQPK QEDREFSPRG QHGRRERAGQ
       260        270        280        290        300
EQENEGGNIF SGFTPEFLAQ AFQVDDRQIL QNLRGENESD EQGAIVTVRG
       310        320        330        340        350
GLRILSPDRK RRQQYERPDE EEEYDEDEYE YDEEERQHDR RRGRGSRGSG
       360        370        380        390        400
NGIEETICTA SFKKNIGRNR SPDIYNPQAG SLKTANELNL LILRWLGLSA
       410        420        430        440        450
EYGNLYRNAL FVPHYNTNAH SIIYALRGRA HVQVVDSNGD RVFDEELQEG
       460        470        480        490        500
HVLVVPQNFA VAGKSQSENF EYVAFKTDSR PSIANLAGEN SFIDNLPEEV
       510        520        530
VANSYGLPRE QARQLKNNNP FKFFVPPSEQ SLRAVA
```

(B) 3D structure crystal structure (DOI:10.2210/pdb3c3v/pdb) of Ara h 3 monomer showing target peptide locations

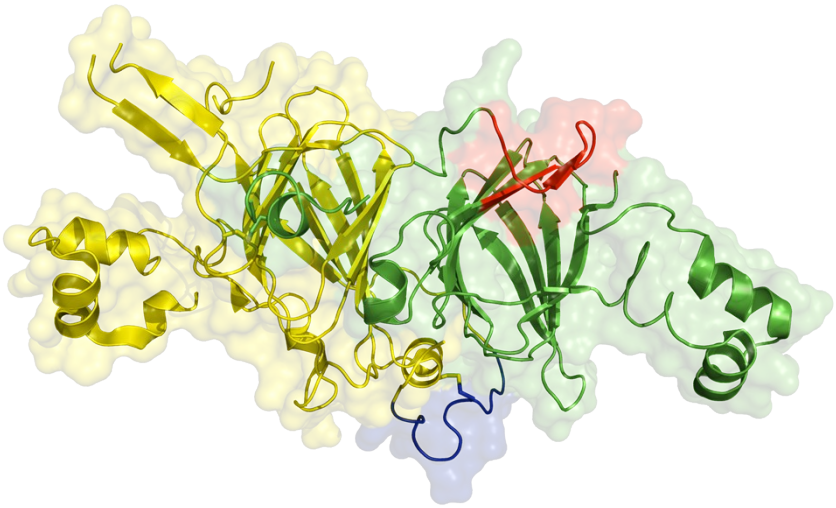


Figure S-**4** Peptide targets for Prolamins: Ara h 2, 6 and 7 were mapped on to their respective protein sequences (A-C) and for Ara h 6 only its 3D structure (Bii).
(A) Ara h 2 protein sequence for isoform Q6PSU2 (UniProt)is shown with the signal peptide sequence in italics and target peptides **Arah2(Q6PSU2)**$^{103\text{-}115}$ in blue and **Arah2(Q6PSU2)**$^{147\text{-}155}$ in red; (B) (i) Ara h 6 protein sequence for isoform Q647G9 (UniProt) is shown with the signal peptide sequence in italics and target peptide **Arah6(Q647G9)**$^{136\text{-}144}$ in red with associated nmr derived structure (DOI:10.2210/pdb1w2q/pdb) showing van der Waal's space-filling at 80% transparency and (ii) cartoon structures modelled using The PyMOL Molecular Graphics System, Version 1.5.0.4 Schrödinger, LLC.; and (C) Ara h 7 protein sequence for isoform B4XID4 (UniProt) is shown with the signal peptide sequence in italics and target peptide **Arah7B4XID4**$^{143\text{-}151}$ in red.

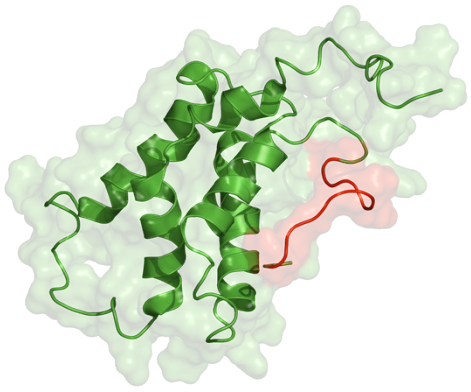
(A) Ara h 2 UniProtKB - Q6PSU2 (CONG7_ARAHY)

```
        10          20          30          40          50
MAKLTILVAL ALFLLAAHAS ARQQWELQGD RRCQSQLERA NLRPCEQHLM
        60          70          80          90         100
QKIQRDEDSY GRDPYSPSQD PYSPSQDPDR RDPYSPSPYD RRGAGSSQHQ
       110         120         130         140         150
ERCCNELNEF ENNQRCMCEA LQQIMENQSD RLQGRQQEQQ FKRELRNLPQ
       160         170
QCGLRAPQRC DLEVESGGRD RY
```

(B) (i) Ara h 6 UniProtKB - Q647G9 (CONG_ARAHY)

```
        10          20          30          40          50
MAKSTILVAL LALVLVAHAS AMRRERGRQG DSSSCERQVD RVNLKPCEQH
        60          70          80          90         100
IMQRIMGEQE QYDSYDIRST RSSDQQQRCC DELNEMENTQ RCMCEALQQI
       110         120         130         140
MENQCDRLQD RQMVQQFKRE LMNLPQQCNF RAPQRCDLDV SGGRC
```

(ii) 3D structure derived by nmr (DOI:10.2210/pdb1w2q/pdb) of Ara h 7 protein with target peptides highlighted



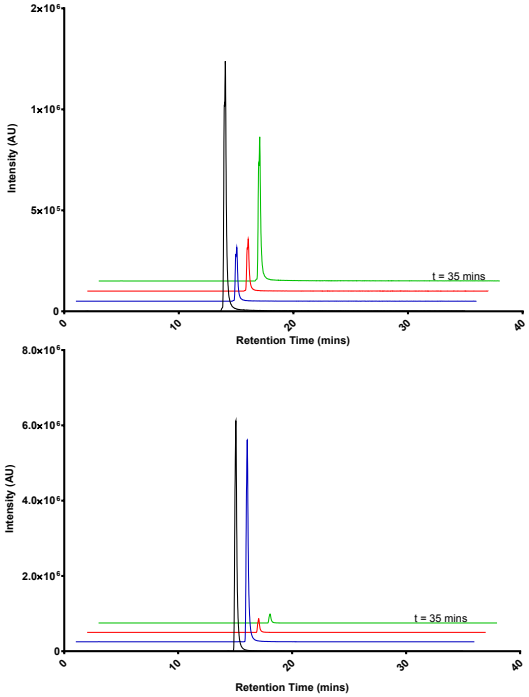
(C) Ara h 7 UniProtKB - B4XID4 (B4XID4_ARAHY)

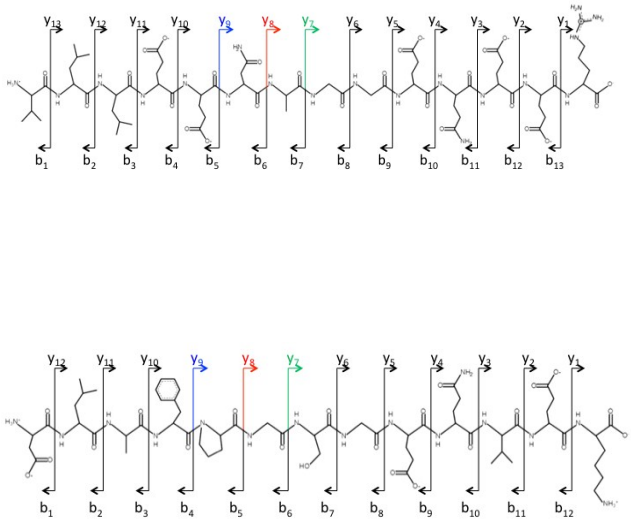
```
        10          20          30          40          50
MVKLSILVAL LGALLVVASA TRWDPDRGSR GSRWDAPSRG DDQCQRQLQR
        60          70          80          90         100
ANLRPCEEHI RQRVEKEQEQ EQDEYPYIQR GSRGQRPGES DEDQEQRCCN
       110         120         130         140         150
ELNRFQNNQR CMCQALQQIL QNQSFRFQQD RSQLHQMERE LRNLPQNCGF
       160
RSPSRCDLSS RTPY
```

c) Development of multiple-reaction monitoring (MRM) experiments for peanut allergens

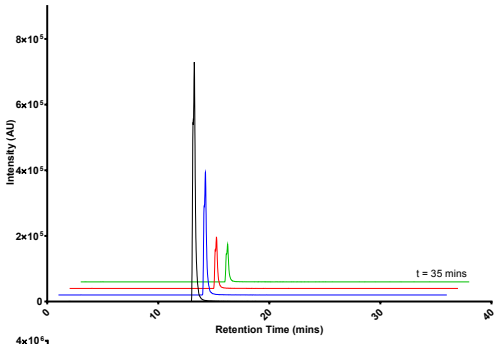
Figure S-5 Ion chromatograms of isotopically labelled peptide targets using equivalent femtomoles loaded onto the column showing the contribution of each transitions ion intensity to the total with inserted peptide structures showing fragmentation patterns for b and y ions;

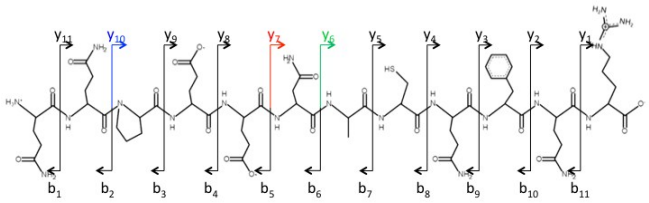
I. Ara h 1; (i) Arah1(P43237)$^{329-342}$;
y9 (blue), y8 (red), y7 (green), with total in black, and (ii) Arah1(P43237)$^{555-577}$; y9 (blue), y8 (red), y7 (green), y6 (pink) with total in black.

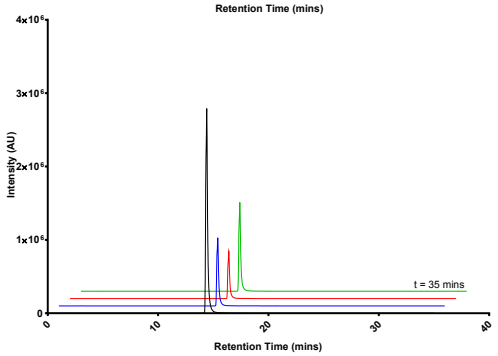


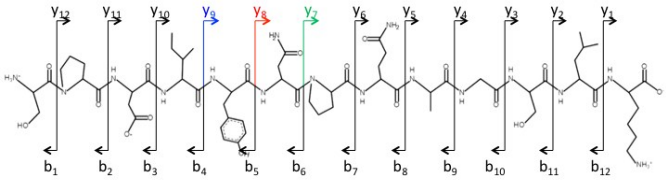


II. Ara h 3 (i) Arah3(Q647H4)$^{25-41}$;
y10 (blue), y7 (red), y6 (green) with total in black, and (ii) Arah3(Q647H4)$^{372-384}$; y9 (blue), y8 (red), y7 (green) with total in black.

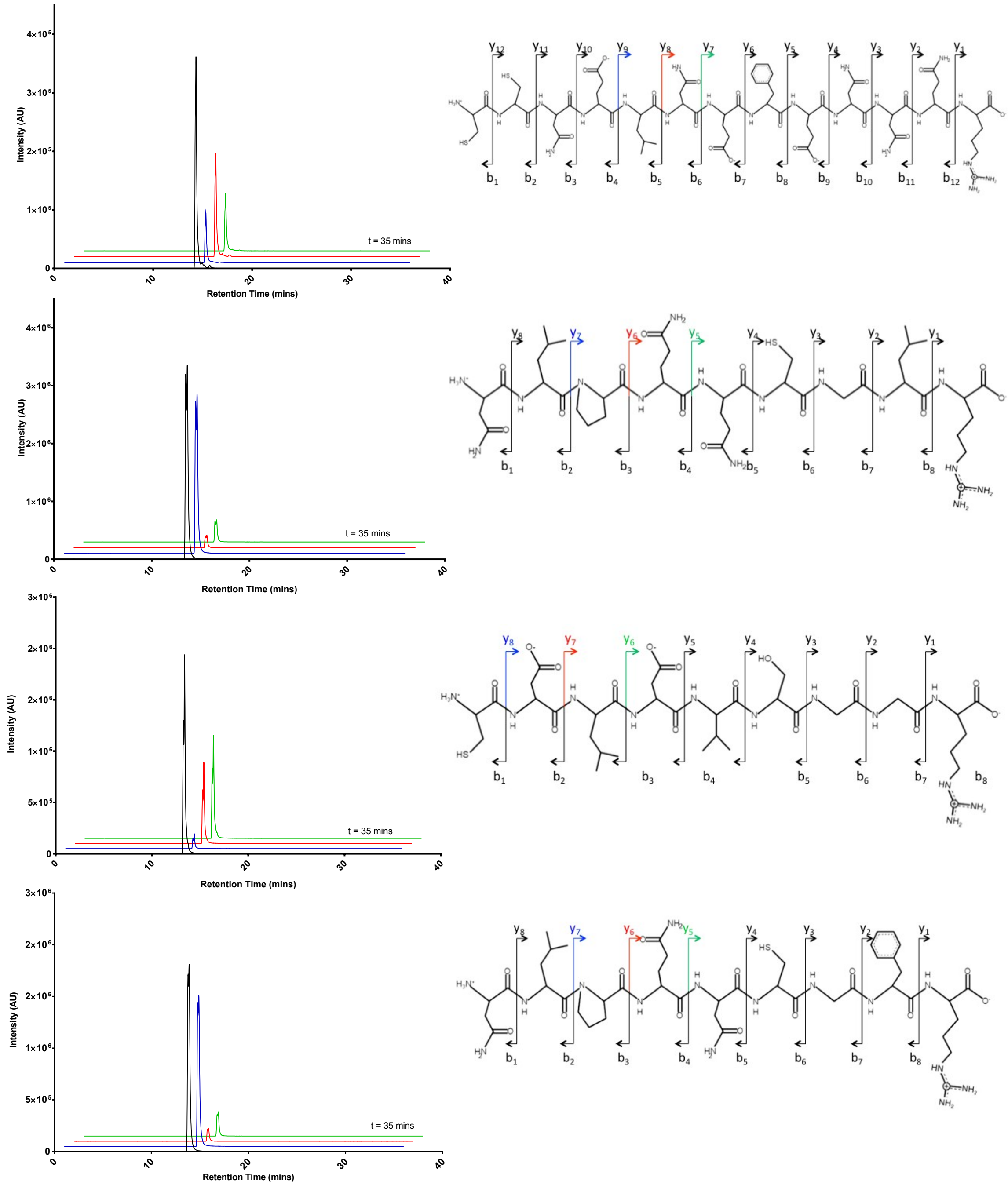








III. Arah2(Q6PSU2)$^{103\text{-}115}$;
y9 (blue), y8 (red), y7 (green), with total in black, (ii) Arah2(Q6PSU2)$^{147\text{-}155}$; y7 (blue), y6 (red), y5 (green), (iii) Arah6(Q647G9)$^{136\text{-}144}$; y8 (blue), y7 (red), y6 (green), and (iv) Arah7(B4XID4)$^{143\text{-}151}$; y7 (blue), y6 (red), y5 (green).



d) Effect of thermal processing on peptide target quantification

Figure S-6 Effect of thermal processing on the level of individual allergens present in peanuts.

The mass of different allergens present, per g of extracted protein was determined by MRM analysis. Samples were extracted in triplicate (1=O, 2=X, 3=Δ) and each extract analysed in triplicate. (A) Ara h 1: Using a Mr of 61.7 kDa for the Ara h 1 subunit these data indicate that the raw peanuts contained between 24.2-30.4 mg (Arah1(P43237)329-342 and 6.8-9.6 mg (Arah1(P43237)555-577 which corresponds to estimates that Ara h 1 comprises 2.4-3.0% of seed protein, (B) Ara h 3: Using a Mr of 59.6 kDa for the Ara h 3 these data indicate that the raw peanuts contained between 17.5-21.5 mg Arah3(Q647H3)25-41 and 82-87.9 mg of Arah3(Q647H3)372-384. This corresponds to estimates that Ara h 3 comprises 8.2-8.8% of seed protein and (C) Ara h 2,6,7: Using a Mr of 17.9 kDa for the Ara h 2, 14.8 kDa for Ara h 6 and 17.4 kDa for Ara h 7 these data indicate that the raw peanuts contained between 11.8-18.2 mg of Arah2(Q6PSU2)103-115, 11.8-16 mg of Arah2(Q6PSU2)147-155, 5.2-6.7 mg of Ara h 6 and 2.7-7.5 mg Ara h 7. This corresponds to estimates that Ara h 2, 6 and 7 collectively comprise 2.0-3.1% of seed protein.

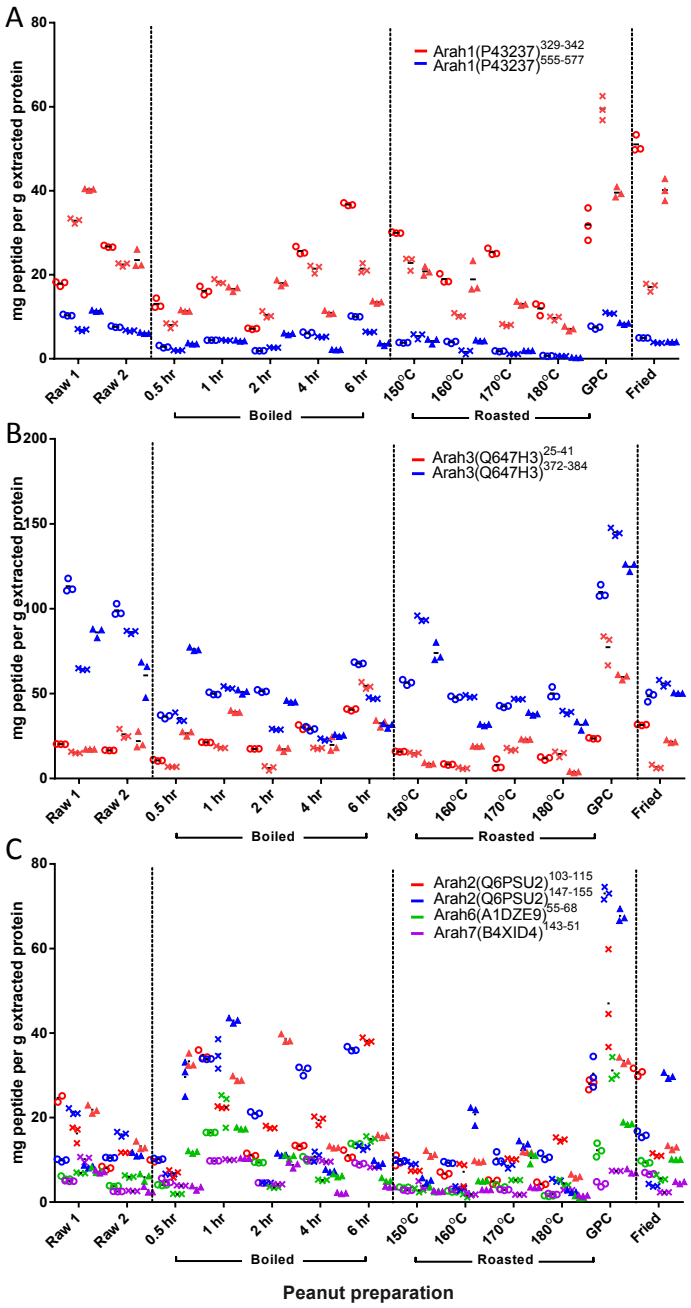


Figure S-7 1D-PAGE analysis of thermally processed peanuts extracts prepared from buffer (5).

Defatted peanut samples were extracted with buffer 5 (7 M urea, 2 M thiourea, 2 % (w/v) CHAPS, 1 % (w/v) DTT, pH 8.8) at 60 ^{o}C for 15 mins in a sonicating water bath with vortexing every 5 mins. Protein was determined by 2D-Quant Assay Kit and 5 µg protein loaded per lane as follows; lane 1 - Raw 1; lane 2 - raw 2 (blanched); lane 3 - boiled 30 mins; lane 4 - boiled 1 hour; lane 5 - boiled 2 hours; lane 6 - boiled 4 hours; lane 7 - boiled 6 hours; lane 8 - roasted 150 ^{o}C; lane 9 - 160 ^{o}C; lane 10 - 170 ^{o}C; lane 11 - 180 ^{o}C; lane 12 - commercial flour and lane 13 – fried. Protein composition was determined by 1D-PAGE (5µg protein per well) on a 4-12% precast gradient gel (Invitrogen, Paisley, UK) using the NuPAGE® gel system under reducing conditions (200V for 40 min) and visualised using a Coomassie-based stain (Simply Blue Safe Stain, Invitrogen, Paisley, UK). The key allergenic proteins have been identified using published motilities (Marsh et al., 2008) are highlighted in red boxes. Relative molecular weights determined using Mark12^{TM} unstained standards (Myosin 200 kDa, β-galactosidase 116.3 kDa, Phosphorylase B 97.3 kDa, BSA 66.3 kDa, Glutamic dehydrogenase 55.4 kDa, Lactate dehydrogenase 36.5 kDa, Carbonic anhydrase 31 kDa, Trypsin inhibitor 21.5 kDa, Lysozyme 14.4 kDa, Aprotinin 6 kDa, Insulin B chain 3.5 kDa and insulin A chain2.5 kDa).