Unravelling Procyanidin effect on Gliadin Digestion and Immunogenicity

Dias, Ricardo[1]; Francisco, Telmo[1]; Pereira, Catarina[1]; Mateus, Nuno[1]; de Freitas, Victor[1] ; Pérez-Gregorio, Rosa[1]

[1]*LAQV-REQUIMTE Departamento de Química e Bioquímica, Faculdade de Ciências da Universidade do Porto, Rua do Campo Alegre 687, 4169-007 Porto, Portugal.*

e-mail: maria.gregorio@fc.up.pt

Supplementary Material

The gastrointestinal digestion of gluten proteins has a major impact on their immunological properties. Besides, the interaction with dietary components such as polyphenols could potentially interfere with the structure of proteins, reduce the number of cleavage sites made available to proteolytic enzymes and change the diversity of peptides that are released to the human gut. This study aims at unravelling the impact of a common food polyphenol, the procyanidin dimer B3, in the:

i) ability to bind to gluten proteins and/or digestive enzymes under in vitro human digestion physiological conditions;
ii) proteolytic activity of trypsin and chymotrypsin;
iii) sequencing of released peptides while mimicking a human in vitro gastrointestinal digestion;
iv) immunogenicity of released peptides.

Supplementary material includes information relative to the overall identification and characterization of proteins from gluten fractions and the identification of proteins remaining in supernatant after the binding selectivity on molecular interaction assay - (SDS-PAGE).

1. **Characterization of proteins from gluten fractions.**
   Proteins from gliadin fractions were characterized as previously reported(Perez-Gregorio, Días, Mateus, & De Freitas, 2017). Furthermore, other physicochemical parameters clearly differentiate the composition among gliadin fractions as is displayed in **Table 1S**.

Gliadin Fraction A						Gliadin Fraction B						Gliadin Fraction C					
Protein Code	Hydropathicity (GRAVY)	Aliphatic index	Major AA	Charge		Protein Code	Hydropathicity (GRAVY)	Aliphatic index	Major AA	Charge		Protein Code	Hydropathicity (GRAVY)	Aliphatic index	Major AA	Charge	
				-	+					-	+					-	+
D2KFH2	-0.485	82.84	(Q) 22.4%	(D + E): 4	(R + K): 5	D2KFH2	-0.485	82.84	(Q) 22.4%	(D + E): 4	(R + K): 5	A0A290XYT7	-1.687	18.32	(Q) 36.9 %	(D + E): 1	(R + K): 3
D2KFG8	-0.345	80.85	(Q) 26.3%	(D + E): 5	(R + K): 4	I3QPI2	-0.233	85.95	(Q) 16.7%	(D + E): 2	(R + K): 4	R9XSW3	-0.679	82.30	(Q) 33.7 %	(D + E): 5	(R + K): 4
Q3LGB3	-0.221	89.87	(Q) 23.4%	(D + E): 2	(R + K): 5	A0A023WGX2	-1.155	47.12	(Q) 35.9 %	(D + E): 3	(R + K): 2	C0KEI2	-1.466	41.9	(Q) 40.5 %	(D + E): 7	(R + K): 8
D2KFH0	-0.208	82.54	(Q) 17.8%	(D + E): 7	(R + K): 8	Q9M4M0	-0.562	82.46	(Q) 25.4 %	(D + E): 5	(R + K): 5	Q5MFQ4	-0.936	61.06	(Q) 36.7 %	(D + E): 6	(R + K): 8
D2KFG9	-0.336	83.62	(Q) 26%	(D + E): 6	(R + K): 5	K7X1I3	-0.999	65.87	(Q) 35.2 %	(D + E): 7	(R + K): 4	R9XUS6	-0.991	55.98	(Q) 37.3 %	(D + E): 5	(R + K): 6
V9P737	-0.308	89.87	(Q) 22.1%	(D + E): 2	(R + K): 6	Q41529	-1.055	65.84	(Q) 38.5 %	(D + E): 5	(R + K): 7	A0A2U8JD15	-1.571	28.60	(Q) 41.1 %	(D + E): 6	(R + K): 3

R4JAP5	-0.588	75.75	(Q) 30.7%	(D + E): 4	(R + K): 6	A0A0B5JD20	-1.834	21.67	(Q) 50.3 %	(D + E): 9	(R + K): 8	Q402I5	-1.820	29.71	(Q) 53.1 %	(D + E): 9	(R + K): 7
Q00M55	-0.521	81.69	(Q) 31.7%	(D + E): 3	(R + K): 7	Q9XGF0	-0.257	86.10	(Q) 26 %	(D + E): 1	(R + K): 4	A0A2U8JD21	-1.798	30.28	(Q) 53.6 %	(D + E): 8	(R + K): 5
F8SGN3	-0.482	84.44	(Q) 31.0%	(D + E): 4	(R + K): 7	O22116	-0.060	103.77	(Q) 26.4 %	(D + E): 2	(R + K): 6	A0A1W6C2J9	-1.457	29.50	(Q) 32.1 %	(D + E): 20	(R + K): 20
B2Y2Q7	-0.507	81.69	(Q) 31.0%	(D + E): 3	(R + K): 7	A0A2U8JD31	-1.665	23.72	(Q) 42.3 %	(D + E):5	(R + K): 3	B8PSA6	-1.504	29.87	(Q) 34.2 %	(D + E): 31	(R + K): 24
B2Y2S3	-0.507	81.69	(Q) 31.0%	(D + E): 3	(R + K): 7	Q8W3V0	-1.170	50.54	(Q) 41.1 %	(D + E):6	(R + K): 7	Q52JL2	-1.458	32.22	(Q) 34.1 %	(D + E): 10	(R + K): 22
Q9XGF0	-0.257	86.10	(Q) 26.0%	(D + E): 1	(R + K): 4	A0A2U8JD20	-1.180	28.60	(Q) 53.1 %	(D + E):8	(R + K): 6	Q84UY5	-1.664	23.67	(Q) 37.5 %	(D + E):10	(R + K): 16
Q38LF4	-0.794	46.19	Thr (T) 21.7%	(D + E): 4	(R + K): 80	E4W506	-1.441	31.26	(Q) 32.7 %	(D + E):22	(R + K): 21	W0C8N8	-1.580	20.73	(Q) 34 %	(D + E): 16	(R + K): 21

						V9TP40	-1.468	31.96	(Q) 34. 3 %	(D + E):20	(R + K): 22	A0A060MZP 1	-1.847	16.41	(Q) 38. 5 %	(D + E): 17	(R + K): 15
						Q41553	-1.608	27.77	(Q) 34. 6 %	(D + E):30	(R + K): 26	T2HRF3	-1.601	29.04	(Q) 35. 1 %	(D + E): 29	(R + K): 26
x ¯	**-0.43**	**80.55**	-	-	-	**x ¯**	**-0.93**	**59.67**	-	-	-	**x ¯**	**-1.43**	**37.24**	-	-	-

Table 1S. Physico-chemical characterization of proteins from Gliadin Fractions according to their hydropathicity, aliphatic index and charge https://web.expasy.org/protparam

1. Identification of proteins remaining in supernatant after molecular binding.

The electrophoretic profile of proteins (gluten proteins, trypsin and chymotrypsin) were characterized to analyse the molecular interaction in a ternary mixture of gastrointestinal proteases, gluten proteins and procyanidin dimer B3 (**section 2.4** in manuscript). After that, the proteins remaining in supernatant were analysed according to the following instructions:

Trypsin Autolysis. Commercial Trypsin (1 mg) was mixed with 1 mL 0.1 M NH4HC03, 1M Urea and 0.1 mM CaCl, (pH 7.5). The mixture was allowed to stand at 37ºC overnight. Peptides resulting from autolysis were cleaned-up after incubation with Thermo Scientific Pierce C18 Spin Tips following the manufacturer instructions and further analysed by MALDI-TOF. Results obtained in experimental analysis by MALDI-TOF were correlated with theoretical autolysis peaks from MASCOT search engine database.

In-gel protein identification. The protein spots a to f (**Figure 1**-lines 10-12) were excised from SDS-PAGE gels and digested by in-gel trypsin procedure. Gel pieces (3 mm), thoroughly washed with 10 µL of deionized water and dried, were decolorized with a 50 mM ammonium bicarbonate solution and acetonitrile (ACN) (50:50, v/v) (four times for 15 min). After liquid removal the gel pieces were dehydrated for 30min by 30 µL of ACN. The dehydrated gel pieces were incubated with a solution of 50 mM ammonium bicarbonate/ ACN (50:50, v/v) with gentle agitation for 30 min, and briefly dehydrated with ACN. The wash step was repeated twice before the gel pieces were dried. Proteins in the gel pieces were digested with 5 µL of trypsin (10 ng/µL) in 25 mM ammonium bicarbonate at 37 °C overnight. A 5 µL volume of urea (1 M, pH 8) as the denaturing agent and calcium chloride (1 mM) were added to the solution to promote the enzyme action. After the digestion, peptides were extracted with 5 µL of a 50 mM ammonium bicarbonate solution for 15 min and 5 µL of an ACN/formic acid (FA; 1% in water) solution (50:50, v/v). After 15 min the solution containing the extracted peptides was collected, cleaned-up with Thermo Scientific Pierce C18 Spin Tips following the manufacturer instructions and further analysed by MALDI-TOF and nano LC-MS/MS.

MALDI-TOF peptides characterization. Peptide samples from in-gel digestion and trypsin autolysis were analyzed by MALDI-TOF. Briefly, 2 µL of sample were mixed with 2 µl of matrix solution containing 10 mg/mL alpha-cyano hydroxycinammic acid in acetonitrile, water and trifluoroacetic acid (30:69.9:0.1, v/v). A 2 µL aliquot was further applied onto a stainless steel target plate (MTP 394 target plate polished steel BC, Bruker Daltonik GmbH (Germany)) by triplicates and air-dried at room temperature.

Mass spectra were automatically acquired on an UltrafleXtream MALDI-TOF-TOF mass spectrometer (BrukerDaltonics, Germany) operating in reflector positive ion detection mode with laser SmartBeam-III and under FlexCompass 1.4 software control (BrukerDaltonics, Germany). The mass spectra (range from 100 to 3500 Da) of 1000 laser shots were accumulated with 2000 Hz of frequency, for each sample, and further analysed in the FlexAnalysis software.

Nano LC-MS/MS proteins identification. The chromatographic separation of peptides was performed following the methods described in **section 2.7** in manuscript. Spectra were submitted to Proteome Discoverer 1.4 software (Thermofisher Scientific) using the SEQUEST search engine to peptide sequencing. Bovine Trypsin and chymotrypsin as well as the list of the already identified gluten proteins were included in a FASTA file used for identification purposes.

1.1. Results obtained from trypsin autolysis

In-gel digestion peptides spectra were firstly characterized by MALDI-TOF and further compared with the bovine trypsin autolysis. Bovine trypsin autolysis was performed in order to avoid any interaction in the sequencing analysis. Protein spots c, f and h were identified as trypsin after comparing the spectra autolysis pattern with the isolated standard, as displayed in **Table 2S** and **Figure 1S**.These results were further confirmed by bottom up proteomics (nano-LC-MS/MS analysis). Peptides with high or medium confidence levels were highlighted in green and yellow, respectively.

```
        10          20          30          40          50
MKTFIFLALL GAAVAFPVDD DDKIVGGYTC GANTVPYQVS LNSGYHFCGG
        60          70          80          90         100
SLINSQWVVS AAHCYKSGIQ VRLGEDNINV VEGNEQFISA SKSIVHPSYN
       110         120         130         140         150
SNTLNNDIML IKLKSAASLN SRVASISLPT SCASAGTQCL ISGWGNTKSS
       160         170         180         190         200
GTSYPDVLKC LKAPILSDSS CKSAYPGQIT SNMFCAGYLE GGKDSCQGDS
       210         220         230         240
GGPVVCSGKL QGIVSWGSGC AQKNKPGVYT KVCNYVSWIK QTIASN
```

Appendix 1S. Alpha-trypsin from bovine pancreas (TRY1_BOVIN) sequencing.

Cleavage	Theoretical Mascot MS		MALDI-MS m/z				Sequence
	Mono.	Avg.	Standard	c	f	h	
110-111	259.19	259.35	261.373	261.411	261.369	261.357	LK
157-159	362.20	362.49	361.344	361.381	361.327	361.364	CLK
238-243	632.31	632.67	636.185	636.232	636.166	636.143	QTIASN
64-69	658.38	658.76	657.595	657.651	657.576	657.555	SGIQVR
112-119	804.41	804.86	842.334	842.384	842.299	842.273	SAASLNSR
221-228	905.50	906.05	ND	907.268	ND	ND	NKPGVYTK
160-169	1019.50	1020.17	1043.926	1043.984	1043.883	1043.843	APILSDSSCK
229-237	1110.55	1111.33	ND	ND	1108.716	ND	VCNYVSWIK
146-156	1152.57	1153.25	1157.958	1158.004	1157.862	1157.836	SSGTSYPDVLK
207-220	1432.71	1433.65	1439.785	1439.834	1439.679	1439.688	LQGIVSWGSGCAQK
191-206	1494.61	1495.61	ND	ND	ND	ND	DSCQGDSGGPVVCSGK
70-89	2162.05	2163.33	2164.696	2164.632	2164.553	2164.458	LGEDNINVVEGNEQFISASK
170-190	2192.99	2194.47	2196.426	2196.450	2196.246	2196.222	SAYPGQITSNMFCAGYLEGGK
90-109	2272.15	2273.60	2268.763	2268.800	2268.545	2268.557	SIVHPSYNSNTLNNDIMLIK
120-145	2551.24	2552.91	2556.486	2556.442	2556.230	2556.217	VASISLPTS…LISGWGNTK

Table 2S. Identification of major peptides resulting from bovine trypsin autolysis

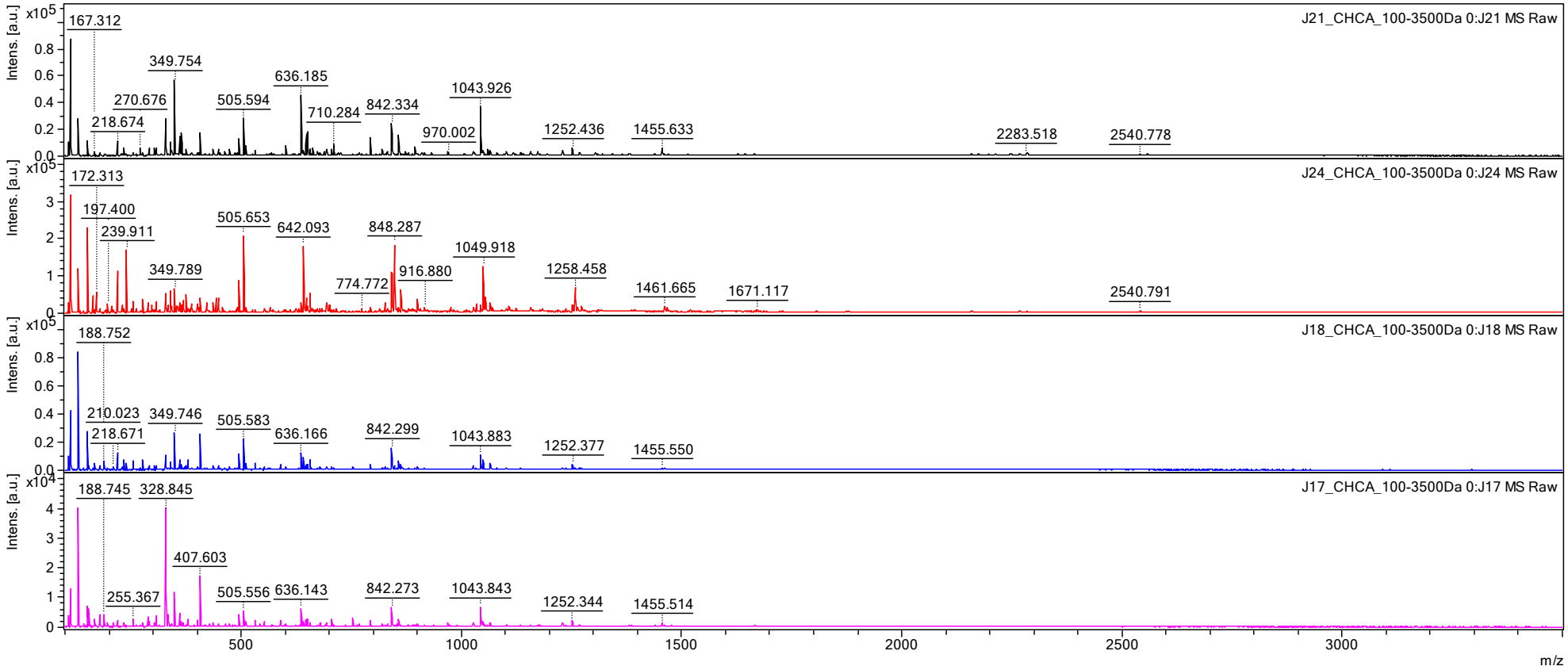
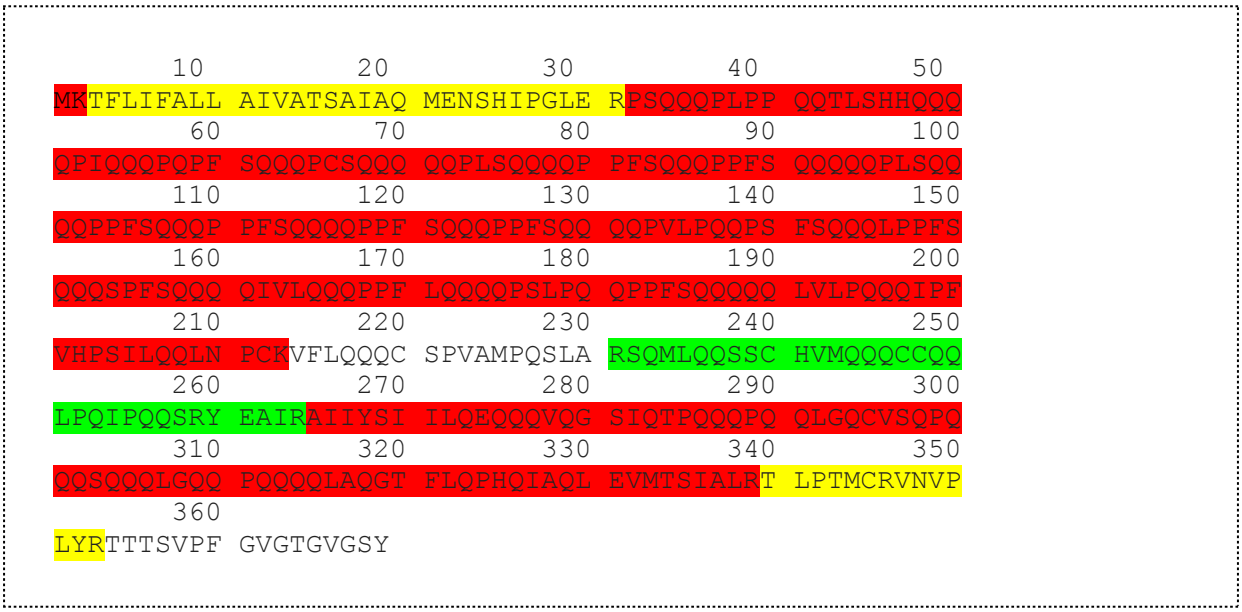


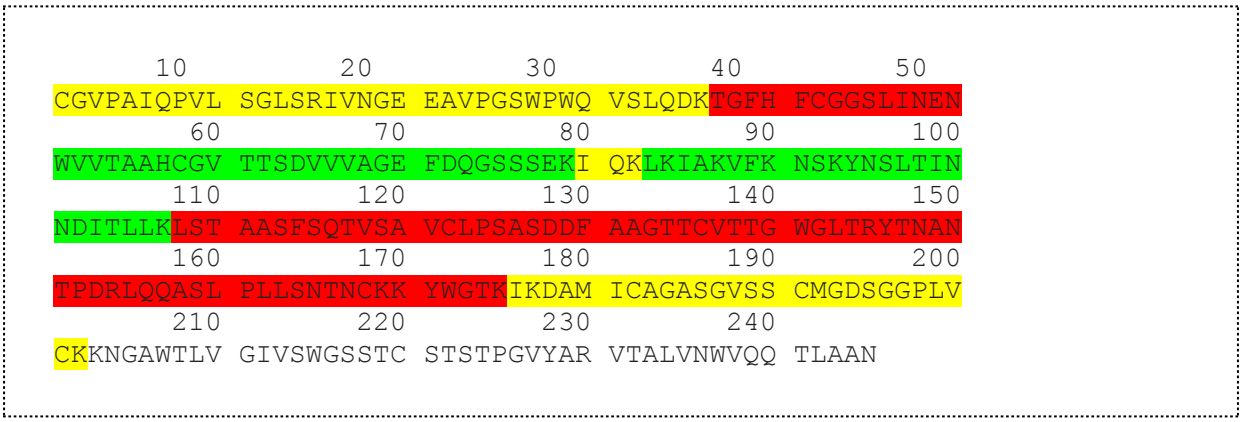
Figure 1S. MS Spectra of trypsin autolysis highlighted in black and in-gel digestion of spots c (red), f (blue) and h (pink).

1.2. Identification of protein spots by bottom up proteomics

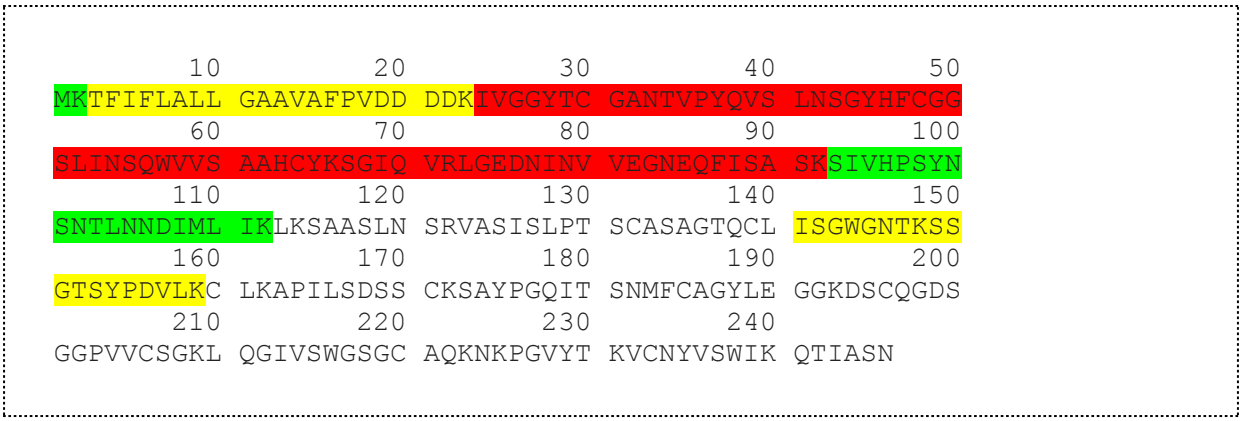
In-gel digestion peptides spectra were also sequenced by nano-LC-MS/MS in order to achieve results that are more accurate. As displayed in Appendix 2S to Appendix 9S, spots b, e and g were identified as trypsin; spots c, f and h as chymotrypsin and spots a and d are part of the low-molecular weight gliadin subunit - *B2Y2S3* and *B2Y2Q7*, respectively, highlighting a slightly digestion of Gliadin Fraction A during the incubation prior to analysis. Peptides are displayed in green, yellow or red according to the high medium or low scores, respectively. It should to be noticed that some peptides were low-confidence sequenced probably given the in-gel protein fixation, which difficult the protein extraction but this barrier was overcome by achieving high proteins coverage.



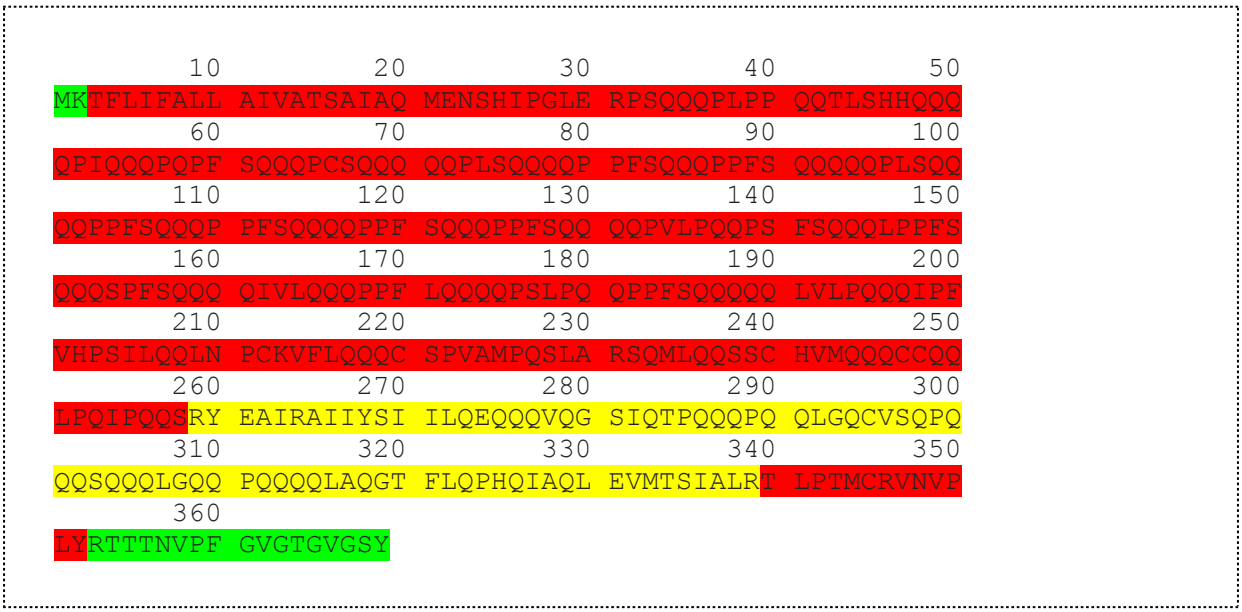
Appendix 2S. Protein spot a sequencing results. Peptides are displayed in green, yellow or red according to the high medium or low scores, respectively.



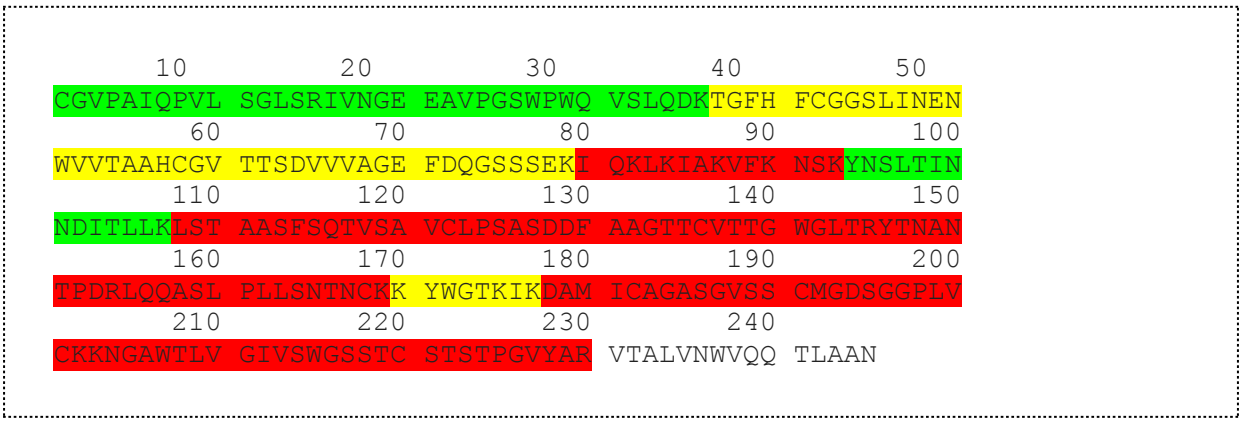
Appendix 3S. Protein spot b sequencing results. Peptides are displayed in green, yellow or red according to the high medium or low scores, respectively.



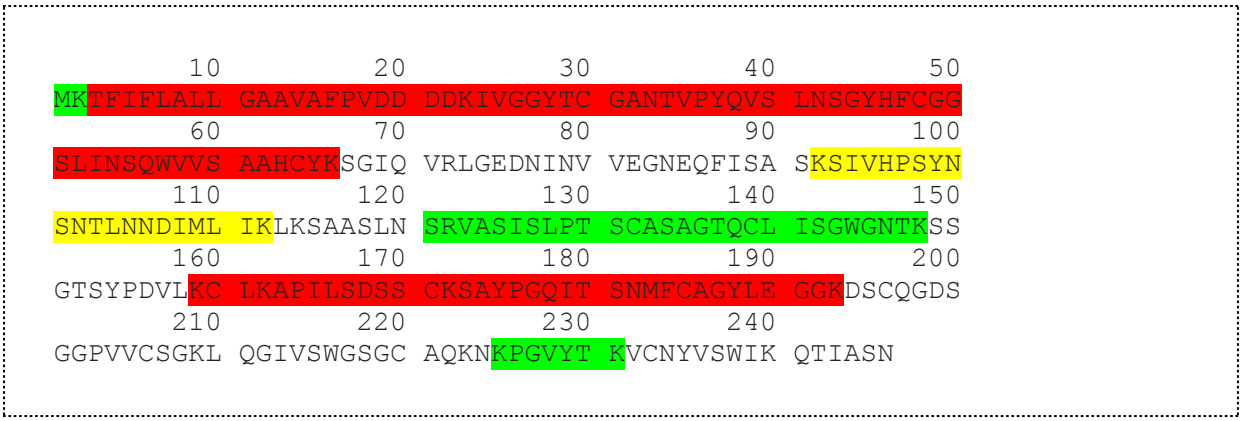
Appendix 4S. Protein spot c sequencing results. Peptides are displayed in green, yellow or red according to the high medium or low scores, respectively.



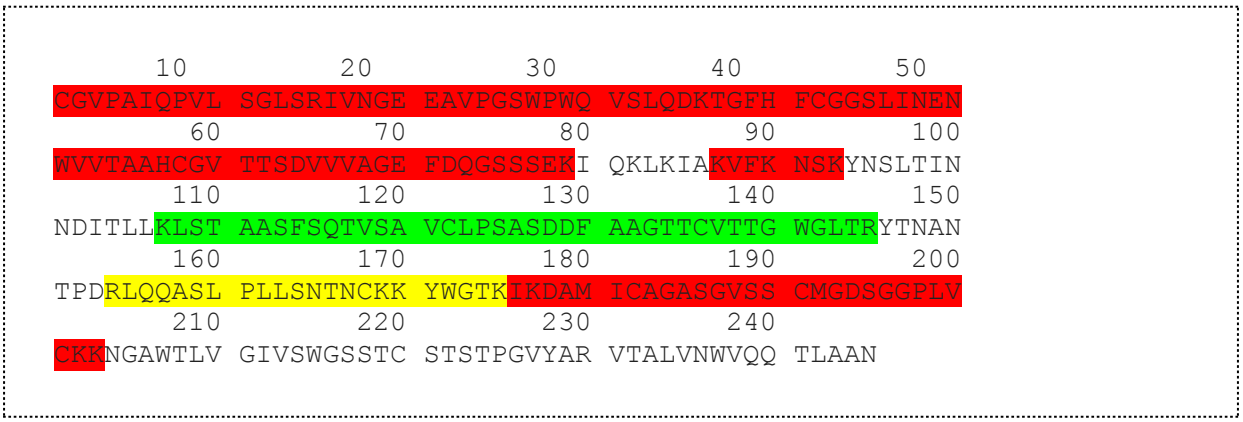
Appendix 5S. Protein spot d sequencing results. Peptides are displayed in green, yellow or red according to the high medium or low scores, respectively.



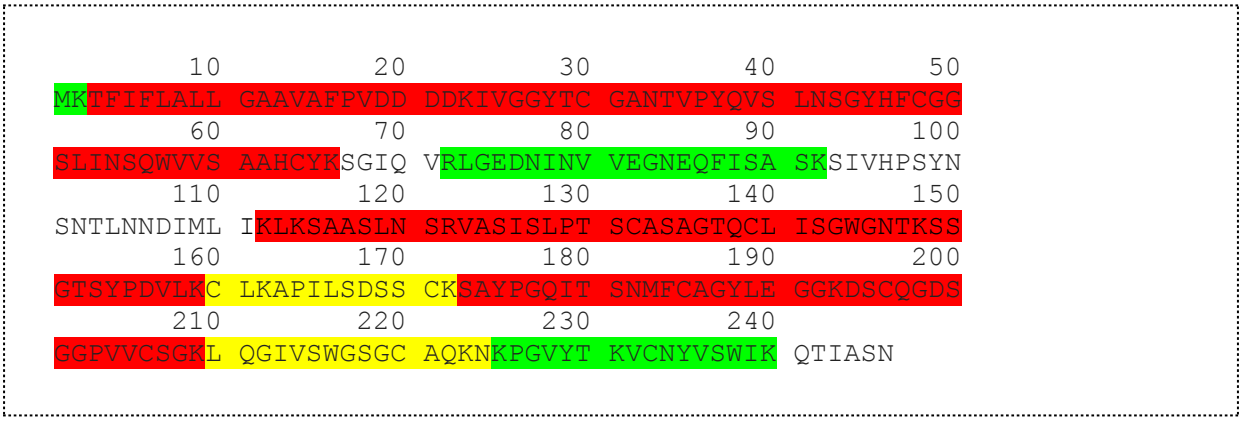
Appendix 6S. Protein spot e sequencing results. Peptides are displayed in green, yellow or red according to the high medium or low scores, respectively.



Appendix 7S. Protein spot f sequencing results. Peptides are displayed in green, yellow or red according to the high medium or low scores, respectively.



Appendix 8S. Protein spot g sequencing results. Peptides are displayed in green, yellow or red according to the high medium or low scores, respectively.



Appendix 9S. Protein spot h sequencing results. Peptides are displayed in green, yellow or red according to the high medium or low scores, respectively

References

Perez-Gregorio, M. R., Días, R., Mateus, N., & De Freitas, V. (2017). Chromatographic and mass spectrometry analysis of wheat flour prolamins, the causative compounds of celiac disease. *Food and Function, 8*(8), 2712-2721. doi:10.1039/c7fo00266a