

Stability of the major allergen Brazil nut 2S albumin (Ber e 1) to physiologically relevant *in vitro* gastrointestinal digestion

F. Javier Moreno¹, Fred A. Mellon¹, Martin S. J. Wickham¹, Andrew R. Bottrill² and E. N. Clare Mills¹

¹ Institute of Food Research, Norwich Research Park, Norwich, UK

² John Innes Centre, Norwich Research Park, Norwich, UK

Keywords

2S albumin; digestion; food allergy; mass spectrometry; Brazil nut

Correspondence

F. J. Moreno, Fundación AZTI,
Txatxarramendi ugarte a z/g, 48395
Sukarrieta, Bizkaia, Spain
Fax: +34 946870006
Tel: +34 946029410
E-mail: jmoreno@suk.azti.es

(Received 3 August 2004, revised 29
October 2004, accepted 5 November 2004)

doi:10.1111/j.1742-4658.2004.04472.x

The major 2S albumin allergen from Brazil nuts, Ber e 1, was subjected to gastrointestinal digestion using a physiologically relevant *in vitro* model system either before or after heating (100 °C for 20 min). Whilst the albumin was cleaved into peptides, these were held together in a much larger structure even when digested by using a simulated phase 1 (gastric) followed by a phase 2 (duodenal) digestion system. Neither prior heating of Ber e 1 nor the presence of the physiological surfactant phosphatidylcholine affected the pattern of proteolysis. After 2 h of gastric digestion, ≈25% of the allergen remained intact, ≈50% corresponded to a large fragment of M_r 6400, and the remainder comprised smaller peptides. During duodenal digestion, residual intact 2S albumin disappeared quickly, but a modified form of the 'large fragment' remained, even after 2 h of digestion, with a mass of ≈5000 Da. The 'large fragment' comprised several smaller peptides that were identified, by using different MS techniques, as deriving from the large subunit. In particular, sequences corresponding to the hypervariable region (Q37–M47) and to another peptide (P42–P69), spanning the main immunoglobulin E epitope region of 2S albumin allergens, were found to be largely intact following phase 1 (gastric) digestion. They also contained previously identified putative T-cell epitopes. These findings indicate that the characteristic conserved skeleton of cysteine residues of 2S albumin family and, particularly, the intrachain disulphide bond pattern of the large subunit, play a critical role in holding the core protein structure together even after extensive proteolysis, and the resulting structures still contain potentially active B- and T-cell epitopes.

2S storage albumins occur in a diverse range of plant seeds, are members of the prolamin superfamily [1] and constitute one of the most important major plant food allergens that sensitize via the gastrointestinal (GI) tract [2]. Among the tree nuts, Brazil nut is frequently associated with immunoglobulin E (IgE)-mediated food allergy [3], the 2S albumin, known as Ber e 1, being the major allergen [4].

2S albumins are considered to be structurally homologous, typically heterodimeric (small and large subunits of ≈4000 and 9000 M_r , respectively) globular proteins. They have a conserved skeleton of cysteine residues (typical of members of the prolamin superfamily), which form four intermolecular disulphide bonds that hold the two subunits together and contribute to their stability and compactness [5]. This rigid

Abbreviations

GI, gastrointestinal; IgE, immunoglobulin E; SGF, simulated gastric fluid; PtdCho, egg l-phosphatidylcholine.

structure is thought to be responsible for the stability of the 2S albumins to proteolytic attack. Thus, following SDS/PAGE analysis, 2S albumins from mustard [6] and Brazil nut [7] have been shown to be resistant to pepsin in simulated gastric fluid at pH 1.2 for longer than 60 and 15 min, respectively.

It has been postulated that for an allergen sensitizing an individual via the GI tract, it must have properties which preserve its structure from degradation in the GI tract (such as resistance to low pH, bile salts and proteolysis), thus allowing enough allergen to survive in a sufficiently intact form to be taken up by the gut and sensitize the mucosal immune system [8–11]. Consequently, it has been proposed that resistance of proteins to pepsin digestion in the stomach is a marker for potential allergenicity [6]. Protein digestibility (measured as resistance to pepsin) is also one of the relevant parameters used for assessing the allergenic potential of novel proteins [12].

In this study, the resistance to digestion of a single 2S albumin isoform (ExPASy entry P04403), in either a native or a heated form, was assessed by using an *in vitro* digestion model system employing two physiologically relevant stages to mimic the passage of food through the stomach (phase 1) into the duodenum (phase 2). The role of the physiological surfactant phosphatidylcholine (PtdCho), which is secreted by the gastric mucosa and also occurs in the bile, was also investigated. Finally, the allergen fragments that resist pepsinolysis were identified by using a combination of mass spectrometric techniques, including RP-HPLC-ESI-MS and MALDI-TOF, as well as nanoelectrospray Q-TOF MS/MS, in order to sequence the peptides.

Results and Discussion

In vitro digestion of Brazil nut 2S albumin, Ber e 1

Gastric digestion (phase 1)

The 2S albumin (Ber e 1) was found to be very resistant to pepsinolysis, with a prominent band evident on SDS/PAGE after 2 h of digestion (Fig. 1A). No difference was observed between native or preheated (at either neutral or acid pH) 2S albumin phase 1 digests, and this was not affected by the presence of PtdCho (data not shown). However, as digestion proceeded, the Ber e 1 band showed a slightly faster electrophoretic mobility in all cases, although no smaller peptides were evident upon SDS/PAGE. Following reduction, the large M_r 9000 and small M_r 3000 subunits of the undigested protein were both still evident after 120 min of

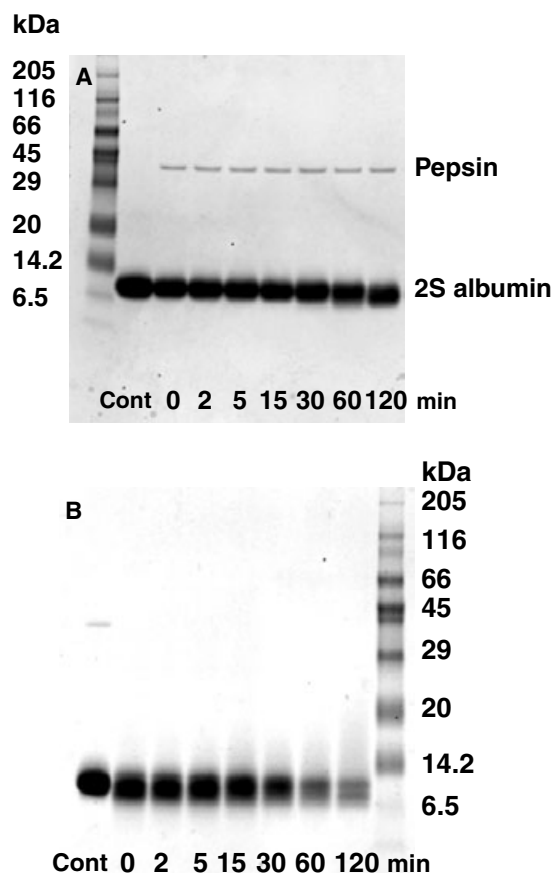


Fig. 1. SDS/PAGE analyses showing (A) the gastric digestion (phase 1) and (B) the gastric and duodenal digestion (phases 1 + 2) of 2S albumin native under nonreducing conditions.

phase 1 digestion, with a very faint band running below the large subunit that was evident after reduction (data not shown).

HPLC analysis of peptides indicated that the profiles were essentially identical from native (Fig. 2) and from preheated (data not shown) Ber e 1, and when digestions were performed in the presence (data not shown) or absence (Fig. 2) of PtdCho. The intact 2S albumin was resolved as a single peak of 42.5 min retention time (Fig. 2A). The first peptides appeared after 15 min of digestion, and an incomplete conversion of the intact protein into a poorly resolved peak with a shorter retention time of 38.5 min took place. This 38.5 min peak (termed 'large fragment') became the main component after 120 min of digestion (Fig. 2D) and probably corresponds to the faster running species observed by SDS/PAGE (Fig. 1A). A protein column was then used to improve the resolution of higher molecular weight species (Fig. 3). Intact 2S albumin was completely resolved from the 'large fragment' formed as consequence of digestion (Fig. 3B).

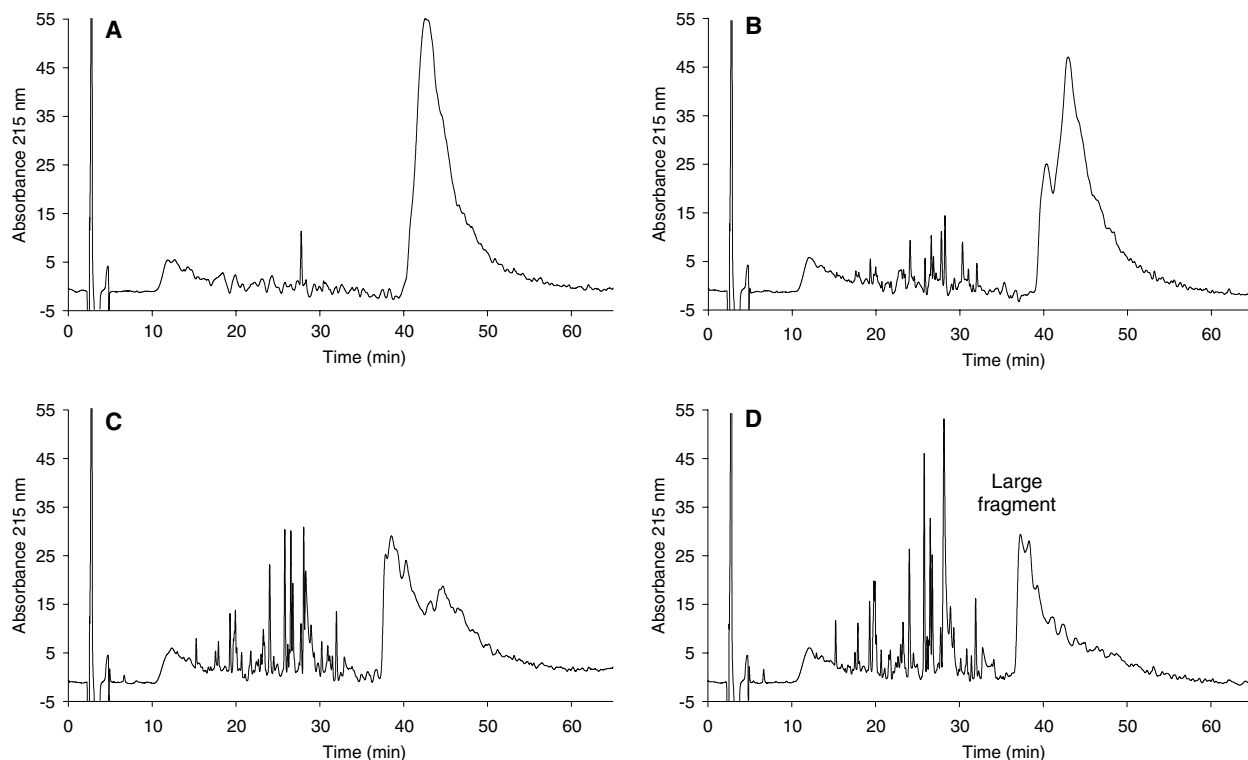


Fig. 2. RP-HPLC patterns using a peptide (90 Å pore size) column of nonreduced samples corresponding to native gastric digested (phase 1) 2S albumin. (A) 0 min; (B) 15 min digestion; (C) 60 min digestion; and (D) 120 min digestion.

Assuming that the UV absorbance was equal for all species, analysis of peak areas was used to determine the yield of peptides in the HPLC profile. This showed that $\approx 25\%$ of the allergen remained intact, $\approx 50\%$ corresponded to the 'large fragment' and the remainder comprised small peptides. Following reduction of the digestion products, HPLC analysis showed the characteristic large (peak 7) and small (peak 5) sub-units of the native 2S albumin at the start of digestion (Fig. 3C). After 120 min, some of the same peptides were observed as under nonreducing conditions (Fig. 3B,D), indicating that these are 'free' peptides and not covalently linked to the core 2S albumin structure. The 'large fragment' observed in the absence of a reducing agent was lost, indicating that it comprised a number of disulphide linked peptides (Fig. 3D).

Duodenal digestion (phase 2)

After 2 h of gastric digestion, the pH was increased and trypsin and chymotrypsin were added with bile salts in order to simulate a duodenal environment (phase 2). No noteworthy differences in digestion patterns were found between native and preheated 2S albumin and the presence or absence of PtdCho (data

not shown). Even after 2 h of gastric digestion followed by 2 h of duodenal digestion, a weak band of slightly lower molecular weight than the undigested 2S albumin was detected by SDS/PAGE (Fig. 1B). This band probably corresponds to the broad peak that eluted between 33 and 40 min on peptide HPLC (Fig. 4). The peptide profile was essentially the same after 5 min and 120 min of duodenal digestion (Fig. 4A,D), and changes were only observed in the broad peak. In addition, polypeptide digests observed by SDS/PAGE under reducing conditions had lower molecular weights than those found during the gastric digestion, indicative of further proteolysis (data not shown). Protein HPLC showed that intact 2S albumin, remaining after phase 1 digestion, disappeared quickly at the beginning of the duodenal digestion, although the broad peak corresponding to the 'large fragment' observed after phase 1 digestion remained (Fig. 5A). This peak decreased in area and broadened owing to the formation of a range of new fragments as digestion advanced (Fig. 5B). As for the phase 1 digests, when analysed in the presence of a reducing agent this 'large fragment' peak was lost, indicating that it comprised several smaller disulphide-linked polypeptides (Fig. 5C).

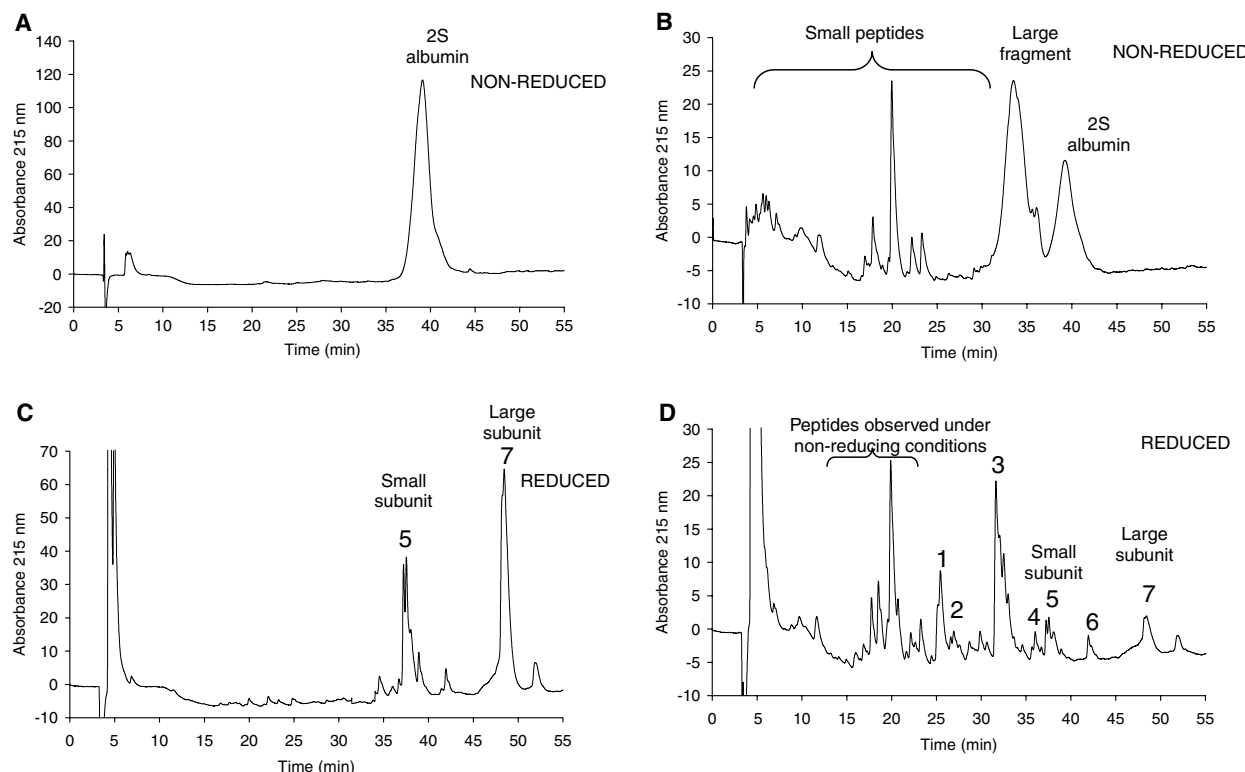


Fig. 3. RP-HPLC patterns using a protein (300 Å pore size) column of native gastric digested (phase 1) 2S albumin. (A) 0 min; (B) 120 min digestion nonreduced; (C) 0 min; (D) 120 min digestion reduced. Labelled peaks are described in the text.

Identification of peptides resulting from digestion

Gastric digestion (phase 1)

RP-HPLC-ESI-MS analysis of an intact 2S albumin peak (Fig. 3B) showed the presence of four molecular masses (12 212.1, 12 125.8, 11 980.0 and 11 504.0) corresponding to the intact 2S albumin, together with ragged C and N-termini, as previously shown [13]. Such post-translational modification is typical for 2S albumins from different plant species such as rapeseed [14–17] and castor bean [18]. As expected, following reduction two peaks (5 and 7, Fig. 3D) were found to correspond to the intact small and large subunits, including the ragged C-termini (Table 1) [13].

Without reduction, the 'large fragment' observed after 2 h of gastric digestion (Figs 2D and 3B) comprised three molecular masses of > 6 kDa (6368.4, 6483.3 and 6236.8), as determined by RP-HPLC-ESI-MS. Constituent peptides in the 'large fragment' were characterized by RP-HPLC-ESI-MS and MALDI-TOF, following reduction (Table 1, Fig. 3D). A good correlation between the molecular masses obtained by ESI-MS and MALDI-TOF was obtained, although some peptides could not be identified by MALDI-TOF

as their masses were outside the mass range scanned. The additional peptides observed on reduction of gastric digests result from the M_r 6400 'large fragment' observed under nonreducing conditions (Fig. 3B,D). The main peptide peak, 3, comprised three different molecular masses that probably correspond to a C-terminal peptide AENIPSRCNLSPMRCMPMGGS(54–73), derived from the large subunit, with some of the C-terminal residue deletions observed in the intact protein. Further confirmation of this identification was obtained by nanoelectrospray Q-TOF MS/MS sequencing which verified the presence of the following peptides: AENIPSRCNLSPMRCMPMGGS(54–73), AENIPSRCNLSPMRCMPMG(54–72) and AENIPSRCNLSPMRCMPMG(54–71).

Peptide peak 1 contained one mass, and peak 2 contained two masses that probably correspond to peptides derived from DESCRCGLRMM(20–31) of the large subunit (Table 1). Pepsinolysis removed either the N-terminal Asp or the C-terminal Met, giving rise to peptides ESCRCGLRM(21–30) and ESCRCGLRMM(21–31), respectively. These removals imply differences of 246 and 115 atomic mass units; such variations were also found in the peak corresponding to

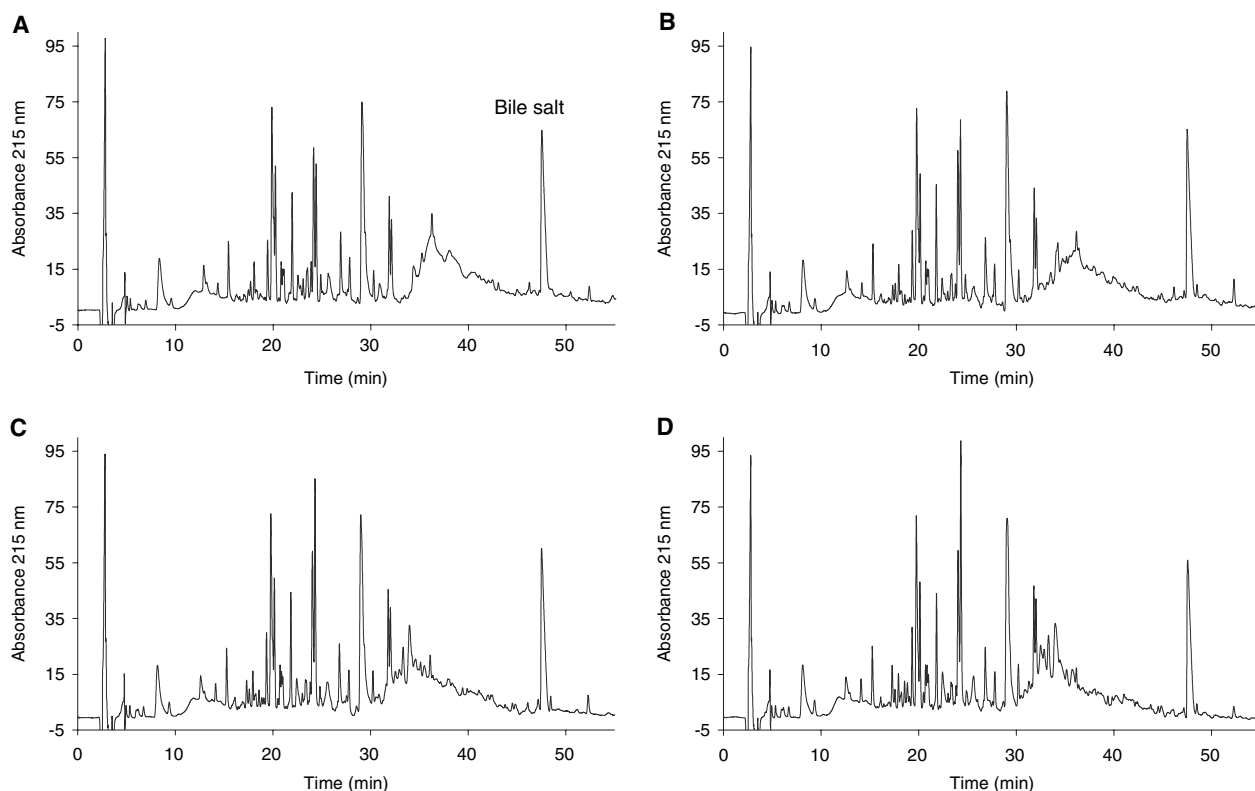


Fig. 4. RP-HPLC patterns using a peptide (90 Å pore size) column of nonreduced samples corresponding to native 2S albumin subjected to gastric (120 min) and duodenal digestion (phases 1 + 2) for (A) 5 min; (B) 30 min; (C) 60 min; and (D) 120 min.

the large fragment under nonreducing conditions (Table 2), again supporting the conclusion that these peptides make up part of the 'large fragment'. Other minor masses were also found and assigned to peptides MSECCEQLEG(9–18) (peak 2), MSECCEQLEGMDESCRCEGLR(9–29) (peak 4), CEGLRMMMMRMQQEEMQPRGEQ(25–46) (peak 4) and PRGEQMRRMMRLAENIPSRCNLSPMRCP(42–69) (peak 6) (Table 1). The diversity in molecular masses found after reduction suggests that the 'large fragment' is not a single unique combination of disulphide-linked peptides but rather a complex mixture. Nevertheless, these data indicate that peptides MSECCEQLEG(9–18), DESCRCEGLRMM(20–31) and AENIPSRCNLSPMRCPMGGS(54–73) (of total mass 4675 Da) are probably the main components of this fragment. Figure 6A shows the potential cleavage sites of pepsin defined by the PEPTIDE CUTTER tool of ExPASy (<http://us.expasy.org/tools/peptidecutter/>). The predicted C-terminal fragment of the small subunit originating from pepsinolysis SHCRMVYMRQQMEES(15–28) would have a molecular mass of 1815. When added to the other three peptides described above, it would give rise to a fragment of total mass 6484 Da (taking into

account disulphide bond formation), which corresponds closely to the major molecular mass found in the 'large fragment' peak under nonreducing conditions (Table 2).

Peptide HPLC-ESI-MS of gastric digestion (phase 1) showed the presence of a wide range of small peptides eluting between 14 and 33 min (Fig. 2), with molecular masses within the range 400–1100 Da. These peptides were too small to be analysed by MALDI-TOF. Overall, excluding those masses that might match with peptides resulting from pepsin, trypsin and chymotrypsin autolysis, three peptides with retention times of 14–19 min could be tentatively assigned on the basis of mass as being derived from the small subunit. Tentative matches for nine peptides with retention times of 20–32 min suggest that they are derived from the large subunit (Table 2).

Duodenal digestion (phase 1)

During phase 2 digestion, the 'large fragment' (Fig. 5A,B) gave several masses when analysed by HPLC-ESI-MS; the most abundant masses were of 5755 Da and 5739 Da (Table 3), and there were several

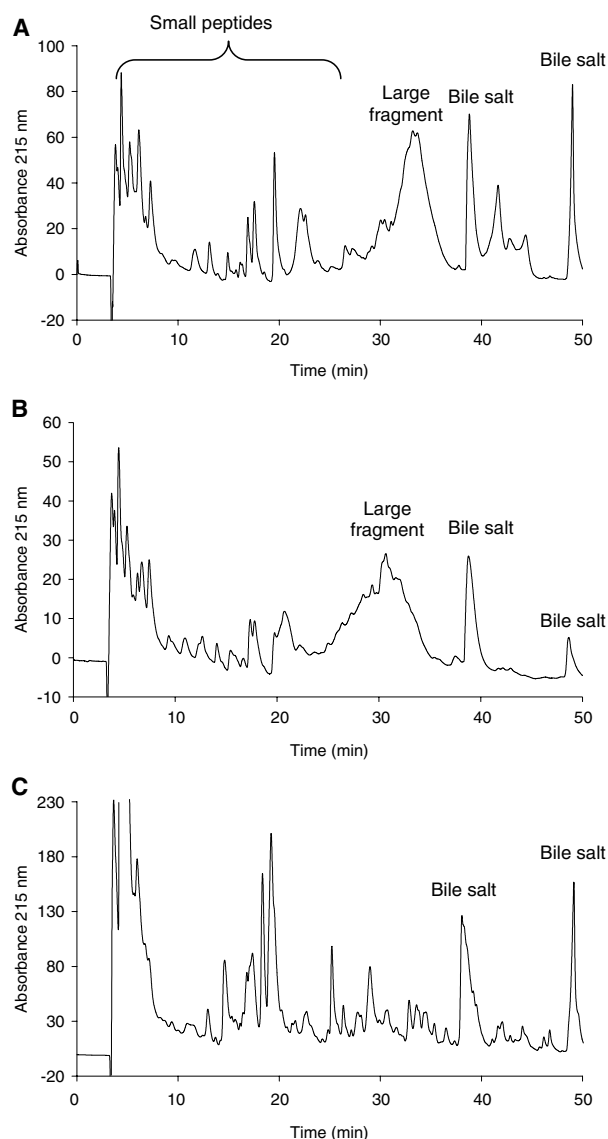


Fig. 5. RP-HPLC patterns using a protein (300 Å pore size) column of native 2S albumin subjected to gastric (120 min) and duodenal digestion (phases 1 + 2) for (A) 5 min and (B) 120 min (non reduced); and for (C) 15 min (reduced).

minor molecular masses at ≈ 5 kDa, in good agreement with results obtained by SDS/PAGE (Fig. 1). This suggests that the 'large fragment' observed in phase 2 digestion comprises a complex mixture of polypeptides similar to those found during the phase 1 digestion but with some differences. Duodenal digestion of the 'large fragment' therefore resulted in a reduction of mass from ≈ 6200 – 6500 to ≈ 5000 – 5700 . This reduction in size probably results from a loss of peptides of 500–1500 Da. Peptide AENIPSRCNLSPMRCMPMGGS(54–73) (peak 3), a major component of the 'large

fragment', disappeared following duodenal digestion (Figs 3D and 5C). Taking into consideration its potential tryptic and chymotryptic cleavage sites, different peptides would be generated, including SPM(64–66) and GCS(71–73), as well as a free arginine residue, which would imply a loss of 728 Da. This corresponds to the mass difference between the phase 1 'large fragment' (6483.3 Da) and the phase 2 'large fragment' (5755 Da). Further identification (by MS) of the constituent peptides of the large fragment, during the duodenal digestion and under reducing conditions, was unsuccessful. This may have been caused by the multi-component medium for phase 2, including bile salts, lipase, colipase, enzymatic inhibitor, etc., which interfered with the ionization of these large peptides.

Such problems were not encountered for peptide HPLC-ESI-MS of duodenal digestion (phase 2), which showed the presence of one new (penta)peptide derived from the small subunit, whilst 10 new peptides were found to be consistent with being derived from the large subunit sequence (Table 3).

General discussion

Resistance to digestion in the gastrointestinal tract is thought to be one of the factors that may contribute to the allergenic potential of food proteins by allowing sufficient intact (or a large fragment of) protein to be taken up by the gut and sensitize the mucosal immune system. The 2S albumin family has been described as highly stable to proteolysis and thermal denaturation owing to its compact 3D structure, which is dominated by the pattern of cysteine residues [7,19,20]. In this study, the Brazil nut 2S albumin allergen, Ber e 1, exhibited a similar behaviour and, following *in vitro* gastric digestion, $\approx 25\%$ of the allergen remained intact, whereas 50% corresponded to a 'large fragment' (M_r 6400) comprising mainly peptides matching with the large chain linked together by intrachain disulphide bridges. Figure 6A shows the position of some peptides identified as resistant to *in vitro* gastric digestion in the Brazil nut 2S albumin structure. From the data presented here, it is evident that the conserved skeleton of cysteine residues and, particularly, the intrachain disulphide bonds of the large chain, play an important role in holding the core protein structure together, even after extensive proteolysis. On the basis of mass spectrometric analysis it can be postulated that the 'large fragment' mostly comprises peptides from the large subunit, suggesting that the large chain was more resistant to proteolytic attack than the small chain. The fact that 2S albumin digestion was not affected by preheating at 100 °C, at either acid or

Table 1. Brazil nut 2S albumin (Ber e 1) polypeptides following gastric (phase 1) digestion for 120 min under reducing conditions and identified by RP-HPLC-ESI-MS by using a protein 300 Å pore size column and MALDI-TOF MS. Peaks are as described in Fig. 3D.

Subunit	Peak no.	Retention time (min)	Molecular masses observed by RP-HPLC-ESI-MS	Molecular masses observed by MALDI-TOF	Sequence assigned by using ExPASy P04403
Small	5	37.55	3618.1 3530.2	3616.8 3529.7	QEECREQMQRQQMLSHCRMVYMRQQMEES(1–28) ^a QEECREQMQRQQMLSHCRMVYMRQQMEE(1–27) ^a
Large	1	25.02	1183.4	Not determined	ESCRCEGLRM(21–30)
			1429.6	Not determined	DESCRCEGLRMM(20–31) or MDESCRCEGLRM(19–30)
	3	31.75	1314.6		ESCRCEGLRMM(21–31)
			1128.2		MSECCEQLEG(9–18) or SECCEQLEG(10–19)
			2120.0	2119.9	AENIPSRNLSMPMRCPMGG(54–73)
			1976.0	1976.0	AENIPSRNLSMPMRCPMG(54–71)
	4	35.91	2033.2	2032.1	AENIPSRNLSMPMRCPMGG(54–72)
			2408.2	2407.9	MSECCEQLEGMDSCRCEGLR(9–29) or SECCEQLEGMDSCRCEGLRM(10–30)
	6	42.23	2729.0	2728	CEGLRMMMMRMQQEEMQPRGEQ(25–46)
			3330.1	3329.6	PRGEQMRMMMMRLAENIPSRNLSMPMRCP(42–69)
	7	48.48	8601.2 8513.0 8457.2	Not determined	PRRGMEPHMSECCE...RCNLSMPMRCPMGG(1–73) PRRGMEPHMSECCE...RCNLSMPMRCPMGG(1–72) PRRGMEPHMSECCE...RCNLSMPMRCPMG(1–71)

^a Conversion of N-terminal glutamine to pyroglutamic acid.**Table 2.** Peptide profile of Brazil nut 2S albumin (Ber e 1) after 120 min gastric (phase 1) digestion alone, determined by RP-HPLC-ESI-MS using a peptide 90 Å pore size column. Peaks are described in Fig. 2D.

Subunit	Retention time (min)	Mass	Putative sequence assigned by using ExPASy P04403
Small	14.41	434.5 ^a	MRQ(21–23)
	16.75	623.4 ^a	QMEES(24–28)
	19.09	562.5 ^a	MQRQ(8–11) or MRQQ(21–24)
Large	20.05	820.6 ^b	RMQQEE(34–39)
	22.73	745.7 ^a	PRRGME(1–6)
	24.03	824.4 ^b	MMRMQ(32–37)
	24.03	593.4 ^b	MRRM(47–50) or RRRM(48–51) or RMMR(49–52)
	25.42	1105.7 ^b	EMQPRGEQ(39–47)
	25.77	889.3 ^b	QQEEMQPR(36–42)
	27.85	1110.8 ^b	PRRGMEPHM(1–9)
	31.23	568.5 ^b	RMMM(29–32), MMRM(31–34) or MMRM(32–35)
	32.35	862.7 ^b	RMMML(48–53)
	40.76	12 212.1	
Intact protein		12 125.8 ^c	
		11 980.0 ^c	
		11 504.0 ^c	
Uncertain origin	38.94	6483.3	
		6368.4	
		6236.8	

^a Peptides resulting from specific cleavage of pepsin. ^b Peptides resulting from nonspecific cleavage of pepsin. ^c Ragged C- and N-termini.

neutral pH, can be attributed partly to its thermostability with minimal and reversible changes at the level of the secondary structure and partly to its disulphide bonded structure [7,13].

It is interesting to note that all the IgE-binding epitopes characterized to date in 2S albumin allergens are

located in the large chain. Therefore, a common IgE epitope has been described in the large chain of 2S albumins from yellow and oriental mustard *Sin a 1* and *Bra j 1* [21,22] whilst Robotham *et al.* [23] determined one major IgE epitope that corresponded to the large chain of 2S albumin from walnut (*Jug r 1*). This

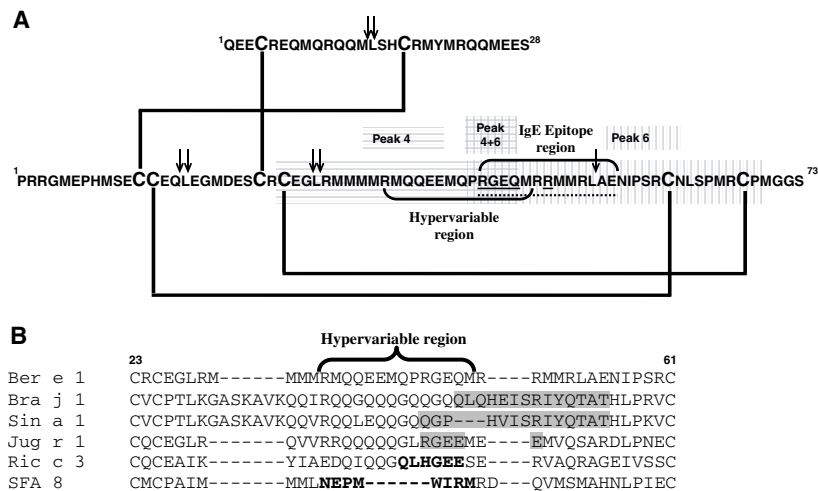


Fig. 6. Position of major gastric (phase 1) resistant peptides in the Brazil nut 2S albumin (Ber e 1). (A) Potential cleavage sites of pepsin are indicated with arrows. Major resistant peptides are shaded; peaks 4 and 6 are as described in Fig. 3D. Amino acids which would coincide with the position of the known epitopes of 2S albumin from walnut (solid line) and mustard (dotted line) are underlined. (B) Alignment of the hypervariable region and immunoglobulin E (IgE) epitopes (shaded) of 2S albumin from different species (Ber e 1, Brazil nut; Bra j 1, oriental mustard; Sin a 1, yellow mustard; Jug r 1, English walnut; Ric c 3, castor bean; SFA 8, sunflower seeds) by using T-COFFEE [36]. The hyper-variable regions of Ric c 3 and SFA-8 (bold) were taken from the 3D structure determined by NMR methods [24,37]. Numbering is given according to the primary structure of Ber e 1.

Table 3. Peptide profile of Brazil nut 2S albumin (Ber e 1), following combined phase 1 (gastric) digestion for 120 min followed by phase 2 (duodenal) digestion for 60 min, as determined by RP-HPLC-ESI-MS using a peptide 90 Å pore size column. Peaks are as described in Fig. 4C.

Subunit	Retention time (min)	Mass	Putative sequence assigned by using ExPASy P04403
Small	15.71	434.4 ^c	MRQ(21–23)
	18.31	691.4 ^a	EQMQR(6–10)
Large	20.13	562.3 ^c	MQRQ(8–11) or MRQQ(21–24)
	13.72	855.6 ^b	RMMMMR(29–34)
	19.27	490.3 ^b	SPMR(64–67) or MRLA(51–54)
	19.27	620.3 ^a	GEQMR(44–48)
	22.73	586.4 ^b	PRGEQ(42–46) or QPRGE(41–45) or NIPSR(56–60)
	29.75	701.2 ^b	GMEPHM(4–9)
	30.79	1037.4 ^b	EQLEGMDSE(14–22)
	32.44	419.4 ^b	LRM(28–30) or MRL(51–53)
	32.70	412.2 ^b	MMM(30–32) or MMM(31–33)
	32.70	568.5 ^c	RMMM(29–32) or MMMR(31–34) or MMRM(32–35)
	33.31	826.9 ^b	MMMRMQ(31–36)
	34.26	1110.5 ^c	PRRGMEPHM(1–9)
	34.26	1013.9 ^b	RRGMEPHM(2–9)
Uncertain origin	38–39	5739.0	
		5755.0	

^a Peptides resulting from specific cleavage of trypsin/chymotrypsin. ^b Peptides resulting from nonspecific cleavage of trypsin/chymotrypsin. ^c Peptides obtained during gastric digestion (phase 1).

main IgE epitope region albumin (Fig. 6B) is located in a large peptide PRGEQMRRMMRLAENIPSRC NLSPMRCP(42–69) found following phase 1 digestion (Table 1, Fig. 6A). Although the IgE epitopes from these plant species have different amino acid sequences, they are all located in the same hypervariable region, which forms a very flexible loop between helices III and IV [20,24]. These helical regions contain Cys12, Cys13 and Cys25 residues (position numbers given according to the 2S albumin Brazil nut sequence), which are

involved in the formation of the intrachain disulphide bonds in the large subunit. These cysteine residues are present in several peptides that were identified in this study as being very resistant to proteolysis (Table 1, Fig. 6A). The hypervariable region of the Brazil nut 2S albumin corresponds to the fragment QEEMQPR GEQM(37–47) according to Monsalve *et al.* [25], which was also found to be largely intact (except for the C-terminal methionine) following gastric (phase 1) digestion (peak 4, Table 1, Fig. 6A).

Recently, Stickler *et al.* [26], by using synthetic peptides, determined the location of four immunodominant CD4⁺ T-cell epitopes in the unprocessed precursor of the Brazil nut 2S albumin. One of these epitopes matched with the signal and propeptide regions, and therefore would not be present in the mature protein, but the remainder corresponded to the large chain, with two also containing cysteine residues 12, 13, 23 and 25. It is therefore possible that the 'large fragment' identified in this study survives gastric and duodenal digestion and contains sufficient immunologically active structures (T-cell and B-cell epitopes) to potentially either sensitize an individual or elicit an allergic reaction. Further studies are underway to characterize the IgE binding to Brazil nut 2S albumin digestion products. This stresses the importance of studying their digestibility in physiologically relevant conditions and, in the case of structurally related allergen families, the elucidation of the 3D structure could help to gain a better understanding of their intrinsic allergenic properties.

Experimental procedures

Purification of Brazil nut 2S albumin (Ber e 1)

The main 2S albumin (Ber e 1) isoform (Expasy entry P04403) was purified to homogeneity by using gel filtration chromatography and gradient chromatofocusing on an anion-exchange column and then characterized by using proteomic techniques as described by Moreno *et al.* [13]. Ber e 1 was digested either before or after preheating at 100 °C for 20 min in 10 mM sodium phosphate buffer, pH 7, or 0.15 M NaCl, pH 2.5, adjusted with 1 M HCl (simulated gastric fluid, SGF). After heating, the samples were immediately cooled in ice.

In vitro gastric and duodenal models

Preparation of phospholipid vesicles

Egg L-PtdCho, grade 1, was obtained from Lipid Products (South Nutfield, Redhill, Surrey, UK) at 99% purity. The storage solvent was removed first under rotary evaporation and then under vacuum overnight in the absence of oxygen

(under nitrogen). The dry PtdCho was dispersed in warmed SGF by sonication at 5 °C (10 min set at 30% full power, 9/10 power cycle) using a Status Ultrasonic (Avestin, Canada) US200 homogenizer fitted with a TT13 titanium flat tip. Phospholipid vesicles were collected and filtered through Millex-HA 0.45 µm mixed cellulose (Millipore, Billerica, MA, USA) to remove titanium particles.

In vitro gastric digestion (phase 1)

Digestions were performed in either the presence or absence of PtdCho. In the former, the PtdCho solution was replaced by SGF, pH 2.5. Control samples, with no enzyme additions, were also analysed. 2S albumin was dissolved in SGF (5.55 mg·mL⁻¹), mixed with PtdCho vesicle solution (1 : 1.2, v/v) and the pH was adjusted to 2.5 with 1 M HCl, if necessary. After incubation at 37 °C for 15 min, a solution of pepsin (EC 3.4.23.1) 0.32% (w/v) in SGF, pH 2.5 (Sigma, Poole, Dorset, UK; product No. P 6887; activity: 3640 U·mg⁻¹ of protein calculated using haemoglobin as the substrate), was added at an approximately physiological ratio of enzyme/substrate (1 : 20, w/w); 182 U pepsin·mg⁻¹ of 2S albumin. This gave a final volume of 3.5–4 mL and a final concentration of 6.3 mM PtdCho and of 2.5 mg·mL⁻¹ 2S albumin in the final phase 1 digestion mix. The digestion was performed at 37 °C in an incubator with moderate agitation, and aliquots, which were withdrawn from a single digestion mixture, were taken at 0, 2, 5, 15, 30, 60 and 120 min for further analysis. The digestion was stopped by raising the pH to 7.5 by the addition of 40 mM ammonium bicarbonate (BDH, Poole, Dorset, UK).

In vitro duodenal digestion (phase 2)

In vitro duodenal digestion was performed by using 120-min gastric digests as the starting material. Although it has been described that the pH of the duodenum may vary within the range 5–7 [27–29], the most accurate range seems to be 6–6.5 [30–34]. Therefore, the pH of the digests was adjusted to 6.5 and the following were added (a) a bile salt mixture containing equimolar quantities (0.125 M) of sodium taurocholate (Sigma) and glycodeoxycholic acid (Calbiochem, La Jolla, CA, USA), (b) 1 M CaCl₂ (BDH), (c) 0.25 M Bistris, pH 6.5 (Sigma), (d) porcine pancreatic lipase (EC 3.1.1.3; 20 µL per 10 mL of total volume) (0.1% w/v; Sigma product no. L-0382; activity 25 600 U·mg⁻¹ of protein), and (e) porcine colipase (40 µL per 10 mL of total volume) (0.055%, w/v; Sigma product no. C3028) [35]. Finally, solutions of trypsin (EC 3.4.21.4; 0.1% w/v; Sigma product no. T 7418; activity: 13 800 U·mg⁻¹ of protein using *N*-benzoyl-L-arginine ethyl ester as the substrate) and α-chymotrypsin (EC 3.4.21.1; 0.4% w/v; Sigma product no. C 7762; activity 44 U·mg⁻¹ of protein using *N*-benzoyl-L-tyrosine ethyl ester as the substrate) in SGF, pH 7.0, were

prepared and added at approximately physiological ratios of 2S albumin (as denoted by the initial concentration in phase 1)/trypsin/ chymotrypsin, 1 : 400 : 100 (w/w/w); 1 mg/34.5 U/0.44 U. This gave the following final phase 2 digestion mix: 5.8 mM PtdCho, 2.3 mg·mL⁻¹ 2S albumin, 7.4 mM bile salts, 9.2 mM CaCl₂ and 24.7 mM Bistris. The digestion was performed at 37 °C and aliquots were taken at 0, 2, 5, 15, 30, 60 and 120 min for further analysis. The digestion was stopped either by heating at 80 °C for 5 min or by adding a solution of Bowman–Birk trypsin-chymotrypsin inhibitor from soybean (Sigma product no. T9777), at a concentration calculated to inhibit twice the amount of trypsin and chymotrypsin present in the digestion mix.

SDS/PAGE analysis

Samples taken at different stages of the digestion were analysed by SDS/PAGE. Digests (20 µL) were added to 17.5 µL of ultrapure water and to 12.5 µL of 4× NuPAGE[®] lithium dodecyl sulfate sample buffer [40% (w/v) glycerol, 0.1 M Tris/HCl buffer, pH 8.5, 8% (w/v) lithium dodecyl sulfate, 0.075% (w/v) Serva Blue G250 and 0.025% (w/v) Phenol Red, pH 8.5; Invitrogen, Carlsbad, CA, USA] and heated at 70 °C for 10 min. When required, samples were reduced with 0.5 M dithiothreitol. Samples (10 µL) were loaded onto a 12% polyacrylamide NuPAGE[®] Novex Bistris precast gel. A continuous buffer system (50 mL of 20× NuPAGE[®] Mes SDS running buffer with 950 mL of ultrapure water) was used. Gels were run for 35 min at 120 mA per gel and 200 V and then stained using a Colloidal Blue Staining Kit (Invitrogen). Marker proteins were: aprotinin (*M_r* 6500), α-lactalbumin (*M_r* 14 200), trypsin inhibitor (*M_r* 20 000), carbonic anhydrase (*M_r* 29 000), ovalbumin (*M_r* 45 000), BSA (*M_r* 66 000), β-galactosidase (*M_r* 116 000) and myosin (*M_r* 205 000) (Sigma).

RP-HPLC-ESI-MS

Digested 2S albumin samples (50 µL) were applied to either a peptide (Phenomenex Jupiter Proteo 90 Å pore size, 4 µm particle size, 250 × 4.6 mm internal diameter) or protein (Phenomenex Jupiter 300 Å pore size, 5 µm particle size, 250 × 4.6 mm internal diameter) column coupled to a Jasco PU-1585 triple pump HPLC equipped with an AS-1559 cooled autoinjector, CO-1560 column oven and UV-1575 UV detector (Jasco Ltd, Great Dunmow, Essex, UK). The HPLC was, in turn, attached to a Micromass Quattro II triple quadrupole mass spectrometer (Micromass, Manchester, UK). 2S albumins were eluted by using 0.1% (w/v) trifluoroacetic acid in double-distilled water as solvent A and 0.085% (w/v) trifluoroacetic acid in double-distilled water/acetonitrile (10 : 90, v/v) as solvent B. The column was equilibrated with 1% (v/v) solvent B. The elution was performed as follows: 0–5 min, 1% (v/v) solvent B in isocratic mode, and then as a linear gradient by increasing the

concentration of solvent B from 1% (v/v) to 50% (v/v) in 55 min. The HPLC column temperature was maintained at 25 °C and the autoinjector at 4 °C. The 1 mL·min⁻¹ mobile phase flow exiting the HPLC column was split by using an ASI 600 fixed ratio splitter valve (Presearch, Hitchin, Herts, UK) so that ≈ 200 µL·min⁻¹ entered the mass spectrometer; the remainder of the flow was diverted to the UV detector (215 nm monitored). The flow split was monitored by using a Humonics Optiflow 1000 flowmeter (Sigma) coupled to the outflow of the UV cell.

Mass spectra were obtained in positive ion electrospray mode by using a Micromass Z-sprayTM ion source. The electrospray probe was operated at 3.46 kV and at a cone voltage of 35 V. The source and desolvation temperatures were 120 °C and 300 °C, respectively. The nitrogen nebulizing and drying gas flow rate were optimized at 15 L·h⁻¹ and 500 L·h⁻¹, respectively. The mass range *m/z* 300–2200 was scanned every 5 s in continuum mode, with an interscan time of 0.2 s. Data were processed by using MASSLYNX 3.4 software (Micromass). Search against a database (ExPASy, <http://us.expasy.org/>) of expected proteolysis fragments deduced from the known Brazil nut 2S albumin sequence (no. P04403) was performed using the following search parameters (a) peptide masses were stated to be monoisotopic, and (b) the mass tolerance was maintained at 0.5 Da.

MALDI-TOF-MS

Prior to analysis, gastric (phase 1) digests were subjected to microdialysis against 10 mM ammonium bicarbonate overnight at 2 °C by using the Micro Dispodialyzer membrane cut-off of 1000 Da (Harvard Apparatus Inc., Holliston, MA, USA). 2S albumin (50 µL, 0.125 mg) was reduced with 10 mM dithiothreitol (50 µL) dissolved in 10 mM ammonium bicarbonate and incubated at 65 °C for 30 min. The protein digest was acidified and spotted directly onto a thin layer of matrix on a stainless steel target plate. The matrix consisted of four parts of a saturated solution of α-cyano-4-hydroxycinnamic acid in acetone mixed with one part of a 1 : 1 (v/v) mixture of acetone/isopropanol containing 10 mg·mL⁻¹ nitrocellulose. Analysis was carried out using a Reflex III MALDI-TOF mass spectrometer (Bruker UK Ltd, Coventry, UK). A nitrogen laser was used to desorb/ionize the matrix/analyte material, and ions were detected in positive ion reflectron mode. Spectra were obtained over the *m/z* range 1610–8430 and calibrated using peptide standards obtained from Sigma (bombesin, adrenocorticotrophic hormone clip 1–17 and clip 18–39, somatostatin and insulin). The acceleration voltage was set to 25 kV, the reflection voltage to 28.7 kV, the ion source acceleration voltage to 21.1 kV, and the reflector-detector voltage to 1.65 kV. Peptide mass fingerprints were searched as described above. Peptides resulting from autolysis of the proteases were observed by analysis control digests to which only enzymes (and no Ber e 1) was added.

Sequencing by nanoelectrospray Q-TOF MS/MS

Prior to analysis, microdialysis was carried out as described above. The resulting sample was diluted 100-fold by using a solution of 50% (v/v) acetonitrile containing 0.2% (v/v) formic acid, and sprayed from the tip of a borosilicate nanospray needle directly into the Z-sprayTM ion source of a quadrupole TOF mass spectrometer (Micromass). The nanospray probe was operated at 0.8 kV and at a cone voltage of 35 V. Mass spectra were acquired over the *m/z* range 50–1600. The most intense multiply charged ions were, in turn, selected for collision-induced dissociation. Fragment-ion spectra (MS/MS spectra) were acquired over the *m/z* range 50–2500 and were sequenced by using the PepSeq *de novo* sequencing algorithm (Micromass).

Acknowledgements

This research was supported by a Marie Curie Fellowship of the European Community programme 'Quality of Life and Management of Living Resources' under contract number QLK1-CT-2001–51997 and the BBSRC competitive strategic grant to IFR. MALDI-TOF and Q-TOF mass spectrometry was carried out in the Joint IFR-JIC Proteomics Facility which is funded in part by BBSRC (JREI grant numbers: JRE10832, JE412701, JE412631) and by grants from Syngenta and Unilever.

References

- Kreis M, Forde BG, Rahman S, Mifflin BJ & Shewry PR (1985) Molecular evolution of the seed storage proteins of barley, rye and wheat. *J Mol Biol* **183**, 499–502.
- Mills ENC, Madsen C, Shewry PR & Wichers HJ (2003) Food allergens of plant origin – their molecular and evolutionary relationships. *Trends Food Sci Technol* **14**, 145–156.
- Ewan PW (1996) Clinical study of peanut and nut allergy in 62 consecutive patients: new features and associations. *Br Med J* **312**, 1074–1078.
- Pastorello EA, Farioli L, Pravettoni V, Ispano M, Conti A, Ansaloni R, Rotondo F, Incorvaia C, Bengtsson A, Rivolta F, Trambaioli C, Previdi M & Ortolani C (1998) Sensitization to the major allergen of Brazil nut is correlated with the clinical expression of allergy. *J Allergy Clin Immunol* **102**, 1021–1027.
- Shewry PR, Napier JA & Tatham AS (1995) Seed storage proteins: structures and biosynthesis. *Plant Cell* **7**, 945–956.
- Astwood JD, Leach JN & Fuchs RL (1996) Stability of food allergens to digestion *in vitro*. *Nat Biotechnol* **14**, 1269–1273.
- Murtagh GJ, Archer DB, Dumoulin M, Ridout S, Matthews S, Arshad SH & Alcocer MJC (2003) *In vitro* stability and immunoreactivity of the native and recombinant plant food 2S albumins Ber e 1 and SFA-8. *Clin Exp Allergy* **33**, 1147–1152.
- Metcalfe DD, Astwood JD, Townsend R, Sampson HS, Taylor SL & Fuchs RL (1996) Assessment of the allergenic potential of foods derived from genetically engineered crop plants. *Crit Rev Food Sci Nutr* **36**(S), S165–S186.
- Taylor SL & Hefle SL (2001) Will genetically modified foods be allergenic? *Curr Rev Allergy Clin Immunol* **107**, 765–771.
- Taylor SL (2002) Protein allergenicity assessment of foods produced through agricultural biotechnology. *Annu Rev Pharmacol Toxicol* **42**, 99–112.
- Mills ENC, Moreno J, Sancho A, Jenkins JA & Wichers HJ (2004) Processing approaches to reducing the allergenicity of foods. In *Proteins in Food Processing* (Yada R, ed.), pp. 396–418. Woodhead Publishing Ltd, Cambridge, UK.
- FAO/WHO (2002) *Report of the Third Session of the Codex ad hoc Intergovernmental Task Force on Foods Derived from Biotechnology*. Yokohama, Japan. Food and Agriculture Organization of the United Nations (FAO), Rome, Italy.
- Moreno FJ, Jenkins JA, Mellon FA, Rigby NM, Robertson JA, Wellner N & Mills ENC (2004) Mass spectrometry and structural characterization of 2S albumin isoforms from Brazil nuts (*Bertholletia excelsa*). *BBA-Proteins Proteomics* **1698**, 175–186.
- Monsalve RI, Menéndez-Arias L, López-Otín C & Rodríguez R (1990) β -turns as structural motifs for the proteolytic processing of seed proteins. *FEBS Lett* **263**, 209–212.
- Monsalve RI, López-Otín C, Villalba M & Rodríguez R (1991) A new distinct group of 2S albumins from rapeseed. *FEBS Lett* **295**, 207–210.
- Murén E, Ek B & Rask L (1995) Processing of the 2S storage protein pronapin in *Brassica napus* and in transformed tobacco. *Eur J Biochem* **227**, 316–321.
- Gehrig PM, Krzyzaniak A, Barciszewski J & Biemann K (1996) Mass spectrometric amino acid sequencing of a mixture of seed storage proteins (napin) from *Brassica napus*, products of a multigene family. *Proc Natl Acad Sci USA* **93**, 3647–3652.
- da Silva JG Jr, Machado OLT, Izumi C, Padovan JC, Chait BT, Mirza UA & Greene LJ (1996) Amino acid sequence of a new 2S albumin from *Ricinus communis* which is part of a 29-kDa precursor protein. *Arch Biochem Biophys* **336**, 10–18.
- Ampe C, Van Damme J, de Castro LAB, Sampaio MJAM, Van Montagu M & Vandekerckhove J (1986) The amino-acid sequence of the 2S sulphur-rich proteins

- from seeds of Brazil nut (*Bertholletia excelsa* H.B.K.). *Eur J Biochem* **159**, 597–604.
- 20 Pantoja-Uceda D, Bruix M, Giménez-Gallego G, Rico M & Santoro J (2003) Solution structure of allergenic 2 S albumins. *Biochem Soc Trans* **30**, 919–924.
- 21 Menéndez-Arias L, Domínguez J, Moneo I & Rodríguez R (1990) Epitope mapping of the major allergen from yellow mustard seeds, *Sin a I*. *Mol Immunol* **27**, 143–150.
- 22 Monsalve RI, González de la Peña MA, Menéndez-Arias L, López-Otín C, Villalba M & Rodríguez R (1993) Characterization of a new oriental-mustard (*Brassica juncea*) allergen, *Bra j 1E*: detection of an allergenic epitope. *Biochem J* **293**, 625–632.
- 23 Robotham JM, Teuber SS, Sathe SK & Roux KH (2002) Linear IgE epitope mapping of the English walnut (*Juglans regia*) major food allergen, *Jug r 1*. *J Allergy Clin Immunol* **109**, 143–149.
- 24 Pantoja-Uceda D, Bruix M, Jiménez-Gallego G, Rico M & Santoro J (2003) Solution structure of RicC3, a 2S albumin storage protein from *Ricinus communis*. *Biochemistry* **42**, 13839–13847.
- 25 Monsalve RI, Villalba M & Rodríguez R (2001) Allergy to mustard seeds: the importance of 2S albumins as food allergens. *Internet Symp Food Allergens* **3**, 57–69.
- 26 Stickler M, Mucha J, Estell D, Power S & Harding F (2003) A human dendritic cell-based method to identify CD4⁺ T-cell epitopes in potential protein allergens. *Environ Health Persp* **111**, 251–254.
- 27 Benn A & Cooke WT (1971) Intraluminal pH of duodenum and jejunum in fasting subjects with normal and abnormal gastric or pancreatic function. *Scand J Gastroenterol* **6**, 313–317.
- 28 Friend DR (1991) Colon-specific drug delivery. *Adv Drug Delivery Rev* **7**, 149–199.
- 29 Dressman JB, Amidon GL, Reppas C & Shah VP (1998) Dissolution testing as a prognostic tool for oral drug absorption: immediate release dosage forms. *Pharm Res* **15**, 11–22.
- 30 Evans WJ & Pierce AG (1982) Interaction of phytic acid with the metal-ions, copper (II), cobalt (II), iron (III), magnesium (II), and manganese (II). *J Food Sci* **47**, 1014–1015.
- 31 Denigris SJ, Hamosh M, Kasbekar DK, Fink CS, Lee TC & Hamosh P (1985) Secretion of human gastric lipase from dispersed gastric glands. *Biochim Biophys Acta* **836**, 67–72.
- 32 O'Connor CJ & Wallace RG (1985) Physico-chemical behavior of bile salts. *Adv Colloid Interface Sci* **22**, 1–111.
- 33 Fallingborg J, Christensen LA, Jacobsen BA, Ingemann Nielsen M, Rasmussen HH, Abildgaard K & Rasmussen SN (1994) Effect of olsalazine and mesalazine on intraluminal pH of the duodenum and proximal jejunum in healthy humans. *Scand J Gastroenterol* **29**, 498–500.
- 34 Ambler J, Baker E, Bentley D, Brown L, Butler K, Butler P, Farr D, Dunnet K, Le Grand D, Hayler J *et al.* (1999) Studies towards the identification of potent, selective and bioavailable thrombin inhibitors. *Bioorg Med Chem Lett* **9**, 737–742.
- 35 Rich GT, Faulks RM, Wickham MSJ & Fillery-Travis A (2003) Solubilization of carotenoids from carrot juice and spinach in lipid phases. II. Modeling the duodenal environment. *Lipids* **38**, 947–956.
- 36 Notredame C, Higgins DG & Heringa J (2000) T-Coffee: a novel method for fast and accurate multiple sequence alignment. *J Mol Biol* **302**, 205–217.
- 37 Pantoja-Uceda D, Shewry PR, Bruix M, Tatham AS, Santoro J & Rico M (2004) Solution structure of a methionine-rich 2S albumin from sunflower seeds: relationship to its allergenic and emulsifying properties. *Biochemistry* **43**, 6976–6986.