DOI 10.1002/mnfr.200700524 Mol. Nutr. Food Res. 2008, *52*, S272 – S285

Research Article

Purification and characterisation of a panel of peanut allergens suitable for use in allergy diagnosis

Justin Marsh[1]*, Neil Rigby[2]*, Klaus Wellner[2], Gerald Reese[3], André Knulst[4], Jaap Akkerdaas[5], Ronald van Ree[5], Christian Radauer[6], Alison Lovegrove[1], Ana Sancho[2], Clare Mills[2], Stefan Vieths[3], Karin Hoffmann-Sommergruber[6] and Peter R. Shewry[1]

[1] Rothamsted Research, Harpenden, Hertfordshire, UK
[2] Institute of Food Research, Norwich, UK
[3] Paul-Ehrlich-Institut, Langen, Germany
[4] Department of Dermatology/Allergology, University Medical Centre Utrecht, Utrecht, The Netherlands
[5] Department of Experimental Immunology, Academic Medical Center, University of Amsterdam, Amsterdam, The Netherlands
[6] Department of Pathophysiology, Medical University of Vienna, Vienna, Austria

Peanut is a major cause of type 1 hypersensitive reactions including anaphylaxis. This results from the presence of a number of protein allergens, six of which are being studied as part of the EU FP6 EuroPrevall programme. These are Ara h 1 (7S globulin), Ara h 2, Ara h 6 (2S albumins), Ara h 3/4 (11S globulins) and Ara h 8 (Bet v 1 homologue). Methods for the purification of Ara h 1, Ara h 3/4, Ara h 2 and Ara h 6 from peanut seeds and for the production of recombinant Ara h 8 in *Escherichia coli* are described with spectroscopic analyses being used to confirm that they are authentically folded. *N*-terminal sequencing of the proteins purified from peanut seeds also revealed details of the differences between isoforms and their generation by proteolytic processing within the seed. Preliminary IgE binding studies of the purified allergens confirmed that they retained their immunological properties indicating their suitability for use in allergy diagnosis.

Keywords: Allergens / Characterisation / Peanuts / Purification

Received: December 10, 2007; revised: April 10, 2008; accepted: April 19, 2008

1 Introduction

Peanut (*Arachis hypogaea*), also called groundnut, is a major crop in the tropics and warm temperate regions [1]. This is due to its high contents of oil and protein and its wide adaptability which allow it to be cultivated in over 80 countries. It is a member of the genus *Arachis* which also includes a number of wild species and is native to South America. Although most species of *Arachis* are diploids, cultivated peanut is a tetraploid, with A and B genomes thought to be related to those present in wild diploid species [1].

Whole peanuts are consumed either intact after roasting or boiling, as peanut butter and in a range of other products, particularly confectionery [2], while peanut oil is widely used in food preparations and in creams and cosmetics. However, peanut is also a major cause of severe type 1 hypersensitive reactions [3, 4] with the highest potential of any food allergen to induce anaphylactic shock [5, 6]. Peanut allergy affects children as well as adults and persists into adulthood in the majority of patients, unlike allergies to eggs and milk.

A number of allergens have been characterised from peanut and in all cases these belong to protein families which include allergens in other foods. The major allergens are seed storage proteins of the 7S globulin (Ara h 1) [7–11] and 2S albumin (Ara h 2) [12] families with the latter family also containing two minor 2S albumin allergens (Ara h 6 and Ara h 7) [9, 13]. Agglutinin (lectin) [14], 11S storage globulin (Ara h 3/4) [8, 13, 15], profilin (Ara h 5) [13], Bet v 1-related (Ara h 8) [16] and oleosin [17, 18] proteins have

Correspondence: Dr. Peter R. Shewry, Rothamsted Research, Harpenden, Hertfordshire AL5 2JQ, UK
E-mail: peter.shewry@bbsrc.ac.uk
Fax: +44-1582-763010

Abbreviations: CD, circular dichroism; **CV,** column volume; **FT-IR,** fourier transform infrared; **IMAC,** immobilised metal affinity chromatography; **MWCO,** M_r cut off

* These authors contributed equally to this work.

also been described as minor allergens. IgE crossreactivity between peanut allergens and related seed storage proteins from lupins and soya has also been frequently observed, although its clinical significance is rather low [19]. Cross-reactivity between Bet v 1 from birch pollen and the homologous protein in peanut has also recently been shown [20]. Most of the peanut allergens are seed storage proteins which comprise heterogeneous mixtures of closely related components. This poses a challenge for their purification and characterisation, and for the preparation of pure fractions for allergy diagnosis. The present paper therefore provides detailed methods for the purification and characterisation of the peanut allergens Ara h 1 (7S globulin), Ara h 2, Ara h 6 (2S albumin), Ara h 3/4 (11S globulin) and Ara h 8 (Bet v 1). Classical biochemical methods were used to purify the abundant albumin and globulin storage proteins. In contrast, the less abundant and highly labile Ara h 8 protein was prepared by recombinant expression in *Escherichia coli*, following a different strategy to that reported previously by Mittag *et al.* [20]. The panel of allergens prepared by these methods is suitable for allergy diagnosis, and is being used for this purpose within the EU FP6 EuroPrevall programme.

2 Materials and methods

Red shelled peanuts were obtained from local suppliers, representing the type of peanut consumed widely in Europe. After shelling, the skins were removed and the nuts ground in liquid nitrogen using a Waring blender with a stainless steel beaker. The meal was then defatted by stirring with 5 – 10 vol of hexane for 1 h, recovered by filtration on a glass sinter and the defatting repeated at least once. The meal was then allowed to dry in the air.

2.1 Purification of native peanut 2S albumins – nAra h 2 and nAra h 6

Proteins were extracted by stirring 40 g of defatted peanut meal in 400 mL of 50 mM Tris/HCl, pH 8.2, at 4°C, overnight. The resulting suspension was clarified by centrifugation (10 000 × *g* for 10 min, 4°C) and the supernatant filtered through four layers of cheesecloth prior to ammonium sulphate fractionation. The fraction precipitated between 40 and 80% ammonium sulphate saturation at 4°C, was collected by centrifugation (10 000 × *g* for 20 min, 4°C) and resuspended in 200 mL of 20 mM Tris/HCl, pH 8.0 containing 1 mM EDTA, prior to dialysis against 10 L of 20 mM Tris/HCl, pH 8.0 (overnight, 4°C) using 3.5 kDa MW cut off (MWCO) dialysis tubing (Pierce, Rockford, IL, USA). The dialysate was clarified by centrifugation (30 000 × *g* for 30 min, 4°C) and the supernatant passed through a 0.2 µm filter (Whatman, Maidstone, Kent) prior to chromatography.

The crude extract was loaded onto a Superdex 75 column (bed dimension = 10 × 150 mm^2, GE Healthcare, Little Chalfont, UK) attached to a 2138 Uvicord S detector (GE Healthcare). The column was equilibrated and eluted with 20 mM Tris/HCl, pH 8, at a flow rate of 0.75 mL/min. The absorbance of the eluate was monitored at 280 nm and peak fractions analysed by SDS-PAGE. Those containing components of approximately 10 – 25 kDa were pooled, and concentrated by ultrafiltration with an Amicon Ultra-15 filter with a MWCO of 5.0 kDa (Millipore, Hertfordshire, UK). Fractions from gel filtration were loaded onto an anion exchange column (HQ-Poros, 4.6 × 100 mm^2, PerSeptive Biosystems, MA, USA) attached to a BioCad Sprint HPLC system (Applied Biosystems, Cheshire, UK), equilibrated with 20 mM Tris/HCl, pH 8. Bound proteins were eluted at a flow rate of 5 mL/min using a step gradient as follows: five column volumes (CV) of 20 mM Tris/HCl, pH 8; 15 CV of 20 mM Tris/HCl, pH 8, 62.5 mM NaCl; five CV of 20 mM Tris/HCl, pH 8, 125 mM NaCl and five CV of 20 mM Tris/HCl, pH 8, 250 mM NaCl. Ion-exchange fractions were dialysed against water and concentrated by ultrafiltration with an Amicon Ultra-15 filter with an MWCO of 5.0 kDa (Millipore). They were then loaded onto a Vydac C18 preparative RP-HPLC column attached to a Beckman 126 solvent module and a 166 detector. The bound proteins were eluted with an ACN gradient from 15 – 40% using 1:2000 v/v, trifluoracetic acid (TFA), water as buffer A and 1:2000 v/v, TFA, ACN as buffer B. Peaks eluting at 25 and 30 min were shown to be enriched in Ara h 2 and Ara h 6 by SDS-PAGE and were pooled prior to dialysis against water and concentrated by ultrafiltration as above.

2.2 Purification of native peanut 7S globulins – nAra h 1

Ara h 1 was purified essentially as described by Eiwegger *et al.* [21] with the following modifications: defatted peanut meal (10 g) was extracted by stirring with water (1:5 w/v) containing 6.2 mM sodium azide and protease inhibitor (2 tablets/100 mL, Roche Complete Protease inhibitor tablet, Roche Diagnostics, Sussex, UK); for 1 h at 20°C. The extract was centrifuged at 2000 × *g* for 10 min at 20°C, the supernatant collected and recentrifuged at 34 000 × *g* for 40 min at 10°C. Solid ammonium sulphate was added to a final concentration of 70% saturation at 0°C and precipitation allowed to proceed for 1 h at 0°C in an ice bath. The supernatant was collected by centrifugation at 40 000 × *g* for 40 min at 1 – 2°C and dialysed overnight against buffer (1:60 v/v, 20 mM Tris, 500 mM NaCl, 3 mM NaN_3, pH 7.5) at 4°C using 12 kDa MWCO dialysis tubing. Aliquots (10 mL) of the supernatant were applied at 2 mL/min to a column (1.6 × 10 cm^2) of ConA Sepharose-4B affinity chromatography medium (GE Healthcare) attached to a BioCad Sprint HPLC system (Applied Biosystems) and equilibrated with 20 mM Tris-HCl pH 7.5 containing

500 mM NaCl, 3 mM NaN_3. After washing, bound protein was eluted using equilibration buffer containing 0.4 M methyl mannopyranoside. Peak fractions containing partially purified peanut 7S globulins (Ara h 1) were collected, pooled and concentrated by ultrafiltration using an Amicon stirred cell equipped with a regenerated cellulose membrane with a 10 kDa MWCO membrane (Millipore). Aliquots (2 mL) of the crude Ara h 1 preparation were then subjected to gel permeation chromatography using a Superdex Preparative Grade S200 column (GE Healthcare) attached to a BioCad Sprint HPLC system (Applied Biosystems) equilibrated in 50 mM sodium phosphate buffer pH 7.0, containing 200 mM NaCl and eluted with a flow rate of 1 mL/min. Pooled fractions were dialysed in regenerated cellulose dialysis tubing with a 12 kDa exclusion limit at 4°C against 0.5 M $(NH_4)_2CO_3$ before being freeze dried and stored at −20°C until required.

2.3 Purification of native peanut 11S globulins – nAra h 3/4

Defatted peanut meal (30 g) was stirred for 2 h at 20°C with 10 mL/g of 20 mM Tris, pH 7.2, buffer containing 15.4 mM sodium azide and protease inhibitor (two Roche Complete Protease inhibitor tablets (Roche Diagnostics) *per* 100 mL).

The extract was centrifuged at $2000 \times g$ for 10 min at 20°C and the supernatant removed and recentrifuged at $30\,000 \times g$ for 40 min at 10°C. Aliquots (1 mL) were applied to an anion exchange column (Poros HQ20, 4.6×100 mm^2, PerSeptive Biosystems, MA, USA) attached to an AKTA BASIC FPLC™ system (GE Biosciences) equilibrated in 20 mM Tris pH 7.2 using a flow rate of 4 mL/min. Bound protein was eluted using a 0–600 mM NaCl gradient in equilibration buffer. Fractions identified as containing peanut 11S globulins by SDS-PAGE were pooled and concentrated by ultrafiltration using an Amicon stirred cell equipped with a regenerated cellulose membrane with a 10 kDa MWCO (Millipore). Aliquots (2 mL) of this material was then subjected to gel permeation chromatography on a Superdex Preparative Grade S200 column (GE Healthcare) attached to a BioCad Sprint HPLC system (Applied Biosystems) equilibrated in 50 mM Tris-HCl pH 7.5, containing 200 mM NaCl. SDS-PAGE of the peak fraction demonstrated the presence of traces of Ara h 1. This was removed by passing the partially purified peanut 11S globulins through a column of ConA Sepharose (1.6×8 cm^2) equilibrated with 50 mM Tris-HCl pH 7.5, containing 200 mM NaCl. The unbound fraction containing highly purified Ara h 3/4 was concentrated by ultrafiltration using an Amicon stirred cell equipped with a regenerated cellulose membrane with a 10 kDa MWCO and dialysed in regenerated cellulose dialysis tubing with a 12 kDa MWCO at 4°C against 0.5 M $(NH_4)_2CO_3$ before being freeze dried and then stored at −20°C until required.

2.4 Analytical gel-filtration, SDS-PAGE and protein determination

Protein concentrations were determined by calibration against BSA standards using the Pierce BCA assay.

Analytical gel-permeation chromatography was carried out using a column (1.6×60 cm^2) of preparative grade Superdex S200 equilibrated and eluted with 25 mM Tris/HCl pH 7.4 containing 200 mM NaCl, 3 mM sodium azide. The column was calibrated with a set of gel filtration MW standards (BioRad, Hertfordshire, UK) and the absorbance monitored at 280 nm.

SDS-PAGE analysis was performed using two systems. For 2S albumins, 4–12% gradient gels were run in MES buffer (2-(*N*-mopholino)ethanesulphonic acid). Reduced samples were prepared by heating with 50 mM 1,4-dithioerythritol. Reference markers of known MW (mark 12 unstained standard (Invitrogen, CA, USA) or Pageruler Prestained protein ladder (Fermentas, Ontario, Canada)) were run on the same gel. Gels were fixed in methanol/water/acetic acid (40:50:10 by volume), stained using Colloidal Coomassie (Invitrogen) and scanned using a GS800 calibrated densitometer and Quantity One software (BioRad). Globulins (7S and 11S) were separated by SDS-PAGE as above, but the gels were stained using SimplyBlue safestain (Invitrogen).

2.5 Generation of recombinant peanut Bet v 1 homologue – rAra h 8

A synthetic gene construct Nde_I-MRGS-Ara_h_8–6xHis-Stop-Stop-Xho_I (GeneArt, Regensburg, Germany), optimised for expression in *E. coli* was cloned into the pET23(+) vector (Merck, Darmstadt, Germany). According to Bujard *et al.* [22] the MRGS sequence (ATG–AGA–GGA–TCC) represents a strong initiation and enhancer motif.

Single colonies were grown in LB medium containing 50 μg/mL of ampicillin and 30 μg/mL of kanamycin at 28°C for 2–4 h. Protein expression was induced by raising the temperature to 42°C and continued overnight. Cells were harvested by centrifugation (20 min, $5000 \times g$, 4°C), resuspended in lysis buffer (50 mM NaH_2PO_4, 300 mM NaCl, 20 mM imidazole, 8 M urea) and stored at −80°C overnight. The lysate was thawed at 37°C in a waterbath and centrifuged at $20\,000 \times g$ at room temperature (RT) for 1 h. The supernatant was filtered (5 μm followed by 1.2 μm cellulose acetate filter (Sartorius, Goettingen, Germany) and diluted 1:2 with lysis buffer without urea to reduce the urea concentration to 4 M. The recombinant Ara h 8 was purified *via* immobilised metal affinity chromatography (IMAC) using metal chelate spin columns (Sartorius) loaded with Ni^{2+} ions and equilibrated twice with equilibration buffer (50 mM NaH_2PO_4, 300 mM NaCl, 10 mM imidazole). After binding to the column the protein was

refolded by several washing steps (50 mM NaH_2PO_4, 300 mM NaCl, 20 mM imidazole) with decreasing amounts of urea. Recombinant Ara h 8 was eluted stepwise with buffer containing 250 mM, 500 mM and 1 M imidazole. Protein in eluted fractions was detected by SDS-PAGE and CBB staining (Pierce) and pooled. Attempts to concentrate rAra h 8 failed, using either Vivaspins (MWCO 10 000), precipitation with acetone, precipitation with ammonium sulphate, dialysis against 20% PEG 200 000 or addition of 1% v/v sucrose, glycerol or sorbitol before concentration with Vivaspins. The SEC purification step could not be performed since the concentration could not be raised above approximately 350 μg/mL.

2.6 *N*-terminal sequencing

N-terminal sequences were determined using an Applied Biosystems Procise 491 sequencer (Applied Biosystems). nAra h 1 (50 pmol) was adsorbed on a Prosorb cartridge and subjected to sequence analysis. Alternatively, Ara h 2, Ara h 6 and Ara h 3/4 were subjected to SDS-PAGE and blotted onto PVDF membranes. In order to reduce the glycine background signal in the first sequencing cycle, the electrotransfer was performed in 10 mM sodium 3-(cyclohexylamino)-propanesulfonate buffer, pH 11.0 containing 10% v/v methanol. Membranes were stained with 0.1% w/v CBB R250 in 50% v/v methanol, 1% v/v acetic acid for 1 min, destained with several changes of 50% v/v methanol, 1% v/v acetic acid and extensively washed with water. Individual bands were excised and subjected to *N*-terminal sequencing. Subsequently, sequence data were compared with the protein databases using the BLAST program (www.ncbi.nlm.nih.gov/blast/).

2.7 Circular dichroism (CD) and fourier transform infrared (FT-IR) spectroscopy

FT-IR spectra of freeze-dried purified globulins Ara h 1 and Ara h 3/4 were obtained by averaging 128 scans at 2 cm^{-1} resolution on a BioRad FTS 175C FT-IR spectrometer equipped with a Specac ‚Golden Gate' ATR module.

Far-UV CD spectra (270 – 190 nm) of Ara h 1, Ara h 3/4, Ara h 2 and Ara h 6 were recorded using a Jasco J-710 Spectropolarimeter (Jasco, Tokyo, Japan) with a 0.1 or 0.2 cm path length quartz cell (Hellma UK, Essex, UK). Protein was prepared in 10 mM phosphate buffer, pH 7.6, by dialysing a solution against the measurement buffer (2 × 500 mL for 30 min) using a mini dialysis unit (3.5 kDa membrane, Perbio Science UK, Northumberland, UK). Spectra represent the average of nine accumulations collected at 50 nm/min with a 2 s time constant, 0.5 nm resolution and sensitivity of ±100 mdeg. Protein concentrations were determined by measurement of the UV spectra of the dialysed protein solutions and by using experimentally determined extinction coefficients at 280 nm. The mean molar ellipticity *per* residue (with respect to moles of amide bond) was calculated using CD Pro.

Recombinant Ara h 8 was dialysed against 10 mM KH_2PO_4/K_2HPO_4 buffer (pH 7.4), and the concentration adjusted to 5.2 μM. CD spectroscopy was performed on a J-180 S spectropolarimeter (Jasco) with constant nitrogen flushing at 20°C. For wavelength analysis, Ara h 8 was scanned with a step width of 0.2 nm and a band width of 1 nm. The spectral range was 185 – 255 nm at 50 nm/min. Ten scans were accumulated and the mean residue molar ellipticity *per* residue was calculated.

2.8 Patients and sera

Sera ($n = 6$; serum nos. 3 – 8, Fig. 8) were obtained from patients with established food allergy to peanut according to convincing case history, positive peanut RAST (≥3) and positive skin prick tests. Upon consuming peanuts the patients experienced generalised symptoms ranging from skin reactivity to hypotension and in one case anaphylaxis (serum no. 8). In addition, sera ($n = 2$, serum nos: 1, 2, Fig. 8) from patients with peanut and birch pollen allergy (RAST ≥ 3) were identified. For control experiments, sera from healthy persons ($n = 5$) were identified and pooled. The sera were stored at –20°C until use.

2.9 IgE ELISA

IgE ELISA was performed as described previously [23]. Briefly, microtiter plates (Nunc Maxisorp, Nalge Nunc International, Roskilde, Denmark) were coated with 1 μg of protein *per* well and nonspecific binding blocked by incubation with Tris-buffered saline (TBS) containing 0.5% v/v Tween 20 and 3% w/v nonfat dried milk. After five washing steps, sera (1:4 dilution) were applied to the plates and incubated overnight at 4°C. After washing, plates were incubated with a 1:1000 diluted alkaline phosphatase (AP)-conjugated mouse anti-human IgE antibody (BD Pharmingen, San Diego, CA, USA) for 2 h at RT. Bound IgE was detected after adding disodium *p*-nitrophenyl phosphate as substrate (Sigma – Aldrich, Steinheim, Germany) and the OD measured at 405 nm. Buffer and sera of three nonallergic subjects were used as controls. OD values were counted positive if they exceeded the mean OD of the negative controls plus three SDs.

2.10 IgE immunoblotting

For immunoblotting purified nAra h 1, nAra h 2 and nAra h 6, nAra h 3/4 and rAra h 8 (1 μg each) were subjected to 15% SDS-PAGE. The proteins were then electrotransferred to nitrocellulose (NC) membranes. After blocking, strips were incubated with sera from peanut allergic patients (1:5 dilution in TBST, 0.5% BSA) overnight at 4°C and bound IgE was detected by ^{125}I-labelled rabbit α-human IgE (1:40

dilution in TBST, 0.5% BSA; MALT Allergy System Isotope Reagent, IBL Hamburg, Germany). Buffer and normal human serum (NHS) were used for control experiments. For immunoblot inhibition of total protein extracts sera were pooled (pool 1 serum nos. 3 – 8 and serum pool 2: serum nos 1 and 2). IgE binding of either serum pool 1 or serum pool 2 to total peanut protein extract was performed as described above. In parallel, serum pool 1 was preincubated with purified Ara h 1 (200 µg/mL), Ara h 3/4 (200 µg/mL) and Ara h 2 and Ara h 6 (100 µg/mL each), respectively. Serum 2 was preincubated with rAra h 8 (200 µg/mL). Subsequently, residual IgE binding activity to blotted peanut proteins was investigated as described above.

3 Results and discussion

3.1 Purification of allergens

3.1.1 Native 2S albumins – nAra h 2 and nAra h 6

SDS-PAGE of fractions from gel filtration chromatography (Fig. 1a) of a crude extract of peanut meal showed that Ara h 2 and Ara h 6 were concentrated in several minor peaks. These peaks were collected from a number of runs, bulked and concentrated to a final volume of 150 mL. SDS-PAGE of this fraction showed the presence of three major bands, two of approximately 17 kDa which were assumed to correspond to Ara h 2 and Ara h 2.02 [24] and a third of about 15 kDa which was shown by MS to correspond to Ara h 6 [25]. The combined fractions from gel-filtration were therefore separated by anion exchange chromatography with elution with a stepped salt gradient (Fig. 1b). SDS-PAGE showed that the fractions from this separation were highly enriched in Ara h 6 (Peak 1) and in Ara h 2 and Ara h 2.02 (peaks 2 and 3). After dialysis and concentration, these enriched fractions were separated by RP-HPLC using a Vydac C18 preparative column (Fig. 1c). All fractions gave a peak at about 25 min which contained two proteins of M_r about 17 kDa corresponding to Ara h 2/2.02 and a peak at about 30 min which contained a major protein of about 15 kDa corresponding to Ara h 6 (Fig. 1d). However, the relative sizes of these two peaks varied in relation to the proportions of the putative Ara h 2/2.02 and Ara h 6 components in the fractions loaded (compare Figs. 1c and d).

3.1.2 Native 7S globulins – nAra h 1

Ara h 1 is glycosylated and was therefore readily purified from defatted peanut meal using a lectin affinity column (with concanavalin A which binds to mannose present in the glycan chain) (Fig. 2a) followed by gel filtration chromatography (Fig. 2b) as described by Eiwegger *et al*. [21]. SDS-PAGE (Fig. 2c) showed a major band of 67 kDa with a minor band of about 33 kDa. Analytical gel permeation chromatography (Fig. 2b inset) showed that the native trimeric protein has a mass of about 235 kDa.

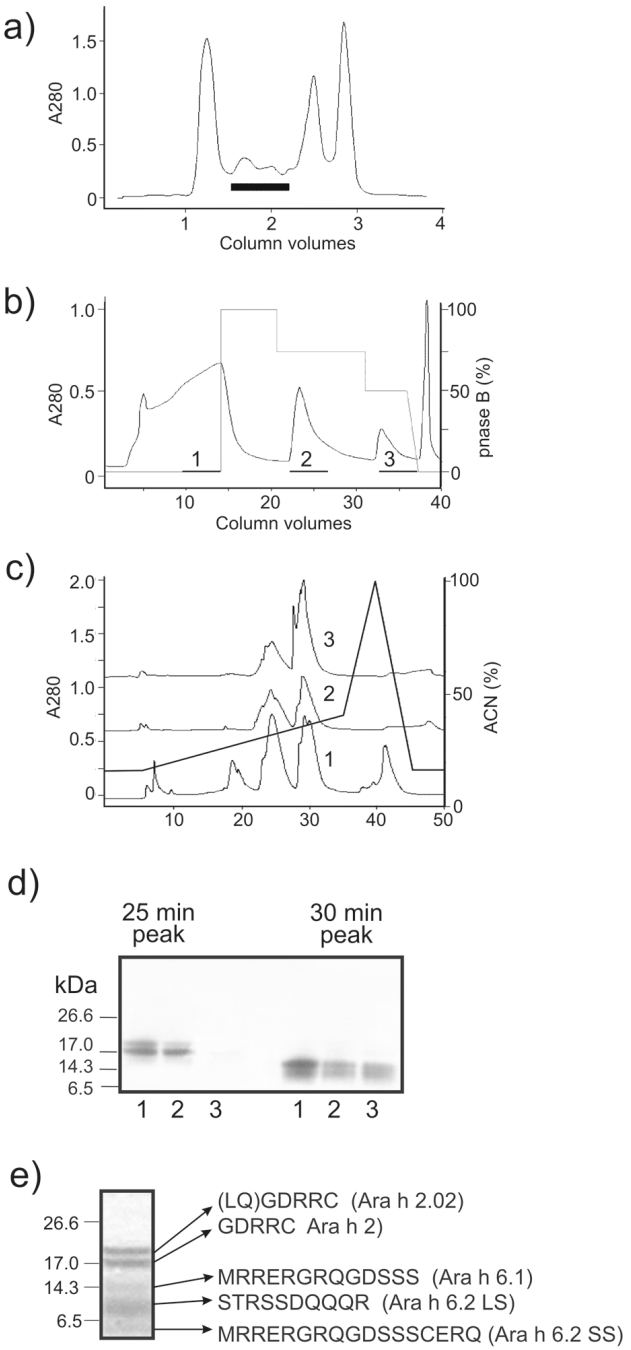

Figure 1. Purification of Ara h 2 and Ara h 6; (a) gel-filtration chromatography of crude extract of peanut meal. The black line indicates the 25 – 10 kDa fraction; (b) anion-exchange chromatography of the fraction from gel filtration chromatography (part a). The black lines indicate the pooled peaks 1 – 3; (c) RP-HPLC of fractions 1, 2 and 3 from anion-exchange chromatography (part b); (d) SDS-PAGE of the RP-HPLC (part c) peaks eluted at 25 and 30 min; (e) SDS-PAGE of a preparation of Ara h 2 and Ara h 6 showing the *N*-terminal sequences of the major bands; LS, large subunit; SS, small subunit.

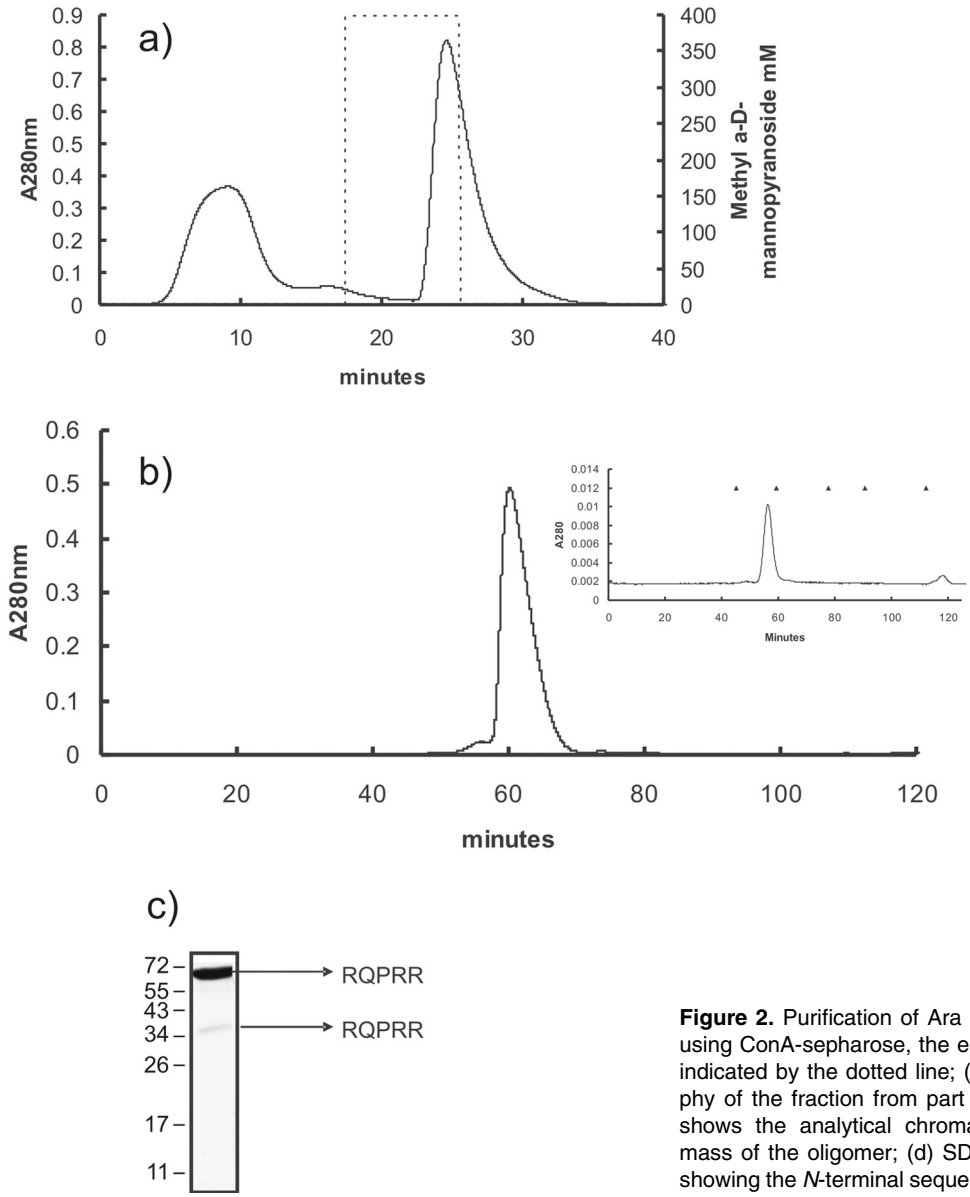

Figure 2. Purification of Ara h 1; (a) affinity chromatography using ConA-sepharose, the eluted peak containing Ara h 1 is indicated by the dotted line; (b) gel-permeation chromatography of the fraction from part a on Superdex S200; the inset shows the analytical chromatogram used to calculate the mass of the oligomer; (d) SDS-PAGE of the purified fraction showing the *N*-terminal sequences of the major bands.

3.1.3 Native 11S globulins – nAra h 3/4

Peanut 11S globulins, Ara h 3/4, were purified from defatted peanut flour by anion exchange chromatography (Fig. 3a) followed by gel filtration (Fig. 3b) and analytical size exclusion chromatography (Fig. 3b inset). SDS-PAGE under reducing conditions of fractions containing Ara h 3/4 showed two groups of bands which corresponded to the *N*-terminal acidic (about 40 – 43 kDa) and *C*-terminal basic (about 20 kDa) subunits (Fig. 3c). The acidic and basic subunits are linked by a single interchain disulphide bond and migrate as single bands of about 45 – 58 kDa when separated by SDS-PAGE under nonreducing conditions these subunits (not shown). Analytical gel permeation chromatography gave a mass of approximately 262 kDa for the native hexameric protein (Fig. 3b inset).

3.1.4 Recombinant Ara h 8 (Bet v 1 homologue)

Recombinant Ara h 8 was synthesised in *E. coli* with a histidine tag (Fig. 4a) to allow purification using IMAC. SDS-PAGE of the purified protein showed a single major band migrating at about 20 kDa (Fig. 4b).

3.2 Sequence analysis

3.2.1 Native 2S albumins – nAra h 2 and nAra h 6

The Ara h 2 fraction described above, comprised two major components with masses of about 17 kDa. *N*-terminal sequencing of the slower component gave a mixture of two sequences, LQGDRRC and GDRRC (*i.e.* lacking the first two amino acids). The faster component also had the truncated *N*-terminal sequence GDRRC (Fig. 1e). These match

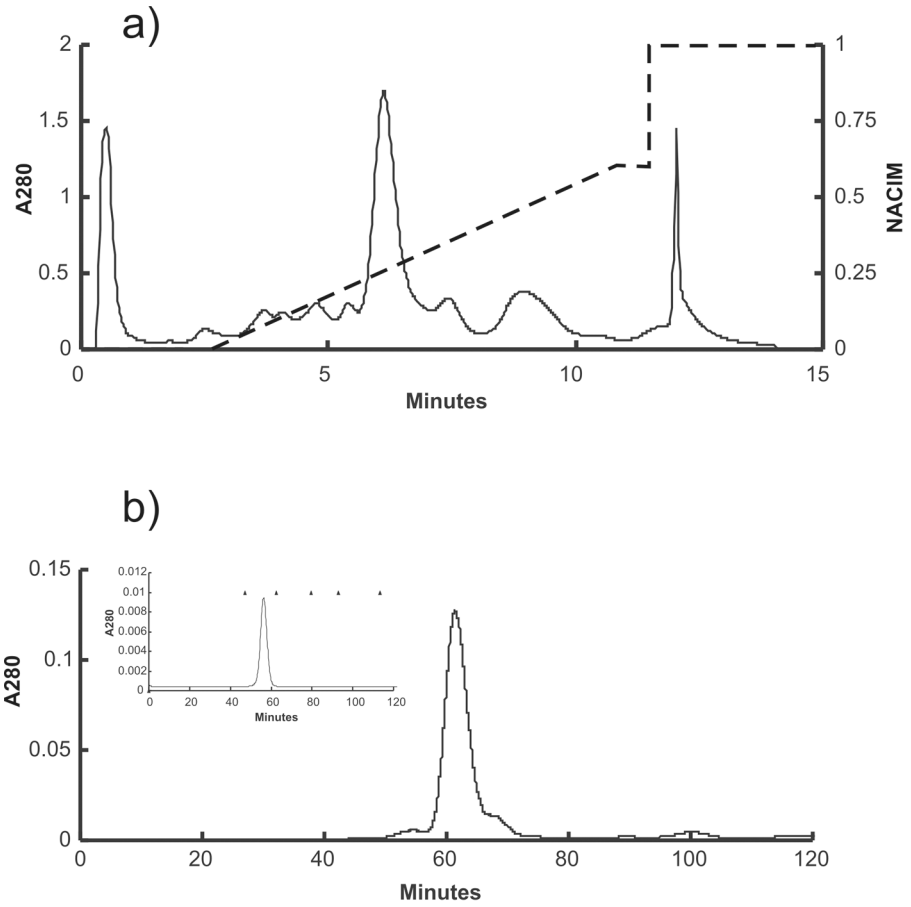

Figure 3. Purification of Ara h 3/ 4; (a) anion exchange chromatography of crude peanut extract. Peak fractions containing the putative Ara h 3/4 were pooled; (b) gel permeation chromatography on Superdex S200 of the fraction from anion exchange chromatography; the inset shows the analytical chromatogram that was used to calculate the mass of the oligomer; (c) SDS-PAGE of the purified fraction showing the *N*-terminal sequences of the major bands.

the published sequences of Ara h 2 components (Fig. 5a); the longer sequence corresponding to Ara h 2.02 and the shorter sequence to Ara h 2 [24]. It was therefore concluded that the slow and fast 17 kDa bands in our preparation corresponded to Ara h 2.02 and Ara h.2, respectively. However, the *N*-terminal sequences determined for our components start five (Ara h 2.02) and seven (Ara h 2) residues downstream of the *N*-termini reported in other studies [24, 25], which correspond to the predicted sites of signal peptide cleavage (Fig. 5a). This indicates that further *N*-terminal processing of our preparations has occurred, presumably as part of the vacuolar processing of the proteins in the developing seed.

The reduced Ara h 6 preparation contained three major bands which were also analysed by *N*-terminal sequencing (Fig. 1e). The major 15 kDa component had the *N*-terminal sequence STRSS, corresponding to residues 69 – 73 of Ara h 6 (Swiss-Prot entry Q647G9) (Fig. 5b). This presumably corresponds to the unprocessed form of Ara h 6 which comprises a single subunit [25]. We have designated this as Ara h 6.1 in Figs. 1d and 5b. The *N*-terminal sequences of the two faster bands demonstrate that these correspond to the large and small subunits of a processed form of Ara h 6, as described recently [26]. We designate this form as Ara h 6.2 (NCBI entry: AF092846). Searches of sequence databases show two sequences encoding forms of Ara h 6, which dif-

(a)

```
                 M  R  G  S  G  V  F  T  F  E  D  E  I  T  S  T  V  P  P  A
Arah8_wtDNA      ATGAGAGGATCCGGCGTCTTCACTTTCGAGGATGAAATCACCTCCACCGTGCCTCCGGCC 60
Arah8_GeneArt    ATGAGAGGATCCGGCGTGTTTACGTTTGAAGATGAAATCACCAGCACCGTTCCGCCGGCG 60
                 ********************** ** ** ** ** ************  ****** ** *****

                  K  L  Y  N  A  M  K  D  A  D  S  I  T  P  K  I  I  D  D  V
Arah8_wtDNA      AAGCTTTACAATGCTATGAAGGATGCCGACTCCATCACCCCTAAGATTATTGATGACGTC 120
Arah8_GeneArt    AAACTGTATAACGCCATGAAAGATGCCGATAGCATTACCCCGAAAATCATTGATGACGTG 120
                 ** ** ** ** ** ***** ********   *** ***** ** ** ************

                  K  S  V  E  I  V  E  G  N  G  G  P  G  T  I  K  K  L  T  I
Arah8_wtDNA      AAGAGTGTTGAAATTGTTGAGGGAAACGGTGGTCCCGGAACCATCAAGAAACTCACCATT 180
Arah8_GeneArt    AAAAGCGTGGAAATTGTTGAAGGCAACGGCGGTCCCGGCACCATTAAAAAACTGACCATC 180
                 ** ** ** ************ ** ***** ***** ** ***** ** ***** *****
                  V  E  D  G  E  T  K  F  I  L  H  K  V  E  S  I  D  E  A  N
Arah8_wtDNA      GTCGAGGATGGAGAAACCAAGTTTATCTTGCACAAGGTGGAGTCAATAGATGAGGCCAAC 240
Arah8_GeneArt    GTGGAAGATGGCGAAACCAAATTCATCCTGCACAAAGTGGAAAGCATTGATGAAGCGAAC 240
                 ** ** ***** ******** ** *** ******* *****     ** ***** ** ***
                  Y  A  Y  N  Y  S  V  V  G  G  V  A  L  P  P  T  A  E  K  I
Arah8_wtDNA      TATGCATACAACTACAGCGTTGTTGGAGGAGTGGCTCTGCCTCCCACGGCGGAGAAGATA 300
Arah8_GeneArt    TATGCCTATAACTATAGCGTGGTGGGCGGTGTGGCGCTGCCCGCCGACCGGCGAAAAATT 300
                 ***** ** ***** ***** ** ** ** ***** ***** ** ** ** ** ** **
                  T  F  E  T  K  L  V  E  G  P  N  G  G  S  I  G  K  L  T  L
Arah8_wtDNA      ACATTTGAGACAAAGCTGGTTGAAGGACCCAACGGAGGATCCATTGGGAAGCTTACTCTC 360
Arah8_GeneArt    ACCTTCGAAACCAAACTGGTGGAAGGCCCCGAACGGCGGCAGCATTGGCAAACTGACCCTG 360
                 ** ** ** ** ** ***** ***** ** ***** **   ****** ** ** ** **
                  K  Y  H  T  K  G  D  A  K  P  D  E  E  E  L  K  K  G  K  A
Arah8_wtDNA      AAGTACCACACCAAAGGAGATGCAAAGCCAGATGAGGAAGAGTTGAAGAAGGGTAAGGCC 420
Arah8_GeneArt    AAATATCATACCAAAGGCGATGCCAAACCCGATGAAGAAGAACTGAAAAAAGGCAAAGCG 420
                 ** ** ** ******** ***** ** ** ***** *****  **** ** ** ** **
                  K  G  E  G  L  F  R  A  I  E  G  Y  V  L  A  N  P  T  Q  Y
Arah8_wtDNA      AAGGGTGAAGGTCTCTTCAGGGCTATTGAGGGTTACGTTTTGGCCAACCCTACTCAATAT 480
Arah8_GeneArt    AAAGGCGAAGGCCTGTTTCGCGCCATTGAAGGCTATGTGCTGGCGAACCCGACCCAGTAT 480
                 ** ** ***** ** **  * ** ***** ** ** **  **** ***** ** ** ***
                  H  H  H  H  H  H  -  -
Arah8_wtDNA      ..................TAG... 483
Arah8_GeneArt    CATCATCATCACCACCACTAATAA 504
                                  **
```

(b)

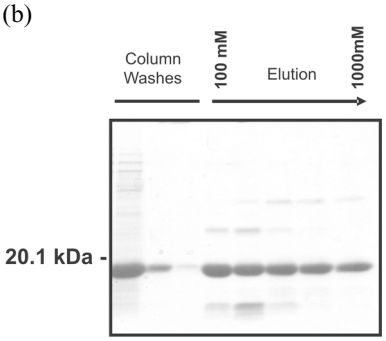

Figure 4. Expression and characterisation of the recombinant Ara h 8; (a) the sequences of the wild type Ara h 8 gene (Arah8_wtDNA) and protein and of the optimised gene (Arah8_GeneArt) used for expression of recombinant Ara h 8; (b) purification by IMAC. The *N*-terminal sequence determined by direct sequencing (MRGSGVFT) is underlined.

fer in the presence of aspartate (D) or asparagine (N) two residues upstream of the *N*-terminus of the large subunit of Ara h 6.2 (see arrowheads in Fig. 5b). Bernard *et al*. [26] suggested that the form with aspartate was the precursor of Ara h 6.2, and cleavage adjacent to aspartate residues certainly occurs in the 2S albumins of Arabidopsis catalysed by an aspartyl proteinase [27, 28]. However, cleavage adjacent to asparagine residues is a more frequent event during the processing of vacuolar proteins, including 2S albumins (reviewed in ref. [29]), and is catalysed by a class of cysteine proteinases called legumains (reviewed in ref. [30]). We therefore consider it more likely that the asparagine-containing form of Ara h 6 (Ara h 6.2) is the predominant form undergoing processing into small and large subunits, as shown in Fig. 4b. In either case a dipeptide (IR or FG) must also be lost from the *N*-terminus of the large subunit [26].

3.2.2 Native 7 S globulin – nAra h 1

N-terminal sequencing of purified nAra h 1 gave the sequence R(RGPRR)QPRR (Fig. 2c). This matches amino acid residues 74 – 78 of the Ara h 1 protein encoded by

a)

```
                                        ↓
Arah2.02          1 MAKLTILVALALFLLAAHASARQQWELQGDRRCQSQLERANLRPCEQHLM     50
                    ||||||||||||||||||||||||||||||||||||||||||||||||||
Arah2             1 MAKLTILVALALFLLAAHASARQQWELQGDRRCQSQLERANLRPCEQHLM     50

Arah2.02         51 QKIQRDEDSYGRDPYSPSQDPYSPSQDPDRRDPYSPSPYDRRGAGSSQHQ    100
                    ||||||||||.||||||||            ||||||||||||||||||||
Arah2            51 QKIQRDEDSYERDPYSPSQ------------DPYSPSPYDRRGAGSSQHQ     88

Arah2.02        101 ERCCNELNEFENNQRCMCEALQQIMENQSDRLQGRQQEQQFKRELRNLPQ    150
                    ||||||||||||||||||||||||||||||||||||||||||||||||||||
Arah2            89 ERCCNELNEFENNQRCMCEALQQIMENQSDRLQGRQQEQQFKRELRNLPQ    138

Arah2.02        151 QCGLRAPQRCDLEVESGG    168
                    ||||||||||||:|||||
Arah2           139 QCGLRAPQRCDLDVESGG    156
```

b)

```
                                        ↓
Arah6.1           1 MAKSTILVALLALVLVAHASAMRRERGRQGDSSSCERQVDRVNLKPCEQH     50
                                        |||||||||||||||||||||||||.|||||||||
Arah6.2           1                     AHASAMRRERGRQGDSSSCERQVDGVNLKPCEQH     34

                                     ▼
Arah6.1          51 IMQRIMGEQEQYDSYDIRSTRSSDQQQRCCDELNEMENTQRCMCEALQQI    100
                    ||||||||||||||||:..|||||||||||||||||||||||||||||||||
Arah6.2          35 IMQRIMGEQEQYDSYNFGSTRSSDQQQRCCDELNEMENTQRCMCEALQQI     84
                                        ▲

Arah6.1         101 MENQCDRLQDRQMVQQFKRELMNLPQQCNFRAPQRCDLDVSGGRC    145
                    |||||||.|||||||.|||||||||||||.|||||||||||||||
Arah6.2          85 MENQCDGLQDRQMVQHFKRELMNLPQQCNFGAPQRCDLDVSGGRC    129
```

c)

```
                                           ↓
Arah1.1           1 MRGRVSPLMLLLGILVLASVSATQAK-SPY-RKTENPCAQRCLQSCQQEP     48
                    ||||||||||||||||||||||||.|| ||| :||||||||||||||||||
Arah1.2           1 MRGRVSPLMLLLGILVLASVSATHAKSSPYQKKTENPCAQRCLQSCQQEP     50

Arah1.1          49 DDLKQKACESRCTKLEYDPRCVYD----TGATNQRHPPGERTRGRQPGDY     94
                    |||||||||||||||||||||||    ||.||||.|||||||||||||||
Arah1.2          51 DDLKQKACESRCTKLEYDPRCVYDPRGHTGTTNQRSPPGERTRGRQPGDY    100

Arah1.1          95 DDDRRQPRREEGGRWGPAEPREREREEDWRQPREDWRRPSHQQPRKIRPE    144
                    |||||||||||||||||.||||||||||||||||||||||||||||||||||
Arah1.2         101 DDDRRQPRREEGGRWGPAGPREREREEDWRQPREDWRRPSHQQPRKIRPE    150
```

d)

```
                                        ↓
Arah3/4.43        1 MGKLLALSVCFCFLVLGASSISFRQQPEENACQFQRLNAQRPDNRIESEG     50
                    |.||||||:|||.|||||||::|||..|||.||||||||||||||||||||
Arah3/4.38        1 MAKLLALSLCFCVLVLGASSVTFRQGGEENECQFQRLNAQRPDNRIESEG     50

                     ▼
Arah3/4.43      346 SRGSGNGIEETICTASFKKNIGRNRSPDIYNPQAGSLKTANELQLNLLIL    395
                    .||..||||||||:||.|||:||:.:|||||||||||::.||  |:|.||
Arah3/4.38      321 RRGYKNGIEETICSASVKKNLGRSSNPDIYNPQAGSLRSVNE--LDLPIL    368
                     ▲
```

Figure 5. Comparison of the *N*-terminal sequences determined for the allergen preparations with the sequences of corresponding genes and proteins reported in the literature. In all parts the predicted signal peptide is shown in italics and the signal peptide cleavage site indicated by an arrow above the sequence. (a) Ara h 2, the sequences of Ara h 2 and Ara h 2.02 are from GenPept accessions AAK96887 and AAN77576, respectively. The grey boxes indicate the *N*-termini determined for our preparations. (b) Ara h 6, the sequences of Ara h 6.1 and Ara h 6.2 are from Swiss-Prot entry Q647G9 and NCBI entry AF092846, respectively. The grey boxes indicate the *N*-termini determined for our preparations. Putative processing sites (aspartate, D and asparagine, N residues) are indicated by arrowheads. (c) Ara h 1, the sequences of the *N*-terminal parts of Ara h 1.1 and Ara h 1.2 are from Swiss-Prot entries P43237 and P43238, respectively. The grey boxes indicate the *N*-termini determined for our preparations. The open box indicates the *N*-terminal sequence determined by [31]. (d) Ara h 3/4, the sequences of the *N*-terminal parts of the large acidic (top) and small basic (bottom) subunits are shown, taken from Swiss-Prot entry Q8LKN1 for Ara h 3, Ara h 4.43 and Q0GM57 for Ara h 3, Ara h 4.38. The grey boxes indicate the *N*-termini determined for our preparations. The asparagine residues adjacent to the *N*-terminus of the basic subunit (arrowhead) indicate that processing is probably catalysed by a legumain type proteinase.

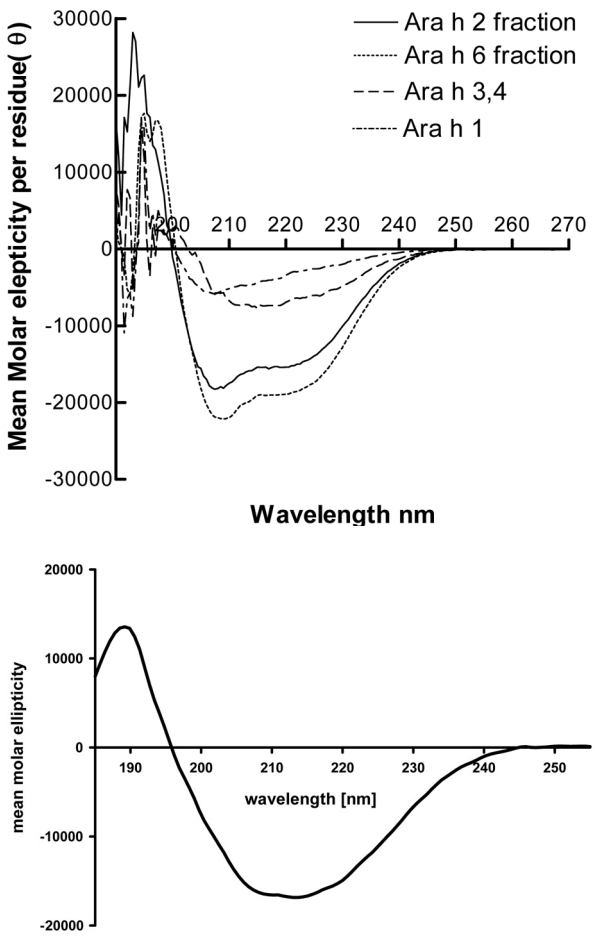

Figure 6. Far-UV CD spectroscopy of purified peanut allergens; (a) natural Ara h 1, Ara h 2, Ara h 3/4 and Ara h 6; (b) recombinant Ara h 8.

Swiss-Prot entry P43237 and residues 80 – 84 of that encoded by Swiss-Prot entry P43238, respectively (Fig. 5c). However, comparison with the *N*-terminal sequence reported by Wichers *et al*. [31] shows that additional processing has occurred. Wichers determined the *N*-terminal sequence of Ara h 1 as RS/H-PPGERTRG, which is present either 48 residues within the mature Ara h 1.1 isoform or 52 residues within the mature Ara h 1.2 isoform, assuming that signal peptide cleavage occurs after Ala 25 (as shown in Fig. 5c). The sequence RQPRR occurs a further 19 residues within the protein purified by Wichers *et al*. [31] (Fig. 5c). This may indicate that our preparation of Ara h 1 had been processed in two stages, or that the precise pattern of processing differs between peanut lines.

The *N*-terminal sequence of the 33 kDa band was also RQPRR showing that it was derived from processing of a larger precursor protein. It is notable that a subunit of similar size (33 kDa) present in pea vicilin (7S globulin) is also derived from processing of a 50 kDa precursor protein, and has the same *N*-terminal sequence (reviewed in ref. [32]).

3.2.3 Native 11s globulins – nAra h 3/4

The preparation of Ara h 3/4 contained two major groups of bands corresponding to the acidic and basic subunits of the 11S globulin. *N*-terminal sequencing (Fig. 3c) showed that the 43 and 38 kDa bands corresponded to acidic subunits of Ara h 3/4 (Swiss-Prot entries Q8LKN1 and QOGM57, respectively) (Fig. 5d). These subunits were also identified as isoforms of Ara h 3/4 [33]. However, the *N*-terminal sequence of the 43 kDa component (SFRQQ) differed from that previously reported [33] in the absence of an isoleucine residue (Fig. 5d). Both subunits also lack an *N*-terminal dipeptide (SS) which is predicted to be present based on the consensus rules for signal peptide cleavage. The inclusion of proteinase inhibitors in the extraction protocol suggests that this trimming occurred in the seed, as observed for 2S albumins.

The *N*-terminal sequence of the 24 kDa band was GIEETI, which corresponds to the *N*-termini of the basic subunits encoded by both QOGM57 [34] and QQLKN1 (Fig. 5d) and to the basic subunits identified by Boldt *et al*. [33]. This band may therefore comprise a mixture of at least two basic subunits which are associated with the acidic 43 and 38 kDa subunits. It can also be noted that the *N*-terminus of the basic subunit is adjacent to an asparagine residue, suggesting that the processing is catalysed by a legumain proteinase as described for other legume 11S globulins [30].

3.2.4 Recombinant Ara h 8 (Bet v 1 homologue)

The *N*-terminal sequence of the rAra h 8 preparation was MRGSGVFT, which corresponds to the MRGS-tag followed by residues 2 to 5 of the Ara h 8 protein (Uniprot entry Q6VT83; Fig. 4a).

3.3 Spectroscopic analysis of secondary structure

An assessment was made of the folded state of purified allergens using CD spectroscopy (Figs. 6a and b) with additional characterisation of the 7S and 11S globulin allergens (Ara h 1 and Ara h 3/4) being performed using FT-IR spectroscopy since this is especially informative about the presence of β-sheet structures associated with aggregated states (Fig. 7). The noisy CD spectra below 200 nm resulted from the absorbance of Cl^- ions which the proteins required for solubility. The two globulins, Ara h 1 and Ara h 3/4 had much lower overall molar ellipticities than the 2S albumins, Ara h 2, and Ara h 6; their spectra being similar to those of the homologous globulins from hazelnut (Rigby, N. M., *et al*., The purification and characterisation of allergenic hazelnut seed storage proteins, submitted to *Mol. Nutr. Food Res*.) and soyabean [35, 36]. The Ara h 1 spectrum had a positive maximum at around 190 nm and a negative minimum at 205 nm, the molar elipticity gradually increasing until reaching around zero at around 240 nm. The Ara h 3/4 spectrum had a positive maximum at 195 nm and a shal-

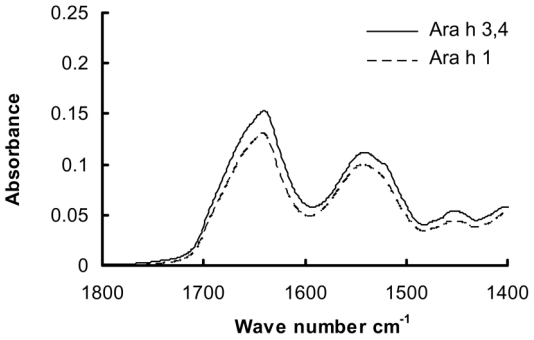

Figure 7. FT-IR spectroscopy of Ara h 1 and Ara h 3/4.

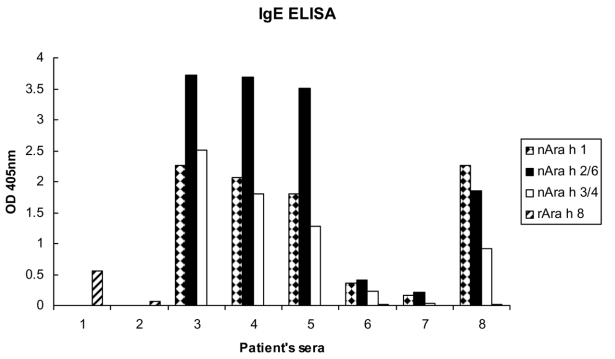

Figure 8. IgE ELISA of purified peanut allergens. Purified nAra h 1, nAra h 2/6, nAra h 3/4 and rAra h 8 were tested for IgE binding using sera from eight peanut allergic patients.

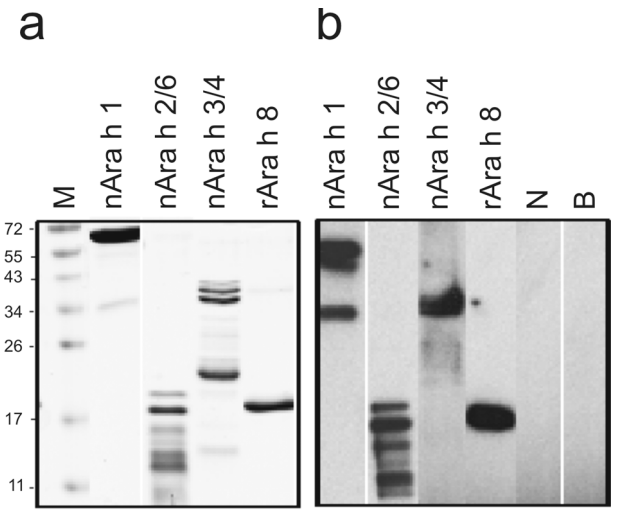

Figure 9. Coomassie stained SDS-PAGE gel (A) and IgE immunoblot (B) of peanut allergens.
The IgE immunoblot analysis in B was performed using selected sera as used in the IgE ELISA (Figure 8): serum no. 6 for nAra h 1; serum no. 8 for nAha h 2/6 and nAra h 3/4 and serum no. 2 for rAra h 8. Bound IgE antibodies were detected by ^{125}I labelled anti-human IgE antibodies. Normal human serum (lane N) and buffer (lane B) omitting the serum were tested with total peanut extract as controls.

low negative minimum between 210 and 220 nm. These spectral characteristics indicate a largely unordered structure and this conclusion was supported by FT-IR spectroscopy; both fractions giving an amide I band shape with a maximum at 1645 cm^{-1} indicating that the proteins largely comprised unordered structures. However, a shoulder at 1657 cm^{-1} indicated that they also contained some α-helical structures while the broad band shape and a smaller shoulder around 1690 cm^{-1} may indicate the presence of β-sheet structures.

As might be anticipated, Ara h 2 and Ara h 6 had very similar CD spectra with strong positive maxima at 190 nm (Ara h 2) and 198 nm (Ara h 6). These spectra are consistent with the α-helical structure (as expected for 2S albumins) with a pronounced double negative minimum at 209 and 222 nm. However, this negative minimum was more intense for Ara h 6 than for Ara h 2.

Ara h 8 is a member of the Bet v 1 family and its CD spectrum is very similar to those of Bet v 1 itself [37] and other recombinant homologues from fruit, such as the allergenic Pru av 1 from cherry [38]. However, some subtle differences are present. Thus, Ara h 8 gave a strong positive maximum, which while being almost identical in intensity to that in the spectrum of Pru av 1 was at 190 nm instead of 195 nm. The negative minimum had a very similar wavelength span and mean molar ellipticity to that of Pru av 1, between 209 and 217 nm. These data indicate the recombinant protein is folded, adopting a primarily β-sheet conformation.

3.4 Immunological characterisation of allergens

The purified fractions (including a mixture of Ara h 2, Ara h 6 and Ara h 3/4, respectively) were tested against a panel of sera from eight patients with peanut allergy using a fluid phase assay (Fig. 8). Two sera (nos.1 and 2) reacted with Ara h 8 but not with any of the other allergen preparations. These patients had also diagnosed birch pollen allergy. In contrast, the remaining six sera reacted with all of the allergen preparations except Ara h 8, with five of the sera reacting most strongly with the Ara h 2, Ara h 6 (2S albumin) preparations (Fig. 8).

The same protein preparations were also used for immunoblotting, using single selected sera for each preparation (Fig. 9). Serum no. 6 used for Ara h 1 (7S globulin) displayed strong IgE reactivity to the major band of 67 kDa which was clearly visible on the stained gel separation of the fraction (compare Figs. 8a and b) but also with a band of slightly faster mobility which was not apparent on the stained gel and probably corresponded to a minor isoform, and to the small vicilin subunit of 33 kDa which stained only weakly.

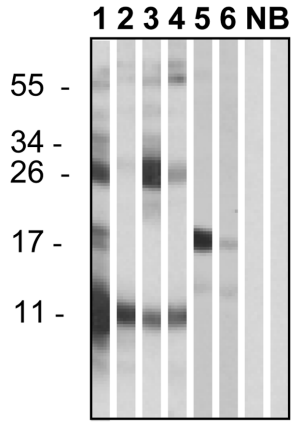

Figure 10. IgE immunoblot inhibition on total peanut protein extract: IgE binding of serum pool 1 (serum nos. 3–8) to total peanut protein extract (lane 1), after preincubation with purified Ara h 1 (lane 2), purified Ara h 2, Ara h 6 (lane 3), purified Ara h 3/4 (lane 4), respectively. In parallel, serum pool 2 (serum nos. 1 and 2) was tested for IgE binding to total peanut protein extract (lane 5) and after preincubation with purified rAra h 8 (lane 6). Detection of bound IgE and controls (lines N and B) were performed as described above.

Similarly, serum no. 8 showed IgE reactivity to the full range of subunits of Ara h 2, Ara h 6 which could be observed on the stained gel, but in this case the intensity of reaction corresponded broadly to the relative staining intensity. In contrast, the same serum (no. 8) used for the Ara h 3/4 (11S globulin) blot reacted only with the large acidic subunits, and not with the small basic subunits. Finally, serum no. 2 displayed strong IgE reactivity to blotted rAra h 8, the major band present in the preparation (Fig. 9). Furthermore, a rabbit antiserum raised against the related soyabean allegen (Gly m 4) showed crossreactivity to rAra h 8 in immunoblotting (results not shown). Pools of sera with IgE reactivity to Ara h 1, Ara h 3/4, Ara h 2/6 (pool 1) and to Ara h 8 (pool 2) were used for inhibition assays. The reactivity to total peanut protein extract (Fig 10 lane 1) was then determined after pre-incubation with individual protein preparations. Pre-incubation with Ara h 1 resulted in reduced binding of serum pool 1 to the 7S globulin (Ara h 1) bands at about 67 kDa and 34 kDa but also reduced binding to Ara h 3/4 components at 40–43 kDa, due to cross-reactivity (Fig. 10, lane 2). Similarly, pre-incubation with Ara h 3/4 resulted in reduced binding of serum pool 1 to the 11S globulin bands at 40–43 kDa and also reduced binding to the 7S globulin band at 34 kDa (Fig. 10, lane 4). Pre-incubation with Ara h 2 and Ara h 6 resulted in reduced binding of serum pool 1 to the corresponding 2S albumin bands in the range of 11–17 kDa (Fig. 10, lane 3). A band at 17 kDa was recognised by serum pool 2 (Fig. 10, lane 5) but the binding to this was significantly reduced on pre-incubation with rAra h 8 (Fig. 10, lane 6). Cross-inhibition between Ara h 1 and Ara h 3/4 was also observed in ELISA experiments (results not shown).

4 Concluding remarks

The 2S albumin, 7S globulin and 11S globulin fractions from peanut are all mixtures of subunits encoded by small multigene families: this is reflected in the composition of the purified fractions reported here and was verified by SDS-PAGE analysis and *N*-terminal sequencing. Furthermore, the complexity may be increased by the occurrence of post-translational processing, which takes place in the vacuole and is frequently catalysed by a specific class of enzymes (called legumains) which cleave adjacent to asparagine residues. This processing is difficult to replicate *in vitro* and this, combined with the abundance and stability of the proteins, means that purification from seeds is the preferred approach to prepare standard fractions for characterisation and allergy diagnosis. The complete spectrum of isoforms and differentially processed proteins therefore needs to be purified and analysed, in order to define the contributions of the individual forms to the overall activity of the fraction. In contrast, the Bet v 1 homologue Ara h 8 is present in small amounts and is labile to denaturation during purification. It is therefore most readily produced as a correctly folded and immunologically active recombinant protein. However, the inhibition assay showed that the immunologcial properties of the recombinant protein were equivalent to those of its natural counterpart.

The preliminary immunological studies reported here show several results of interest. Firstly, it is clear that reaction to the pollen-related Ara h 8 allergen occurs in a different group of patients to those reacting with the major seed protein allergens Ara h 1, Ara h 2, Ara h 3/4 and Ara h 6. This presumably reflects the sensitisation of these patients by birch or other tree pollens and by peanuts in food, respectively. Secondly, it is clear that subunits of the major seed protein allergens may differ in their allergenicity. Most notably, the large acidic subunits of the 11S globulin allergens Ara h 3/4 appear to be allergenic but not the small basic subunits of these proteins. A similar difference has been observed with the related 11S globulin allergen (Cor a 9) of hazelnut (Rigby, N. M., *et al*., The purification and characterisation of allergenic hazelnut seed storage proteins, submitted to *Mol. Nutr. Res*.). It should be noted that these two subunits are synthesised as a single precursor protein in the developing seed and remain associated in the hexameric 11S globulin structure until digested in the GI tract. Hence, the differences in allergenicity do not reflect quantitative differences in terms of exposure to the patient.

Recently, Ramos *et al*. [39] have shown that *A. hypogaea* cv Georgia Green has two genes encoding Ara h 2 and Ara h 2.02 and three genes encoding Ara h 6. The gene encoding Ara h 2 was assigned to the A genome, while that encoding

Ara h 2.02 was assigned to the B genome. Two of the Ara h 6 genes are closely linked to the Ara h 2.02 gene on the B genome. The third Ara h 6 gene is on the A genome. However, none of the Ara h 6 genes encodes an asparagine residue at position 66. The Ara h 6.2 isoform included in Fig. 5 was derived from a Virginia-type cultivar (NCBI accession: AF092846) and commercially packaged peanuts, which were used in our experiments, are generally of the Virginia-type due to their larger size. The post-translational cleavage that we observed may therefore result from cleavage at an asparagine residue introduced by a mutation present only in Virginia-type cultivars.

Taken together the purification protocols for peanut allergens described here have enabled us to obtain a better insight into the range of immunoreactive forms that are present which should in turn lead to a better understanding of the relevant IgE epitopes. Our studies of the properties of the proteins will also facilitate the development of improved tools for allergy diagnosis.

The work described was carried out under the EU FP6 project "The Prevalence, Cost and Basis of Food Allergy Across Europe", EuroPrevall (Contract No. 51400). Rothamsted Research and IFR receive grant-aided support from the Biotechnology and Biological Sciences Research Council (BBSRC) of the UK. We want to thank Gerlinde Hofstetter for excellent technical assistance.

The authors have declared no conflict of interest.

5 References

[1] Singh, A. K., Groundnut, in: Smartt, J., Simmonds, N. W. (Eds.), *Evolution of Crop Plants*, 2nd Edn., Longman Scientific and Technical, Harlow 1995, pp. 246 – 250.

[2] Salunkhe, D. K., Chavan, J. K., Adsule, R. N., Kadam, S. S. (Eds.), *World Oilseeds – Chemistry, Technology and Utilization*. Van Norstrang Reinhold, New York 1992.

[3] Emmett, S. E., Angus, F. J., Fry, J. S., Lee, P. N., Perceived prevalence of peanut allergy in Great Britain and its association with other atopic conditions and with peanut allergy in other household members. *Allergy* 1999, *54*, 380 – 385.

[4] Sicherer, S. H., Morrow, E. H., Sampson, H. A., Dose-response in double-blind, placebo-controlled oral food challenges in children with atopic dermatitis. *J. Allergy Clin. Immunol.* 2000, *105*, 582 – 586.

[5] Yunginger, J. W., Sweeney, K. G., Sturner, W. Q., Giannandrea, L. A., *et al*., Fatal food-induced anaphylaxis. *JAMA* 1988, *260*, 1450 – 1452.

[6] Sampson, H. A., Mendelson, L., Rosen, J. P., Fatal and near-fatal anaphylactic reactions to food in children and adolescents. *N. Engl. J. Med*. 1992, *327*, 380 – 384.

[7] Burks, A. W., Williams, L. W., Helm, R. M., Connaughton, C. *et al*., Identification of a major peanut allergen, Ara h I, in patients with atopic dermatitis and positive peanut challenges, *J. Allergy Clin. Immunol*. 1991, *88*, 172 – 179.

[8] Koppelman, S. J., Wensing, M., Ertmann, M., Knulst, A. C., Knol, E. F., Relevance of Ara h 1, Ara h 2 and Ara h 3 in peanut-allergenic patients, as determined by immunoglobulin E Western blotting. basophil-histamine release and intracutaneous testing: Ara h 2 is the most important peanut allergen. *Clin. Exp. Allergy* 2004, *34*, 583 – 590.

[9] Koppelman, S. J., de Jong, G. A. H., Laaper-Ertmann, M., Peeters, K. A. B. M., *et al*., Purification and immunoglobulin E-binding properties of peanut allergen Ara h 6: Evidence for cross-reactivity with Ara h 2. *Clin. Exp. Allergy* 2005, *35*, 490 – 497.

[10] Shreffler, W. D., Lencer, D. A., Bardina, L., Sampson, H. A., IgE and IgG_4 epitope mapping by microarray immunoassay reveals the diversity of immune response to the peanut allergen, Ara h 2. *J. Allergy Clin. Immunol.* 2005, *116*, 893 – 899.

[11] de Leon, M. P., Drew, A. C., Glaspole, I. N., Suphioglu, C. *et al*., IgE cross-reactivity between the major peanut allergen Ara h 2 and tree nut allergens. *Mol. Immunol*. 2007, *44*, 463 – 471.

[12] Burks, A. W., Williams, L. W., Connaughton, C., Cockrell, G., *et al*., Identification and characterization of a second major peanut allergen, Ara h II, with use of the sera of patients with atopic dermatitis and positive peanut challenge. *J. Allergy Clin. Immunol*. 1992, *90*, 962 – 969.

[13] Kleber-Janke, T., Crameri, R., Appenzeller, U., Schlaak, M., Becker, W. M., Selective cloning of peanut allergens, including profilin and 2S albumins, by phage display technology. *Int. Arch. Allergy Immunol.* 1999, *119*, 265 – 274.

[14] Burks, A. W., Cockrell, G., Connaughton, C., Guin, J., *et al.*, Identification of peanut agglutinin and soybean trypsin inhibitor as minor legume allergens. *Int. Arch. Allergy Immunol.* 1994, *105*, 143 – 149.

[15] Rabjohn, P., Helm, E. M., Stanley, J. S., West, C. M., *et al.*, Molecular cloning and epitope analysis of the peanut allergen Ara h 3. *J. Clin. Invest.* 1999, *103*, 535 – 542.

[16] Mittag, D., Akkerdaas, J., Ballmer-Weber, B. K., Vogel, L., *et al*., Ara h 8, a Bet v 1-homologous allergen from peanut, is a major allergen in patients with combined birch pollen and peanut allergy. *J. Allergy Clin. Immunol*. 2004, *114*, 1410 – 1417.

[17] Pons, L., Chery, C., Romano, A., Namour, F., *et al*., The 18 kDa peanut oleosin is a candidate allergen for IgE-mediated reactions to peanuts. *Allergy* 2002, *57*, 88 – 93.

[18] Maleki, S. J., Viquez, O., Jacks, T., Dodo, H., *et al*., The major peanut allergen, Ara h 2, functions as a trypsin inhibitor, and roasting enhances this function. *J. Allergy Clin. Immunol*. 2003, *112*, 190 – 195.

[19] Eigenmann, P. A., Burks, A. W., Bannon, G. A., Sampson, H. A., Identification of unique peanut and soy allergens in sera adsorbed with cross-reacting antibodies. *J. Allergy Clin. Immunol*. 1996, *98*, 969 – 978.

[20] Mittag, D., Akkerdaas, J., Ballmer-Weber, B. K., Vogel, L., *et al.*, Ara h 8, a Bet v 1-homologous allergen from peanut is a major allergen in patients with combined birch pollen and peanut allergy. *J. Allergy Clin. Immunol*. 2004, *114*, 1410 – 1417.

[21] Eiwegger, T. T., Rigby, N., Mondoulet, L., Bernard, L., *et al*., Gastro-duodenal digestion products of the major peanut allergen Ara h 1 retain an allergenic potential. *Clin. Exp Allergy*, 2006, *36*, 1281 – 1286.

[22] Bujard, H., Gentz, R., Lanzer, M., Stueber, D., *et al*., A T5 promoter-based transcription-translation system for the analysis of proteins *in vitro* and *in vivo. Meth. Enzymol*. 1987, *155*, 416 – 433.

[23] Ma, Y., Zuidmeer, L., Bohle, B., Bolhaar, S. T. H., *et al.*, Characterization of recombinant Mal d 4 and its application for component-resolved diagnosis of apple allergy. *Clin. Exp. Allergy* 2006, *36*, 1087 – 1096.

[24] Chatel, J. M., Bernard, H., Orson, F. M., Isolation and characterization of two complete Ara h 2 isoforms cDNA. *Int. Arch. Allergy Immunol*. 2003, *131*, 14 – 18.

[25] Suhr, M., Wicklein, D., Lepp, U., Becker, W.-M., Isolation and characterisation of natural Ara h 6: Evidence for a further peanut allergen with putative clinical relevance based on resistance to pepsin digestion and heat. *Mol. Nutr. Food Res.* 2004, *48*, 390 – 399.

[26] Bernard, H., Mondoulet, L., Drumare, E., Paty, E., *et al*., Identification of a new natural Ara h 6 isoform and of its proteolytic product as major allergens in peanut. *J. Agr. Food Chem*. 2007, *55*, 9663 – 9669.

[27] Krebbers, E., Herdies, L., De Clercq, A., Seurinck, J., *et al.*, Determination of the processing sites of an Arabidopsis 2S albumin and characterisation of the complete gene family. *Plant Physiol.* 1988, *87*, 859 – 866.

[28] D'Hondt, K., Bosch, D., Van Damme, J., Goethals, M., *et al*., An aspartic proteinase present in seeds cleaves arabidopsis 2S albumin precursors *in vitro*. *J. Biol. Chem*. 1993, *268*, 20884 – 20891.

[29] Shewry, P. R., Pandya, M. J., in: Shewry, P. R., Casey, R. (Eds.), *The 2S Albumin Storage Proteins*, *Seed Proteins*, 1st Edn., Kluwer Academic Publishers, Dordrecht 1999, pp. 563 – 586.

[30] Muntz, K., Blattner, F. R., Shutov, A. D., Legumains – a family of asparagines-specific cysteine endopeptidases involved in propolypeptide processing and protein breakdown in plants. *J. Plant Physiol*. 2002, *159*, 1281 – 1293.

[31] Wichers, H., De Beijer, T., Huub, F. J., Savelkoul, F. J., van Amerongen, A., The major peanut allergen Ara h 1 and its cleaved-off *N*-terminal peptide; possible implications for peanut allergen detection. *J. Agric. Food Chem.* 2004, *52*, 4903 – 4907.

[32] Gatehouse, J. A., Croy, R. R. D., Boulter, D., The synthesis and structure of pea storage proteins. *CRC Crit. Rev. Plant Sci.* 1984, *1*, 287 – 314.

[33] Boldt, A., Fortunato, C., Conti, A., Petersen, A., *et al*., Analysis of the composition of an immunoglobulin E reactive high molecular weight complex of peanut extract containing Ara h 1 and Ara h 3/4. *Proteomics* 2005, *5*, 675 – 686.

[34] Yan, Y.-S., Lin, X.-D., Zhang, Y.-S., Wang, L., *et al.*, Isolation of peanut genes encoding arachins and conglutins by expressed sequence tags. *Plant Sci*. 2005, *169*, 439 – 445.

[35] Mills, E. N. C., Huang, L., Noel, T. R., Gunning, A. P., Morris, V. J., Formation of thermally induced aggregates of the soya globulin β-conglycinin. *Biochim. Biophys. Acta* 2001, *1547*, 339 – 350.

[36] Mills, E. N. C., Marigheto, N. A., Wellner, N., Fairhurst, S. A. *et al.*, Thermally induced structural changes in glycine, the 11S globulin of soybean (Glycine max) – an *in situ* spectroscopic study. *Biochim. Biophys. Acta* 2003, *1648*, 105 – 114.

[37] Mogensen, J. E., Ferreras, M., Wimmer, R., Petersen, S. V., *et al*., The major allergen from birch tree pollen, Bet v 1, binds and permeabilizes membranes. *Biochemistry* 2007, *46*, 3356 – 3365.

[38] Neudecker, P., Lehmann, K., Nerkamp, J., Haase, T., *et al*., Mutational epitope analysis of Pru av 1 and Api g 1, the major allergens of cherry (*Prunus avium*) and celery (*Apium graveolens*): Correlating IgE reactivity with three-dimensional structure. *Biochem. J.* 2003, *376*, 97 – 107.

[39] Ramos, M. L., Fleming, G., Chu, Y., Akiyama, Y., *et al*., Chromasomal and phylogenetic context for conglutin genes in Arachis based on genomic sequence. *Mol. Gen. Genomics* 2006, *275*, 578 – 592.

