Available online at www.sciencedirect.com
SCIENCE DIRECT®
Protein Expression and Purification 31 (2003) 250–259
Protein Expression & Purification
www.elsevier.com/locate/yprep

High-yield expression in *Escherichia coli*, purification, and characterization of properly folded major peanut allergen Ara h 2

Katrin Lehmann,[a,*] Silke Hoffmann,[c,d] Philipp Neudecker,[a] Martin Suhr,[b] Wolf-Meinhard Becker,[b] and Paul Rösch[a]

[a] *Lehrstuhl Biopolymere, Universität Bayreuth 30, Universitaetsstrasse 30, 95440 Bayreuth, Germany*
[b] *Zentrum für Medizin und Biowissenschaften, Forschungszentrum Borstel, 23845 Borstel, Germany*
[c] *Institut für Physikalische Biologie, Heinrich-Heine-Universität, D-40225 Düsseldorf, Germany*
[d] *Forschungszentrum Jülich, IBI-2, 52425 Jülich, Germany*

Received 7 April 2003, and in revised form 3 June 2003

Abstract

Allergic reactions to peanuts are a serious health problem because of their high prevalence, associated with potential severity, and chronicity. One of the three major allergens in peanut, Ara h 2, is a member of the conglutin family of seed storage proteins. Ara h 2 shows high sequence homology to proteins of the 2S albumin family. Presently, only very few structural data from allergenic proteins of this family exist. For a detailed understanding of the molecular mechanisms of food-induced allergies and for the development of therapeutic strategies knowledge of the high-resolution three-dimensional structure of allergenic proteins is essential. We report a method for the efficient large-scale preparation of properly folded Ara h 2 for structural studies and report CD-spectroscopic data. In contrast to other allergenic 2S albumins, Ara h 2 exists as a single continuous polypeptide chain in peanut seeds, and thus heterologous expression in *Escherichia coli* was possible. Ara h 2 was expressed as Trx–His-tag fusion protein in *E. coli* Origami (DE3), a modified *E. coli* strain with oxidizing cytoplasm which allows the formation of disulfide bridges. It could be shown that recombinant Ara h 2, thus overexpressed and purified, and the allergen isolated from peanuts are identical as judged from immunoblotting, analytical HPLC, and circular dichroism spectra.

Keywords: Allergy; Major peanut allergen; 2S albumin; α-helical protein

Peanuts have eminent nutritional properties, but can cause severe, type I hypersensitivity reactions [1]. The potency of peanuts to induce an anaphylactic shock is very high, and peanut allergy is more often associated with fatal-induced anaphylaxis than any other food allergy [2,3]. Additionally, in contrast to allergies caused by other foods like milk and eggs, allergic reactions to peanuts tend to persist for the whole life in sensitized individuals [4,5].

Several different peanut allergens have been isolated, and the three major allergens Ara h 1, Ara h 2, and Ara h 3 could be assigned to the vicilin, conglutin, and glycinin families of seed storage proteins, respectively. Ara h 2 is recognized by serum IgE from over 90% of a peanut allergic patient population [6,7].

Ara h 2 and the minor peanut allergens Ara h 6 and Ara h 7 show high sequence homology to proteins of the conglutin family from various *dicotyledonous* plants [8]. Interestingly, 2S albumin seed storage proteins from soybean, walnut, Brazil nut, oilseed rape, castor bean, and mustard are known to bind IgE and induce allergic reactions, and therefore 2S albumins can be regarded as pan allergens in seeds [9–13]. 2S albumins are typically synthesized as large prepro-proteins, and after cotranslational removal of the signal peptide, the precursor is cleaved into two subunits linked by disulfide bonds [14]. These proteins show a common, strictly conserved pattern of eight cysteine residues which also appears in other small seed proteins not sequence-related to 2S albumins, such as nonspecific lipid transfer proteins (ns-LTPs) and the hydrophobic protein from soybean (HPS), which both consist of a single uncleaved polypeptide chain [15,16].

* Corresponding author. Fax: +49-921-553544.
E-mail address: katrin.lehmann@uni-bayreuth.de (K. Lehmann).

1046-5928/$ - see front matter © 2003 Elsevier Science (USA). All rights reserved.
doi:10.1016/S1046-5928(03)00190-6

Although this 2S albumin related protein group contains many highly potent allergens and many studies explored the allergenic epitopes of Ara h 2 and other allergenic 2S albumins using immunological as well as mutational methods, only one global fold of a 2S albumin with reasonable sequence identity to Ara h 2, napin BnIb from rape seed, has been published so far [17].

In structural studies of ns-LTPs, HPS, and the napin 2S albumin the proteins had to be purified from natural sources, commonly from plant seeds, as suitable high-yield heterologous expression and purification system were not available. Purification from natural sources, however, has several serious drawbacks: Seeds contain various protein isoforms with nearly identical biochemical and physical properties. Thus, these isoforms cannot easily be separated, and the protein samples are too inhomogeneous for high resolution structural studies.

Expression of 2S albumins is not straightforward as these proteins are modified post-translationally, and four specific disulfide bonds are formed. Additional problems of *Escherichia coli* heterologous expression systems are the different codon usages in eukaryotes and prokaryotes that may cause ineffective protein expression and the lack of post-translational modification mechanisms in bacteria. In the case of 2S albumins, this only allows the expression of the precursor forms in *E. coli*. Recently, the restrictions of the poor codon usage have been overcome for the peanut proteins Ara h 1, Ara h 2, and Ara h 6 in our laboratory using a special *E. coli* BL21 strain (BL21-CodonPlus(DE3)-RIL) which carries extra copies of the necessary tRNA genes [18]. Most of the produced Ara h 2, however, was deposited in inclusion bodies, most likely caused by improper disulfide bond formation occurring in the reducing bacterial cytoplasm.

Ara h 2 shows up to 34% sequence identity to members of the 2S albumin seed storage proteins, but in contrast to allergenic 2S albumins it exists mainly as one continuous peptide chain with a molecular mass of 17.5 kDa in the peanut seed. Therefore, the heterologously expressed protein from *E. coli* (rAra h 2) can directly be compared to that isolated from peanut (nAra h 2). In the present study, we report the bacterial expression of authentically folded Ara h 2, which was compared to peanut Ara h 2 by CD, immunoblotting, and HPLC.

Methods

Construction of the expression vector pET-32a–Trx–(His)6–PreS–Ara h 2 and pET-16b-Ara h 2

For the construction of the expression vectors for mature Ara h 2 protein a full-length Ara h 2 isoform cloned into pET-16b was used as a template. The Ara h 2 sequence corresponding to the GenBank entry (Accession No. L77197), except for the nucleotides covering the N-terminal leader sequence (LTILVALALFLLAAHASA), was initially transferred into the cloning vector pBluescript KS by conventional PCR techniques and verified by DNA sequencing. For the construction of pET-16b-Ara h 2 this sequence was cloned in frame into the T7 promoter driven expression vector pET-16b (Novagen) by *Nco*I/*Bpu*1102I restriction sites, yielding a tagless Ara h 2 expression vector. The restriction sites were incorporated into the Ara h 2 gene by PCR using ARA-5′ sense primer 5′-gga gga gaa ttc acc atg gcg cgt cag cag tgg gaa ctc caa gga gac-3′ and ARA-3′ antisense primer 5′-gga gga aag ctt gct cag ctt agt atc tgt ctc tgc cgc cac-3′. For the expression vector pET-32a–Trx–$(His)_6$–PreS–Ara h 2, the Ara h 2 sequence was cloned in frame into the T7 promotor driven expression vector pET-32a (Novagen) by *Bgl*II/*Bpu*1102I restriction sites, and the sequence for an N-terminal PreScission protease (AmershamBiosciences, Freiburg, Germany) cleavage site (Leu-Glu-Val-Leu-Phe-Gln ↓ Gly Pro) was introduced simultaneously upstream of the Ara h 2 coding sequence. The sequences for the restriction and the PreScission protease cleavage sites were introduced by PCR using ARA-PRES-5′ sense primer 5′-gga gga gaa ttc aga tct gct gga agt gct gtt tca ggg ccc gcg tca gca gtg gga act cca agg-3′ and ARA-3′ antisense primer. The resulting vector contains an N-terminal thioredoxin fusion (Trx-tag) and a hexahistidine tag ($(His)_6$-tag). The PreScission protease recognition site introduces two additional amino acids (Gly and Pro) at the Ara h 2 N-terminus after cleavage.

Both expression vectors were used to transform competent *E. coli* BL21-CodonPlus(DE3)-RIL (Novagen, Madision, USA), and *E. coli* Origami(DE3) (Novagen) as well as *E. coli* AD494(DE3) (Novagen), both cotransformed with a modified pUBS520 (pUBS520-Cm). The plasmid which carries a chloramphenicol resistance gene (Cm gene) instead of the original kanamycin resistance gene allows the coexpression of *argU* tRNA [19]. The switch of the selection marker in the pUBS520 plasmid was necessary because of the genomic kanamycin resistance of Origami and AD494. The Cm gene coding sequence was obtained from pLysS plasmid (Novagen) by PCR methods using the pLysS-5′ sense primer 5′-gga gga aag ctt taa ata aga tca cta ccg ggc g-3′ and the pLysS-3′ antisense primer 5′-gga gga gga tcc gcg tag cac cag gcg ttt aag g-3′and cloned into pUBS520 by *Hin*dIII and *Bam*HI restriction sites. The *Hin*dIII recognition site is located in the coding sequence of the kanamycin resistance gene, thus eliminating the kanamycin resistance of the plasmid.

Expression of rAra h 2

Starter cultures of 10–100 ml LB medium containing appropriate antibiotics (Origami(DE3)/pUBS520-Cm:

34 μg/ml chloramphenicol, 20 μg/ml kanamycin, 100 μg/ml ampicillin, and 20 μg/ml tetracyclin; AD494(DE3)/pUBS520-Cm: 34 μg/ml chloramphenicol, 20 μg/ml kanamycin, and 100 μg/ml ampicillin; BL21-CodonPlus(DE3)-RIL: 34 μg/ml chloramphenicol and 100 μg/ml ampicillin) were inoculated with the three different *E. coli* expression strains and grown at 37 °C overnight. For the cultivation of cells in M9 [20] 10 ml medium containing the trace element solution TS2 [21], 2 mM $MgSO_4$, 10 μM Fe(III)citrate, 0.1 mM $CaCl_2$, 80 mg/l Glucose, antibiotics, and 40 μg/ml L-leucine were inoculated with the appropriate *E. coli* strains and cultivated at 37 °C. The cells were diluted 1:6 after 48 h and 1:3 after 72 h and cultivated overnight. Starter cultures were used to inoculate 3 l LB or M9 media containing appropriate antibiotics at an OD_{600} of 0.2–0.3. Cultures were grown to an OD_{600} of 0.8–0.9 and the expression of protein was induced by adding isopropylthiogalactoside (IPTG) to a final concentration of 1 mM. After induction, cultures were incubated at 37 °C for 3 or 4 h. The cell pellet was collected by centrifugation (5000*g*) and frozen.

Purification of rAra h 2 from E. coli

The cells were dissolved in lysis buffer [8 ml/g cell pellet; 20 mM Tris, pH 7.5; one tablet of Protease inhibitor cocktail (Roche, Basel, Switzerland)] and disrupted with lysozyme (100 μg/ml; room temperature, 30 min) and three 30 s sonication steps ($T = 4$ °C: Labsonic U sonicator at 150 W intensity). The extract was clarified by centrifugation for 30 min at 24966*g* and 4 °C.

Trx–Ara h 2 fusion protein overexpressed in *E. coli* Origami(DE3)/pUBS520-Cm/pET-32a–Trx–$(His)_6$–PreS–Ara h 2 was purified from the soluble fraction by Fast Performance Liquid Chromatography on TALON Superflow Resin (Clontech, Heidelberg, Germany) based on the strong binding of the $(His)_6$-tag to immobilized cobalt ions of the resin. After pre-equilibration of the column with 45 ml equilibration buffer (50 mM sodium phosphate buffer, 300 mM NaCl, pH 8, flow rate 2 ml/min), 20 ml of the soluble extract (20 ml) was loaded on a column with 15 ml bed volume (TALON Superflow Resin, flow rate 1 ml/min). The washing step was performed with 75 ml equilibration buffer (flow rate 1.5 ml/min). Adsorbed proteins were eluted with increasing concentrations of imidazole (0–1 M imidazole, 175 ml, flow rate 1.5 ml). The fractions containing Trx–Ara h 2 fusion protein were dialyzed (Spectra/Por Membrane, MWCO: 1.000, ROTH, Karlsruhe, Germany) against 5 liter PreScission cleavage buffer (50 mM sodium phosphate and 150 mM NaCl, pH 8) with two buffer changes.

The dialyzed protein was incubated with PreScission protease (0.8 U/100 μg protein, Amersham–Pharmacia Biotech, Freiburg, Germany) for 30–60 min at room temperature. To remove the Trx-tag and the $(His)_6$-tag as well as uncleaved fusion protein the sample was loaded (1 ml/min) on a pre-equilibrated (45 ml, 50 mM sodium phosphate buffer, 300 mM NaCl, pH 8, flow rate 2 ml/min) TALON resin. The column was washed with 75 ml equilibration buffer (flow rate 1.5 ml/min), and adsorbed Trx–$(His)_6$ and uncleaved protein were eluted with a gradient of imidazole (120 ml, 50 mM sodium phosphate buffer, 300 mM NaCl, 0–1 M imidazole, pH 8, flow rate 1.5 ml/min). Ara h 2 found in the flow-through fractions was pooled, dialyzed (Spectra/Por Membrane, as above) twice against 5 liter water for more than 2 h, lyophilized, and stored at 4 °C. For circular dichroism (CD) spectroscopic analysis traces of PreScission protease, which was used as GST fusion protein, were removed by affinity chromatography using glutathione sepharose resin (GSTrap column, Amersham Biosciences, Freiburg, Germany).

Purification of Ara h 2 from peanut

Unshelled peanuts were purchased from a local grocery, ground in a mortar with a pestle, and stored at 4 °C. Peanut protein extracts were prepared as described earlier [22] with minor modifications. About 5 g ground peanuts was suspended in 100 ml buffer, Tris, 20 mM, pH 8.2, and stirred for 2 h at room temperature. The soluble fraction was collected by centrifugation (5 min, 5000*g*, 4 °C). Residual traces of fat and insoluble particles were removed by subsequent centrifugation (30 min; 24966*g*; 4 °C). Extracts were clarified with folded filters (Ø 240 mm, Schleicher & Schuell, Dassel, Germany) and immediately used for further purification.

About 67 ml crude extract was loaded (flow rate 1 ml/min) onto Q-Sepharose Fast Flow Resin (30 ml; Amersham Biosciences, Freiburg, Germany) equilibrated (flow rate 2 ml/min) with 120 ml Tris, 20 mM, pH 8. After washing the column with 90 ml Tris, 20 mM, pH 8 (flow rate 1.5 ml/min), two constant salt gradients of 60 ml (0–100 mM NaCl in Tris, 20 mM, pH 8), and 450 ml (0.1–1 M NaCl in Tris, 20 mM, pH 8) were applied to elute the proteins (flow rate 1.5 ml/min). Fractions were analyzed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS–PAGE; 19% polyacrylamide), pure protein fractions were pooled and dialyzed twice against 5 liter water for more than 2 h and stored at 4 °C. Concentrations of samples containing rAra h 2 were determined by 280 nm absorbance using the native extinction coefficient $\varepsilon_{280} = 9126\ M^{-1}\ cm^{-1}$ [23].

Analytical reversed-phase HPLC (RP-HPLC)

Chromatographic analysis of rAra h 2 and nAra h 2 was performed using a Kontron HPLC system (Kontron Instruments, Osterode, Germany), including a diode array detector (detection at 220 and 280 nm) on an

C18-RCM 8 × 10 module (Waters, 8 × 10 mm, Eschborn, Germany). Elution was carried out with a constant gradient of 0–56% acetonitrile in 0.1% trifluoroacetic acid (TFA) over 50 min (flow rate 1.5 ml/min).

Immunoblotting

After separation by SDS–PAGE (19% (w/v) acrylamide), proteins were transferred to a nitrocellulose membrane (pore size 0.45 μm, Sartorius). Blotting was performed for 2 h at 28 mA using a Semi-Phor Semi-Dry transfer unit (Amersham Biosciences, Freiburg, Germany) and an MPC 2 × 3000 (MWG Biotech, Ebersberg, Germany) power supply. Transferred proteins were visualized by reversible Ponceau S (0.2% in acetic acid) staining. For further immunochemical detection the membrane was blocked by incubation in 0.1 M Tris-buffered saline (pH 7.4, TBS), 0.05% (v/v) Tween 20 (TBST), for 1 h under gentle agitation. Subsequently, the membrane was incubated in an appropriate dilution of the primary antibody. Patient serum JG3 was diluted 1:20 in TBST. Rabbit antiserum raised against rAra h 2 was applied in $1{:}1 \times 10^6$ dilution. Specific binding of patient IgE or rabbit antibodies was detected using a monoclonal mouse antibody anti-human-IgE (diluted 1:2000; Allergopharma, Reinbek, Germany) or a goat antibody anti-rabbit-IgG (1:5000; Dianova, Hamburg, Germany) as secondary antibodies, both conjugated with alkaline phosphatase. Bound antibody was visualized after repeated washings with TBST by incubation with nitroblue tetrazolium/5-bromo-4-chloro-3-indolyl phosphate (NBT/BCIP) as substrate in AP-buffer (0.1 M Tris-buffered saline, pH 9.5).

Gel electrophoresis

Routine monitoring of protein purity and apparent molecular mass was controlled by denaturating SDS–PAGE (19% (w/v) polyacrylamide, [24]). Samples were diluted with Roti-Load 1 reducing sample buffer (ROTH, Karlsruhe, Germany) and denaturated for 5–10 min at 95 °C.

N-terminal sequencing

Peanut extract was separated by SDS–PAGE and blotted onto a polyvinylidene fluoride (PVDF) membrane. Relevant protein bands were excised and subjected to N-terminal protein sequencing by sequential Edman degradation.

CD

CD studies were performed on a Jasco 600 spectropolarimeter (Japan Spectroscopic, Groß-Umstadt, Germany) equipped with a CDF-426S Peltier temperature control system interfaced with a Julabo F200 water bath (Julabo Labortechnik GmbH, Seelbach, Germany). Far UV CD spectra were recorded at 20 °C and 100 °C in a 1-mm path-length quartz cuvette (200 μl, Hellma, Müllheim, Germany) at protein concentrations of 5–10 μM. Samples were dissolved in 10 mM potassium phosphate, pH 7, and spectra were recorded with a scanning speed of 20 nm/min, and eight scans were averaged. Buffer spectra were subtracted. The results are expressed as mean residue molar ellipticity $[\Theta]_{MRW} = \Theta/(c \cdot d \cdot N)$; Θ, observed ellipticity; N, number of amino acid residues; d, optical path length; and c, protein concentration.

Results and discussion

Expression levels of rAra h 2 in E. coli hosts

Ara h 2 cDNA includes a high number of codons, especially arginine codons, rarely used in *E. coli*. Additionally, the protein contains eight cysteine residues, which are arranged in disulfide bonds. Cystine containing proteins tend to be expressed in bacteria with reducing cytoplasm in a mis-folded form and to accumulate in inclusion bodies [25]. However, disulfide bonds can form in cells that are defective for certain components of the reducing pathway [26,27]. We used *E. coli* Origami(DE3) with a *trxB*/*gor* double mutation and *E. coli* AD494(DE3) with a *trxB* mutation. It could be earlier shown in our laboratory that *E. coli* Origami(DE3) is a suitable host for the heterologous expression of a properly folded and accurately disulfide linked extracellular protein with several cysteines [28]. To enhance the expression rate of Ara h 2 in these strains an additional plasmid (pUBS520-Cm) carrying argU tRNA genes was transferred into Origami(DE3) and AD494(DE3).

Ara h 2 expression levels varied with strain (*E. coli* BL21-CodonPlus(DE3)-RIL, *E. coli* Origami(DE3)/pUBS520-Cm, *E. coli* AD494(DE3)/pUBS520-Cm) as well as with expression vector (pET 32a, pET-16b; Fig. 1). pET 32a was used to express the Ara h 2 sequence fused with thioredoxin. We used thioredoxin fusion tags because they do not only tend to enhance the solubility of proteins normally produced in an insoluble form [29] but also appear to catalyze the formation of disulfide bonds in the cytoplasm of *trxB* mutants like *E. coli* Origami(DE3) and AD494(DE3) [30]. Expression of Ara h 2 without affinity tag or fusion partner was performed with pET-16b vector. The expression of tagless Ara h 2 as pET-16b construct, which was very efficient using *E. coli* BL21-CodonPlus(DE3)-RIL cells (Fig. 1A), was rather poor with strains possessing an oxidative cytoplasm like *E. coli* AD494(DE3)/pUBS520-S (Fig. 1B). Thioredoxin–Ara h 2 fusion protein could

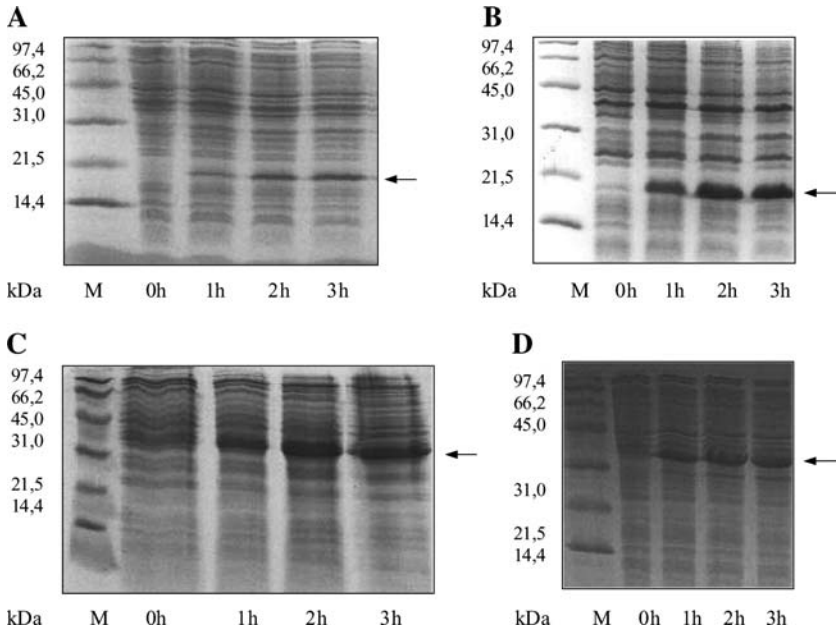

Fig. 1. Comparison of expression yields of rAra h 2 with *E. coli* BL21-CodonPlus(DE3)-RIL-pET16b (A), *E. coli* AD494(DE3)/pUBS520M-pET16b (B), *E. coli* AD494(DE3)/pUBS520M-pET32a (C), and in *E. coli* OrigamiTM(DE3)/pUBS520M-pET32a (D). Thioredoxin–Ara h 2 fusion protein (C,D) was expressed from pET-32a vector and tagless rAra h 2 (A,B) was expressed from pET-16b vector. AD494(DE3) and Origami(DE3) contain the extra plasmid pUBS520M carrying argU tRNA genes. Expression was performed over a period of 4 h at 37 °C. Aliquots of total cell lysates were collected before induction (0 h) and at 1, 2, and 3 h after IPTG induction, analyzed by 19% SDS–PAGE, and visualized by Coomassie blue staining.

be expressed strongly in the disulfide bridge promoting strains AD494(DE3) (Fig. 1C) and *E. coli* Origami(DE3) (Fig. 1D) supplemented with pUBS520-Cm, with the expression level in *E. coli* Origami(DE3) being slightly higher than in *E. coli* AD494(DE3).

The ratio of Ara h 2 in the soluble and inclusion body fraction after cell disruption varied extremely with *E. coli* strain and vector (Table 1). In the case of tagless expression using the pET-16b vector in combination with each strain more than 90% of Ara h 2 was deposited in inclusion bodies. In contrast, approximately 50% of thioredoxin tagged Ara h 2 was found in the soluble fraction after expression in the disulfide promoting strains. Thus, *E. coli* Origami(DE3) as well as AD494(DE3) in combination with pET 32a and pUBS520-Cm are suitable for the production of a rAra h 2 with properties similar to the nAra h 2. *E. coli* Origami(DE3) was used for further experiments.

Table 1
Ratio of Ara h 2 in the soluble and inclusion body fraction using different strain vector combinations

Strain	Vector	Inclusion bodies	Soluble
E. coli BL21-Codon-Plus(DE3)-RIL	pET 16b	>95%	<5%
E. coli AD494(DE3)	pET 16b	90%	10%
	pET 32a	50%	50%
E. coli Origami (DE3)	pET 16b	>95%	<5%
	pET 32a	50%	50%

Equal amounts of the soluble and inclusion body fraction were analyzed by 19% SDS–PAGE and visualized by Coomassie blue staining. The ratio of Ara h 2 in the two fractions was estimated.

Purification and yield of rAra h 2

After cell lysis of *E. coli* Origami(DE3)/pUBS520-Cm/pET-32a–Trx–$(His)_6$–PreS–Ara h 2 the soluble fraction was purified with Co^{2+} immobilized metal affinity chromatography (IMAC) under native conditions (Fig. 2A). $(His)_6$-tagged thioredoxin–Ara h 2 fusion protein was eluted with 200 mM imidazole. The average yield of eluted fusion protein was about 46 mg/l culture in LB-medium and 30 mg/l in M9-medium. $(His)_6$-tagged thioredoxin could be removed from Ara h 2 using PreScission protease. Cleavage was very efficient and specific, no secondary cleavage products were detectable. Ara h 2 could be separated from $(His)_6$-tagged thioredoxin fusion partner by a second Co^{2+}-IMAC (Fig. 2B). Eluted fractions were analyzed by SDS–PAGE (Fig. 2C, lanes 10–13). As Ara h 2 contains no polyhistidine-tag after cleavage, it appears in the column flow-through (Fig. 2C, lanes 15–19). About 19 mg/l culture in LB medium and 11 mg/l culture in minimal medium of pure rAra h 2 were thus obtained.

One step purification of nAra h 2

Ara h 2 was purified from peanut extract with a one-step procedure without defatting steps with organic

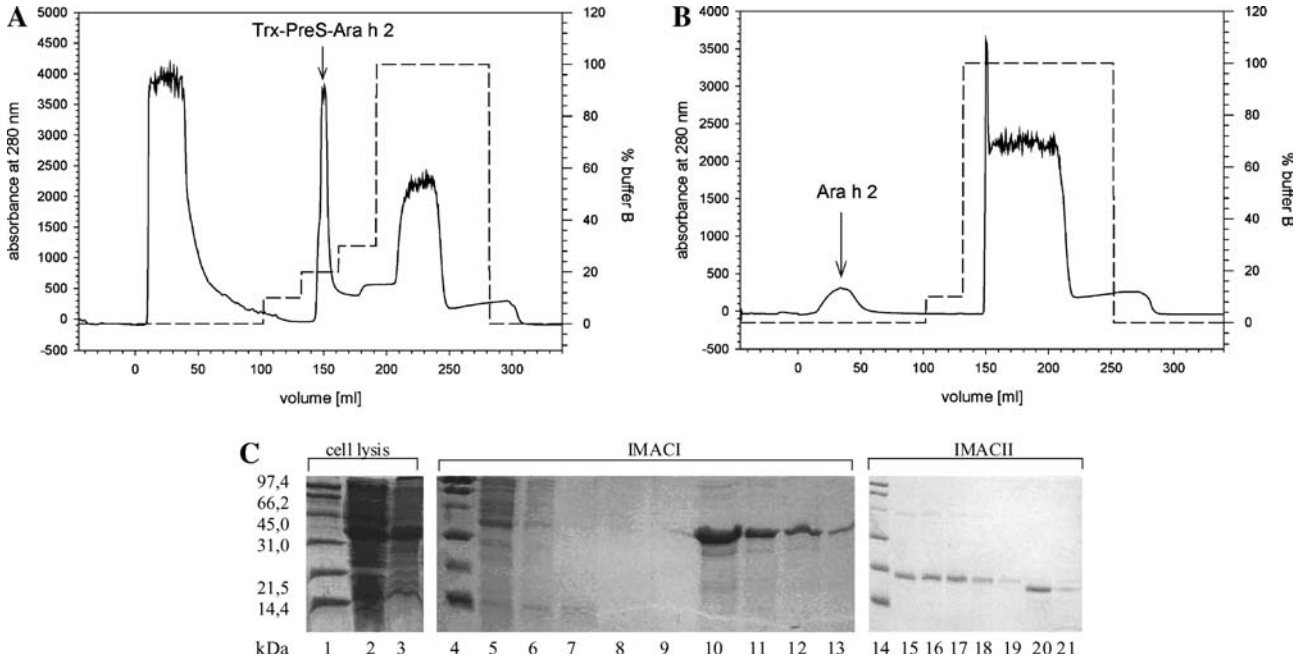

Fig. 2. IMAC purification of rAra h 2. TALON Superflow Resin column; buffer A, 50 mM sodium phosphate and 300 mM NaCl; buffer B, 50 mM sodium phosphate, 300 mM NaCl, and 1 M imidazole; monitoring absorbance at 280 nm. Buffer A was used as equilibration and wash buffer. (A) First IMAC for purification of His_6–Ara h 2–thioredoxin fusion protein (Trx–$(His)_6$–Ara h 2). Soluble extract was loaded onto TALON Superflow Resin. Elution of bound His_6 fusion protein was obtained by increasing concentration of buffer B (0–1 M imidazole; dashed line). (B) Second IMAC after tag removal by PreScission protease cleavage. Cleaved fusion protein was loaded onto TALON Superflow Resin. Tagless rAra h 2 was found in the flow-through and Trx–$(His)_6$ was eluted with 100% buffer B (dotted line). (C) Samples from purification of rAra h 2. 1, 4, and 14, molecular weight standard (BIO-RAD); 2, insoluble pellet after cell lysis; 3, soluble supernatant after cell lysis; 5 and 6, flow-through from first IMAC; 7–9 wash fraction from first IMAC; 10–13, Trx–$(His)_6$–Ara h 2; 15–19, column flow-through of second IMAC (tagless Ara h 2); 20–21, His_6-tagged thioredoxin. Samples were separated with 19% SDS–PAGE and visualized by Coomassie blue staining.

solvents, since these conditions could denature proteins [22]. N-terminal sequencing of two closely migrating bands from the cell-free extract revealed the sequence RQQXELQGDRR, identical to the previously published N-terminus of Ara h 2 [6,7] as well as that deduced from the cDNA sequence (GenBank Accession No. L77197). Ara h 2 was purified from the extract with anion exchange chromatography which was eluted at 150 mM NaCl (Fig. 3A). Eluted fractions were analyzed by SDS–PAGE (Fig. 3B), and fractions 5 and 6 contained very pure Ara h 2. As observed earlier, Ara h 2 appears as a double band in SDS–PAGE stained with Coomassie brilliant blue or immunoblots [6].

IgE recognition of nAra h 2 and rAra h 2

We analyzed protein from either source at different preparation stages by Coomassie blue staining (Fig. 4A) and immunoblots (Figs. 4B and C). To establish immunological equivalence of rAra h 2 and nAra h 2 isolated from peanut, we performed immunoblot experiments using rabbit antiserum raised against rAra h 2. A dominant band corresponding to thioredoxin–Ara h 2 fusion protein (M_r 38 kDa) was detected in *E. coli* cell free extract (Fig. 4B, lane 1). Rabbit antiserum showed a clear reaction to purified Ara h 2 fusion protein (Fig. 4B, lane 2) and tagless Ara h 2 (M_r 18 kDa, Fig. 4B, lane 3). Additional bands of lower intensity recognized by rabbit antiserum may stem from by-products in the rAra h 2 used for the immunization of rabbits. Strong reactions to rabbit antiserum were also observed with peanut extract and Ara h 2 purified from peanut (Fig. 4B, lanes 4 and 5).

Furthermore, we performed immunoblot experiments using serum derived from a peanut allergic patient. Strong IgE binding to nAra h 2 as well as to rAra h 2 was detected (Fig. 4C). The two bands of nAra h 2 have identical immunological properties. Patients' IgE and rabbit antiserum recognize both bands equally well.

HPLC analysis of nAra h 2 and rAra h 2

Reverse phase HPLC (RP-HPLC) experiments were carried out to evaluate the purity of the Ara h 2 samples and the identity of the disulfide pattern of natural and recombinant Ara h 2. Separation of proteins in RP-HPLC is mainly based on the different hydrophobicities of the sample components. As a modified disulfide pattern can result in a changed distribution of hydrophobicity of the protein, the same elution profiles of two proteins indicate identity of disulfide linkage. We used rAra h 2 in its oxidized, reduced, and reoxidized form to demonstrate

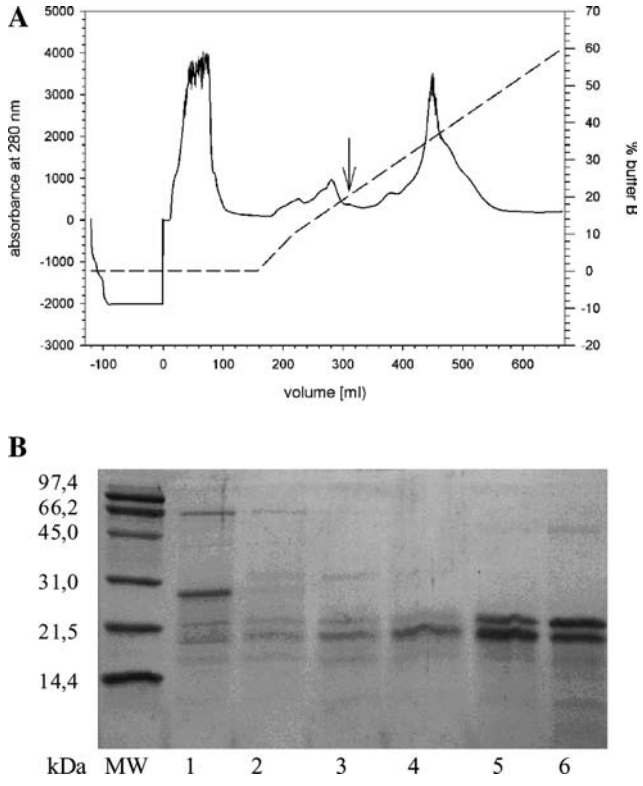

Fig. 3. nAra h 2 purification by anion exchange chromatography, Q-Sepharose Fast Flow column, buffer A, 20 mM Tris, pH 8; buffer B, 20 mM Tris and 1 M NaCl, pH 8; monitoring absorbance at 280 nm. Buffer A was used as equilibration and wash buffer. (A) Crude extract of peanut seeds was loaded onto Sepharose Fast Flow. Bound proteins were eluted using constant gradients of buffer B (0–0.1 M NaCl, 60 ml; 0.1–1 M NaCl, 450 ml; dashed line). nAra h 2 is marked with an arrow. (B) Fractions eluted from column. MW, molecular weight standard (Bio-Rad); 1–6, nAra h 2 containing fractions. Samples were separated with 19% SDS–PAGE.

that different disulfide linkages in Ara h 2 cause different retention times in RP-HPLC using a C18-column.

rAra h 2 expressed in an *E. coli* Origami(DE3)/pUBS520-Cm host with oxidized cytoplasm was tested for free thiol groups. No reaction was observed with DTNB (5,5′-Dithio-bis(2-nitrobenzoic acid)) using a 0.5 mg/ml solution of Ara h 2, indicating that all cysteines are oxidized and involved in disulfide bridges.

Recombinant Ara h 2 could be reduced completely with 10 mM DTT, resulting in one single peak at a different retention time compared to the oxidized form (Figs. 5A and B). Reoxidation was performed by removal of the reducing agent by buffer exchange with an NAP-10 column or dialysis. After this procedure a variety of different species could be detected by RT-HPLC analysis (Fig. 5C). These results indicate that reoxidation of reduced disulfide bridges leads to the formation of disulfide isomers in the case of Ara h 2. Obviously, recovery of a uniform and native disulfide connection after reduction cannot be easily obtained with Ara h 2.

Whether this disulfide pattern is identical to that of nAra h 2 purified from peanut was analyzed by additional RP-HPLC experiments. rAra h 2 eluted with a retention time of 32.06 min from the C18-column (Fig. 5D). nAra h 2 eluted at nearly the same retention time (32.12 min) as a main peak and two weak additional peaks at later retention times from isoforms or contaminations (Fig. 5E). Co-injection of rAra h 2 and nAra h 2 resulted in a strong single peak (Fig. 5F). In contrast to the results of the reoxidation experiments, which show a very heterogeneous peak pattern, it can be supposed that natural and recombinant Ara h 2 which elute in one single peak have identical disulfide pattern.

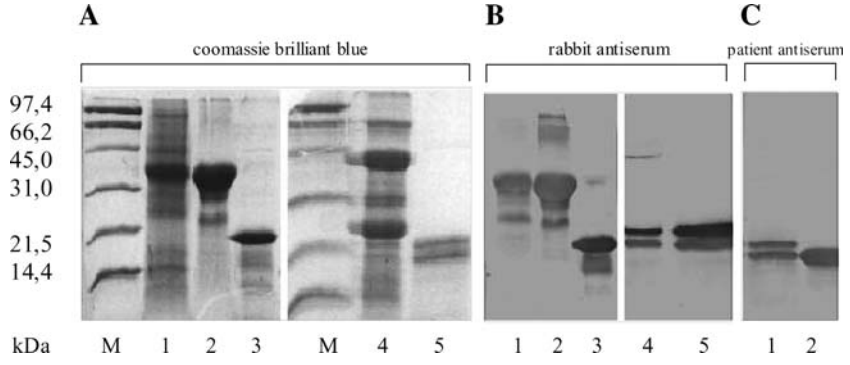

Fig. 4. SDS–PAGE and immunoblots of nAra h 2 and rAra h 2. (A) 19% SDS–PAGE of *E. coli* cell free extract (1), Trx–$(His)_6$–Ara h 2 fusion protein (2), purified rAra h 2 (3), peanut extract (4), and purified nAra h 2 (5). rAra h 2 appears as single band at a molecular mass of 17–19 kDa and nAra h 2 appears as two closely migrating bands with a mean molecular mass of 17–20 kDa. Bio-Rad molecular standard was used. (B) Immunoblotting of rAra h 2 and nAra h 2 using rabbit antiserum raised against raRa h 2. Rabbit antiserum shows a clear reaction to rAra h 2 fusion protein (1: cell free extract and 2: purified), tagless rAra h 2 (3) and nAra h 2 (4: peanut extract and 5: purified). (C) Immunoblotting of nAra h 2 (1) and rAra h 2 (2) using a peanut allergic patient serum. Specific binding of patient IgE or rabbit antibodies was detected using a monoclonal mouse antibody anti-human-IgE or a goat antibody anti-rabbit-IgG as secondary antibodies both conjugated with alkaline phosphatase. Bound antibody was visualized by incubation with NBT/BCIP.

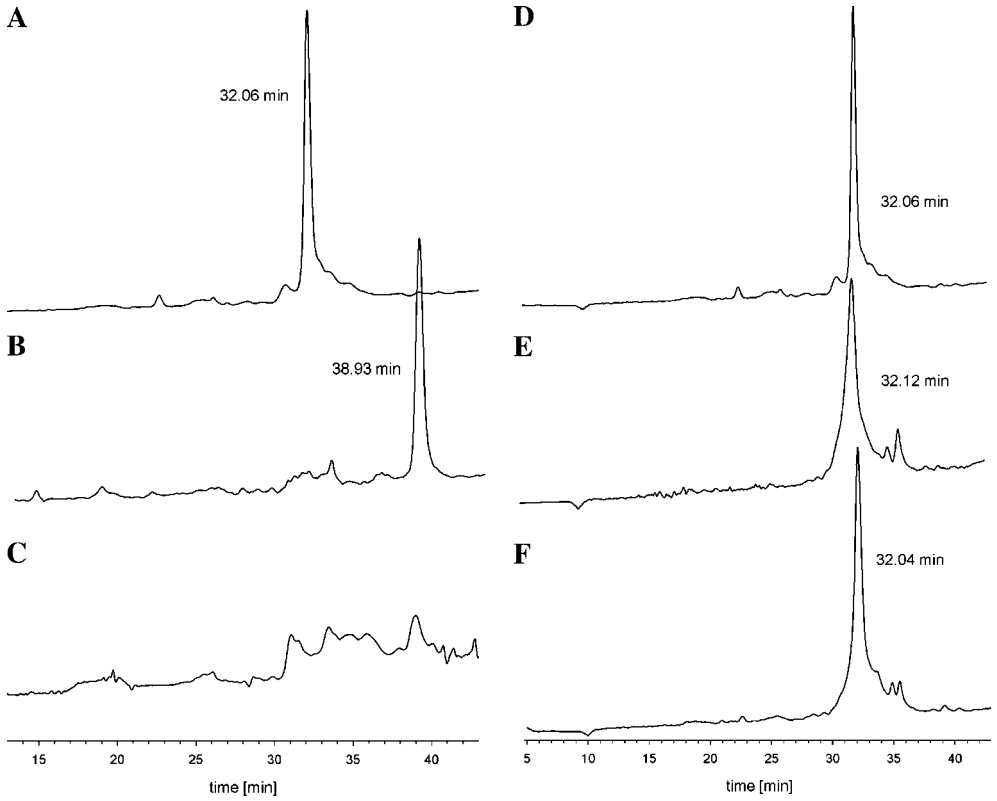

Fig. 5. (A)–(C) Analytical RP-HPLC elution profile of oxidized, reduced, and reoxidized rAra h 2. The bound proteins were eluted using a constant gradient (0–70% in 50 min) of 80% acetonitrile in aqueous 0.1% TFA (buffer B). C18-Column, RCM 8 × 10 module (Waters, 8 × 10 mm); 1.5 ml/min; monitoring absorbance at 220 nm. (A) rAra h 2 without reducing agent. (B) rAra h 2 in the presence of 10 mM DTT. (C) rAra h 2 after removal of DTT by buffer exchange with a NAP-10 column. (D)–(F) HPLC profiles of rAra h 2 (D), nAra h 2 (E), and co-injected rAra h 2/nAra h 2 (F).

CD comparison of nAra h 2 and rAra h 2

To compare the secondary structure of nAra h 2 and rAra h 2 purified from *E. coli* Origami CD spectra were measured (Fig. 6). nAra h 2 consists of at least two isoforms in an unknown ratio. Therefore, the calculation of an accurate mean residue molar ellipticity per isoform was impossible, and the spectra of nAra h 2 (Fig. 6A) and rAra h 2 (Fig. 6B) can only be compared qualitatively. The two CD-spectra are nearly identical and slight differences may be due to the isoform composition of nAra h 2. The secondary structure of nAra h 2 and of rAra h 2 appears to be identical. The CD spectra of Ara h 2 from 190 to 260 nm show two minima at 208–210 and 222 nm and a maximum at 190–193 nm, typical for largely α-helical proteins. In contrast to the previously published results, we have absolutely no indications to β-strands [31]. In Addition, the consensus secondary structure prediction with J-Pred (http://jura.ebi.ac.uk:8888) also shows five α-helices [a1 (14–19), a2 (23–32), a3 (65–74), a4 (84–98), and a5 (103–113)], three short [(20–22), (75–83), and (99–102)] and a long region (33–64) between helices a2 and a3 without regular secondary structure, but no β-strands.

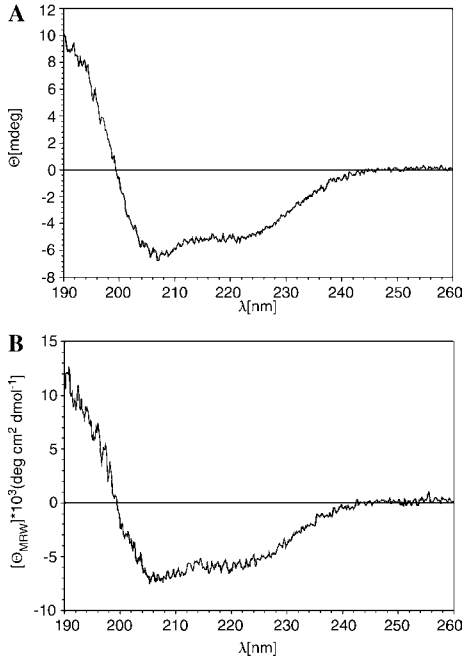

Fig. 6. Far-UV (190–260 nm) CD spectra of nAra h 2 (A) and rAra h 2 (B), 10 mM potassium phosphate, pH 7.

Conclusions

The present strategy is clearly suitable for recombinant expression in *E. coli* of the major peanut allergen Ara h 2 and highly likely other members of this protein family with native fold in quantities sufficient for structural studies at atomic resolution by X-ray crystallography and multidimensional NMR-spectroscopy. The intact secondary structure, disulfide pattern, and immunological reactivity of the recombinant protein were verified using the native protein from peanut. Preliminary CD spectroscopic data indicate that it is a predominantly α-helical protein. Steps towards the determination of the high-resolution structure of Ara h 2 are currently in progress in our laboratory.

Acknowledgments

We thank N. Herz for expert technical assistance. We gratefully acknowledge B. Lindner for MALDI-TOF mass analysis. This work was supported by grants from the Deutsche Forschungsgemeinschaft (Ro617/11-1/-2) to P.R and (Be730/3-2) to W.M.B.

References

[1] B.S. Bochner, L.M. Lichtenstein, Anaphylaxis, N. Engl. J. Med. 324 (1991) 1785–1790.
[2] J.W. Yunginger, K.G. Sweeney, W.Q. Sturner, L.A. Giannandrea, J.D. Teigland, M. Bray, P.A. Benson, J.A. York, L. Biedrzycki, D.L. Squillace, Fatal food-induced anaphylaxis, Jama 260 (1988) 1450–1452.
[3] H.A. Sampson, L. Mendelson, J.P. Rosen, Fatal and near-fatal anaphylactic reactions to food in children and adolescents, N. Engl. J. Med. 327 (1992) 380–384.
[4] F.M. Atkins, S.S. Steinberg, D.D. Metcalfe, Evaluation of immediate adverse reactions to foods in adult patients. II. A detailed analysis of reaction patterns during oral food challenge, J. Allergy Clin. Immunol. 75 (1985) 356–363.
[5] S.A. Bock, F.M. Atkins, The natural history of peanut allergy, J. Allergy Clin. Immunol. 83 (1989) 900–904.
[6] A.W. Burks, L.W. Williams, C. Connaughton, G. Cockrell, T.J. O'Brien, R.M. Helm, 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. 90 (1992) 962–969.
[7] J.S. Stanley, N. King, A.W. Burks, S.K. Huang, H. Sampson, G. Cockrell, R.M. Helm, C.M. West, G.A. Bannon, Identification and mutational analysis of the immunodominant IgE binding epitopes of the major peanut allergen Ara h 2, Arch. Biochem. Biophys. 342 (1997) 244–253.
[8] T. Kleber-Janke, R. Crameri, U. Appenzeller, M. Schlaak, W.M. Becker, Selective cloning of peanut allergens, including profilin and 2S albumins, by phage display technology, Int. Arch. Allergy Immunol. 119 (1999) 265–274.
[9] A.W. Burks Jr., J.R. Brooks, H.A. Sampson, Allergenicity of major component proteins of soybean determined by enzyme-linked immunosorbent assay ELISA, immunoblotting in children with atopic dermatitis and positive soy challenges, J. Allergy Clin. Immunol. 81 (1988) 1135–1142.
[10] S.S. Teuber, K.C. Jarvis, A.M. Dandekar, W.R. Peterson, A.A. Ansari, Identification and cloning of a complementary DNA encoding a vicilin-like proprotein, jug r 2: from english walnut kernel *Juglans regia*, a major food allergen, J. Allergy Clin. Immunol. 104 (1999) 1311–1320.
[11] E.S. Gander, K.O. Holmstroem, G.R. De Paiva, L.A. De Castro, M. Carneiro, M.F. Grossi de Sa, Isolation, characterization, and expression of a gene coding for a 2S albumin from *Bertholletia excelsa* Brazil nut, Plant Mol. Biol. 16 (1991) 437–448.
[12] R.I. Monsalve, M.A. Gonzalez de la Pena, L. Menendez-Arias, C. Lopez-Otin, M. Villalba, R. Rodriguez, Characterization of a new oriental-mustard *Brassica juncea* allergen, Bra j IE: detection of an allergenic epitope, Biochem. J. 293 (1993) 625–632.
[13] L. Menendez-Arias, I. Moneo, J. Dominguez, R. Rodriguez, Primary structure of the major allergen of yellow mustard *Sinapis alba* L. seed, Sin a I, Eur. J. Biochem. 177 (1988) 159–166.
[14] M.L. Ericson, J. Rodin, M. Lenman, K. Glimelius, L.G. Josefsson, L. Rask, Structure of the rapeseed 1.7 S storage protein, napin, and its precursor, J. Biol. Chem. 261 (1986) 14576–14581.
[15] B. Heinemann, K.V. Andersen, P.R. Nielsen, L.M. Bech, F.M. Poulsen, Structure in solution of a four-helix lipid binding protein, Protein Sci. 5 (1996) 13–23.
[16] S. Odani, T. Koide, T. Ono, Y. Seto, T. Tanaka, Soybean hydrophobic protein. Isolation, partial characterization, and the complete primary structure, Eur. J. Biochem. 162 (1987) 485–491.
[17] M. Rico, M. Bruix, C. Gonzalez, R.I. Monsalve, R. Rodriguez, ^{1}H NMR assignment and global fold of napin BnIb, a representative 2S albumin seed protein, Biochemistry 35 (1996) 15672–15682.
[18] T. Kleber-Janke, W.M. Becker, Use of modified BL21DE3. *Escherichia coli* cells for high-level expression of peanut allergens affected by poor codon usage, Protein Expr. Purif. 19 (2000) 419–424.
[19] U. Brinkmann, R.E. Mattes, P. Buckel, High-level expression of genes in *Escherichia coli* is dependent on the availability of the dnaY gene product, Gene 85 (1989) 109–114.
[20] J. Sambrook, E.F. Fritsch, T. Maniatis, Molecular Cloning, A Laboratory Manual, Cold Spring Harbor, New York, 1989.
[21] O. Meyer, H.G. Schlegel, Biology of aerobic carbon monoxide-oxidizing bacteria, Annu. Rev. Microbiol. 37 (1983) 277–310.
[22] S.J. Koppelman, R.A. Vlooswijk, L.M. Knippels, M. Hessing, E.F. Knol, F.C. van Reijsen, C.A. Bruijnzeel-Koomen, Quantification of major peanut allergens Ara h 1 and Ara h 2 in the peanut varieties Runner, Spanish, Virginia, and Valencia, bred in different parts of the world, Allergy 56 (2001) 132–137.
[23] S.C. Gill, P.H. von Hippel, Calculation of protein extinction coefficients from amino acid sequence data, Anal. Biochem. 182 (1989) 3119–3326.
[24] U.K. Laemmli, Cleavage of structural proteins during the assembly of the head of bacteriophage T4, Nature 227 (1970) 680–685.
[25] G. Georgiou, P. Valax, Expression of correctly folded proteins in *Escherichia coli*, Curr. Opin. Biotechnol. 7 (1996) 190–197.
[26] A.I. Derman, W.A. Prinz, D. Belin, J. Beckwith, Mutations that allow disulfide bond formation in the cytoplasm of *Escherichia coli*, Science 262 (1993) 1744–1747.
[27] W.A. Prinz, F. Aslund, A. Holmgren, J. Beckwith, The role of the thioredoxin and glutaredoxin pathways in reducing protein disulfide bonds in the *Escherichia coli* cytoplasm, J. Biol. Chem. 272 (1997) 15661–15667.
[28] T. Lauber, U.C. Marx, A. Schulz, P. Kreutzmann, P. Rösch, S. Hoffmann, Accurate disulfide formation in *Escherichia coli*:

overexpression and characterization of the first domain HF6478 of the multiple Kazal-type inhibitor LEKTI, Protein Expr. Purif. 22 (2001) 108–112.

[29] E.R. LaVallie, E.A. DiBlasio, S. Kovacic, K.L. Grant, P.F. Schendel, J.M. McCoy, A thioredoxin gene fusion expression system that circumvents inclusion body formation in the *E. coli* cytoplasm, Biotechnology (NY) 11 (1993) 187–193.

[30] E.J. Stewart, F. Aslund, J. Beckwith, Disulfide bond formation in the *Escherichia coli* cytoplasm: an in vivo role reversal for the thioredoxins, EMBO J. 17 (1998) 5543–5550.

[31] M. Sen, R. Kopper, L. Pons, E.C. Abraham, A.W. Burks, G.A. Bannon, Protein structure plays a critical role in peanut allergen stability and may determine immunodominant IgE-binding epitopes, J. Immunol. 169 (2002) 882–887.