The cashew allergens:
a molecular and serological characterisation

Marit Reitsma

Thesis committee

Promotors

Prof. Dr H. J. Wichers
Professor of Immunomodulation by Food
Wageningen University & Research

Prof. Dr H. F. J. Savelkoul
Professor of Cell Biology and Immunology
Wageningen University & Research

Co-promotor

Dr N. W. de Jong,
Assistant professor, Department of Internal Medicine, section Allergology
Erasmus MC, Rotterdam

Other members

Prof. Dr M. A. J. S. van Boekel, Wageningen University & Research
Prof. Dr E. N. C. Mills, University of Manchester, UK
Dr K. Adel-Patient, INRA Food Allergy Laboratory, Gif-sur-Yvette, France
Dr P. V. Jeurink, Danone Nutricia Research, Utrecht

This research has been conducted under the auspices of the Graduate School VLAG.
(Advanced studies in Food technology, Agrobiotechnology, Nutrition and Health Sciences)

The cashew allergens:
a molecular and serological characterisation

Marit Reitsma

Thesis

submitted in fulfilment of the requirements for the degree of doctor
at Wageningen University
by the authority of the Rector Magnificus
Prof. Dr A. P. J. Mol
in the presence of the
Thesis Committee appointed by the Academic Board
to be defended in public
on Monday April 10 2017
at 11 a.m. in the Aula.

Marit Reitsma
“The cashew allergens: a molecular and serological characterization”
186 pages.

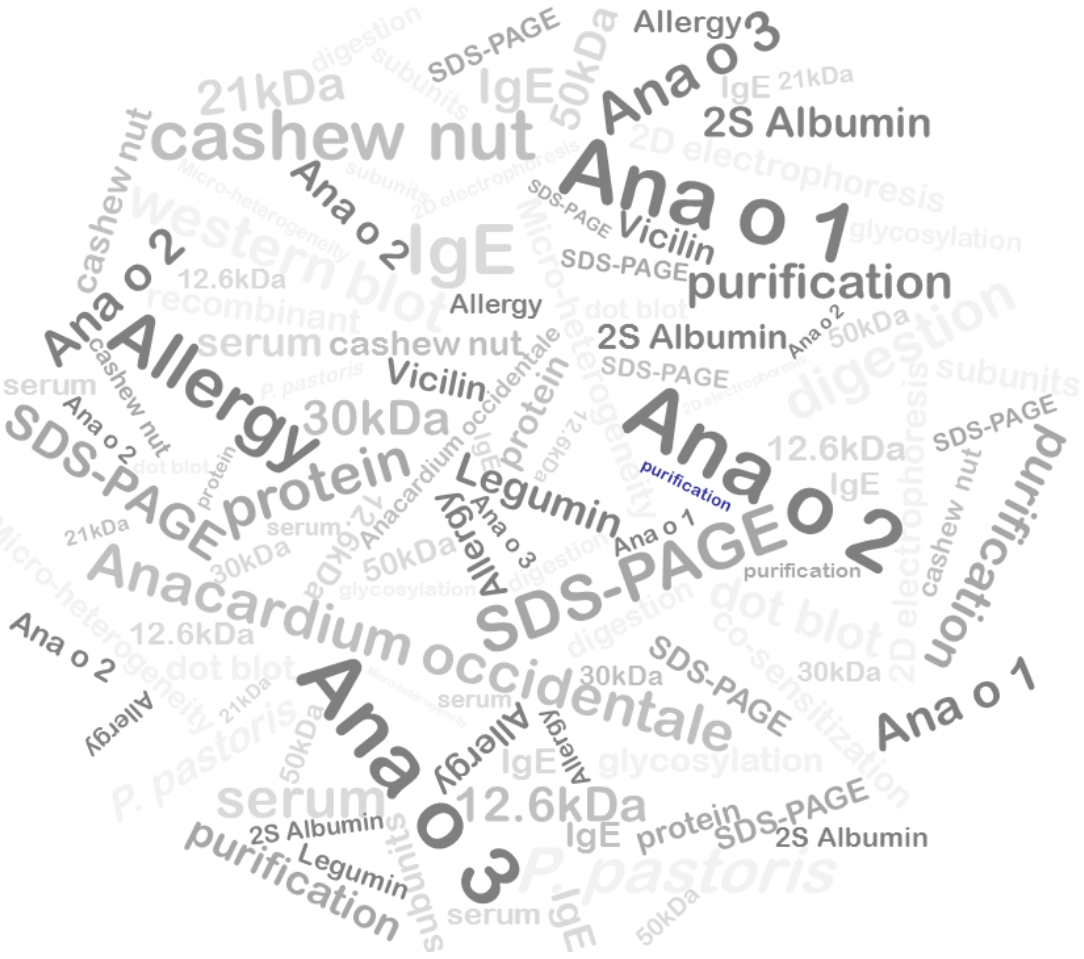
PhD thesis, Wageningen University, Wageningen, the Netherlands (2017)
With references, with summary in English

ISBN: 978-94-6343-078-4
DOI: 10.18174/403658

Table of Contents

Chapter 1

General introduction



Cashew allergy

Cashew allergy is a severe allergy [1-4] of which the prevalence appears to be increasing [6]. As information is scarce on prevalence, allergen characteristics, and cross-reactivity, the IDEAL project was initiated: Improvement of Diagnostic mEthods for ALlergy assessment of cashew allergy in children. In this project the clinical side of cashew allergy was combined with molecular studies in order to get a broad picture of cashew nut allergy and its responsible allergens. This was realised by involving three tertiary care centres for food allergy (Erasmus MC Rotterdam, University Medical Centre Groningen, and Reinier de Graaf Gasthuis) to perform Double Blind Placebo Controlled Food Challenges (DBPCFC) in children with suspected cashew allergy.

This thesis research was performed as part of this IDEAL project and focussed on the molecular biological studies. This thesis describes the study of cashew nut (*Anacardium occidentale*) proteins, focussed on its known allergens: Ana o 1, Ana o 2 and Ana o 3. These proteins are studied with regards to purification, effects of heat treatments, cross-reactivity, and IgE binding.

Allergy

The self-reported prevalence of food allergy in Europe is approximately 6%. However, the prevalence of food challenge-confirmed food allergy is below 1% [7]. Type I hypersensitivity, or IgE-mediated allergy, can be divided into three phases: the sensitization, stimulation, and effector phase [8], see Figure 1.1. During the sensitization phase an allergen (protein) enters the body, for example via ingestion, and this allergen is taken up and processed by an Antigen Presenting Cell (APC). The APC presents the allergen to $CD4^{+}$-T cells through their MHCII molecule and T Cell Receptor (TCR). Once these cells differentiate into Th2 cells, they can activate B cells by a combination of cytokines (IL4, IL13) and co-stimulatory receptor binding (CD40-CD40Ligand, MHCII-B-cell receptor). The activated B cells start to produce and release IgE, which binds to mast cells and basophils due to the presence of high affinity IgE receptors (FcεRI) expressed on these cells [8, 9]. During the second phase, the allergen again enters the body and binds and crosslinks two receptor-bound IgE molecules, thereby inducing the third phase. In this third phase, allergic mediators such as histamine and other inflammatory mediators are released from the mast cells and basophils, inducing the symptoms typical for an allergic reaction (e.g. itchiness, bronchial spasm, etc.) [10].

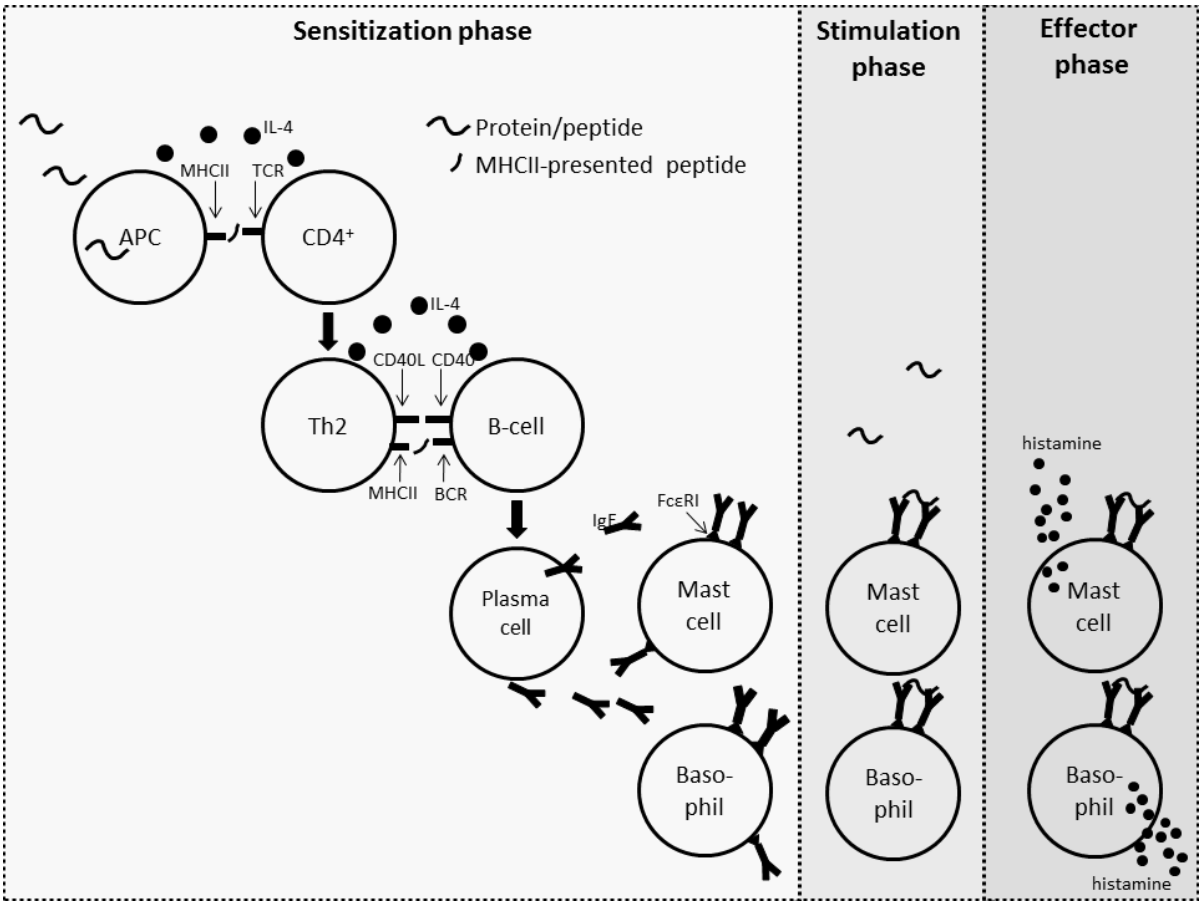


Figure 1.1 *Food allergy: sensitization, stimulation and effector phase. Based on [9].*

As explained above, in order to develop an allergy, a protein or peptide must come into contact with an antigen presenting cell. For food allergy the most obvious contact between food proteins and cells is through the digestive system. Proteins can be transported over the gut by several methods: for example paracellular transport [12], or endocytosis via M-cells [12] or enterocytes [13] could explain the ability of food proteins/peptides to reach immune cells and initiating the sensitization or stimulation phase. The 2S albumins of brazil nut and sesame seed, digested and undigested, can pass the epithelial cell layer intact as shown in a Caco-2 setup, indicating both the resistance to digestion and the capability of intestinal transport of these 2S albumins [14]. Also peanut protein was shown to be transported across the intestinal epithelial layer in an *in vivo* mouse model [15].

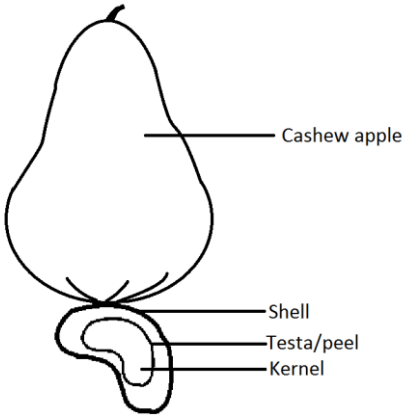


Figure 1.2 *Cashew apple and cashew nut, based on [5].*

Cashew nut

Cashew (*Anacardium occidentale*) originates from Brazil and has been distributed to Mozambique and India by

the Portuguese in the 16th century, after which these nuts were further spread by nature [16]. Cashew trees are easily cultivated, yielding 7 to 11 kilos of cashew nuts per year. The cashew tree is often cultivated by smallholder farms, and cross-pollination occurs freely between trees, leading to high variability between trees with little characterisation regarding to cashew tree varieties [5].

The cashew nut is a kidney-shaped seed which forms below a cashew apple, also named the false fruit of the cashew tree, see Figure 1.2. The edible cashew nut kernel is protected by a peel (testa) and is sheltered inside a shell containing a corrosive liquid named cashew nut shell liquid [5]. In recent years the global production of cashew nuts has been rising with a production of 4.7 million tonnes of raw cashew in 2011 with Vietnam as largest producer [16]. This production is mostly focussed on the cashew nut, but value-added side-products can be obtained from the cashew apple (alcoholic beverages, juice, candy, chutney, jam), testa (poultry feed), and cashew nut shell liquid (ingredient for paint, varnish) [17].

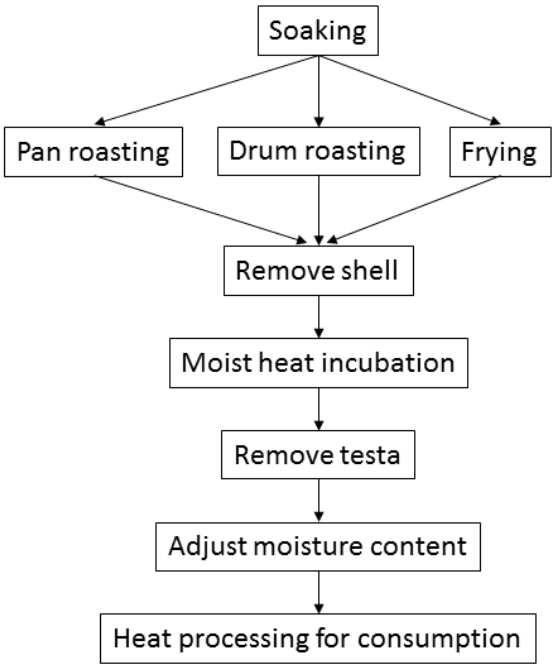


Figure 1.3 *Overview of the steps of cashew nut processing, based on [5].*

As explained by Azam-Ali and Judge [5], cashew nuts require many processing steps to be prepared for consumption. Picked or fallen cashew nuts are soaked to increase the moisture content of the kernel to 9% in order to avoid blackening during the subsequent heating step. This heating step, where the nut is roasted or fried, makes the shell brittle and simplifies taking out the cashew kernel. The shell is most often manually removed, yielding more of the highly desired whole cashew kernels compared to machine cutting of the shell. The removed kernel at this point is still covered by the testa, which is removed after a heating step of 6hr at 70˚C. After subsequent sorting of the cashews based on size, the moisture content is adjusted to 5% by humidification, after which the cashews are shipped and processed (e.g. roasted, fried) for consumption [5], see Figure 1.3.

Clinical aspects

As mentioned by van der Valk *et al*. [18]; allergy to cashew has gained increasing attention in recent years, and is often categorized as a severe allergy compared to other (tree-) nut allergies [1-4]. A retrospective study in a hospital in Sweden showed an increasing number of

cashew-allergic reactions over a period of 10 years (2001-2010) [6]. Another study in Sweden mentioned that in children over 3 years of age, allergic reactions to peanut and tree nuts were the most common food allergies resulting in emergency hospital visits. In total 5% of these food allergic reactions were due to a cashew allergy [19]. In children, most (74%, 58 out of 78) cashew allergic reactions occur below 6 years of age [3].

The prevalence of cashew allergy itself has not been clearly described in literature. Prevalence of tree nut allergy as a group was determined at 1.3% based on self-reported lifetime allergy and 0.5% based on oral food challenges [20]. Within this group of tree nut allergy the percentage of cashew allergic individuals lies somewhere between 5% [21] and 55% [22] (see Table 1.1). The prevalence of tree nut allergy as a group seems to be increasing in children as based on self-reported allergies [23].

Table 1.1 *Cashew allergy prevalence as calculated from several studies.*

Year	Region	Population	% cashew allergic of tree nut allergic	Methods	Source
1995-1997	USA	Children with acute reactions to peanut and or tree nuts	55% (11/20)	Questionnaire upon hospital allergy treatment	[22]
1997	USA	Members of the Food Allergy and Anaphylaxis Network, patients of allergists (89% <18 year old)	20%	Questionnaire	[24]
1997	USA	General US population	7% (8/118)	Telephone survey	[25]
2002	USA	General US population	44% (36/82)	Telephone survey	[26]
2003-2006	The Netherlands	Adults with suspected food allergy	20%	Questionnaire upon suspected food allergy	[27]
2005	USA	3-21 years old with tree nut allergy	30% (34/115)	Diagnosed by history, skin prick test or IgE levels	[4]
2005	France	Schoolchildren	5% (1/19)	Questionnaire	[21]
2005	UK	7-10 year old children	19% (8/43)	Skin prick test	[28]
2008	USA	General US population	35% (29/84)	Telephone survey	[29]
2016	Europe	0- <18 year old children	30% (78/256)	Medical history	[3]

Considering the symptoms upon an allergic reaction to cashew, certain dissimilarities were found between multiple studies. A study on 42 cashew-allergic children in France showed that most suffered from cutaneous symptoms (56%), 25% displayed respiratory reactions and 7% displayed asthmatic reactions [30]. Hourihane *et al.* showed that 14 out of 29 (48%) cashew allergic patients experienced wheezing and 11 out of 29 (38%) experienced collapse or feeling faint [31]. Maloney *et al.* show 84% of cashew allergic reactions result in cutaneous symptoms, 70% in respiratory symptoms and about 65% in gastrointestinal symptoms [32]. On the other hand in the UK, from 47 cashew-allergic children 98% showed cutaneous symptoms, 32% gastrointestinal and 40% suffered from wheezing [2]. These data

show high variability in the perceived symptoms. An explanation for these differences might be found in a difference in study population or different methods of observation.
Despite the difference in symptoms observed between different studies, all studies agree on the severity of this allergy. The severity of cashew allergy can be explained by two factors: the high risk of a severe allergic reaction and the low amount of cashew required in order to develop this reaction. The high risk associated with cashew allergy has been revealed by Davoren *et al*. In a retrospective study in Australia it was shown that although peanut allergy was more prevalent, incidence of anaphylaxis upon cashew ingestion (20 out of 27 cashew allergic reactions developed into anaphylaxis, 74%) was higher than peanut induced anaphylaxis (54 out of 177, 31%) [1]. In another study, cashew-allergic children were matched and compared to peanut-allergic children. This comparison showed that cashew allergy resulted more often in severe reactions, such as lower airway narrowing, than peanut allergy. Next to this, upon emergency treatment, adrenalin was more often administered to cashew allergic than peanut allergic children [2].
The low threshold required for an allergic reaction was described by Hourihane *et al.* stating that 14 out of 29 (48%) cashew-allergic patients experienced an allergic reaction after exposure to cashew via smelling/touching or tasting but not eating [31].

Known allergens in cashew: Ana o 1, Ana o 2 and Ana o 3

The majority of a cashew nut is composed of lipids (44%), but the second largest component in cashew is protein (19%) [33]. In this protein fraction the three known allergens from cashew are present: Ana o 1, Ana o 2 and Ana o 3. The nucleotide and amino acid sequences of these three major allergens are depicted in Figure 1.4.

Ana o 1

Ana o 1 is a 7S globulin of 50kDa. The cDNA coding sequence of Ana o 1 has been described by Wang *et al.* after preparing a cDNA library from cashew in late maturation, subsequent expression in *E. coli* and screening by serum from cashew allergic patients. For Ana o 1, two cDNA sequences were identified: Ana o1.0101 and Ana o1.0102, which differ by a single nucleotide [34]. As a 7S globulin, Ana o 1 is a seed storage protein that is part of the cupin protein family. Based on other cupins, Ana o 1 is expected to contain a double stranded α-helix [10] and, like other 7S globulins, is expected to form a trimer of 150kDa in native state [38].
Ana o 1 has been noted as a major allergen when recombinantly produced Ana o 1 bound 50% of patient sera (10/20). Upon subsequent epitope studies, 11 linear epitopes [34] and one conformational epitope [39] were observed.

Ana o 2

Ana o 2, also named anacardein or cashew major protein, is an 11S globulin that also belongs to the cupin protein family [10]. In general, 11S globulins are the most prevalent proteins in many seeds [38] and when examining cashew protein extracts this also seems the case for cashew [40]. 11S globulins are commonly hexameric proteins composed of 60kDa subunits. These subunits each contain an acidic large subunit of about 40kDa and a basic small subunit of about 20kDa linked together by a single disulphide bond [41]. In Ana o 2, these size descriptions are correct as the large subunit has been determined at 33kDa [35] and the small subunit at 20kDa [42, 43]. On denaturing SDS-PAGE also a minor band of 53kDa was tentatively identified as Ana o 2 [35].

Ana o 1: Ana o 1.0101, vicilin-like protein [*Anacardium occidentale*] GenBank: AAM73730.2 [34]

```
1    mgpptkfsfs lflvsvlvlc lgfalakidp elkqckhqck vqrqydeqqk eqcvkeceky
61   ykekkgrere heeeeeewgt ggvdepsthe paekhlsqcm rqcerqeggq qkqlcrfrcq
121  erykkergqh nykreddede dedeaeeede npyvfededf ttkvkteqgk vvllpkftqk
181  skllhaleky rlavlvanpq afvvpshmda dsiffvswgr gtitkilenk resinvrqgd
241  ivsissgtpf yianndenek lylvqflrpv nlpghfevfh gpggenpesf yrafsweile
301  aalktskdtl eklfekqdqg timkaskeqi ramsrrgegp kiwpfteest gsfklfkkdp
361  sqsnkygqlf eaeridyppl ekldmvvsya nitkggmsvp fynsratkia ivvsgegcve
421  iacphlsssk sshpsykklr arirkdtvfi vpaghpfatv asgnenleiv cfevnaegni
481  rytlagkkni ikvmekeake lafkmegeev dkvfgkqdee fffqgpewrk ekegrade
```

Ana o 1: Ana o 1.0102, vicilin-like protein, partial [*Anacardium occidentale*] GenBank: AAM73729.1 [34]

```
1    pptkfsfslf lvsvlvlclg falakidpel kqckhqckvq rqydeqqkeq cvkecekyyk
61   ekkgrerehe eeeeewgtgg vdepsthepa ekhlsqcmrq cerqeggqqk qlcrfrcqer
121  ykkergqhny kreddedede deaeeedenp yvfededftt kvkteqgkvv llpkftqksk
181  llhalekyrl avlvanpqaf vvpshmdads iffvswgrgt itkilenkre sinvrqgdiv
241  sissgtpfyi anndenekly lvqflrpvnl pghfevfhgp ggenpesfyr afsweileaa
301  lktskdtlek lfekqdqgti mkaskeqvra msrrgegpki wpfteestgs fklfkkdpsq
361  snkygqlfea eridypplek ldmvvsyani tkggmsvpfy nsratkiaiv vsgegcveia
421  cphlsskss hpsykklrar irkdtvfivp aghpfatvas gnenleivcf evnaegniry
481  tlagkkniik vmekeakela fkmegeevdk vfgkqdeeff fqgpewrkek egrade
```

Ana o 2: Ana o 2, partial [*Anacardium occidentale*] GenBank: AAN76862.1 [35]

```
1    lsvcflilfh gclasrqewq qqdecqidrl dalepdnrve yeagtveawd pnheqfrcag
61   valvrhtiqp nglllpqysn apqliyvvqg egmtgisypg cpetyqapqq grqqgqsgrf
121  qdrhqkirrf rrgdiiaipa gvahwcyneg nspvvtvtll dvsnsqnqld rtprkfhlag
181  npkdvfqqqq qhqsrgrnlf sgfdtellae afqvderlik qlksednrgg ivkvkddelr
241  virpsrsqse rgseseeese dekrrwgqrd ngieetictm rlkenindpa radiytpevg
301  rlttlnslnl pilkwlqlsv ekgvlyknal vlphwnlnsh siiygckgkg qvqvvdnfgn
361  rvfdgevreg qmlvvpqnfa vvkrareerf ewisfktndr amtsplagrt svlggmpeev
421  lanafqisre darkikfnnq qttltsgess hhmrdda
```

Ana o 3: 2s albumin [*Anacardium occidentale*] GenBank: AAL91665.1 [36]

```
1    makfllllsa favlllvana siyraiveve edsgreqscq rqfeeqqrfr ncqryvkqev
61   qrggrynqrq eslreccqel qevdrrcrcq nleqmvrqlq qqeqikgeev relyetasel
121  pricsispsq gcqfqssy
```

Figure 1.4 *Protein sequence of Ana o 1, Ana o 2 and Ana o 3 with the difference between Ana o 1.0101 and Ana o 1 1.0102 indicated in bold and italic, IgE-binding epitopes [34, 35, 37] underlined.*

To the recombinant version of Ana o 2, 13 out of 21 (62%) cashew-allergic patient IgE could bind, indicating Ana o 2 to be a major allergen. Ana o 2 contains at least 22 linear epitopes which are spread out over the entire protein [35]. When these linear epitopes were compared to linear epitopes on other 11S globulin proteins, four "hot spots" were identified, IgE binding epitopes at overlapping positions in aligned protein sequences. One of these hotspots in Ana o 2 was identified to be shielded within the monomeric subunits of Ana o 2, requiring denaturation of the protein before exposure of this epitope [44]. Next to these linear epitopes at least one conformational epitope on Ana o 2 has been discovered [43, 45]. This conformational epitope was studied using a monoclonal mouse antibody that has been shown to inhibit human IgE binding to a conformational epitope on Ana o 2. This conformational epitope consists of a protein segment of 24 amino acids containing β-strands and a short helical segment on the large subunit of Ana o 2, which connects to the small subunit of Ana o 2 [43, 45].

Ana o 3

Ana o 3 belongs, as a 2S albumin, to the prolamin superfamily of proteins. These proteins are mostly seed storage proteins and are commonly small (7-16kDa) with multiple inter-chain disulphide bonds and four α-helices [10]. Also the cDNA sequence of Ana o 3 has been determined by Robotham *et al.* and the recombinant Ana o 3 has been used to produce goat anti-Ana o 3 antibodies that can be used to purify native Ana o 3. This showed native Ana o 3 to be a 12.598kDa 2S albumin which probably undergoes posttranslational modification as the protein size based on the cDNA sequence was predicted to be 16.335kDa [37]. Based on literature this protein is most probably proteolytically cleaved at the C-terminus into a small and a large subunit that stay associated by four disulphide bonds [41, 46]. The large subunit of this protein is present in three isoforms of 6, 8 and 10kDa [37]. The small subunit is not mentioned in experimental studies, perhaps because it is too small to be observed on SDS-PAGE or simply because it was never looked for. On western blot 21 out of 26 (81%) patient sera bound to rAna o 3, confirming Ana o 3 to be a major allergen. In Ana o 3, 8 linear epitopes have been identified, of which some show high similarity with Jug r 1 from walnut and sesame seed 2S albumin [37].

Purified and recombinantly produced cashew allergens

The cashew nut allergens Ana o 1, 2 and 3 have been recombinantly produced in *E. coli* by Wang and Robotham *et al*. [34, 35, 37]. Cashew-derived Ana o 1, 2, and 3 have been identified using these recombinant allergens by inhibition blotting. Inhibition blotting was done by pre-incubating serum from a cashew-allergic person with e.g. recombinant Ana o 1, and observing on western blot which protein bands from cashew are no longer bound by IgE, thereby identifying native Ana o 1. *E. coli* derived recombinant Ana o 1 is a protein of 55 and

65kDa (two clones, different in start site, differing 73 amino acids) [34], rAna o 2 is 52kDa and also forms a dimer of 120kDa [35], rAna o 3 is a 14kDa protein [37].

Of the three allergens, only Ana o 3 has been purified from the cashew nut [42]. In this protocol Mattison *et al.* used a sodium phosphate gradient on a ceramic hydroxyapatite column after defatting of the cashews and precipitation of the protein extract. In this article neither the yield nor the purity of the purified Ana o 3 allergen was mentioned. For Ana o 2 a purification protocol has been described [47], however, as mentioned by Teuber *et al*. this fraction is not immunologically pure [48]. No protocol for the purification of Ana o 1 has been described in literature.

Allergen protein purification starting from the food source or after generating recombinant proteins, warrants a discussion on subsequent application. The recombinant production of proteins is, in general, easier to standardise, often yields higher amounts of pure protein (depending on the expression system used), and contamination of one allergen with the other is much less likely to occur compared to protein purification from the food source. Purification from the food source directly, on the other hand, results in a native protein with correct post-translational modifications like protein folding, disulphide bridges, glycosylation, etc. When multiple isoforms of a single protein are present in the food source purification might extract multiple (but not necessarily all) isoforms, while recombinant protein expression only produces one isoform unless multiple colonies of the different protein isoforms are prepared. Also some technical specifications can prompt different choices as protein purification and recombinant expression both require different specialist equipment, and besides, for recombinant protein expression the protein/cDNA sequence of the protein of interest should be known, otherwise first a cDNA library should be prepared. Lastly the effect of the food matrix when studying heat treatment effects, protein digestion or protein transport characteristics is missed for recombinantly produces proteins.

Heat treatments and *in vitro* digestion

Cashews are eaten in processed form: most commonly the raw nuts are heated for 20-35min at 150°C in order to remove the shell, afterwards they can be roasted or fried. Roasted cashews (120°C or 160°C for 20min) are most often eaten in the US while in the Netherlands mostly fried cashews (93°C increased to 135°C in 35-40min or 150°C -160°C for 1-3min) are consumed [49]. Because cashews are always heat-treated before consumption, and as this might affect the structure and function of the proteins, several studies have examined the effect of processing on cashew proteins. Most of these studies have been performed using SDS-PAGE and western blotting techniques, using either patient IgE or polyclonal anti-cashew antibodies, thereby mostly focussing on (linear) epitopes.

Studies focussing on roasted cashew show varying results fluctuating from no difference in antibody binding [50], to decreased [51] or even increased [52] antibody binding (patient IgE [50, 51] or polyclonal IgG [51, 52]) upon roasting of cashew nuts. The discrepancies between these studies can to some extent be caused by the increased solubility of Ana o 3 in roasted cashew [51]. When the antibodies used, bind specifically to Ana o 3, this binding will be increased in extracts from roasted cashew nuts simply due to the higher amount of Ana o 3 present in the roasted cashew sample.

Besides roasting, also more harsh heat treatments like γ-irradiation and autoclaving have been applied to cashew. γ-Irradiation followed by frying or blanching did not affect binding of a polyclonal antibody, while γ-irradiation followed by autoclaving or roasting did reduce binding of anti-cashew IgG as observed on western blot [53]. The required dose of γ-irradiation will, however, lead to inedible cashews [54].
A study by Venkatachalam *et al.* showed a thorough comparison of cashew subjected to many different heat treatments by using rabbit, goat and mouse antibodies. For Ana o 1, a decrease in antibody binding after roasting, blanching, microwaving and autoclaving was observed but not after γ-irradiation, indicating the probable loss of a conformational epitope within Ana o 1 upon heating. Another antibody showed only decreased binding to Ana o 1 after prolonged autoclaving of the cashew nuts [55]. Ana o 2 remained immunologically stable despite microwaving, roasting, blanching or γ-irradiation. Only autoclaving for 20min could decrease antibody binding slightly [55]. The immunoreactivity of Ana o 2 towards two monoclonal mouse antibodies could chemically be reduced by SDS but not by guanidinium HCl or urea despite that the protein conformation of Ana o 2 was changed by all three chemicals [56]. Ana o 3 appeared more sensitive to processing steps such as roasting, autoclaving and blanching [55]. However, for Ana o 3 this was assessed using a monoclonal antibody that potentially binds a conformational epitope [55], while for Ana o 2 the antibodies used in the various studies probably target linear epitopes [56], making it difficult to compare the stability of these two allergens. It is expected that Ana o 3, like other 2S albumins, is highly resistant to heat processing [46]. Considering pH stability, all three allergens were stable in the middle-pH range but were unstable at the extreme pH of 1 and 13 [55].

Lastly, the effect of sulphite, a chemical that has been shown to disrupt disulphide bonds, has been tested on cashew extracts [52]. On both western blot and dot blot, a decrease in patient-IgE binding was observed after sulphite treatment of both raw and roasted cashew [52]. However, such a treatment cannot be applied to whole cashews, and, as already mentioned by the authors, sodium sulphite itself can cause allergic reactions.

Besides the effect of heat and processing treatments on cashew nut proteins, also the *in vitro* digestibility of these proteins has been studied. This *in vitro* digestion showed Ana o 1 and the 30kDa subunit of Ana o 2 to be partially susceptible to both pepsin and trypsin digestion. However, Ana o 3 and the 20kDa subunit of Ana o 2 were not completely digested by either pepsin (0.8U 2hr) or trypsin (20U 2hr) alone, nor by pepsin (0.8U 30min) followed by trypsin (2U 30min). Ana o 3 pre-treated with DTT to break disulphide bonds was more susceptible to pepsin digestion, as based on reduced visibility on SDS-PAGE and reduced (but not completely obstructed) IgE binding in inhibition ELISA testing [42]. The effect of *in vitro* digested cashew on the *in vivo* allergic response has also been studied. Mice sensitized to native cashew showed reduced allergic reactions upon intraperitoneal injected pepsin-digested cashew compared to exposure to non-digested cashew protein. Next to this, immunotherapy using a pepsin-digested cashew protein extract could reduce the allergic reaction upon cashew challenges in mice, indicating a possible role for pepsin-digested cashew in immunotherapy [57].

To summarise, as could be expected from 7S globulins, 11S globulins, and 2S albumins [10] Ana o 1, Ana o 2 and Ana o 3 are highly resistant to heat treatments and quite resistant to *in vitro* digestion. The IgE-immunoreactivity of Ana o 2 and 3 can be decreased by sulphite treatments and intense γ-irradiation. However, applying these treatments in the cashew nut industry is not feasible as sodium sulphite itself might cause allergic reactions and cannot be applied to whole cashew nuts, and as the use of high γ-irradiation would result in non-edible cashews [54].

Cross-allergenicity with other Anacardiaceae and other tree nuts

Cashew is a tree nut belonging to the family of Anacardiaceae together with other edible plants and trees like sumac, mango, pistachio, and pink pepper [58]. Cross-allergenicity between cashew and other Anacardiaceae and tree nuts has been studied to some extent. Especially the possibility of cross reactivity between cashew and pistachio has been presented by several groups [59-63] and part of this cross reactivity might be explained by cross reactivity between Pis v 3 and Ana o 1 [62], and sequence homology between Pis v 2 and Ana o 2 and between Pis v 1 and Ana o 3 [63]. Cashew allergen Ana o 3 even showed a higher specificity in correctly identifying pistachio allergy by serum IgE, compared to using a pistachio extract [64]. Also *in vivo* cross-reactivity between cashew and pistachio has been noted. In this study, cashew-immunotherapy in cashew-sensitized mice resulted in decreased allergic reaction towards both cashew and pistachio [65]. Also the role of walnut, a tree nut but not a member of the Anacardiaceae, was studied here. Mice sensitized to both cashew and walnut showed a diminished allergic reaction towards cashew after receiving walnut immunotherapy [65]. The other way around, when walnut was introduced

into the diet of cashew-sensitized mice, the mice showed elevated walnut-specific IgE levels and developed an allergic reaction towards walnut [66].
Besides walnut and pistachio, also studies with mango have been performed. Mango as member of the Anacardiaceae is suspected to cross-react with both cashew and pistachio as based on inhibition RAST results [67-69].
Studies on the possibility of cross reactivity between cashew and peanut are not unexpected as 20-30% of peanut allergic individuals are also allergic to one or more tree nuts [23]. However, studies using cashew and peanut inhibition ELISAs [50, 70], basophil activation test (BAT) [70], and studies on epitope homology of Ana o 1 and Ara h 1 [34] did not substantiate a claim of cross allergenicity between peanut and cashew. No cross-reacivity was found by inhibition ELISA and inhibition western blot between roasted cashew and rAra h 2 [71]. Epitope homology between the 11S globulins Ara h 3 (peanut), Cor a 9 (hazelnut), Jug r 4 (walnut) and Ana o 2, however, has been suggested [72]. Overall, cross reactivity between cashew and peanut is not confirmed, which is not unexpected as peanut is neither a tree nut nor part of the Anacardiaceae family. This notion is further substantiated by the finding that *in vivo* it has been shown that peanut-allergic persons, who are also tree nut allergic, are less likely to be allergic to cashew than peanut-tolerant persons who are tree nut allergic [32].

Cashew apple and cashew pollen

Besides allergens in the cashew itself and cross-reactive allergens in related plants, also some research has been performed on allergenic proteins in other parts of the cashew tree: the cashew apple, pollen and the shell. Despite the limited exposure to these products outside cashew cultivation areas, it is still worth mentioning.
Presence of allergens in the cashew apple similar to the allergens present in cashew nuts has been described by Comstock *et al*. In this study cashew-allergic patients' sera showed IgE binding to proteins from cashew apple juice concentrate on western blot, a reaction that could be inhibited (7.5, 20, 25, 50, 60kDa bands) for some sera when pre-incubated with cashew nut extract [73]. Upon western blotting with anti-Ana o 1 and anti-Ana o 2 monoclonal antibodies, several proteins from cashew apple juice could be detected (45kDa anti-Ana o 1; 37 and 46kDa anti-Ana o 2) [73]. Besides cashew apple allergy, one can also be allergic to cashew tree pollen. In India, a country with many cashew-plantations, 65 patients with allergic bronchial asthma were tested for cashew-pollen sensitisation by a bronchial provocation test and a skin prick test. Of these 65, patients 20 were found to be allergic to pollen of the *Anacardium occidentale* tree by both tests [74]. The last study deals with the shell around the cashew kernel. A defatted protein extract was made from cashew nut shell, showing presence of multiple proteins of 14–97kDa as visualised by SDS-PAGE [75]. In addition, the oil present in this shell can cause contact dermatitis [76].

Detection by immunoassay

As for now, cashew allergy cannot be cured, cashew allergic patients should avoid cashews, even trace amounts. Several methods for the detection of cashew nut allergens in food products have been developed, e.g. ELISA, immunoblotting, (RT-)PCR, dipstick, and mass spectrometry [77]. As described by van *Hengel* [77] each of these methods has it's pro's and con's. For example, PCR analysis detects DNA or RNA sensitively, serving as an indicator but does not directly detect the proteins. The dipstick method is a fast method to detect proteins but is not quantitative, and mass spectrometry is very specific but requires highly specialised equipment and trained personnel [77].

Specifically for the detection of cashew nut in food products, methods have been developed. First of all an ELISA assay has been developed for the detection of the so-called "cashew major protein" which mainly exists of Ana o 2. The assay is able to detect levels of 0.02ppm cashew major protein, and is suitable for use in most food matrixes (e.g. wheat flour, rice cereal, chocolate cookies) and compatible with several spices (e.g. salt, brown sugar, cardamom). The assay did show diminished sensitivity when combined with milk chocolate, raisin bran cereal, cinnamon and nutmeg. Upon heat-processing of cashew the detection was diminished as well, especially roasting of the cashew at 170°C for 20min diminished the sensitivity of this ELISA assay [78]. Gaskin and Taylor also developed a cashew specific ELISA with high sensitivity (1µg cashew/g of product) but with significant cross reactivity towards pistachio and, to a certain degree, also towards hazelnut [79]. Secondly, a mass spectrometry method has been developed, detecting Ana o 2 and Ana o 3 in the sub-ppm range, thereby being declared as more sensitive than ELISA assays [80]. Lastly, also PCR assays detecting DNA, have been developed targeting Ana o 3, detecting 2mg/kg [81] and 0.005% of the total food weight [82].

Aim and outline of the thesis

Diagnostic procedures in allergy are partly based on the detection of serum-IgE for the offending allergen. Very important drawbacks of this type of testing, are the often limited clinical relevance of its outcomes with often high numbers of false-positive or false-negative test results. The aim of this thesis is a better chemical identification and functional characterisation of the major cashew nut allergens Ana o 1, 2 and 3. This characterisation will contribute to the development of more sensitive and reliable diagnostic procedures for the early detection of cashew allergy in young children. This will allow improved disease monitoring and future therapy for this 'vulnerable' group.

In the **second chapter** of this thesis a review is provided on intestinal protein transport, focussing on the difference between sensitised versus non-sensitised persons. This difference in intestinal protein transport between sensitised and non-sensitised persons could indicate a possible mechanism of how allergens could come into contact with immune cells and cause the deveopment of a food allergy. The experimental chapters of this thesis are focussed on cashew nut allergy, studying the cashew nut allergens Ana o 1, Ana o 2 and Ana o 3. In **chapter 3,** a purification method is described for Ana o 1, Ana o 2 and Ana o 3 from blanched cashew nuts, and their protein characteristics such as IgE-binding and glycosylation are described. For Ana o 3 epitope mapping experiments have been performed, and mass spectrometry was done to study the N- and C-termini of Ana o 3. Next, in **chapter 4,** a comparison was made between cashews of different origins which have been subjected to different heat treatments. This was done as all cashew studies described in literature have been performed using store-bought cashews, from unknown origin and with doubtfull information on pre-treatments (e.g. heating steps applied in order to remove the shell) while being sold and used as " raw" cashew. Therefore, in chapter 4, 8 different origins of cashew were compared for their protein composition. This allowed us to assess whether or not data from literature can be universally used in cashew protein research or whether it is specific to a cashew of a particular origin. Furthermore, the difference between different origins and the influence of heat treamtens on the cashews' protein composition, protein glycation/glycosylation, digestibility, and IgE binding was studied. In **Chapter 5,** the recombinant production of Ana o 1, 2 and 3 from *Pichia pastoris* is described. This was done as recombinantly expressed proteins are usually more easily purified in greater quantity and with no chance of allergen contamination between the purified Ana o 1, Ana o 2 and Ana o 3. The cloning and purification of these three allergens, as well as characterization on 1D and 2D electrophoresis, western blot, and inhibition blot, is described in this chapter.
The results of these 5 chapters are discussed in **chapter 6**. The newly obtained knowledge on Ana o 1, 2 and 3 is summarized. Also, extra information regarding the importance of cross-reactivity with other nuts and Anacardiaceae family members is provided, and the relevance of these results for the clinical field is discussed.

References

[1] Davoren, M., Peake, J., Cashew nut allergy is associated with a high risk of anaphylaxis. Arch Dis Child 2005, 90, 1084-1085.

[2] Clark, A. T., Anagnostou, K., Ewan, P. W., Cashew nut causes more severe reactions than peanut: case-matched comparison in 141 children. Allergy 2007, 62, 913-916.

[3] Grabenhenrich, L. B., Dolle, S., Moneret-Vautrin, A., Kohli, A., et al., Anaphylaxis in children and adolescents: The European Anaphylaxis Registry. J. Allergy Clin. Immun. 2016, 137, 1128-1137 e1121.

[4] Fleischer, D. M., Conover-Walker, M. K., Matsui, E. C., Wood, R. A., The natural history of tree nut allergy. J. Allergy Clin. Immun. 2005, 116, 1087-1093.

[5] Azam-Ali, S. H., Judge, E. C., Food and Agriculture Organization of the United Nations 2001.

[6] Johnson, J., Malinovschi, A., Alving, K., Lidholm, J., et al., Ten-year review reveals changing trends and severity of allergic reactions to nuts and other foods. Acta paediatrica 2014, 103, 862-867.

[7] Nwaru, B. I., Hickstein, L., Panesar, S. S., Muraro, A., et al., The epidemiology of food allergy in Europe: a systematic review and meta-analysis. Allergy 2014, 69, 62-75.

[8] Multidisciplinary approaches to allergies, Springer 2012.

[9] Larche, M., Akdis, C. A., Valenta, R., Immunological mechanisms of allergen-specific immunotherapy. Nat. Rev. Immunology 2006, 6, 761-771.

[10] Mills, E. N. C., Madsen, C., Shewry, P. R., Wicher, H. J., Food allergens of plant origin - their molecular and evolutionary relationships. Trends Food Sci Tech 2003, 14, 145-156.

[11] Lack, G., Update on risk factors for food allergy. J. Allergy Clin. Immun. 2012, 129, 1187-1197.

[12] Price, D., Ackland, L., Suphioglu, C., Nuts 'n' guts: transport of food allergens across the intestinal epithelium. Asia Pacific allergy 2013, 3, 257-265.

[13] Yu, L. C., The epithelial gatekeeper against food allergy. Pediatrics and neonatology 2009, 50, 247-254.

[14] Moreno, F. J., Rubio, L. A., Olano, A., Clemente, A., Uptake of 2S albumin allergens, Ber e 1 and Ses i 1, across human intestinal epithelial Caco-2 cell monolayers. J. Agr. Food Chem. 2006, 54, 8631-8639.

[15] Chambers, S. J., Wickham, M. S., Regoli, M., Bertelli, E., et al., Rapid in vivo transport of proteins from digested allergen across pre-sensitized gut. Biochem Biophys Res Commun 2004, 325, 1258-1263.

[16] Dendena, B., Corsi, S., Cashew, from seed to market: a review. Agron Sustain Dev 2014, 34, 753-772.

[17] Nair, K. P., The agronomy and economy of important tree crops of the developing world, Elsevier 2010.

[18] van der Valk, J. P., Dubois, A. E., Gerth van Wijk, R., Wichers, H. J., de Jong, N. W., Systematic review on cashew nut allergy. Allergy 2014, 69, 692-698.

[19] Vetander, M., Helander, D., Flodstrom, C., Ostblom, E., et al., Anaphylaxis and reactions to foods in children--a population-based case study of emergency department visits. Clin. Exp. Allergy 2012, 42, 568-577.

[20] Nwaru, B. I., Hickstein, L., Panesar, S. S., Roberts, G., et al., Prevalence of common food allergies in Europe: a systematic review and meta-analysis. Allergy 2014, 69, 992-1007.

[21] Rance, F., Grandmottet, X., Grandjean, H., Prevalence and main characteristics of schoolchildren diagnosed with food allergies in France. Clin. Exp. Allergy 2005, 35, 167-172.

[22] Sicherer, S. H., Burks, A. W., Sampson, H. A., Clinical features of acute allergic reactions to peanut and tree nuts in children. Pediatrics 1998, 102, e6.

[23] McWilliam, V., Koplin, J., Lodge, C., Tang, M., et al., The Prevalence of Tree Nut Allergy: A Systematic Review. Curr. Allergy Asthma R. 2015, 15, 54.

[24] Sicherer, S. H., Furlong, T. J., Munoz-Furlong, A., Burks, A. W., Sampson, H. A., A voluntary registry for peanut and tree nut allergy: characteristics of the first 5149 registrants. J. Allergy Clin. Immun. 2001, 108, 128-132.

[25] Sicherer, S. H., Munoz-Furlong, A., Burks, A. W., Sampson, H. A., Prevalence of peanut and tree nut allergy in the US determined by a random digit dial telephone survey. J. Allergy Clin. Immun. 1999, 103, 559-562.

[26] Sicherer, S. H., Munoz-Furlong, A., Sampson, H. A., Prevalence of peanut and tree nut allergy in the United States determined by means of a random digit dial telephone survey: a 5-year follow-up study. J. Allergy Clin. Immun. 2003, 112, 1203-1207.

[27] Le, T. M., Lindner, T. M., Pasmans, S. G., Guikers, C. L., et al., Reported food allergy to peanut, tree nuts and fruit: comparison of clinical manifestations, prescription of medication and impact on daily life. Allergy 2008, 63, 910-916.

[28] Roberts, G., Peckitt, C., Northstone, K., Strachan, D., et al., Relationship between aeroallergen and food allergen sensitization in childhood. Clin. Exp. Allergy 2005, 35, 933-940.

[29] Sicherer, S. H., Munoz-Furlong, A., Godbold, J. H., Sampson, H. A., US prevalence of self-reported peanut, tree nut, and sesame allergy: 11-year follow-up. J. Allergy Clin. Immun. 2010, 125, 1322-1326.

[30] Rance, F., Bidat, E., Bourrier, T., Sabouraud, D., Cashew allergy: observations of 42 children without associated peanut allergy. Allergy 2003, 58, 1311-1314.

[31] Hourihane, J. O., Harris, H., Langton-Hewer, S., Kilburn, S. A., Warner, J. O., Clinical features of cashew allergy. Allergy 2001, 56, 252-253.

[32] Maloney, J. M., Rudengren, M., Ahlstedt, S., Bock, S. A., Sampson, H. A., The use of serum-specific IgE measurements for the diagnosis of peanut, tree nut, and seed allergy. J. Allergy Clin. Immun. 2008, 122, 145-151.

[33] Venkatachalam, M., Sathe, S. K., Chemical composition of selected edible nut seeds. J. Agr. Food Chem. 2006, 54, 4705-4714.

[34] Wang, F., Robotham, J. M., Teuber, S. S., Tawde, P., et al., Ana o 1, a cashew (Anacardium occidental) allergen of the vicilin seed storage protein family. J. Allergy Clin. Immun. 2002, 110, 160-166.

[35] Wang, F., Robotham, J. M., Teuber, S. S., Sathe, S. K., Roux, K. H., Ana o 2, a major cashew (Anacardium occidentale L.) nut allergen of the legumin family. Int. Arch. Allergy A. Imm. 2003, 132, 27-39.

[36] Robotham, J. M., Wang, F., Seamon, V., Teuber, S. S., et al., Ana o 3, an important cashew nut (Anacardium occidentale L.) allergen of the 2S albumin family. J. Allergy Clin. Immun. 2005, 115, 1284-1290.

[37] Robotham, J. M., Wang, F., Seamon, V., Teuber, S. S., et al., Ana o 3, an important cashew nut (Anacardium occidentale L.) allergen of the 2S albumin family. J. Allergy Clin. Immun. 2005, 115, 1284-1290.

[38] Astwood, J. D., Silvanovich, A., Bannon, G. A., Vicilins: a case study in allergen pedigrees. The J. Allergy Clin. Immun. 2002, 110, 26-27.

[39] Guan, X., Noble, K. A., Tao, Y., Roux, K. H., Sathe, S. K., Young, N. L., Marshall, A. G., Epitope mapping of 7S cashew antigen in complex with antibody by solution-phase H/D exchange monitored by FT-ICR mass spectrometry. J. Mass Spectrometry 2015, 50, 812-819.

[40] Hummel, M., Wigger, T., Hoper, T., Westkamp, I., Brockmeyer, J., Simple, Rapid, and Selective Isolation of 2S Albumins from Allergenic Seeds and Nuts. J. Agr. Food Chem. 2015.

[41] Mills, E. N. C., Shewry, P. R., Plant food allergens, Blackwell science Ltd 2004.

[42] Mattison, C. P., Grimm, C. C., Wasserman, R. L., In vitro digestion of soluble cashew proteins and characterization of surviving IgE-reactive peptides. Mol. Nutr. Food Res. 2014, 58, 884-893.

[43] Xia, L., Willison, L. N., Porter, L., Robotham, J. M., et al., Mapping of a conformational epitope on the cashew allergen Ana o 2: a discontinuous large subunit epitope dependent upon homologous or heterologous small subunit association. Mol Immunol 2010, 47, 1808-1816.

[44] Robotham, J. M., Hoffman, G. G., Teuber, S. S., Beyer, K., Sampson, H. A., Sathe, S. K., Roux, K. H., Line ar IgE-epitope mapping and comparative structural homology modeling of hazelnut and English walnut 11S globulins. Mol. Immunol. 2009, 46, 2975-2984.

[45] Robotham, J. M., Xia, L., Willison, L. N., Teuber, S. S., et al., Characterization of a cashew allergen, 11S globulin (Ana o 2), conformational epitope. Mol Immunol 2010, 47, 1830-1838.

[46] Moreno, F. J., Clemente, A., 2S Albumin Storage Proteins: What Makes them Food Allergens? The open biochemistry journal 2008, 2, 16-28.

[47] Sathe, S. K., SzeTao, K. W. C., Wolf, W. J., Hamaker, B. R., Biochemical characterization and in vitro digestibility of the major globulin in cashew nut (Anacardium occidentale). J. Agr. Food Chem. 1997, 45, 2854-2860.

[48] Teuber, S. S., Sathe, S. K., Peterson, W. R., Roux, K. H., Characterization of the soluble allergenic proteins of cashew nut (Anacardium occidentale L.). J. Agr. Food Chem. 2002, 50, 6543-6549.

[49] Masthoff, L. J., Hoff, R., Verhoeckx, K. C., van Os-Medendorp, H., et al., A systematic review of the effect of thermal processing on the allergenicity of tree nuts. Allergy 2013, 68, 983-993.

[50] de Leon, M. P., Glaspole, I. N., Drew, A. C., Rolland, J. M., et al., Immunological analysis of allergenic cross-reactivity between peanut and tree nuts. J. Clin. Exp. Allergy 2003, 33, 1273-1280.

[51] C. P. Mattison, Y. B.-M., B. Vant-Hullc, A. M. Vargasa, R. L. Wassermand, C. C. Grimm, Heat-induced alterations in cashew allergen solubility and IgE binding. Tox. Reports 2016, 244-251.

[52] Mattison, C. P., Desormeaux, W. A., Wasserman, R. L., Yoshioka-Tarver, M., et al., Decreased Immunoglobulin E (IgE) Binding to Cashew Allergens following Sodium Sulfite Treatment and Heating. J. Agr. Food Chem. 2014, 62, 6746-6755.

[53] Su, M., Venkatachalam, M., Teuber, S. S., Roux, K. H., Sathe, S. K., Impact of gamma-irradiation and thermal processing on the antigenicity of almond, cashew nut and walnut proteins. J Sci Food Agr 2004, 84, 1119-1125.

[54] Mexis, S. F., Kontominas, M. G., Effect of gamma-irradiation on the physicochemical and sensory properties of cashew nuts (Anacardium occidentale L.). Lwt-Food Sci Technol 2009, 42, 1501-1507.

[55] Venkatachalam, M., Monaghan, E. K., Kshirsagar, H. H., Robotham, J. M., et al., Effects of processing on immunoreactivity of cashew nut (Anacardium occidentale L.) seed flour proteins. J. Agr. Food Chem. 2008, 56, 8998-9005.

[56] Kshirsagar, H. H., Fajer, P., Sharma, G. M., Roux, K. H., Sathe, S. K., Biochemical and spectroscopic characterization of almond and cashew nut seed 11S legumins, amandin and anacardein. J. Agr. Food Chem. 2011, 59, 386-393.

[57] Kulis, M., Macqueen, I., Li, Y., Guo, R., et al., Pepsinized cashew proteins are hypoallergenic and immunogenic and provide effective immunotherapy in mice with cashew allergy. J. Agr. Food Chem. 2012, 130, 716-723.

[58] USDA, 2016, https://www.usda.gov/

[59] Noorbakhsh, R., Mortazavi, S. A., Sankian, M., Shahidi, F., et al., Pistachio allergy-prevalence and in vitro cross-reactivity with other nuts. Allergology international 2011, 60, 425-432.

[60] Parra, F. M., Cuevas, M., Lezaun, A., Alonso, M. D., et al., Pistachio nut hypersensitivity: identification of pistachio nut allergens. Clni. Exp. Allergy 1993, 23, 996-1001.

[61] Goetz, D. W., Whisman, B. A., Goetz, A. D., Cross-reactivity among edible nuts: double immunodiffusion, crossed immunoelectrophoresis, and human specific igE serologic surveys. Ann. Allerg. Asthma Im. 2005, 95, 45-52.

[62] Willison, L. N., Tawde, P., Robotham, J. M., Penney, R. M. t., et al., Pistachio vicilin, Pis v 3, is immunoglobulin E-reactive and cross-reacts with the homologous cashew allergen, Ana o 1. Clin. Exp. Allergy 2008, 38, 1229-1238.

[63] Ahn, K., Bardina, L., Grishina, G., Beyer, K., Sampson, H. A., Identification of two pistachio allergens, Pis v 1 and Pis v 2, belonging to the 2S albumin and 11S globulin family. Clin. Exp. Allergy 2009, 39, 926-934.

[64] Savvatianos, S., Konstantinopoulos, A. P., Borga, A., Stavroulakis, G., et al., Sensitization to cashew nut 2S albumin, Ana o 3, is highly predictive of cashew and pistachio allergy in Greek children. J. Allergy Clin. Immun. 2015, 136, 192-194.

[65] Kulis, M., Li, Y., Lane, H., Pons, L., Burks, W., Single-tree nut immunotherapy attenuates allergic reactions in mice with hypersensitivity to multiple tree nuts. J. Allergy Clin. Immun. 2011, 127, 81-88.

[66] Kulis, M., Burks, A. W., Effects of a pre-existing food allergy on the oral introduction of food proteins: findings from a murine model. Allergy 2015, 70, 120-123.

[67] Fernandez, C., Fiandor, A., Martinez-Garate, A., Martinez Quesada, J., Allergy to pistachio: crossreactivity between pistachio nut and other Anacardiaceae. Clin. Exp. Allergy 1995, 25, 1254-1259.

[68] Funes, E., Milán, J. M., Pagán, J. A., López, J. D., Garcia, F. J., Negro, J. M., Hernández, J., Polo, F., Rico, P., Allergy to Anacardiaceae. Identification of allergens. Alergol. Immunol. Clin 1999, 14, 82-89.

[69] Besler, M., Paschke, A., Rodriguez, J., Internet Symposium on Food Allergens, www.food-allergens.de 2001.

[70] de Leon, M. P., Drew, A. C., Glaspole, I. N., Suphioglu, C., et al., Functional analysis of cross-reactive immunoglobulin E antibodies: peanut-specific immunoglobulin E sensitizes basophils to tree nut allergens. Clin. Exp. Allergy 2005, 35, 1056-1064.

[71] 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.

[72] Barre, A., Jacquet, G., Sordet, C., Culerrier, R., Rouge, P., Homology modelling and conformational analysis of IgE-binding epitopes of Ara h 3 and other legumin allergens with a cupin fold from tree nuts. Mol Immunol 2007, 44, 3243-3255.

[73] Comstock, S. S., Robotham, J. M., Tawde, P., Kshirsagar, H., et al., Immunoglobulin E-reactive proteins in cashew (Anacardium occidentale) apple juice concentrate. J. Agr Ffood Chem. 2008, 56, 5977-5982.

[74] Fernandes, L., Mesquita, A. M., Anacardium occidentale (cashew) pollen allergy in patients with allergic bronchial asthma. J. Allergy Clin. Immun. 1995, 95, 501-504.

[75] Yuliana, M., Truong, C. T., Huynh, L. H., Ho, Q. P., Ju, Y. H., Isolation and characterization of protein isolated from defatted cashew nut shell: Influence of pH and NaCl on solubility and functional properties. Lwt-Food Sci Technol 2014, 55, 621-626.

[76] Lampe, K. F., Dermatitis-producing Anacardiaceae of the Caribbean area. Clinics in dermatology 1986, 4, 171-182.

[77] van Hengel, A. J., Food allergen detection methods and the challenge to protect food-allergic consumers. Anal Bioanal Chem 2007, 389, 111-118.

[78] Wei, Y., Sathe, S. K., Teuber, S. S., Roux, K. H., A sensitive sandwich ELISA for the detection of trace amounts of cashew (Anacardium occidentale L.) nut in foods. J. Agr. Food Chem. 2003, 51, 3215-3221.

[79] Gaskin, F. E., Taylor, S. L., Sandwich Enzyme-Linked Immunosorbent Assay (ELISA) for Detecti on of Cashew Nut in Foods. J Food Sci 2011, 76, T218-T226.

[80] Sealey-Voyksner, J., Zweigenbaum, J., Voyksner, R., Discovery of highly conserved unique peanut and tree nut peptides by LC-MS/MS for multi-allergen detection. Food chemistry 2016, 194, 201-211.

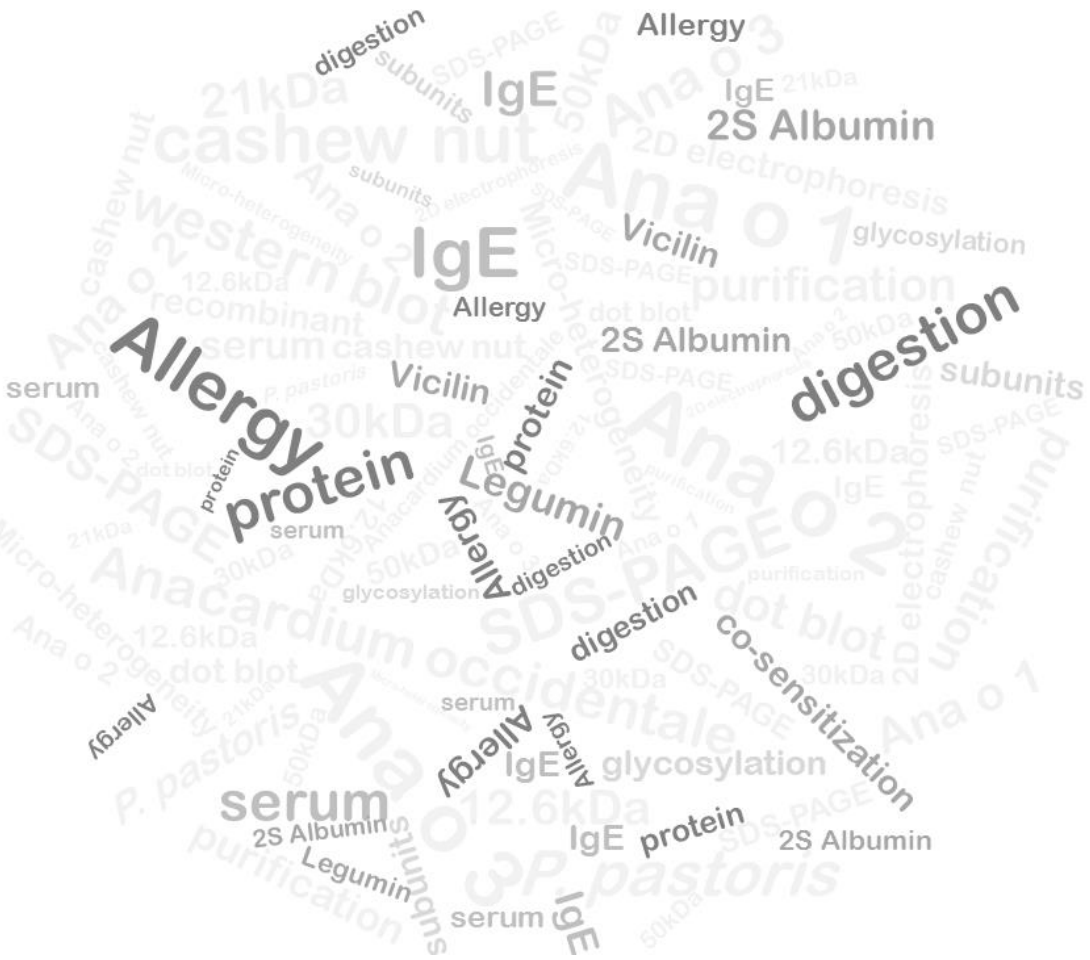
[81] Ehlert, A., Hupfer, C., Demmel, A., Engel, K. H., Busch, U., Detection of Cashew Nut in Foods by a Specific Real-time PCR Method. Food Anal Method 2008, 1, 136-143.

[82] Cheng, F., Wu, J., Zhang, J., Pan, A., et al., Development and inter-laboratory transfer of a decaplex polymerase chain reaction assay combined with capillary electrophoresis for the simultaneous detection of ten food allergens. Food Chem. 2016, 199, 799-808.

Chapter 2

Protein transport across the small intestine in food allergy

Marit Reitsma, Joost Westerhout, Harry J. Wichers, Heleen M. Wortelboer, Kitty C. M. Verhoeckx



Published in Molecular Nutrition and Food Research, 2014, 58, 194–205

Abstract

In view of the imminent deficiency of protein sources for human consumption in the near future, new protein sources need to be identified. However, safety issues such as the risk of allergenicity are often a bottleneck, due to the absence of predictive, validated and accepted methods for risk assessment. The current strategy to assess the allergenic potential of proteins focuses mainly on homology, stability and cross-reactivity, although other factors such as intestinal transport might be of added value too. In this review, we present an overview of the knowledge of protein transport across the intestinal wall and the methods currently being used to measure this. A literature study reveals that protein transport in sensitised persons occurs para-cellularly with the involvement of mast cells, and trans-cellularly via enterocytes, while in non-sensitised persons micro-fold cells and enterocytes are considered most important. However, there is a lack of comparable systematic studies on transport of allergenic proteins. Knowledge of the multiple protein transport pathways and which model system can be useful to study these processes may be of added value in the risk assessment of food allergenicity.

Introduction

In the near future, a shortage of protein sources for human consumption is foreseen and, therefore, alternative and sustainable protein sources (e.g. insects and algae) are now being explored for the production of food and feed. However, safety issues, such as the risk of allergenicity of novel proteins are often a bottleneck in bringing these products to the food market due to the absence of predictive, validated and accepted methods for risk assessment. New proteins or genetically modified foods are currently assessed for their allergenic potential using an allergenicity assessment strategy advised by the Food and agriculture Organization and the WorldHealth Organization. This strategy is based on a weight of evidence approach that recognises that no single endpoint can be used to predict human allergenic potential [1, 2] and focuses on characterising the protein/gene source, amino acid sequence homology to known allergens, in vitro cross-reactivity with known allergens and protein stability in a static pepsin digestion model [3]. However, these methods are mostly subjective and no guidance on procedures and interpretation of the outcome is available as yet. For example, several reviews and studies have indicated that in vitro digestion using the simulated gastric fluid method, protein digestion by pepsin in acidic conditions, is not always a good predictor of allergenicity [4,5]. A comparative study by Fu *et al*. found no correlation between simulated gastric and intestinal fluid stability of allergenic and non-allergenic protein and between major and minor allergens [6]. Furthermore, it was shown (own experience) in a dynamic digestion model (TNO's intestinal model for the gastrointestinal tract) that even less stable proteins might reach the intestinal tract in an intact, at least immunogenic, form. We hypothesised that transport of food proteins and peptides across the gastrointestinal barrier is needed to induce sensitisation or to elicit an allergic reaction, making it an important parameter in allergy research, next to digestion of these proteins. Therefore, we collected information on how proteins are transported in both sensitised persons and non-sensitised persons. This information can be used to estimate whether intestinal passage of protein or immune-reactive protein fragments can be incorporated into a new method for risk assessment of allergenicity of a protein. For this reason a comprehensive literature study was performed using the open-access databases NCBI-PubMed, Scopus and Web of Science to review the current knowledge on the mechanisms of protein transport across the small intestine and which methods are currently being used to study the intestinal absorbance of proteins.

Transport of allergens across the intestinal tract

The gastrointestinal barrier has different functions to fulfil, e.g. absorb nutrients and exclude "unwanted" compounds, such as bacteria and allergens. The intestinal tract is composed of different layers, glycocalyx, luminal mucosa, which are composed of a monolayer of

epithelial cells, lamina propria, which is connective tissue scaffold containing the blood and lymphatic vessels, and the muscularis mucosa (muscle layer) [7]. The lamina propria is highly folded to form villi, increasing the intestinal surface. This surface is covered by a heterogeneous population of epithelial cells, including absorptive intestinal epithelial cells (enterocytes), enteroendocrine cells (L-cells), mucin-secreting goblet cells and microfold cells (M cells) [8].

The secretive and absorptive cells have a highly folded apical membrane, forming uniform micro-villi, express brush border enzymes and membrane-embedded transporter proteins. The epithelial cells are joined at their apical side by tight junctions and other integral membrane proteins (e.g. claudins and occludins) that prevent the passage of macromolecular compounds larger than 600Da [9, 10]. In addition, secretory Ig A (IgA) and mucus (secreted by goblet cells) also restrict the absorption of dietary antigens [11]. Mucus, a mixture consisting mainly of water, glycoproteins (e.g. mucin), lipids and proteins, forms a protective layer of 50–450µm on top of the intestinal epithelial cell layer [12]. Although no reports on the relation between protein absorption and intestinal mucus were found, thickness of the mucus layer can influence the absorption rate of drug compounds [12], making it acceptable to reason that also food protein absorption can be limited or facilitated by the mucus.

Intestinal micro-biota is pivotal to health homeostasis. Although a clear relation between gut micro-biota and allergy has been established many times, the mechanism behind this is not clear yet. As summarised by Gigante *et al*., a difference in gut micro-biota can be seen between atopic (prone to develop allergies) and healthy persons. Furthermore, the gut micro-biota plays a crucial role in oral tolerance induction as germ-free mice do not develop tolerance [13]. Moreover, lack of intestinal bacteria in germ-free mice leads to a decreased intestinal surface, decreased intestinal cell renewal, and a thinner mucus layer [14]. However, no evidence was found on the role of micro-biota in protein transport, but due to the date present, it can be envisioned that micro-biota will play a relevant role.

Para-cellular and trans-cellular transport routes of proteins

Transport of proteins across the intestinal tract depends on size, polarity, shape, aggregation status and 3D structure of the protein and may occur either via the para-cellular route or via trans-cellular routes (Figure 2.1). Para-cellular transport refers to the transfer of compounds through the inter-cellular space between the cells and is regulated by the integrity of tight junctions [19,20]. Para-cellular transport is only considered to occur for small hydrophilic compounds (up to 600Da [9, 10]), but Pauletti *et al*. and Rubas *et al*. described that the maximal radius of a protein, which could pass the intestinal barrier via the para-cellular mechanism was estimated at 15A° (± 3.5kDa) [19, 21]. Proteins transported via the para-cellular route are not exposed to lysosomes in the enterocyte and are therefore not

degraded [15]. In healthy adults, however, paracellular protein/antigen transport is not considered of much importance [22–25]. This is in contrast to already sensitised persons, where the integrity of the tight junctions is decreased due to the presence of mast cells, enhancing the amount of non-degraded protein entering the human body [26] as will be more extensively explained in the section concerning mast cells.

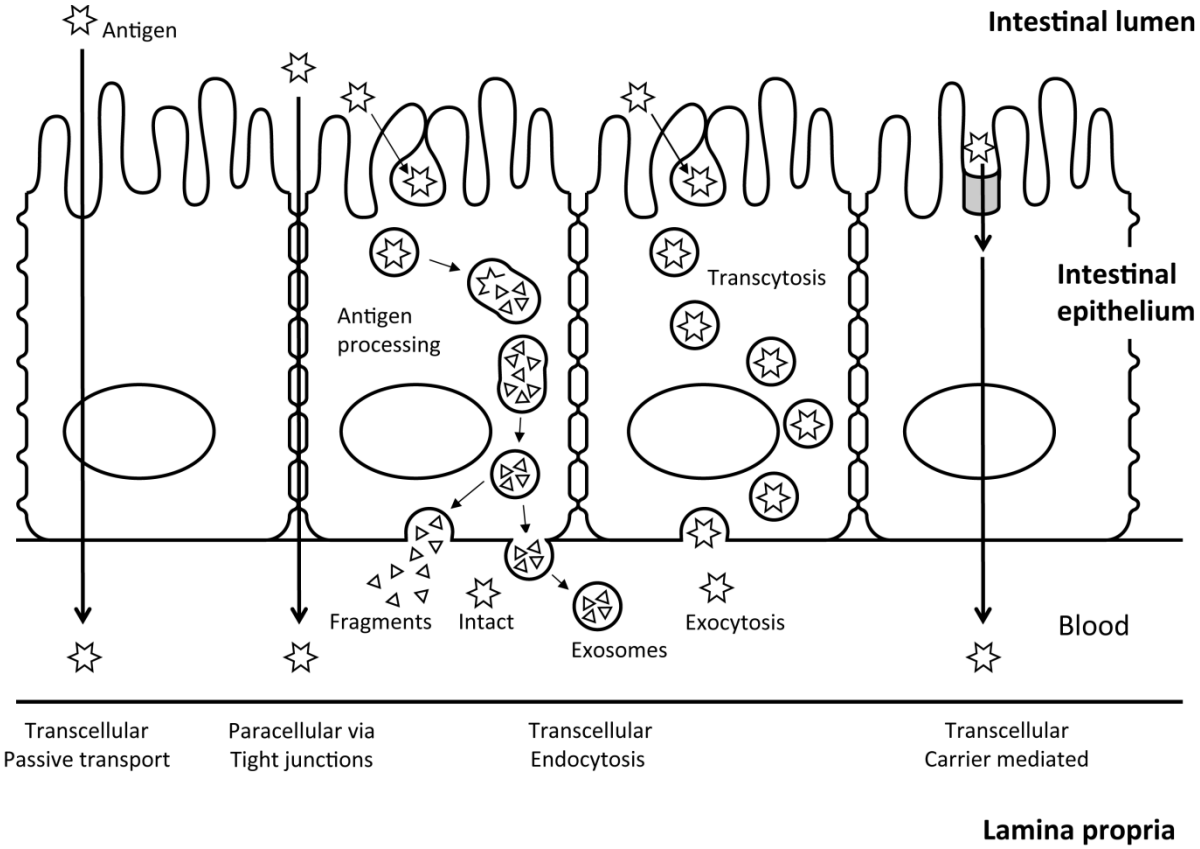


Figure 2.1 *Different passage routes across the intestines.*

The trans-cellular transport route comprises the absorption of compounds via passive diffusion, carrier-mediated transport, or endocytosis (Figure 2.1). As proteins are mostly hydrophilic macro-molecules, passive diffusion across the lipid bi-layer of the cell membrane is considered minimal and will therefore not be discussed in this review. Carrier-mediated transport routes are present in the human intestine for amino-acids [27] and di- and tri-peptides [22], but literature on protein carrier-mediated transport was not found. The main route of trans-cellular protein transport is endocytosis, which is known to occur in different cell types and will be described in the following sections.

Enterocytes

Enterocytes or absorptive intestinal cells are epithelial cells and are the most abundant cells in the intestinal barrier. Due to the formation of tight junctions between two adjacent enterocytes, they are involved in para-cellular transport, but additionally they are also

involved in protein and peptide absorption via the trans-cellular route. According to So *et al*. [28], soluble particles such as proteins can be endocytosed by enterocytes, after which they are transported in small vesicles or larger phagosomes and are either digested in lysosomes or left intact (trans-cytosis) [29], but the fragments as well as the intact proteins may still be allergenic, as was shown by Terpend *et al*. In this study, HT29–9A cells (an enterocytic cell line) were incubated with HRP; 1% HRP was endocytosed by the cells, of which 90% was subsequently hydrolysed intra-cellularly. However, the authors showed that 40% of the hydrolysed protein had a molecular mass of at least 1100Da, which is still large enough to bind MHC-II molecules [30]. In another study, it was estimated that 2% of intact proteins can reach the intestinal lymph and portal circulation under physiological conditions [31].

Belut *et al*. described another trans-cellular pathway of antigen uptake by enterocytic cells, which is facilitated by IgE and only occurs in sensitised persons, when luminal antigen-specific IgE is present [32]. This IgE forms a complex with the antigen and binds to the CD23 receptor, which is over-expressed on the apical membrane of enterocytes in sensitised persons. The CD23 receptor transports the IgE-antigen complex trans-cellularly without lysosomal degradation across the intestinal membrane [9, 33]. The importance of CD23 in this transport system was shown by Bevilacqua *et al*. who exposed the intestine of HRP sensitised and non-sensitised mice to HRP. The HRP sensitised mice showed increased transport of intact HRP, which was eliminated after the addition of anti-CD23 antibodies, whereas the transport of partly degraded HRP was not changed [33]. The involvement of IgE was further demonstrated by Yu *et al*. who passively immunised naive mice by injecting immune serum from a mouse that had been actively sensitised to HRP. IgE depletion of this serum eliminated the induced increase in intestinal trans-epithelial antigen transport [34].

O'Brien *et al*. showed that after antigen uptake with or without subsequent lysosomal degradation, antigens with a size of at least 18–20 amino acids long can be presented on enterocytic MHC-II molecules [35], or secreted into exosomes. Thereafter, the antigen can be presented to T cells by MHCII molecules on the enterocyte, but this rarely happens [36]. It is more likely that antigen-containing exosomes or tolerosomes will present the antigens on their MHC-II molecules to T cells [9, 36, 37], possibly reaching them via pores in the intestinal basement membrane or via the blood circulation. Inflammatory conditions in the intestine may increase exosome secretion from intestinal epithelial cells as was shown in vitro when intestinal epithelial cells increased secretion of exosomes after the addition of the pro-inflammatory cytokine IFN-γ [36]. Exosomes are thought to induce tolerance toward the antigen it contains. However, not all studies confirm this effect [38]. It has been suggested that intestinal epithelial derived MHC-II containing exosomes are tolerogenic in the absence of co-stimulatory molecules, but can activate T cells in presence of co-stimulatory molecules [25].

Besides IgE, also IgA and IgG are involved in enterocytic protein transport. Polymeric IgA and IgG reach the intestinal lumen trans-cellularly via, respectively, the polymeric IgA receptor and the neonatal Fc receptor on enterocytes [39, 40]. However, in contrast to IgE, binding of luminal antigens to IgA prevents their uptake, and binding of IgA to antigens that have already crossed the intestinal tract leads to secretion back into the intestinal lumen. Therefore, antigen-specific IgA is considered to be protective against sensitisation and allergic reactions. Transport of IgG–antigen complexes across the intestines in the neonate seems to be tolerogenic, but in adults this role is less clear [40].

M cells

M cells are specialised epithelial cells that are part of the Peyer's patches in the intestinal tract (Figure 2.2), but they can also be found in the gut independently of these patches [42]. Since M cells contain fewer lysosomes, have a thinner glycocalyx layer compared to enterocytes and do not present membrane-associated enzymes, the chance of compounds to be degraded during trans-cellular transport is low. M cells have been shown to transport proteins, bacteria, viruses and other particles of up to 1µm [43], either via phagocytosis [44], endocytosis [45] or pinocytosis [43]. For example, HRP has been shown to be absorbed by rabbit and piglet tonsil and intestinal Peyer's patch cells [45, 46]. Another study showed that in vivo exposure of murine intestine to IgA and IgG against mouse mammary tumour virus revealed binding and internalisation of both Igs by M cells [47].

The mechanism by which M cells transport proteins and/or antigens has been suggested to occur via a basal pocket, such as invagination in M cells creating "M cell pockets" where B and T cells, macrophages and dendritic cells (DC) appear to be present [48, 49]. The exact function of these pockets is unknown but it is thought that these pockets may shorten the intra-cellular distance for antigens to travel before being displayed to antigen-presenting cells, which then migrate to antigen-specific lymphocytes in underlying lymphoid follicles inducing T-cell proliferation. This process results in the development of IgA-producing B cells, some of which move into the vasculature and then back to the mucosal surfaces, efficiently seeding specific mucosal immunity [43, 50, 51]. Considering antigen transport, most research has been focused on the uptake of antigens by either M cells or enterocytes. The current perception is that particulate or aggregated antigens are taken up by M cells, inducing a local or systemic immune response towards that antigen and inducing the production of IgA. Soluble antigens, on the other hand, are thought to be predominantly absorbed by epithelial cells, leading to suppression of the immune system, induction of tolerance, towards this antigen [11, 52]. A comparison between soluble and aggregated milk proteins (α-lactalbumin and β-lactoglobulin) revealed both in vitro and in vivo an uptake of aggregated milk proteins by Peyer's patches, containing mostly M cells, inducing a stronger

immune reaction than the soluble milk proteins that were predominantly transported into intestinal epithelial cells. However severity of orally induced anaphylactic reactions were greatly impaired by aggregation of the milk proteins, implying that aggregated antigens can induce sensitisation, while allergic reactions following this sensitisation are mainly caused by soluble antigens [53]. However, M cells have been proven to transport not only insoluble, but also soluble peptides and the onset of tolerance versus sensitisation also seems to be influenced by particle size [54, 55].

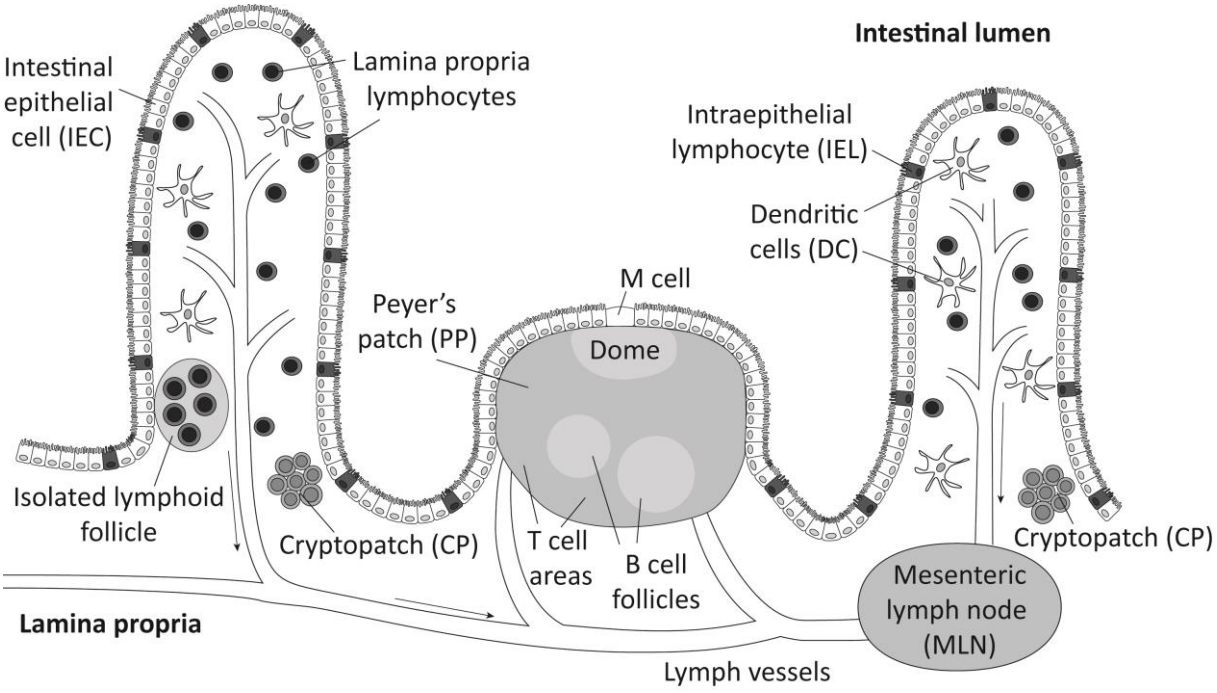


Figure 2.2 *The intestine at a cellular level, based on Spahn and Kucharzik [41].*

Mast cells

Shea-Donohue *et al*. estimated that about 2% of resident intestinal cells are mast cells and during immune responses this amount increases [56]. Mast cells are involved in eliciting symptoms during allergic reactions, but they are also indirectly involved in the transport of allergens. In this respect, food hypersensitivity reactions are known to occur in two phases. The first phase is mast cell independent and antigen-specific with antigen uptake occurring via trans-cytosis by intestinal epithelial cells, which is increased by sensitisation to the allergen. The next phase is mast cell-dependent, increased after sensitisation, and during this phase intestinal permeability to non-specific bystander proteins is also increased [26]. Using colonocytes this increased intestinal permeability was shown to be triggered by IgE-induced tryptase release by mast cells, which activates the protease-activated receptor 2 on colonic epithelial cells, causing redistribution of several tight junction proteins and inducing increased para-cellular permeability to macromolecules [57]. As small intestinal enterocytes

also contain this protease-activated receptor [58], we expect a similar interaction between activated mast cell components and enterocytes in the small intestine.

DCs

DCs constitute a major cellular component of the intestinal lamina propria, and play a pivotal role in the balance between tolerance and sensitisation in the intestine. DCs are antigen-presenting cells and are able to intercalate between intestinal epithelial cells, by virtue of extended dendrites without disrupting the epithelial barrier by opening tight junctions and forming new tight junctional complexes with adjacent epithelial cells, a process that is up-regulated in inflammatory conditions [59]. In this way DCs are able to take up soluble compounds from the lumen, although protein up-take by dendrites is controversial [60] and no literature data of direct protein or peptide sampling by DCs from the intestinal lumen was found.

In addition, Chambers *et al*. showed that intestinal (Peyer's patch) DCs obtained from sensitised and subsequently challenged mice can induce allergen-specific IgE in naive mice in the absence of an allergen challenge [61]. However, in general, intestinal DCs are thought to be more tolerogenic compared to DCs from other organs (e.g. the spleen) as they activate relatively more regulatory T cells [62]. Both the level of DC maturation and the type of DC could be of influence in this "tolerogenic" role; immature DCs are more likely to induce tolerance [36] and plasmacytoid DCs have been shown to decrease sensitisation [63].

Macro-phages

Besides DCs, macro-phages have been shown capable of transporting antigens across the intestinal tract using pseudopodia [64]. After antigen uptake, the antigen-containing macro-phages enter the systemic circulation, leading to systemic reactions towards this antigen. Whether or not Peyer's patch macro-phages transport compounds out of Peyer's patches depends on the compound size; (latex) particles with a diameter <5µm can be transported via macro-phages, while particles with a diameter of 5–10µm mostly remain inside the Peyer's patches, eliciting local mucosal effects [18, 65]. Similar to intestinal DCs, intestinal lamina propria macro-phages display a non-inflammatory phenotype compared to macro-phages from other tissues, e.g. blood monocytes, from which they stem. Unlike other macro-phages, intestinal macro-phages do not express receptors for, among others, LPS, IgG and IL-2, and express low levels of all major pro-inflammatory cytokines, even after phagocytosis, an action which they perform avidly [66]. On top of this non-inflammatory phenotype of intestinal macro-phages, these cells have also been described as critical actors in intestinal tolerance induction to food protein antigens [67].

Goblet cells

The last epithelial cell type that must be discussed is the goblet cell, whose main function is the production of mucin. Recently McDole *et al*. showed in vivo, in healthy persons, that goblet cells transported 10kDa peptides from the intestinal lumen to underlying lamina propria DCs. However, this transport is probably size limited as it was shown that large particles (0.02–1µm) cannot enter via this pathway [68].

From such literature data it can be concluded that protein transport in non-sensitised persons occurs via the M cells and enterocytes, while in sensitised persons the para-cellular route with involvement of mast cells, and trans-cellular route via enterocytes are considered the most important.

Measurement of allergen transport across the intestinal tract

Food proteins need to cross the gastrointestinal tract in order to induce sensitisation or elicit an allergic reaction. Therefore, we assume that the assessment of allergen or protein transport across the intestinal barrier might be a relevant, additional parameter in allergenicity risk assessment. The different techniques to assess intestinal protein transport will be discussed in the following sections. An overview of reviewed transport studies and allergens used is given in Table 2.1.

In vitro assays

Several in vitro methods have been used to study intestinal protein absorption, and an overview is presented in Table 2.1. Various cell lines (e.g. Caco-2 and HT-29) can be used to study the epithelial transport of proteins in vitro. These cell lines are from different origins and species and all have their advantages and disadvantages. The Caco-2 cells, a cell line of human colonic origin, exhibit many properties of small intestinal epithelium as they form a polarised monolayer of well-differentiated columnar absorptive cells expressing a brush border on their apical surface with typical small intestinal enzymes and transporters. The cells form very dense tight junctions, whose tightness resembles more the colonic than small intestinal tissue [8]. HT-29 cells stem from a human colon adenocarcinoma cell line that contains both absorptive and mucus secretive cells [30]. Compared to the Caco-2 cell line or the human intestine, the HT-29 expresses few carriers and is mostly used to study the effect of mucin on transport [80]. One of the major advantages of using cell lines to study the intestinal transport of proteins is the relatively high throughput at which different proteins can be studied. On the other hand, due to the lack of a physiologically relevant environment (e.g. interactions with other cell-types, intestinal fluid composition and mucus presence) [81], prediction of human oral exposure to allergens on the basis of in vitro studies with cell

lines might be prone to errors. Other in vitro methods for measuring intestinal transport of proteins include the Ussing chamber technique and the everted sac technique [75–79, 82].

Table 2.1 *In vitro absorption studies using different techniques, proteins and analysis methods. For each study the transport of intact protein is presented. FITC; Fluorescein isothiocyanate.*

Technique	Species	Protein	Food-related	Intact protein	Analysis	Reference
Caco-2, everted sac	Human, Wistar rat	γ-Conglutin	Lupin	Yes	Immunoblotting	[69]
Caco-2, HT-29	Human	FITC-tetanus toxoid, FITC-ovalbumin	No	Assumed	Laser scanning confocal microscopy for fluorescence intensities	[28]
HT-29	Human	FITC-tetanus toxoid	No	Assumed	Fluorescence microscopy	[70]
Caco-2	Human	Ber e 1, Ses i 1	Brazil nut Sesame seed	Yes	HPLC, ELISA	[98]
Caco-2	Human	β-Lactoglobulin, albumin	Milk	Assumed	^{125}I-radioactive labeling	[72]
HT29-19A	Human	Horseradish peroxidise	No	Yes	HPLC, ^{3}H-labeling	[30]
Caco-2	Human	β-Lactoglobulin, α-lactalbumin	Milk	Yes	HPLC, ^{14}C-labeling	[73]
Ussing chamber	Rabbit	β-Lactoglobulin	Milk	Yes	Elisa, HPLC, SDS-PAGE, ^{14}C-labeling	[74]
Ussing chamber	Rabbit	Horseradish peroxidise	No	Yes	Enzymatic activity, ^{3}H-radioactive labeling	[75]
Ussing chamber	Rabbit	Bovine Serum Albumin	No	Yes	Immunoblotting, ELISA, ^{125}I-radioactive labeling	[76]
Ussing chamber	Human	Horseradish peroxidise	No	Yes	Enzymatic activity, ^{3}H-radioactive labeling	[77]
Ussing chamber	Pig	Horseradish peroxidise	No	Yes	Enzymatic activity	[78]
Ussing chamber	Wistar rat pups	Horseradish peroxidise	No	Yes	^{125}I-radioactive labeling combined with HPLC gel filtration	[79]
Ussing chamber	Human	Horseradish peroxidise	No	Yes	Enzymatic activity, ^{125}I-radioactive labeling	[77]
Introduction of material into the gut lumen (ileal loop)	BALB/c mouse	Ara h 1, Ara h 2, Ara h 3	Peanut	Unknown	Transmission electron microscopy, fluorescence microscopy, detection of protein bodies	[54]

Both techniques have a higher resemblance to the in vivo situation as they require small sections of ex vivo intestinal tissue rather than cell mono-layers. In both the Ussing chamber technique and everted sac technique, a section of the gut is used. In the everted sac method, the intestinal section is everted and both ends are tied after filling the sac with buffer. The 'sausage' is placed into buffer with the compound of interest. With the Ussing chamber technique, the intestinal segment is mounted in an Ussing chamber where one side will be exposed to buffer with the compound of interest (apical or lumen side) and the other side to buffer without the compound of interest (basolateral or mucosal side).
Heyman *et al*. and Majamaa *et al*. both used the Ussing technique with jejunal biopsies from children to study the transport of HRP [77, 82]. However, due to the limited availability of healthy human tissue, intestinal tissue from rabbit [75], piglets [78] or rats [79] is often used. No literature could be found on intestinal protein absorption using the everted sac technique with human tissue. However, the everted sac technique has been applied for investigating the absorption of human γ-globulin and ovalbumin across the ileum of rabbits and guinea pigs [83], as well as γ-conglutin absorption in rats [69]. A disadvantage of everted sac method against the Ussing chamber technique is that with everted sac the muscle layer is still present and this may lead to the underestimation of the protein transport value. A major disadvantage of using ex vivo human tissue is its relatively limited viability and low throughput. Normally, ex vivo tissue is used within 2–3 h, because intestinal oedema and disruption of the villi will occur when tissue is incubated in vitro [84]. The disadvantage of using animal tissue may be the inter-species differences in anatomy, physiology, metabolism, diet and micro-biota, which complicates the extrapolation of data to humans [80, 85]. But an advantage of using animal ex vivo tissue is the possibility to investigate the effect of sensitisation on the intestinal absorption of proteins by immunising the tissue in vivo prior to the in vitro studies [29]. Pigs share more physiological and immunological similarities to humans than rodents, and the use of (mini)pigs is becoming increasingly common in nutritional research [86]. Therefore, we recently developed InTESTineTM, a medium throughput alternative for the Ussing chamber technique using intestinal waste tissue from pigs, incubated on a rocker platform in a high oxygen incubator. In this system, viability of tissue could be retained for 2h and the para-cellular absorption transport resembles that of human intestinal tissue in Ussing chambers (Wortelboer *et al*., unpublished results) and transport of macro-molecular proteins was studied using radioactive labelled proteins (Verhoeckx *et al*., unpublished results).

In vivo assays

The most physiologically relevant way to study intestinal protein absorption is by oral exposure to the protein in vivo, as all types of intestinal cells, blood circulation and mucus are present. Preferably, to elucidate the protein absorption in human, in vivo experiments

are performed in humans, as this is the target species. However, the majority of experimental studies on protein absorption described in literature have been performed in animals (Table 2.2).Consequently, the interpretation of the findings could be difficult due to inter-species differences. Also, the possibility of separating variables that influence protein absorption is limited in in vivo studies [81].

Table 2.2 *In vivo absorption studies using different techniques, proteins and analysis methods. For each studie the transport of intact protein is presented.*

Technique	Species	Protein	Food-related	Intact	Analysis	Reference
Oral exposure	Pig	IgG, β-lactoglobulin, albumin, trypsin	Colostrum	Yes	Immunoblotting	[87]
Oral exposure	Human	α-Lactalbumin	Milk	Assumed	^{125}I-radioactive labeling	[88]
Oral exposure	Human	β-Lactoglobulin	Milk	Assumed	Radioimmunoassay	[89]
Oral exposure	Human	Ovalbumin	Egg	Yes	Gel permeation chromatography, ELISA	[16]
Oral exposure	Brown Norway rat	β-Lactoglobulin	Milk	Assumed	ELISA	[90]
Oral exposure	BALB/c mouse	human serum albumin	No	No	^{125}I-radioactive labeling	[91]
Oral exposure	ddY mouse	ovalbumin	Egg	Yes	ELISA, Immunoblotting	[71]
Oral exposure, everted sac	Guinea pig, hamster, rabbit	albumin, γ-globulin, β_2- γ-globulin	Egg/milk	Assumed	Fluorescence microscopy	[83]
Oral exposure	BALB/c mouse, CD rat rabbit	IgA	No	Assumed	Colloidal gold-conjugation electron microscopy	[47]
Oral exposure	ddY mouse	14-16kDa rice allergen	Rice	Yes	ELISA, Immunoblotting	[92]

The number of in vivo studies of intestinal protein absorption in humans is limited, though we were able to find some examples. Paganelli and Levinsky measured β-lactoglobulin by RIA in the serum of healthy human subjects, indicating that β-lactoglobulin was present in the blood after 0.5 – 3h of milk consumption [89]. Husby and colleagues investigated the absorption of ovalbumin and β-lactoglobulin into the blood of healthy adults by HPLC combined with ELISA. The study revealed that intact ovalbumin was absorbed by seven of eight individuals, whereas β-lactoglobulin was not detected in any of the subjects [93]. Jakobsson *et al*. measured α-lactalbumin from human milk in the serum of breast fed infants by RIA [88]. In the same project, these researchers discovered that the absorption of α-

lactalbumin decreased as age increased, indicating that the intestinal wall becomes less permeable, also mentioned as closure of the gut. A comparable observation was made by Roberton *et al*., who measured the absorption of cows' milk protein β-lactoglobulin into the blood of preterm and term neonates by RIA [94]. This observation appears to be applicable to all species, although the rate at which the decrease in protein absorption occurs is different between species [95].

It is clear that in in vivo protein transport studies many different animal species are used, e.g. rats, mice, guinea pigs, dogs, pigs, calves, cows, steers and sheep [22, 96]. In the absence of a so-called golden standard in predicting or identifying food allergens by use of intestinal transport studies, the use of larger animals, particularly pigs, might be better comparable to the human situation [96].

Protein analysis

A crucial step in protein transport studies is the analytical method to be used to detect the intact protein. In protein transport studies, proteins are often labelled radioactively [30, 72, 75, 76, 79, 91, 97] or fluorescently [28, 70, 83] for relatively easy detection. However, the detection of fluorescent or radioactive labelled proteins in plasma does not necessarily indicate that the protein is absorbed intact [91, 97]. The same is true for using transmission electron microscopy [46, 47, 61, 83], fluorescence microscopy [28, 61, 70, 83], light microscopy [46, 54], ELISA [16, 71, 76, 90, 92, 97, 98] and/or a RIA [88, 89]. A positive result in an ELISA or RIA could be caused by only a specific part of the protein, so even if the protein is degraded intra-cellularly it may still be recognised by the antibody used in these assays. Immunoblotting [69, 71, 87, 92] and high-performance liquid (gel permeation) chromatography [16, 30, 79, 99] are good alternatives, since these techniques separate the proteins by their molecular weight, thus proving the presence of intact protein. Another appropriate method, but unfortunately not always possible, is measuring the functional activity of the protein. For example, Heyman *et al*. used an HRP enzyme activity assay to measure intact HRP, since only intact HRP shows enzyme activity [100]. A more quantitative method, however not applied in the screened literature, is the use of LC coupled to MS. However, the use of these techniques requires special skills and equipment, fragments and intact proteins cannot always be measured simultaneously and the concentration of transported proteins and fragments is often at the lower LOD. Furthermore, the MS spectra of intact proteins are very complex and are difficult to recognise in biological samples containing more than one protein.

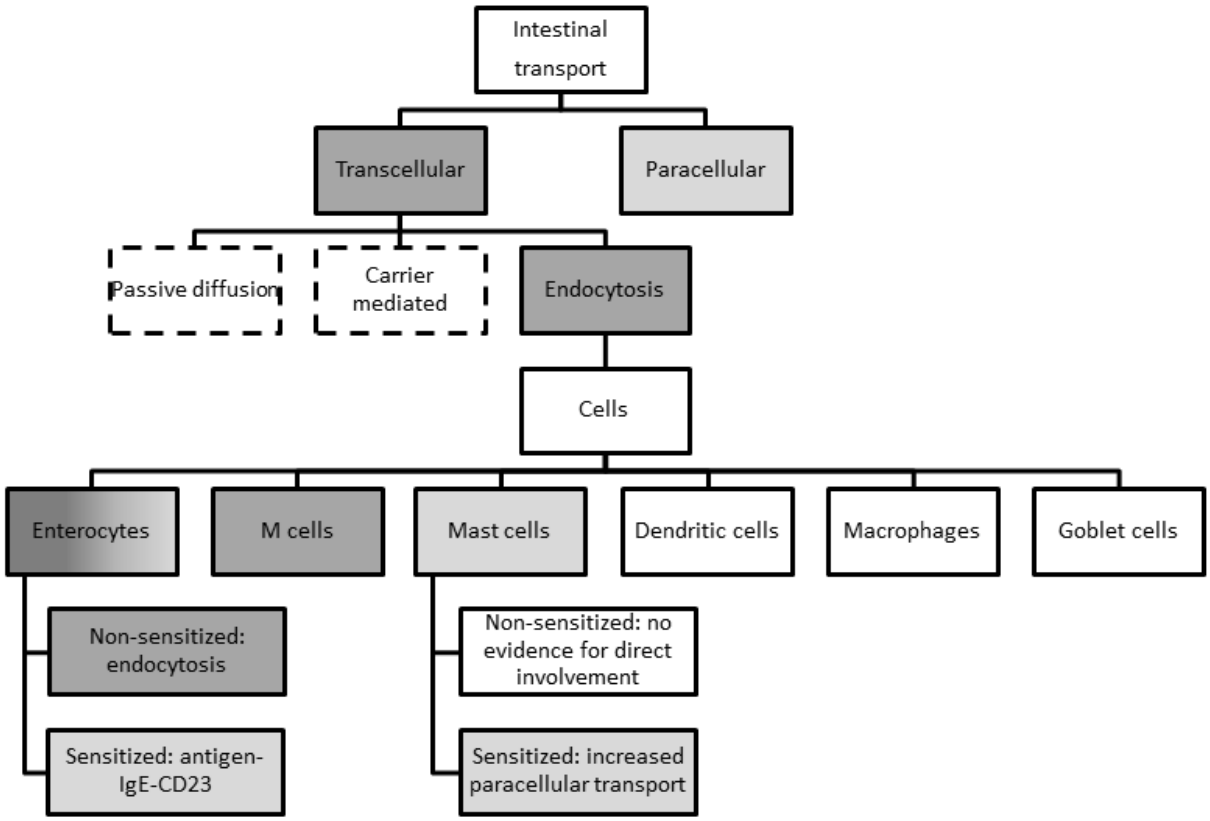


Figure 2.3 *Different mechanisms and cells involved in intestinal protein transport. Boxes indicate they have no (dashed lines), or a possible (no filling) role in intestinal protein transport in sensitised (light grey) or non-sensitised (dark grey) persons.*

Conclusion and discussion

New proteins or genetically modified foods are currently assessed for their allergenic potential using an allergenicity assessment strategy advised by the Food and Agriculture Organization and World Health Organization. This strategy is based on a weight of evidence approach that recognises that no single endpoint can be used to predict human allergenic potential [1, 2].

The goal of this review was to study how protein and immune-reactive protein fragments are transported across the intestinal wall, and whether it can be advised to develop a new model for allergenicity risk assessment based on this protein transport.

Proteins are transported across the intestinal barrier via many different routes and these routes appear to depend on the health status of the individual. The different routes and cells involved are summarised in Figure 2.3. The data indicate that in non-sensitised persons, absorption of proteins via M cells and trans-cellular transport via enterocytes are considered very important, whereas in sensitised persons, proteins can also be transported via the para-cellular transport route with involvement of mast cells, and trans-cellular via IgE-mediated transport via the CD23 receptor on enterocytes. The current perception is that particulate or

aggregated antigens are taken up by M cells, inducing a local or systemic immune response towards that antigen and inducing the production of IgA. Soluble antigens on the other hand are thought to be predominantly absorbed by epithelial cells, leading to suppression of the immune system, induction of tolerance, towards this antigen [11, 52]. For this reason, transport route (M cells or epithelial cells) and/or transported protein size (intact or fragmented) could be possible parameters to predict the allergenic potential of proteins. However, at the moment more research is needed since no comparative study on proteins with different allergenicity was found in the literature and the studies performed on individual allergens cannot be compared due to inter-laboratory differences, differences in transport methods and the analytical methods used.

Enough evidence was found that intact protein can cross the intestinal wall, which indicates that there is a potential risk of immunological sensitisation. This urges the need to develop a good protocol to determine the route of transport. In our opinion, the use of ex vivo intestinal segments or in vivo experiments is crucial next to a reliable instrumental analytical method (e.g. LC-MS) to analyse intact and fragmented proteins preferably simultaneously.

Acknowledgements

The authors thank Marleen Hoozemans for her help with Figure 2.1 and 2.2. The authors are member of Cost-Action INFOGEST.

References

[1] Codex Alimentarius Commission, Alinorm 03/34: Joint FAO/WHO Food Standard Programme, Codex Alimentarius Commission, Twenty-Fifth Session, Rome, Italy 30 June–5 July, 2003. Appendix III, guideline for the conduct of food safety assessment of foods derived from recombinant-DNA plants and Appendix IV, Annex on the assessment of possible allergenicity. 2003, 47–60.

[2] EFSA, EFSA Panel on Genetically Modified Organisms (GMO), Draft Scientific Opinion on the assessment of allergenicity of GM plants and microorganisms and derived food and feed. EFSA J. 2010, 8, 1–168.

[3] FAO/WHO, Evaluation of Allergenicity of Genetically Modified Foods. Report of a Joint FAO/WHO Expert Consultation on Allergenicity of Foods Derived from Biotechnology, Rome, Italy 2001, 22–25 January, 1–18.

[4] Herman, R. A., Woolhiser, M. M., Ladics, G. S., Korjagin, V. A., et al., Stability of a set of allergens and non-allergens in simulated gastric fluid. Int. J. Food Sci. Nutr. 2007, 58, 125–141.

[5] Untersmayr, E., Jensen-Jarolim, E., The role of protein digestibility and antacids on food allergy outcomes. J. Allergy Clin. Immunol. 2008, 121, 1301–1308.

[6] Fu, T., Abbott, U. R., Hatzos, C., Digestibility of food allergens and nonallergenic proteins in simulated gastric fluid and simulated intestinal fluid—a comparative study. J. Agric. Food Chem. 2002, 50, 7154–7160.

[7] Reitsma, S., Slaaf, D. W., Vink, H., van Zandvoort, M. A., oude Egbrink, M. G., The endothelial glycocalyx: composition, functions, and visualization. Pflugers Arch. Eur. J. Phys. 2007, 454, 345–359.

[8] Hillgren, K. M., Kato, A., Borchardt, R. T., In vitro systems for studying intestinal drug absorption. Med. Res. Rev. 1995, 15, 83–109.

[9] Heyman, M., Gut barrier dysfunction in food allergy. Eur. J. Gastroenterol. Hepatol. 2005, 17, 1279–1285.

[10] Kondoh, M., Takahashi, A., Yagi, K., Spiral progression in the development of absorption enhancers based on the biology of tight junctions. Adv. Drug Deliv. Rev. 2012, 64, 515–522.

[11] Sampson, H. A., Food allergy. Part 1: immunopathogenesis and clinical disorders. J. Allergy Clin. Immunol. 1999, 103, 717–728.

[12] Khanvilkar, K., Donovan, M. D., Flanagan, D. R., Drug transfer trough mucus. Adv. Drug. Deliv. Rev. 2001, 48, 173–193.

[13] Gigante, G., Tortora, A., Ianiro, G., Ojetti, V. et al., Role of gut microbiota in food tolerance and allergies. Dig. Dis. 2011, 9, 40–549.

[14] Natividad, J. M. M., Verdu, E. F., Modulation of intestinal barrier by intestinal microbiota: pathological and therapeutic implications. Pharmacol. Res. 2013, 69, 42–51.

[15] Shimizu, M., Food-derived peptides and intestinal functions. Bio. Factors 2004, 21, 43–47.

[16] Husby, S., Jensenius, J. C., Svehag, S. E., Passage of undegraded dietary antigen into the blood of healthy adults. Further characterization of the kinetics of uptake and the size distribution of the antigen. Scand. J. Immunol. 1986, 24, 447–455.

[17] van Ness van Alstyne, E., The Absorption of Protein Without Digestion. Arch. Intern. Med. (chic) 1913, 12, 372–394.

[18] Smith, P. L., Wall, D. A., Gochoco, C. H., Wilson, G., Routes of delivery: case studies. Adv. Drug Deliv. Rev. 1992, 8, 253– 290.

[19] Rubas, W., Cromwell, M. E., Shahrokh, Z., Villagran, J., et al., Flux measurements across Caco-2 monolayers may predict transport in human large intestinal tissue. J. Pharm. Sci. 1996, 85, 165–169.

[20] Morishita, M., Peppas, N. A., Is the oral route possible for peptide and protein drug delivery? Drug Discov. Today 2006, 11, 905–910.

[21] Pauletti, G. M., Gangwar, S., Knipp, G. T., Nerurkar, M. M., et al., Structural requirements for intestinal absorption of peptide drugs. J. Control. Release 1996, 41, 3–17.

[22] Gilbert, E. R., Wong, E. A., Webb, K. E., Jr., Board-invited review: peptide absorption and utilization—implications for animal nutrition and health. J. Anim. Sci. 2008, 86, 2135– 2155.

[23] Heyman, M., Symposium on 'dietary influences on mucosal immunity'. How dietary antigens access the mucosal immune system. Proc. Nutr. Soc. 2001, 60, 419–426.

[24] Keita, A. V., Soderholm, J. D., The intestinal barrier and its regulation by neuroimmune factors. Neurogastroenterol. Motil. 2010, 22, 718–733.

[25] Mallegol, J., van Niel, G., Heyman, M., Phenotypic and functional characterization of intestinal epithelial exosomes. Blood Cell. Mol. Dis. 2005, 35, 11–16.

[26] Berin, M. C., Kiliaan, A. J., Yang, P. C., Groot, J. A., et al., The influence of mast cells on pathways of transepithelial antigen transport in rat intestine. J. Immunol. 1998, 161, 2561–2566.
[27] Thwaites, D. T., Anderson, C. M., The SLC36 family of proton-coupled amino acid transporters and their potential role in drug transport. Br. J. Pharmacol. 2011, 164, 1802– 1816.
[28] So, A. L., Small, G., Sperber, K., Becker, K., et al., Factors affecting antigen uptake by human intestinal epithelial cell lines. Dig. Dis. Sci. 2000, 45, 1130–1137.
[29] Walker, W. A., Isselbacher, K. J., Bloch, K. J., Immunologic control of soluble protein absorption from the small intestine: a gut-surface phenomenon. Am. J. Clin. Nutr. 1974, 27, 1434–1440.
[30] Terpend, K., Boisgerault, F., Blaton, M. A., Desjeux, J. F., et al., Protein transport and processing by human HT29–19A intestinal cells: effect of interferon gamma. Gut 1998, 42, 538–545.
[31] Warshaw, A. L., Walker, W. A., Isselbacher, K. J., Protein uptake by the intestine: evidence for absorption of intact macromolecules. Gastroenterology 1974, 66, 987–992.
[32] Belut, D., Moneret-Vautrin, D. A., Nicolas, J. P., Grilliat, J. P., IgE levels in intestinal juice. Dig. Dis. Sci. 1980, 25, 323–332.
[33] Bevilacqua, C., Montagnac, G., Benmerah, A., Candalh, C., et al., Food allergens are protected from degradation during CD23-mediated transepithelial transport. Int. Arch. Allergy Immunol. 2004, 135, 108–116.
[34] Yu, L. C., Yang, P. C., Berin, M. C., Di Leo, V., et al., Enhanced transepithelial antigen transport in intestine of allergic mice is mediated by IgE/CD23 and regulated by interleukin-4. Gastroenterology 2001, 121, 370–381.
[35] O'Brien, C., Flower, D. R., Feighery, C., Peptide length significantly influences in vitro affinity forMHC class II molecules. Immunome Res. 2008, 4, 6.
[36] Telemo, E., Korotkova, M., Hanson, L. A., Antigen presentation and processing in the intestinal mucosa and lymphocyte homing. Ann. Allergy Asthma Immunol. 2003, 90, 28–33.
[37] van Niel, G., Raposo, G., Candalh, C., Boussac, M. et al., Intestinal epithelial cells secrete exosome-like vesicles. Gastroenterology 2001, 121, 337–349.
[38] Van Niel, G., Mallegol, J., Bevilacqua, C., Candalh, C. et al., Intestinal epithelial exosomes carry MHC class II/peptides able to inform the immune system in mice. Gut 2003, 52, 1690–1697.
[39] Mavris, M., Sansonetti, P., Epithelial cell responses. Best Pract. Res Clin. Gastroenterol. 2004, 18, 373–386.
[40] Berin, M. C., Mucosal antibodies in the regulation of tolerance and allergy to foods. Semin. Immunopathol. 2012. 34, 633–642.
[41] Spahn, T. W., Kucharzik, T., Modulating the intestinal immune system: the role of lymphotoxin and GALT organs. Gut 2004, 53, 456–465.
[42] Jang, M. H., Kweon, M. N., Iwatani, K., Yamamoto, M. et al., Intestinal villous M cells: an antigen entry site in the mucosal epithelium. Proc. Natl. Acad. Sci. USA 2004, 101, 6110–6115.
[43] Miller, H., Zhang, J., Kuolee, R., Patel, G. B. et al., Intestinal M cells: the fallible sentinels? World J. Gastroenterol. 2007, 13, 1477–1486.
[44] Fujimura, Y., Functional morphology of microfold cells (M cells) in Peyer's patches—phagocytosis and transport of BCG by M cells into rabbit Peyer's patches. Gastroenterol. Jpn. 1986, 21, 325–335.
[45] Neutra, M. R., Phillips, T. L., Mayer, E. L., Fishkind, D. J., Transport of membrane-bound macromolecules by M cells in follicle-associated epithelium of rabbit Peyer's patch. Cell Tissue Res. 1987, 247, 537–546.
[46] Gebert, A., Identification of M-cells in the rabbit tonsil by vimentin immunohistochemistry and in vivo protein transport. Histochem. Cell Biol. 1995, 104, 211–220.
[47] Weltzin, R., Lucia-Jandris, P., Michetti, P., Fields, B. N. et al., Binding and transepithelial transport of immunoglobulins by intestinal M cells: demonstration using monoclonal IgA antibodies against enteric viral proteins. J. Cell. Biol. 1989, 108, 1673–1685.
[48] Farstad, I. N., Halstensen, T. S., Fausa, O., Brandtzaeg, P., Heterogeneity of M-cell-associated B and T cells in human Peyer's patches. Immunology 1994, 83, 457–464.
[49] Iwasaki, A., Kelsall, B. L., Unique functions of CD11b+, CD8 alpha+, and double-negative Peyer's patch dendritic cells. J. Immunol. 2001, 166, 4884–4890.
[50] Yamanaka, T., Straumfors, A., Morton, H., Fausa, O. et al., M cell pockets of human Peyer's patches are specialized extensions of germinal centers. Eur. J. Immunol. 2001, 31, 107–117.
[51] Neutra, M. R., Mantis, N. J., Kraehenbuhl, J. P., Collaboration of epithelial cells with organized mucosal lymphoid tissues. Nat. Immunol. 2001, 2, 1004–1009.

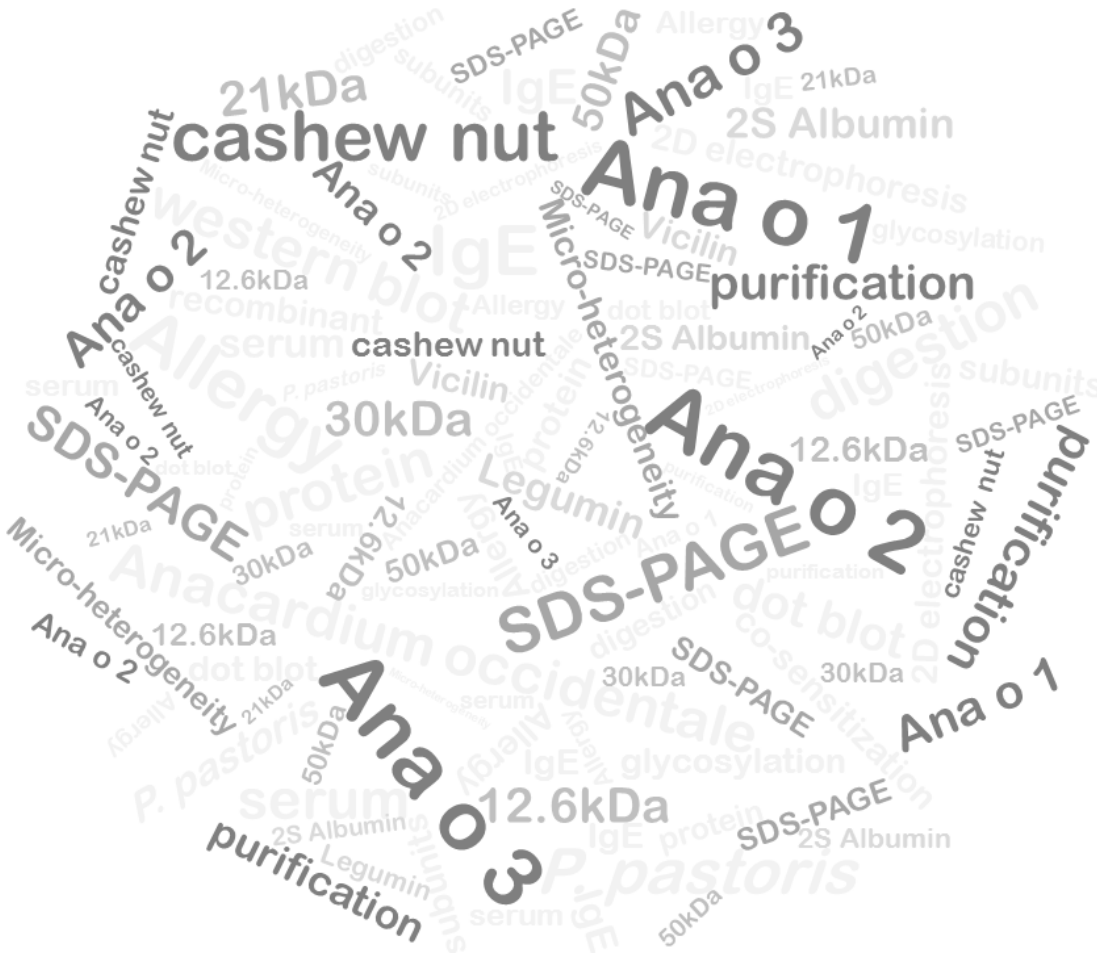
[52] Brandtzaeg, P. E., Current understanding of gastrointestinal immunoregulation and its relation to food allergy. Ann. N Y Acad. Sci. 2002, 964, 13–45.
[53] Roth-Walter, F., Berin, M. C., Anaboldi, P., Escalante, C. R. et al., Pasteurization of milk proteins promotes allegic sensitization by enhancing uptake through Peyer's patches. Allergy 2008, 63, 882–890.
[54] Chambers, S. J., Wickham, M. S., Regoli, M., Bertelli, E. et al., Rapid in vivo transport of proteins from digested allergen across pre-sensitized gut. Biochem. Biophys. Res. Commun. 2004, 325, 1258–1263.
[55] Matsunaga, Y., Wakatsuki, Y., Tabata, Y., Kawasaki, H. et al., Oral immunization with size-purified microsphere beads as a vehicle selectively induces systemic tolerance and sensitization. Vaccine 2000, 19, 579–588.
[56] Shea-Donohue, T., Stiltz, J., Zhao, A., Notari, L., Mast cells. Curr. Gastroenterol. Rep. 2010, 12, 349–357.
[57] Jacobs, C., Yang, P., Darmoul, D., Amadesi, S. et al., Mast cell tryptase controls paracellular permeability of the intestine. J. Biol. Chem. 2005, 280, 36, 31936–31948.
[58] Kong, W., McConalogue, K., Khitin, L. M., Hollenberg, M. D. et al., Luminal trypsin may regulate enterocytes through proteinase-activated receptor 2. Proc. Natl. Acad. Sci. USA 1997, 94, 8884–8889.
[59] Rescigno, M., Urbano, M., Valzasina, B., Francolini, M. et al., Dendritic cells express tight junction proteins and penetrate gut epithelialmonolayers to sample bacteria. Nat. Immunol. 2001, 2, 361–367.
[60] Schulz, O., Jaensson, E., Persson, E. K., Liu, X. et al., Intestinal CD103+, but not CX3CR1+, antigen sampling cells migrate in lymph and serve classical dendritic cell functions. J. Exp. Med. 2009, 206, 3101–3114.
[61] Chambers, S. J., Bertelli, E., Winterbone, M. S., Regoli, M. et al., Adoptive transfer of dendritic cells from allergic mice induces specific immunoglobulin E antibody in naive recipients in absence of antigen challenge without altering the T helper 1/T helper 2 balance. Immunology 2004, 112, 72–79.
[62] Manicassamy, S., Reizis, B., Ravindran, R., Nakaya, H. et al., Activation of beta-catenin in dendritic cells regulates immunity versus tolerance in the intestine. Science 2010, 329, 849–853.
[63] Smit, J. J., Bol-Schoenmakers, M., Hassing, I., Fiechter, D. et al., The role of intestinal dendritic cells subsets in the establishment of food allergy. Clin. Exp. Allergy 2011, 41, 890–898.
[64] Rubas, W., Grass, G. M., Gastrointestinal lymphatic absorption of peptides and proteins. Adv. Drug Deliv. Rev. 1991, 7, 15–69.
[65] Ebel, J. P., A method for quantifying particle absorption from the small intestine of the mouse. Pharm. Res. 1990, 7, 848–851.
[66] Smythies, L. E., Sellers, M., Clements, R. H., Mosteller- Barnum, M. et al., Human intestinal macrophages display profound inflammatory anergy despite avid phagocytic and bacteriocidal activity. J. Clin. Invest. 2005, 115, 66–75.
[67] Mcl Mowat, A., Bain, C. C., Mucosal macrophages in intestinal homeostasis and inflammation. Innate Immun. 2011, 3, 550–564.
[68] McDole, J. R., Wheeler, L. W., McDonald, K. G., Wang, B. et al., Goblet cells deliver luminal antigen to CD103+ dendritic cells in the small intestine. Nature 2012, 483, 345–349.
[69] Capraro, J., Clemente, A., Rubio, L. A., Magni, C. et al., Assessment of the lupin seed glucose-lowering protein intestinal absorption by using in vitro and ex vivo models. Food Chem. 2011, 125, 1279–1283.
[70] Mayer, L., So, L. P., Yio, X. Y., Small, G., Antigen trafficking in the intestine. Ann. N Y Acad. Sci. 1996, 778, 28–35.
[71] Matsubara, T., Aoki, N., Honjoh, T., Mizumachi, K. et al., Absorption, migration and kinetics in peripheral blood of orally administered ovalbumin in a mouse model. Biosci. Biotechnol. Biochem. 2008, 72, 2555–2565.
[72] Puyol, P., Perez, M. D., Sanchez, L., Ena, J. M. et al., Uptake and passage of beta-lactoglobulin, palmitic acid and retinol across the Caco-2monolayer. Biochim. Biophys. Acta 1995, 1236, 149–154.
[73] Caillard, I., Tome, D., Transport of beta-lactoglobulin and alpha-lactalbumin in enterocyte-like Caco-2 cells. Reprod. Nutr. Dev. 1995, 35, 179–188.
[74] Marcon-Genty, D., Tome, D., Kheroua, O., Dumontier, A. M. et al., Transport of beta-lactoglobulin across rabbit ileum in vitro. Am. J. Physiol. 1989, 256, G943–G948.
[75] Ducroc, R., Heyman, M., Beaufrere, B., Morgat, J. L. et al., Horseradish peroxidase transport across rabbit jejunum and Peyer's patches in vitro. Am. J. Physiol. 1983, 245, G54–G58.
[76] Hardin, J. A., Kimm, M. H., Wirasinghe, M., Gall, D. G., Macromolecular transport across the rabbit proximal and distal colon. Gut 1999, 44, 218–225.

[77] Majamaa, H., Isolauri, E., Evaluation of the gut mucosal barrier: evidence for increased antigen transfer in children with atopic eczema. J. Allergy Clin. Immunol. 1996, 97, 985–990.
[78] Keljo, D. J., Hamilton, J. R., Quantitative determination of macromolecular transport rate across intestinal Peyer's patches. Am. J. Physiol. 1983, 244, G637–G644.
[79] Sutas, Y., Autio, S., Rantala, I., Isolauri, E., IFN-gamma enhances macromolecular transport across Peyer's patches in suckling rats: implications for natural immune responses to dietary antigens early in life. J. Pediatr. Gastroenterol. Nutr. 1997, 24, 162–169.
[80] Deferme, S., Annaert, P., Augustijns, P., in: Ehrhardt, C., Kim, K.-J. (Eds.), Drug Absorption Studies, Springer, New York, USA, 2008, pp. 182–215.
[81] Le Ferrec, E., Chesne, C., Artusson, P., Brayden, D. et al., In vitromodels of the intestinal barrier. The report and recommendations of ECVAM Workshop 46. European Centre for the Validation of Alternative methods. Altern. Lab. Anim. 2001, 29, 649–668.
[82] Heyman, M., Boudraa, G., Sarrut, S., Giraud, M. et al., Macromolecular transport in jejunal mucosa of children with severe malnutrition: a quantitative study. J. Pediatr. Gastroenterol. Nutr. 1984, 3, 357–363.
[83] Lecce, J. G., Broughton, C. W., Cessation of uptake of macromolecules by neonatal guinea pig, hamster and rabbit intestinal epithelium (closure) and transport into blood. J. Nutr. 1973, 103, 744–750.
[84] Plumb, J. A., Burston, D., Baker, T. G., Gardner, M. L., A comparison of the structural integrity of several commonly used preparations of rat small intestine in vitro. Clin. Sci. 1987, 73, 53–59.
[85] Nejdfors, P., Ekelund, M., Jeppsson, B.,Westrom, B. R., Mucosal in vitro permeability in the intestinal tract of the pig, the rat, and man: species- and region-related differences. Scand. J. Gastroenterol. 2000, 35, 501–507.
[86] Patterson, J. K., Lei, X. G., Miller, D. D., The pig as an experimental model for elucidating the mechanisms governing dietary influence on mineral absorption. Exp. Biol. Med. (Maywood) 2008, 233, 651–664.
[87] Carlsson, L. C., Westrom, B. R., Karlsson, B. W., Intestinal absorption of proteins by the neonatal piglet fed on sow's colostrum with either natural or experimentally eliminated trypsin-inhibiting activity. Biol. Neonate 1980, 38, 309–320.
[88] Jakobsson, I., Axelsson, I., Juvonen, P., Lindberg, T. et al., Human α-lactalbumin as a marker of macromolecular absorption in early infancy. Acta Paediatr. 1989, 78, 42–47.
[89] Paganelli, R., Levinsky, R. J., Solid phase radioimmunoassay for detection of circulating food protein antigens in human serum. J. Immunol. Methods 1980, 37, 333–341.
[90] Sakamoto, Y., Ohtsuka, T., Yoshida, H., Ohto, K., et al., Time course of changes in the intestinal permeability of foodsensitized rats after oral allergen challenge. Pediatr. Allergy Immunol. 1998, 9, 20–24.
[91] Skogh, T., Overestimate of 125I-protein uptake from the adult mouse gut. Gut 1982, 23, 1077–1080.
[92] Yamada, C., Yamashita, Y., Seki, R., Izumi, H., et al., Digestion and gastrointestinal absorption of the 14–16-kDa rice allergens. Biosci. Biotechnol. Biochem. 2006, 70, 1890–1897.
[93] Husby, S., Jensenius, J. C., Svehag, S. E., Passage of undegraded dietary antigen into the blood of healthy adults. Quantification, estimation of size distribution, and relation of uptake to levels of specific antibodies. Scand. J. Immunol. 1985, 22, 83–92.
[94] Roberton, D. M., Paganelli, R., Dinwiddie, R., Levinsky, R. J., Milk antigen absorption in the preterm and term neonate. Arch Dis. Child. 1982, 57, 369–372.
[95] Lecce, J. G., Absorption of macromolecules by neonatal intestine. Biol. Neonat. 1965, 9, 50–61.
[96] Bischoff, S., Crowe, S. E., Gastrointestinal food allergy: new insights into pathophysiology and clinical perspectives. Gastroenterology 2005, 128, 1089–1113.
[97] Bohe, M., Borgstrom, A., Genell, S., Ohlsson, K., Characterization of radioactivity in plasma after intraduodenal administration of 131I-labelled human cationic trypsin. Scand. J. Gastroenterol. Suppl. 1986, 126, 21–24.
[98] Moreno, F. J., Rubio, L. A., Olano, A., Clemente, A., Uptake of 2S albumin allergens, Ber e 1 and Ses i 1, across human intestinal epithelial Caco-2 cell monolayers. J. Agric. Food Chem. 2006, 54, 8631–8639.
[99] Moreno, F. J., Gastrointestinal digestion of food allergens: effect on their allergenicity. Biomed. Pharmacother. 2007, 61, 50–60.
[100] Heyman, M., Desjeux, J. F., Significance of intestinal food protein transport. J. Pediatr. Gastroenterol. Nutr. 1992, 15, 48–57.

Chapter 3

Purification and characterization of *Anacardium occidentale* (cashew) allergens Ana o 1, Ana o 2, and Ana o 3

Marit Reitsma Shanna Bastiaan-Net, Stefano Sforza, Johanna P. M. van der Valk, Roy van Gerth van Wijk, Huub F. J. Savelkoul, Nicolette W. de Jong, and Harry J. Wichers



J. Agric. Food Chem. 2016, 64, 1191−1201

Abstract

In this study a fast and simple purification procedure for the three known allergens from cashew (7S globulin Ana o 1, 11S globulin Ana o 2, and 2S albumin Ana o 3) is described. The purified allergens are characterized by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE), Western blot, glycoprotein stain, and protein identification. The purified proteins still bind IgE, and this IgE binding varied between different pools of patient serum. Ana o 1 was found to be a glycoprotein. Ana o 3 has been studied more in detail to identify both the small and large subunits, both displaying micro-heterogeneity, and epitope mapping of Ana o 3 has been performed.

Introduction

The prevalence of allergy to tree nuts in Europe is 0.5% (based on oral food challenges) to 1.3% (based on self-reported allergy) [1]. Within this group, the percentage of cashew (*Anacardium occidentale*) allergy is between 7% [2] and 55% [3]. Individuals suffering from cashew allergy can experience cutaneous, respiratory, and gastrointestinal symptoms and, in the most severe case, anaphylactic shock [4–6]. Administration of adrenaline is required more often [7], and the risk of developing an anaphylactic reaction seems higher than in peanut allergic patients [4].

Currently, three major cashew allergens are known: Ana o 1, Ana o 2, and Ana o 3 [8]. Ana o 1 is a 7S vicilin seed storage protein of 50kDa that is expected to form a trimer in the native state [9-10]. The 11S globulin in cashew, Ana o 2, consists of a 33kDa large subunit and a 20kDa small subunit, bound together by disulfide bonds, forming a 53kDa protein [11-12]. In native conditions 11S globulins form hexameric structures of ~360kDa [10]. In literature Ana o 2 is also referred to as anacardein and cashew major protein [13-14]. The third major cashew allergen is Ana o 3, a 2S albumin of 12.6kDa [15]. On the basis of similarity to other 2S albumins, this protein is most probably proteolytically cleaved into an N-terminal small subunit and a C-terminal large subunit that remain associated by four disulfide bonds [10, 16]. Robotham *et al*. reported denatured and reduced Ana o 3 to contain three isoforms of 6, 8, and 10kDa. Each isoform had an N-terminal sequence that corresponded to the large subunit, but no small subunit was described [15].

Reports on isolation and characterization of the cashew allergens are scarce. Ana o 3 has been purified before [17, 18], but neither purity nor yield were mentioned. Occasionally, references to isolated forms of Ana o 2 (anacardein) are made [19, 20]; however, as already indicated by the authors themselves, these fractions are not “immunochemically pure” [20]. When this protocol was repeated in 2010, a purity of >90% was stated, but no protein characterization data like SDS-PAGE or Western blot were shown [13].

In this article we describe a fast and simple method to obtain Ana o 1, Ana o 2, and Ana o 3 with high purity. While purification of Ana o 1 has never been described, the purification of Ana o 2 has never been fully described or characterized. We describe a simple method that allows simultaneous purification of Ana o 1 and Ana o 3. Biochemical and immunochemical analysis of the isolated proteins revealed that the proteins retain their IgE-binding capacity and showed Ana o 1, but not Ana o 2 or Ana o 3, to be a glycoprotein. Ana o 3 has been examined further to identify the large subunit, as well as the so-far unidentified small subunit. Both the small and large subunits display micro-heterogeneity.

Materials and methods

Materials and reagents

Blanched cashews (*Anacardium occidentale*) from Brazil were kindly provided by Intersnack (Doetinchem, The Netherlands). Acetone, ammonium sulfate, and glycine were purchased from Merck (Darmstadt, Germany). Filter paper (595 1/2) was purchased from Whatman (Dassel, Germany). Ammonium bicarbonate, NaCl, NaH2PO4, β-mercaptoethanol, sodium dodecyl sulfate (SDS), Tween-20, bovine serum albumin (BSA), alkaline-phosphatase conjugated goat antirabbit antibody, Sigmafast 5-bromo-4- chloro-3-indolyl phosphate/nitro blue tetrazolium (BCIP/NBT), and dithiothreitol (DTT) were purchased from Sigma-Aldrich (St. Louis, MO). Amicon centrifugal filters were purchased from Merck Millipore (Tullagreen, Ireland). Minisart syringe filters (0.45µm) were purchased from Sartorius Stedim (Goettingen, Germany). Ethylenediaminetetraacetic acid (EDTA), Halt protease inhibitor cocktail (no. 78438), Pierce Coomassie (Bradford) Protein Assay Kit, and Pierce Glycoprotein Staining Kit were purchased from Thermo Scientific (Rockford, IL). NuPAGE lithium dodecyl sulfate (LDS) sample buffer, NuPAGE 1 mm 10% Bis-Tris mini gels, NuPAGE 2-(N-orpholino) ethanesulfonic acid (MES) SDS running buffer, and SimplyBlue SafeStain were purchased from Invitrogen (Carlsbad, CA). Protein molecular weight marker and nitrocellulose membrane were purchased from Bio-Rad (Hercules, CA). Tris was purchased from USB (Cleveland, OH). Methanol was purchased from Actu-All (Randmeer, The Netherlands). Rabbit antihuman IgE was purchased from Dako (Glostrup, Denmark). Eight serum samples were purchased from Plasmalab International (Everett, WA). Ana o 3 peptides for epitope mapping were purchased from Synpeptide (Shanghai, China). The polyester backbone was purchased from GL Precision (San Jose, CA). IgE was purchased from Enzo (Farmingdale, NY). Trypsin was purchased from Promega (Madison, WI).

Obtaining defatted cashew and extraction of total soluble protein

Blanched cashews were ground by mortar and pestle. Ground cashews were defatted twice with acetone, 1:5w/v for 2h at 4°C while stirring, followed by filtration using filter paper and drying for 2h (first defatting step) or overnight (second defatting step). Defatted cashew was suspended in an ammonium bicarbonate buffer, pH7.9 (0.1M ammonium bicarbonate, 0.5M NaCl), 1:5w/v and stirred overnight at 4°C. Subsequently, the samples were centrifuged for 90min at 4000g at 4°C, and the obtained supernatant was used for the purification of Ana o 1 and Ana o 3.

Protein precipitation

Proteins were precipitated in a stepwise fashion. In the first step, 15% w/v ammonium sulfate was added and mixed by stirring for 1h at 4°C; subsequently this was centrifuged for 1h at 4°C in a swinging bucket centrifuge. The process was repeated three more times to yield final ammonium sulfate concentrations of 30%, 45%, and 52.5%w/v. The final supernatant, obtained after the last step of precipitation (52.5%), was diluted four times in Milli-Q water (MQ).

Purification of Ana o 1 using ultrafiltration

Four times 4mL of the diluted final supernatant of ammonium sulfate precipitation was passed through a 30kDa Amicon centrifugal filter by centrifugation, followed by two times 5mL of MQ. Subsequently again four times 4mL of the diluted final supernatant was passed through the same filter. Five mL of MQ was added to the retentate; this retentate was passed through a new 30kDa Amicon centrifugal filter and purified in this filter by 10 times 5min shaking in 5mL of MQ followed by centrifugation. All centrifugation steps were performed for 5min at 4000g at 20°C. The final retentate containing Ana o 1 was suspended in a final volume of 3 mL of MQ, concentrated using a 0.5mL 3Da Amicon centrifugal filter, and stored at −20°C.

Purification of Ana o 3 using ultrafiltration

The filtrate from Ana o 1 purification was concentrated using a 50mL 3kDa Amicon centrifugal filter and stored at −20°C.

Purification of Ana o 2 using gel filtration chromatography

Ground blanched cashew was stirred in ammonium bicarbonate buffer (1:5w/v) overnight at 4°C and centrifuged for 30min at 4000g at 20°C. From the supernatant the top white layer and lipid layer were removed and the remaining supernatant was filtered (0.45µm). Of this filtered supernatant, 300µL was loaded onto a Superdex 200 column (10/300GL, bed volume 24mL) on an Akta purifier (both Amersham Pharmacia Biotech (Uppsala, Sweden)). The proteins were eluted from the column using 15mL elution buffer (50mM NaH_2PO_4, 0.15M NaCl, pH7) at 0.5mL/min; 0.5mL fractions were collected. The fractions were concentrated using a 0.5mL 3kDa Amicon centrifugal filter and stored at −20°C.

Purification of Ana o 3 in the presence of protease inhibitors

Ground blanched cashew was stirred in ammonium bicarbonate buffer (1:5w/v) with added protease inhibitor cocktail and 5mM EDTA for 1h at 4°C and centrifuged for 90min at 4000g

at 4°C. Protein precipitation and purification of Ana o 3 was performed as described above with the adaptation of performing all steps at 4°C.

Protein concentration

Protein concentration was determined using a Coomassie protein assay according to the manufacturer's instructions.

SDS-PAGE

On the basis of a Coomassie protein assay, 20µg of denatured and reduced samples (10% β-mercaptoethanol in NuPAGE LDS sample buffer, 5min, 100°C) were loaded onto NuPAGE 1mm 10% Bis-Tris mini gels, alongside a molecular weight marker (Precision Plus Protein Dual Xtra Standard). Electrophoresis was performed at 160V for 55min in NuPAGE MES SDS running buffer. For SDS-PAGE, the gels were stained with SimplyBlue SafeStain. Imaging and analysis were performed using a Universal Hood III and Image Lab 4.1 software (both Bio-Rad). Nonreducing denaturing SDS-PAGE was performed as described above, but during the sample preparation no β-mercaptoethanol was added.

Western blotting

Transfer of proteins from reducing denaturing SDS-PAGE gel to a nitrocellulose membrane (0.2µm) was performed using a Criterion blotter (Bio-Rad) in a cold tris-glycine buffer (25mM tris, 190mM glycine, 0.1% SDS, 20% methanol) for 36min at 70V. Membranes were washed with TBST (50mM tris, 150mM NaCl, 0.1% Tween-20, pH7.4) and blocked with 3% BSA in TBS (TBST without Tween-20) for 1h at 4°C. After washing with TBST and TBS, the blots were incubated overnight with a pool (to ascertain sufficient coverage of epitopes) of patient serum (group 1) or plasma (group 2) 1:5 diluted in TBS. Subsequently, the blots were washed with TBST and TBS and incubated with rabbit antihuman IgE diluted 1:15.000 in TBS. Again, the blots were washed with TBST and TBS, incubated with alkaline phosphatase conjugated goat anti-rabbit diluted 1:20.000 in TBS, followed by washing with TBST and TBS. Blots were stained with Sigmafast BCIP/NBT for 10min (group 1 and antibody control blot) or 2.5min (group 2). Imaging and analysis were performed using a Universal Hood III and Image Lab 4.1 software. The antibody control blot was performed as described earlier, but instead of serum/plasma the membrane was incubated overnight with TBS. Inhibition Western blotting was performed as described earlier with the exception that the plasma was incubated with either 1mg/mL purified Ana o 1 or Ana o 3 for 4h at room temperature.

Table 3.1 *Clinical characteristics of cashew and tree-nut allergic subjects: self-reported allergies and cashew IgE levels as measured by Immulite.* [a] *positive DBPCFC to cashew as described in van der Valk et al., submitted manuscript.* [b] *sample 1-4 as used in Ana o 3 epitope mapping experiments.* [c] *measured by ImmunoCAP by Plasmalab International.*

Patient # serum pool 1	**Self-reported allergies patients serum pool 1 (blot D1) Dutch children**	**IgE level cashew (kU/l)**	**Patient # plasma pool 2**	**Self-reported allergies patients plasma pool 2 (blot D2) American adults**	**IgE level cashew (kU/l)**
1	Cashew[a], pistachio	23.7	1[b]	Cashew, walnut, peanut, tree pollen, cat, dog, horse	100.0
2	Cashew[a]	21.8	2[b]	Tree nuts, banana, eggplant, dust	56.9
3	Cashew[a], pistachio	10.9	3[b]	Nuts, avocado, corn, raw vegetables, dog, horse, cow	100.0
4	Cashew[a], chicken egg	11.1	4[b]	Tree nuts, pecan, hazelnut, dust, grass, insect venom, alcohol	19.2
5	Cashew[a], kiwi, shrimp	13.0	5	Cashew, brazil nut, almond, hazelnut, peanut, dust	13.2
6	Cashew[a], peanut, hazelnut	38.6	6	Cashew, pistachio, pecan, walnut, almond, hazelnut, macadamia, peanut, coconut, cat, dog, dust	59.6
7	Cashew[a], peanut, hazelnut	27.0	7	Tree nuts, hazelnut, dust	35.5[c]
8	Cashew[a], pistachio, citrus fruit	44.7	8	Cashew, walnut, peanut, shellfish, fish, scotch bloom	26.2
9	Cashew[a]	47.4			
10	Cashew[a], kiwi	51.9			
11	Cashew[a], pistachio, walnut, almond	10.8			
12	Cashew[a], almond, chicken egg, milk	12.7			
13	Cashew[a], cows' milk, chicken egg, pear, peas	48.7			

Serum and plasma

Serum of Western blot group 1 was obtained from 13 cashew-allergic (double-blind placebo-controlled food challenge (DBPCFC) diagnosed) individuals out of the multicentre prospective study "Improvement of diagnostic methods for allergy assessment" with cashew allergy in children as a showcase (trial number NTR3572, medical ethical approval number 2012-125, Erasmus Medical Centre Rotterdam). Plasma of group 2 was obtained from PlasmaLab International. Eight patients with self-reported history of cashew or tree nut allergy were selected based on high cashew-IgE titres. Clinical characteristics of the subjects are described in Table 3.1: self-reported allergies and cashew-IgE levels as determined by

Immulite 2000, F202. Glycoprotein staining. Glycosylation of proteins was checked using a Glycoprotein Staining kit. Samples were run on SDS-PAGE as described above alongside a positive control (horseradish peroxidase, kit content) and a negative control (soybean trypsin inhibitor, kit content). Staining was performed according to the manufacturer's instructions. Imaging and analysis were performed using a Universal Hood III and Image Lab 4.1 software.

Protein identification

Protein identification by MALDIMS/MS was performed by Alphalyse Denmark. Protein samples for protein identification were excised from denaturing reducing Coomassie-stained SDS-PAGE gels. Trypsin-digested protein bands were analysed on a Bruker Autoflex Speed MALDI TOF/TOF. The MS data were blasted against the NCBI and UniProt database.

Additional protein identification analysis of the 10kDa protein band of Ana o 3 was performed by LC-MS/MS analysis. On the basis of the results of Shevchenko *et al*. [21], bands cut from SDS-PAGE gel were washed in 0.2M NH4HCO3 in acetonitrile (1:1) followed by addition of 10mM DTT in 0.1M NH4HCO3 at 56°C for 45min. The gel pieces were incubated in 55mM iodoacetamide (IAA) in 0.1M NH4HCO3 and washed in 0.1M NH4HCO3 in acetonitrile (1:1), and after drying in a SpeedVac, the gel pieces were digested overnight at 37°C in 0.1M NH4HCO3 with sequencing grade modified porcine trypsin (Promega). Gel pieces were incubated with 25mM NH4HCO3 for 15min at 37°C. After addition of acetonitrile the samples were dried and dissolved in 0.1% formic acid. Trypsin-digested peptides were separated on a Thermo Scientific EASY column (3µm bead-packed 10cm C18 column) connected to an Easy-nLC 1000 ultrahigh-pressure system (Thermo Scientific). Peptides were loaded in solvent A (0.1% formic acid in MQ) onto a pre-column (5µm bead-packed 2cm C18 column) prior to separation on the analytical column using a 40min 5–50% linear gradient of solvent B (100% acetonitrile (ACN) in 0.1% formic acid) at a flow rate of 300nL/min. During the gradient, online MS analysis of peptides was performed with a QExactive Plus (Thermo Scientific) mass spectrometer using a nanoelectrospray source. Ionization (2.4kV) was performed using a stainless steel emitter and a heated capillary temperature of 250°C. Full MS scans were acquired over the m/z range 400–1.500 with a mass resolution of 70.000 (at m/z 200). Full scan target was set at 1 × 106 with a maximum fill time of 100ms. The five most intense peaks with charge states 1–4 were fragmented in the higher-energy collisional dissociation (HCD) collision cell with a normalized collision energy of 30%. The mass range was set to 200–2000 with a mass resolution of 17.500 (at m/z 200). The target value for fragment scans was set at 1 × 105, the intensity threshold was kept at 4 × 104, and the maximum allowed accumulation times were 50ms. The peptide match was set to preferred, isolation width was set at 4 and isotope exclusion was on, and the dynamic exclusion was set to 30s. LC-MS/MS data acquired by the Q-Exactive were processed using ProteomeDiscoverer software 1.4 (Thermo Scientific). The fragmentation

spectra were searched against an Anacardiaceae database downloaded from Uniprot using Sequest HT with precursor mass tolerance of 10ppm and fragment mass tolerance of 20mDa. Minimum peptide length was 6 amino acids, and maximum of miscleavages was set to 2. Carbamidomethylation of cysteine was set as fixed modification, and oxidation of methionine was set as variable modification. The percolator was used for false-discovery rate (FDR) estimation using reversed decoy database filtering on q-value below 1% for strict and 5% for relaxed stringency.

Analysis of Ana o 3 by UPLC/PDA/ESI-MS

Identification of the heterogeneity of the N- and C-termini was executed by ultra-performance liquid chromatography UPLC)/photodiode array (PDA)/electrospray ion source (ESI)-MS. Purified native Ana o 3, denatured and reduced Ana o 3 (10mM DTT, 1% SDS, 5min at 100°C), and proteins eluted from de-stained SDS-PAGE (50mM Tris, 0.1mM EDTA, 0.15M NaCl, 0.1% SDS, 1mM DTT, 4h) were filtered over a 0.5 mL 3kDa Amicon centrifugal filter and suspended in 200µL of eluent A (0.1% trifluoroacetyl (TFA) and 1% ACN in water). The solution was analyzed by LC/PDA/ESI-MS by injecting 2µL in an Acquity UPLC separation module equipped with a BEH C4 300Å column (1.7µm, 2.1 × 100mm), interfaced with an Acquity PDA detector and a Synapt G2-Si HD mass spectrometer. Gradient elution was performed between eluent A and B. Eluent A was 1% ACN and 0.1% TFA in water; eluent B was 0.1% TFA in ACN. The gradient ramp was 0–2min isocratic 90% A, 2–12min linear gradient from 90% A to 25% A, 12–15min linear gradient from 25% A to 0.1% A, 15–20min isocratic 0.1% A, plus reconditioning. The flow rate was 0.35 mL/min. PDA detection was performed in the range 200–400nm, with a resolution of 1.2nm and a sampling rate of 40 points/s. MS detection was performed in the full scan mode with the following parameters: positive ion mode, capillary voltage 3kV, source temperature 150°C, cone voltage 40V, desolvation temperature 500°C, cone gas flow (N2) 200L/h, desolvation gas flow (N2) 800L/h, aquisition range 150–4000m/z, scan time 0.3s, and interscan time 0.015s. The UV chromatogram was obtained by extracting the trace measured at 214nm from the PDA full scan trace.

Analysis of Ana o 3 tryptic peptides by UPLC/PDA/ESI-MS

Purified Ana o 3 was denatured and reduced (1% SDS, 10mM DTT, 5min at 100°C) and digested with trypsin for 1h at 37°C. Samples were analysed by LC/PDA/ESI-MS by injecting 2µL in an Acquity UPLC separation module equipped with a Peptide BEH C18 column, 300Å (1.7µm, 2.1mm × 150), interfaced with an Acquity PDA detector and a Synapt G2-Si HD mass spectrometer. Gradient elution was performed between eluent A and B. Eluent A was 1% ACN and 0.1% TFA in water; eluent B was 0.1% TFA in ACN. The gradient ramp was 0–2min

isocratic 97% A, 2–10min linear gradient from 97% A to 78% A, 12–16min linear gradient from 78% A to 70% A, 16–19 min from 70% A to 0.1% A, 19–24min isocratic 0.1% A, plus reconditioning. The flow rate was 0.35mL/min. PDA detection was performed in the range 200–400nm, with a resolution of 1.2nm and a sampling rate of 40 points/s. MS detection was performed in the full scan mode with the following parameters: positive ion mode, capillary voltage 3kV, source temperature 150°C, cone voltage 30V, desolvation temperature 500°C, cone gas flow (N2) 200L/h, desolvation gas flow (N2) 800L/h, acquisition range 100–3000m/z, scan time 0.3s, and interscan time 0.015s. Survey MSe mode was applied with the following parameters: acquisition range 100–3000m/z, ramp high energy from 20.0 to 30.0, TIC threshold 5.0, survey scan time 0.3s. Lock spray mass correction was applied by continuously perfusing Angiotensin II as calibrant into the source during the scan. The data were analysed using BioPharmaLynx 1.3.3 software. Identified peptides were cross-checked manually.

Epitope mapping of Ana o 3

Forty-two peptides of 15 amino acids, with an offset of 3 amino acids, were synthesized by Synpeptide, covering the entire length of Ana o 3 (amino acids 1–138). Peptides were dissolved, and 5µmol was spotted in duplicate on nitrocellulose (0.2µm, placed on a polyester backbone) alongside two positive controls (0.6µg total protein extract and 5µg IgE) and a negative solvent control. The same procedure was followed as explained above for Western blotting with the only difference being that the blots were incubated with plasma from individual patients (group 2) and one blot was incubated with serum from a non-allergic person. The four plasma samples were chosen based on the highest Ana o 3 binding capacity out of the eight plasma samples, as determined by Western blotting (data not shown), for clinical characteristics see Table 3.1. Staining was performed for 30min. After washing in TBS, the average pixel intensity of each duplicate spot was calculated using the Image Lab 4.1 software, and the values for the negative serum were deducted per duplicate spot. Strong binding epitopes were defined as the 30% strongest IgE-binding epitopes.

Results

Protein purification

Purification of Ana o 1 and Ana o 3 was achieved by protein extraction from defatted cashew followed by stepwise ammonium sulfate precipitation. The final supernatant, after 52.5% ammonium sulfate precipitation, was used to purify Ana o 1 and Ana o 3. Ana o 2 was purified by gel filtration chromatography of a non-defatted cashew extract. A flow sheet for the purification of Ana o 1, Ana o 2, and Ana o 3 is shown in Figure 3.1.

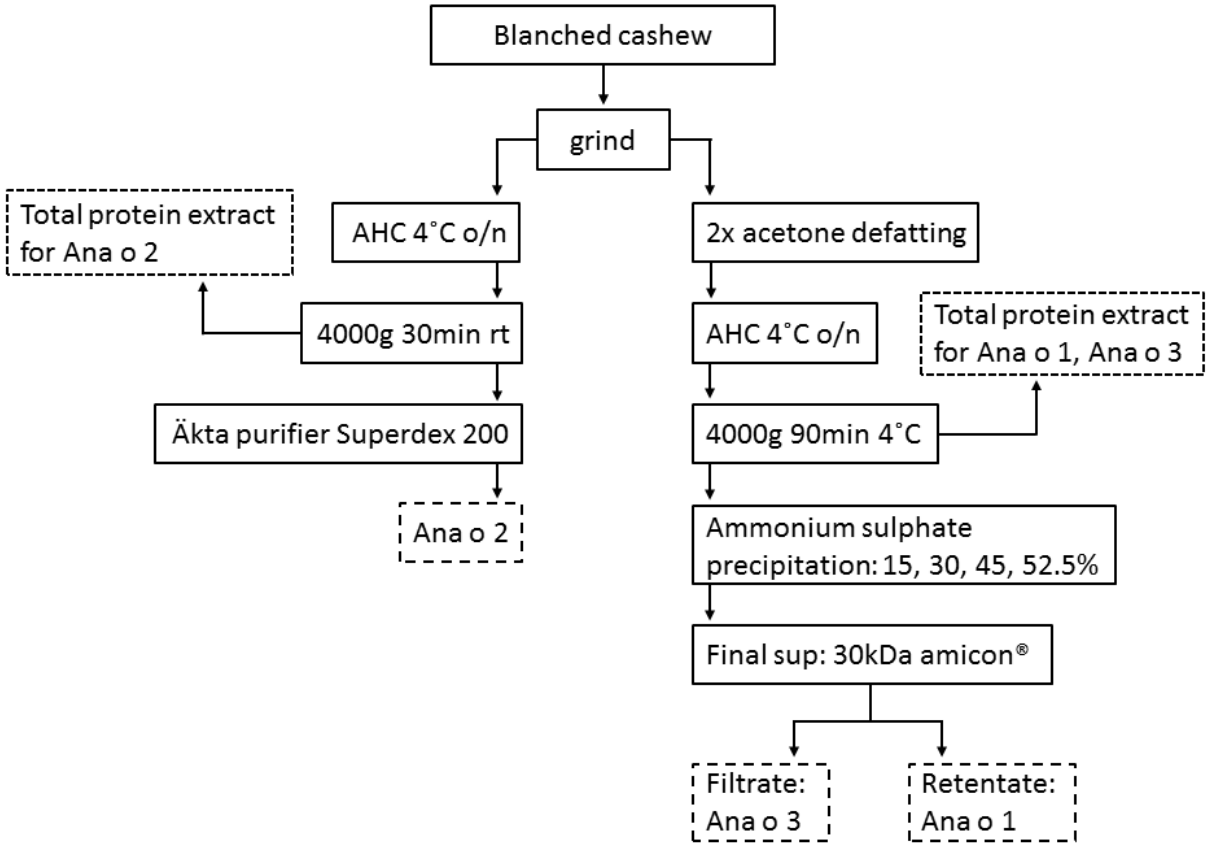


Figure 3.1 *Flow sheet for the purification of Ana o 1, Ana o 2, and Ana o 3. AHC, ammonium bicarbonate buffer; o/n, overnight 16h; rt, room temperature; final sup, final supernatant.*

In Figure 3.2A, various fractions collected during the purification process of cashew allergens are shown on denaturing and reducing SDS-PAGE gel. The two total protein extracts (defatted and non-defatted) contain several proteins next to Ana o 1, Ana o 2, and Ana o 3, with the 11S globulin Ana o 2 as the most prominent protein. Upon stepwise addition of ammonium sulfate to the total protein extract (defatted), to a final concentration of 15, 30, and 45%, generally Ana o 2 together with various high molecular weight proteins precipitated; see Supplementary Information, Figure 3.S1. After the last step of precipitation (52.5%), the final supernatant contains only Ana o 1 and Ana o 3 (Figure 3.2A, lane 3); this last fraction was used to purify these two proteins. Omitting the stepwise precipitation steps by directly adding 52.5% ammonium sulfate resulted in a supernatant containing more impurities (data not shown).

The gel filtration chromatogram from the purification of Ana o 2 showed three peaks (Figure 3.3): peak 1 contains a mixture of several proteins (among others, Ana o 1, Ana o 2, and Ana o 3), peak 2 contains Ana o 1 and Ana o 2, and peak 3 contains the purified Ana o 2 fraction as shown in Figure 3.2A (lane 6).

Protein identification

As specified in Figure 3.2A and Figure 3.S2, on the basis of MALDI-MS/MS, Ana o 1 was confirmed in the 100kDa band (b) and the 50kDa band (c). For Ana o 2 the 53kDa (a), 30kDa (d), and 21kDa bands (f) were confirmed as the complete protein, large subunit, and small subunit, respectively. The 22kDa band (e) was identified as an 11S globulin from pistachio; however, we interpret this to be Ana o 2 from cashew, as 11S from pistachio and Ana o 2 display 49% sequence identity and 65% sequence similarity (EMBOSS Needle, EMBL-EBI). The three bands in lane 7 (10, 8, 6kDa; bands g, h, i, respectively) were identified as Ana o 3.

Purity and yield of the proteins

Purities of the Ana o 1, Ana o 2, and Ana o 3 fractions were determined by applying increased protein concentrations (20, 30, and 40μg) on SDSPAGE (data not shown). Purity of these fractions was calculated as the relative percentage of the desired protein bands versus visible impurities on SDS-PAGE. Purity of Ana o 1 was 95.6% ± 3.3%, purity of Ana o 2 was 92.6% ± 4.6%, and purity of Ana o 3 was 98.5% ± 1.3%. Protein yield of the purified fractions was calculated based on the Coomassie protein assay and SDS-PAGE analysis of several purified fractions. Overall yield is 1% for Ana o 1 (0.33mg per run), 34% for Ana o 2 (1.99mg per run), and 3% for Ana o 3 (5.52 mg per run).

Western blot

In Figure 3.2D, two Western blots performed with two different pools of patient serum/plasma are shown; blot D1 was performed using 13 sera from cashew-allergic children as determined by DBPCFC, while blot D2 was performed using 8 plasma samples from cashew- or tree-nut allergic adults. Less strong IgE-binding to cashew allergens is observed using serum from group 1 (blot D1), compared to plasma from group 2 (blot D2). Blot D1 shows binding of IgE to Ana o 2 (lane 4; 30, 22, and 21kDa), Ana o 3 (lane 5; 10 and 8kDa, not 6kDa), and slightly to Ana o 1 (lane 3; 50kDa). In the total protein extract, the same proteins as well as additional proteins of 80, 70, 53 (full length Ana o 2), 40, and 16kDa appear IgE-reactive. In blot D2 binding of IgE to Ana o 1 (lane 3; 100 and 50kDa), Ana o 2 (lane 4; 30, 22, and 21kDa), and Ana o 3 (lane 5; 10 and 8kDa, not 6kDa) is visible. In the total protein extract, the same proteins as well as additional proteins of 80, 65, 53kDa (Ana o 2), 40, 16, and 12kDa bind IgE. In blot D2, lane 3, a small impurity of 30kDa (Ana o 2) can be observed in the purified Ana o 1 extract. In the third blot, D3, the antibody control blot is shown, demonstrating marginal non-specific binding of the antibodies to the 22–21kDa subunits of Ana o 2. These Western blots indicate that all three purified protein fractions are able to bind IgE.

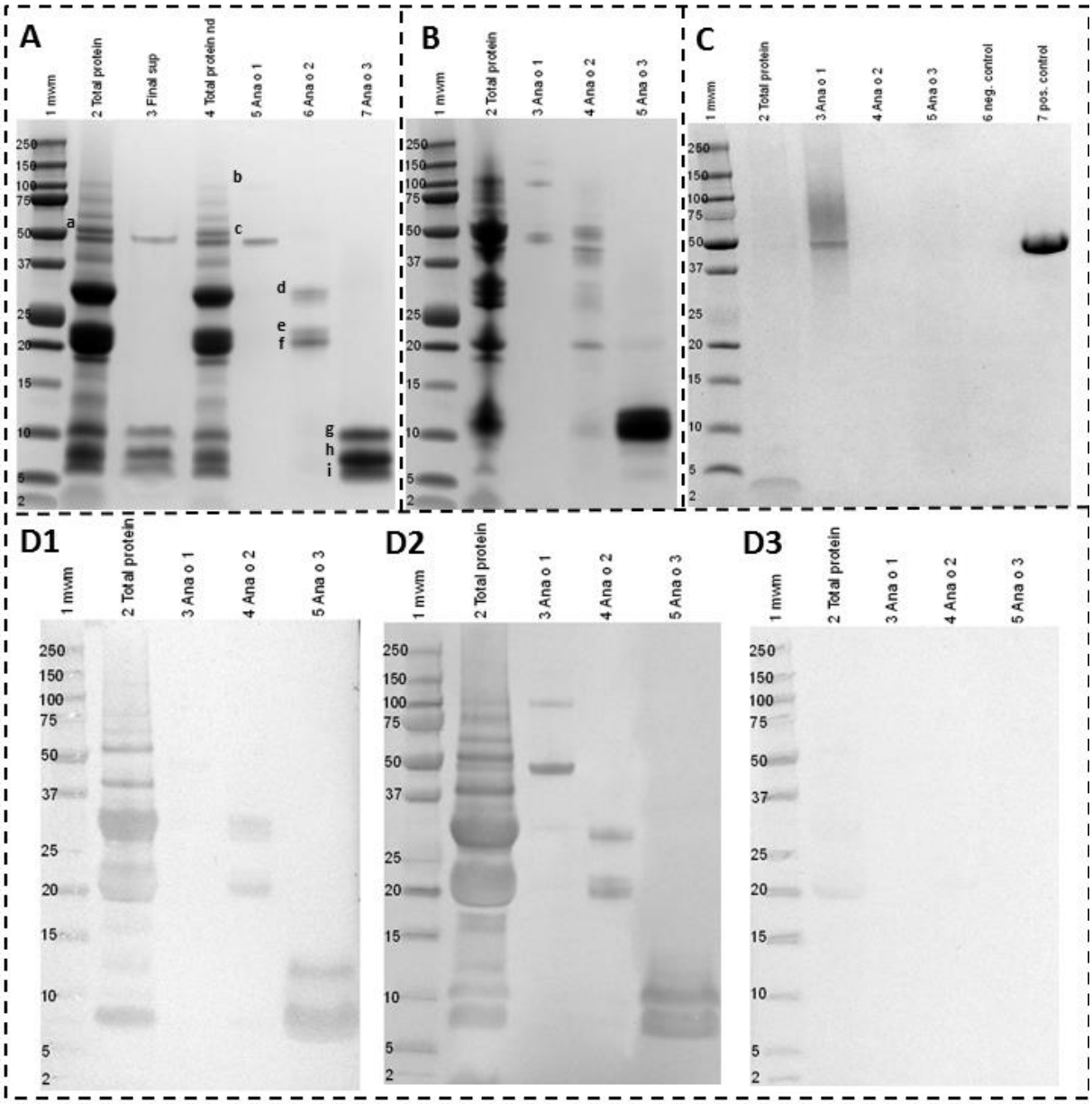


Figure 3.2 *All lanes: mwm, molecular weight marker; Total protein, total protein extract from defatted cashew used for purification of Ana o 1 and Ana o 3; Final sup, final supernatant after 52.5% ammonium sulfate precipitation; Total protein nd, total protein extract from non-defatted cashew used for purification of Ana o 2; neg. control, negative control glycoprotein stain (soybean trypsin inhibitor); pos. control, positive control glycoprotein stain (horseradish peroxidase). (A) Reducing denaturing SDS-PAGE of purified cashew extracts. Letters a–i indicate bands cut for protein identification: (a) 53kDa Ana o 2, (b) 100kDa Ana o 1, (c) 50kDa Ana o 1, (d) 30kDa Ana o 2, (e) 22kDa 11S globulin, (f) 21kDa Ana o 2, (g) 10kDa Ana o 3, (h) 8kDa Ana o 3, (i) 6kDa Ana o 3. (B) Non-reducing denaturing SDS-PAGE of purified cashew extracts, no β-mercaptoethanol added. (C) Glycoprotein stain of purified cashew extracts. (D) Western blots of purified cashew extracts. Blot D1 has been incubated with the serum pool of group 1, blot D2 has been incubated with a plasma pool from group 2, and blot D3 is the antibody control blot that has not been incubated with patient plasma or serum.*

Inhibition blotting, presented in Figure 3.S3, shows that purified Ana o 1 completely inhibited binding to Ana o 1, and Ana o 3 completely inhibited binding to Ana o 3 in the total protein extract. Besides inhibition of binding to Ana o 1, purified Ana o 1 also diminished IgE binding to Ana o 2 and Ana o 3.

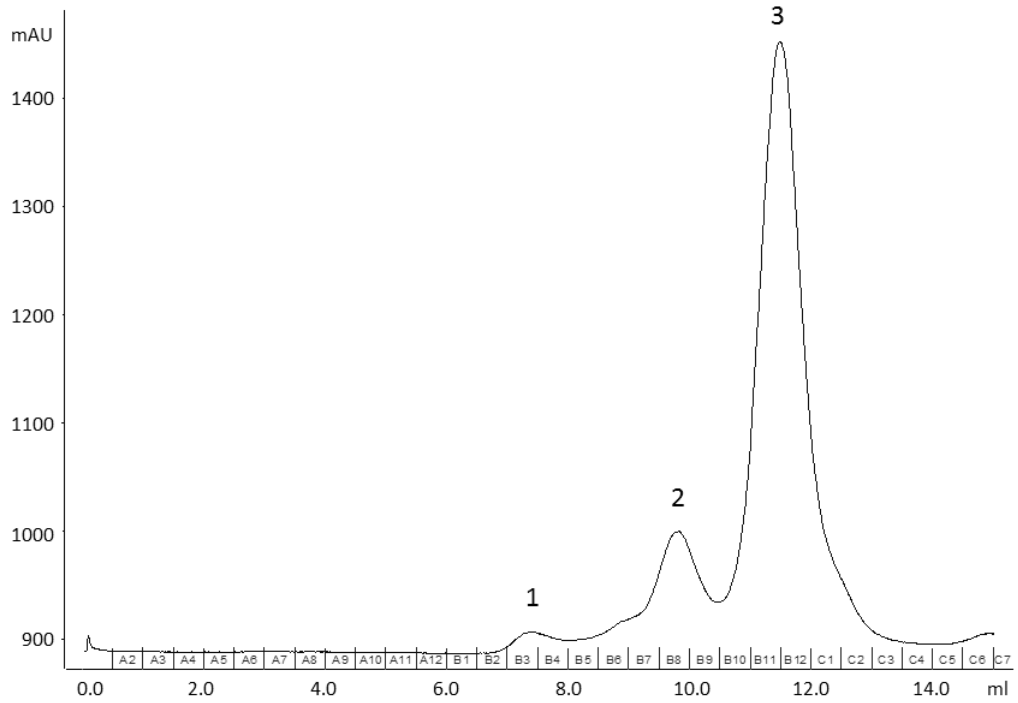


Figure 3.3 *Gel filtration chromatogram of total protein extract of non-defatted cashew used for the purification of Ana o 2.*

Non-reducing denaturing SDS-PAGE

To check the presence of disulfide bridges in the purified allergen fractions, a non-reducing denaturing SDS-PAGE, in the absence of β-mercaptoethanol, was performed. In the Ana o 1 fraction (Figure 3.2B, lane 3) protein bands of 160, 100, and 50kDa can be observed. In lane 4, containing Ana o 2, diffuse bands of 53, 42, 39, 30, 20, and 10kDa are present. Ana o 3 (lane 5) shows an abundant protein band of 12kDa and minor bands of 20 and 5kDa.

Glycosylation of cashew proteins

On the basis of prediction models (NetNGlyc 1.0, Denmark, http://www.cbs.dtu.dk/services/NetNGlyc/), the amino acid sequences of Ana o 1 (amino acid 391, asparagine) and Ana o 3 (amino acid 19, asparagine), but not that of Ana o 2, contain putative N-glycosylation motifs. The predicted glycosylation of Ana o 1 was confirmed by glycoprotein staining (see Figure 3.2C). Both Ana o 2 and Ana o 3 do not seem to be glycosylated, neither in the purified nor in the total protein fraction. In addition an unknown protein of 3kDa, present in the full protein extract, also seems to be glycosylated.

Analysis of Ana o 3

Protein characterization and protein identification of Ana o 3 revealed some discrepancies with data from literature. It was previously reported that all three denatured protein bands from SDS-PAGE represented the large subunit of Ana o 3.15 However, our results identify one of the bands as the small subunit from Ana o 3. Besides, the IgE binding on Western blot to two out of three bands of denatured Ana o 3 indicated the possibility of one of these bands possibly being the small subunit of Ana o 3. To analyse this possibility, both denatured reduced and native Ana o 3 samples were analysed for peptide mass, purity, and annotation of protein sequences using UPLC/PDA/ESI-MS.

Table 3.2 *Molecular masses of denatured Ana o 3 derived from UPLC/PDA/ESI-MS analysis and their tentative sequence identification.* [a]*According to Moreno and Clemente [16],* [b]*According to our data.*

	Tentative sequence annotation within the range 0.5 Da from the theoretical calculated mass	**Experimental molecular mass (Da)**	**Leader peptide** [b]	**Small subunit** [b]	**Large subunit** [b]
Leader peptide	1-20[a] 1-20/28[b]				
Small subunit	21-65[a] 29-65[b]				
Large subunit	66-138[a] 69-137[b]				
Chromatographic Peak 1 (Rt = 5.61)	34-62 / 35-63	3742.7		X	
	34-63 / 35-64 / 36-65	3799.7		X	
	33-63	3886.8		X	
	33-65	4099.9		X	
Chromatographic Peak 2 (Rt = 5.82)	37-64	3514.6		X	
	37-65	3670.7		X	
	22-53	3959.8		X	
	39-69	4016.8		X	X
	33-66	4263.1		X	
	31-66	4507.1		X	
Chromatographic Peak 3 (Rt = 7.60)	2-70 or 56-122 or unknown isoform	8161.7	X	X	X
	1-71 or unknown isoform	8421.8	X	X	X
Chromatographic Peak 4 (Rt= 7.91)	69-136	8084.7			X
	69-137	8171.7			X

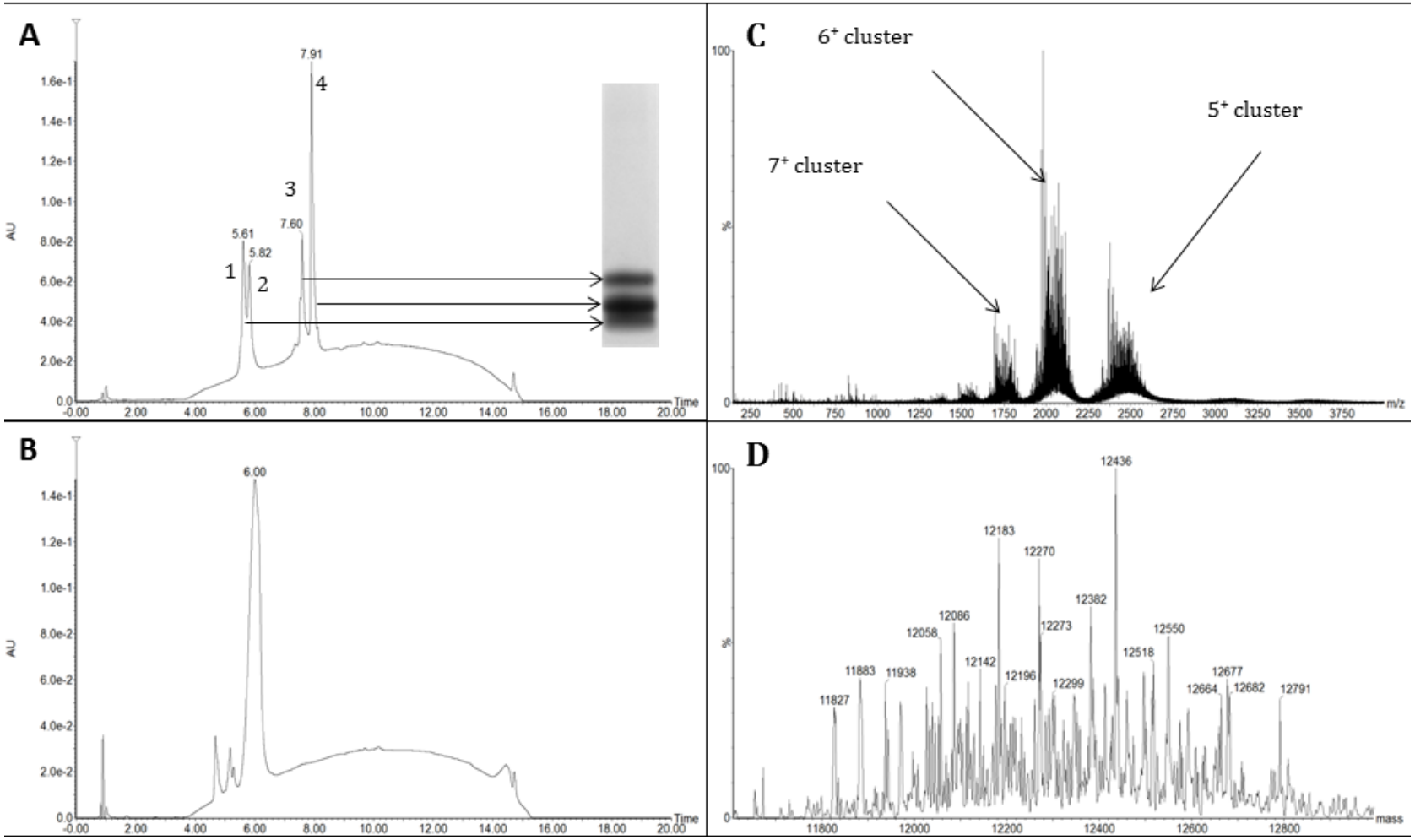


Figure 3.4 *Chromatographic analysis (UV trace at 214nm) of denatured (A) and native (B) Ana o 3. MS spectrum associated with the major peak in native Ana o 3 (full spectrum from 200 to 4000m/z) showing the clusters corresponding to the 7+-charged, 6+-charged, and 5+-charged ions (C), and charge-state deconvolution of this spectrum shown in (C), performed by the MaxEnt software, indicating the actual molecular masses of the main isoforms present (D).*

Denatured and reduced Ana o 3 shows four peaks (Figure 3.4A). Peaks 1 and 2, corresponding to the 6kDa protein band on SDS-PAGE, were compatible with the small subunit of Ana o 3 with molecular weights of 3.7–4.5kDa. Peak 3 (10kDa band) and peak 4 (8kDa band) were both compatible with the large subunit of Ana o 3 with molecular weights of 8.1–8.4kDa. Using both trypsin-digested and nondigested denatured reduced Ana o 3, the peptide sequences from the four peaks were annotated (see Table 3.2). The range of protein molecular weights in the four peaks was found to originate from N- and C-terminal micro-heterogeneity, also called N- or C-terminal clipping (see Table 3.3 and Figure 3.5). Tryptic peptides clearly showed that the small subunit displayed both N- and C-terminal micro-heterogeneity, with its N-terminus starting at amino acid 29–36 and its C-terminus terminating at amino acid 63–65. The large subunit displayed only minimal C-terminal micro-heterogeneity, starting at amino acid 69 and ending at amino acid 136–137. When Ana o 3 was purified using protease inhibitors, the same degree of micro-heterogeneity was obtained (data not shown).

Native Ana o 3 eluted as one major peak of 11.8–12.8kDa (Figure 3.4C and D) with a purity of 88% (Figure 3.4B, purity calculated by UV trace at 214nm). The observed masses fitted with the molecular masses of the small subunit (peaks 1 and 2) linked to the large subunit (peak 4) by two disulfide bridges. Together with the two other predicted disulfide bridges within the large subunit,16 a total of 4 disulfide bridges can be found within native Ana o 3. Besides demonstrating micro-heterogeneity in Ana o 3, the obtained data also indicate the presence of isoforms of the large subunit.

Table 3.3 *Sequence annotation of Ana o 3 tryptic peptides including termini based on molecular mass (<10 ppm error) and fragment identification.*

Peptide m/z (charge state)	Identified peptide sequence	
565.9 (3+)	36-48	*N-termini small subunit* until trypsin cleavage site (amino acid 48)
618.0 (3+)	35-48	
637.0 (3+)	34-48	
666.0 (3+)	33-48	
704.3 (3+)	32-48	
747.3 (3+)	31-48	
790.3 (3+)	30-48	
823.4 (3+)	29-48	
660.4 (2+)	55-65	Trypsin cleavage site (amino acid 55) until
582.3 (2+)	55-64	*C-termini small subunit*
553.8 (2+)	55-63	
759.7 (3+)	69-86	*N-terminus large subunit* until trypsin cleavage site (amino acid 86)
786.4 (2+)	123-137	Trypsin cleavage site (amino acid 123) until
742.8 (2+)	123-136	*C-termini large subunit*

MAKFLLLLSAFAVLLLVANASIYRAIVEVEEDSGREQSCQRQ
FEEQQRFRNCQRYVKQEVQRGGRYNQRQESLRECCQELQ
EVDRRCRCQNLEQMVRQLQQQEQIKGEEVRELYETASELP
RICSISPSQGCQFQSSY

Figure 3.5 *Amino acid sequence (coding region) of Ana o 3, with the small and large subunits indicated as based on UPLC/PDA/ESI-MS analysis. The leader peptide is underlined, the black bars indicate the constant regions, and the light gray bars indicate regions of micro-heterogeneity. A dotted box is placed around the small subunit (first box) and the large subunit (second box) of Ana o 3.*

Epitope mapping of Ana o 3

Using 42 overlapping linear peptides, 12 IgE-binding peptides were detected in Ana o 3 (see Table 3.4). Two IgE-binding peptide regions were identified, with the first consisting of amino acids 10–30 and the second, more strongly IgE-bound region consisting of amino acids 85–108. Besides 4 mildly binding epitopes described in Table 3.4, 2 strongly binding peptides were identified at amino acids 55–69 and 121–135. All strongly binding epitopes were detected in the peptides of the large subunit or containing part of the large subunit.

Table 3.4 *Linear epitope mapping of Ana o 3 using plasma of four individuals (no. 1–4) from group 2; bold peptide numbers indicate a strong IgE-binding epitope, bold amino acid numbers indicate an IgE-binding peptide region; underlined amino acid sequences indicate the small subunit of Ana o 3 (amino acids 29–65), between the leader sequence (1–20) and the large subunit (69–137) of Ana o 3.*

			IgE-binding intensity			
Peptide	**Amino acid #**	**Amino acid sequence**	**1**	**2**	**3**	**4**
1	1-15	MAKFLLLLSAFAVLL				
2	4-18	FLLLLSAFAVLLLVA				
3	7-21	LLSAFAVLLLVANAS				
4	**10-24**	AFAVLLLVANASIYR	+	+	+	++
5	**13-27**	VLLLVANASIYRAIV	+	+		+
6	**16-30**	LVANASIYRAIVEVE	+	+		+
7	19-33	NASIYRAIVEVEEDS				
8	22-36	IYRAIVEVEEDSGRE				
9	25-39	AIVEVEEDSGREQSC				
10	28-42	EVEEDSGREQSCQRQ				
11	31-45	EDSGREQSCQRQFEE				
12	34-48	GREQSCQRQFEEQQR				
13	37-51	QSCQRQFEEQQRFRN				
14	40-54	QRQFEEQQRFRNCQR	++			
15	43-57	FEEQQRFRNCQRYVK				
16	46-60	QQRFRNCQRYVKQEV				
17	49-63	FRNCQRYVKQEVQRG	+	++		
18	52-66	CQRYVKQEVQRGGRY				
19	55-69	YVKQEVQRGGRYNQR	++	+++		+
20	58-72	QEVQRGGRYNQRQES				
21	61-75	QRGGRYNQRQESLRE				
22	64-78	GRYNQRQESLRECCQ				
23	67-81	NQRQESLRECCQELQ				
24	70-84	QESLRECCQELQEVD				
25	73-87	LRECCQELQEVDRRC				
26	76-90	CCQELQEVDRRCRCQ	+		++	+
27	79-93	ELQEVDRRCRCQNLE				
28	82-96	EVDRRCRCQNLEQMV				
29	**85-99**	RRCRCQNLEQMVRQL	++	+++	+	+
30	**88-102**	RCQNLEQMVRQLQQQ	++	++	++	++
31	**91-105**	NLEQMVRQLQQQEQI				
32	**94-108**	QMVRQLQQQEQIKGE	++	++	++	+++
33	97-111	RQLQQQEQIKGEEVR				
34	100-114	QQQEQIKGEEVRELY				
35	103-117	EQIKGEEVRELYETA				
36	106-120	KGEEVRELYETASEL				
37	109-123	EVRELYETASELPRI				
38	112-126	ELYETASELPRICSI	+		+	
39	115-129	ETASELPRICSISPS				
40	118-132	SELPRICSISPSQGC				
41	121-135	PRICSISPSQGCQFQ	+++	+++	+	+
42	124-138	CSISPSQGCQFQSSY				
Total cashew positive control			+++	+++	+++	+++

Discussion

In this study we have purified and characterized the three major cashew allergens Ana o 1, Ana o 2, and Ana o 3. Our method enables the simultaneous purification of Ana o 1 and Ana o 3.

Using precipitation, ultrafiltration, and gel filtration chromatography, we have isolated three cashew allergens with high purity. The purities of the single allergen fractions are approximately 96% for Ana o 1, 93% for Ana o 2, and 99% for Ana o 3 as based on SDS-PAGE and 88% for Ana o 3 based on high-performance liquid chromatography (HPLC) analysis. The only IgE-reactive impurity we observed, as established by Western blotting, was a protein of 30kDa in the Ana o 1 fraction (Figure 3.2, D1). A substantial difference in IgE-binding to Ana o 1 is apparent between the two patient groups (blots D1 and D2), which might be explained by a population difference of the sera. Group 1 (blot D1) consists of Dutch children, and group 2 (blot D2) consists of American adults. Geographical and age differences have been described before to explain differences in allergen binding [22, 23].

For Ana o 3 we observed only IgE binding to the 10 and 8kDa proteins and not to the 6kDa protein. This is consistent with other Western blots performed by us using both various patient pools and individual patient sera. Teuber *et al*., however, did detect IgE binding to the 6kDa protein [20]. Our analyses show that the 6kDa protein is the small subunit of Ana o 3 while both the 10 and 8kDa proteins represent the large subunit. Epitope mapping of Ana o 3 by Robotham *et al*. [15] showed more strongly binding epitopes on the small subunit compared to our data. Table 3.4 shows that our patient IgE binds strongest to the peptides of the large subunit. One mildly IgE-binding peptide region was identified between the presumptive leader sequence15 and the start of the small subunit as determined by UPLC/PDA/ESI-MS analysis. It is unknown if this peptide is present in cashew when eaten; however, IgE binding to leader peptides that are not attached to the mature allergen has been observed before for Ana o [19] and Ana o 2 [11].

On Western blot, besides binding of IgE to Ana o 1, Ana o 2, and Ana o 3, several other proteins were bound in the total protein extract (Figure 3.2, D1 and D2). These allergens of 12, 16, 40, 65, 70, and 80kDa have been observed before [20, 24], but have not been identified.

The yield of purification for Ana o 1 is 1%, for Ana o 2 is 34%, and for Ana o 3 is 3%. The low yield can be explained by the precipitation steps of the total protein extract where Ana o 1 and Ana o 3 are discarded along with the undesirable proteins. However, as the protein content of cashew is 18.8% [25], the amount of cashew needed is relatively low. No isoform selection is expected as inhibition blotting clearly shows complete inhibition of IgE binding to Ana o 1 and Ana o 3 in the total protein extract upon inhibition with, respectively, purified Ana o 1 or Ana o 3. Besides inhibition of IgE binding to Ana o 1, purified Ana o 1 also inhibited binding of IgE to Ana o 3 and partially to Ana o 2. Cross-reactivity between Ana o 1

and Ana o 3 was unexpected, although some cross-reactivity between recombinant Ana o 1 and Ana o 2 has been shown before [11]. Cross-reactivity between 7S globulins, 11S globulins, and 2S albumins has been observed before for peanut [26].

Glycoprotein staining of the cashew protein fractions showed Ana o 1 and an unknown protein of 3kDa to be glycosylated. Both Ana o 2 and Ana o 3 are not glycosylated. Glycosylation of cashew proteins has been studied once before where only one glycosylated protein of ~3kDa was indicated [27]. In their study Sathe *et al*. used a total protein extract that, as can also be observed from our glycoprotein stained gel, possibly did not contain enough Ana o 1 to show a clear band in glycoprotein staining. Another explanation could be a difference in cashew variety used.

Under non-reducing denaturing conditions, Ana o 1 is present as both a dimer of 100kDa and a single protein of 50kDa. Non-reduced denatured Ana o 2 displays some protein bands that are also visible in reducing denaturing SDS-PAGE: the 53kDa full protein and 30 and 21kDasubunits. Ana o 3 seems to form one complex of ~12kDa linked by disulfide bridges, which was confirmed by LC/ESI-MS.

Purified Ana o 3 manifested as a triplet of 10, 8, and 6kDa proteins on SDS-PAGE under denaturing reducing conditions. Native 2S-albumins are post-translationally processed into an N-terminal small subunit of ~5.5kDa and a C-terminal large subunit of ~8.8kDa, held together via cysteine linkages [16]. In an attempt to further identify the protein triplet, we subjected both native and denatured reduced Ana o 3 to LC-MS analysis. This indicated the 10 and 8kDa proteins to be the large subunit of 8.2–8.4 and 8.1–8.2kDa, respectively, and the 6kDa protein to be the small subunit of 3.7–4.5kDa. Tryptic digestion of Ana o 3 indicated the small subunit to span amino acids 29–36 until 63–65, including N- and C-terminal micro-heterogeneity, and the large subunit to span amino acid 69 until 136–137, including C-terminal micro-heterogeneity, differing slightly from the proposed amino acid stretches proposed by Moreno *et al.* [16]. The same degree of micro-heterogeneity was observed when Ana o 3 was purified in the presence of protease inhibitors, indicating this micro-heterogeneity is not an artefact from the method of extraction but is due to proteolytic processes, the presence of different precursors, or the shift in cleavage sites in the cashew itself [16]. Micro-heterogeneity of 2S albumins has been described before for castor bean [28], Brazil nut [29], and sesame [30].

This data differs from the data reported by Robotham *et al*., who described all three protein bands to be isoforms of the large subunit of Ana o 3 with very similar N-terminal sequences, differing in only 1–2 amino acids out of 10 [15]. An explanation for this discrepancy could lay in a difference in cashew variety or storage conditions.

As food allergy is highly prevalent and the incidence of cashew allergy seems to be rising [31], it is relevant to study which proteins are involved. Ana o 1, Ana o 2, and Ana o 3 are the three major allergens in cashew. Using pure allergen fractions allows for determining to

which specific cashew allergen a patient reacts, and they can be used in cross-reactivity studies, where pure allergens can be used for inhibition blotting. Besides, these allergen fractions can also be used for analysis of allergen structure, comparing cashew varieties, and cashew protein processing stability.

The advantage of using purified allergens over the use of recombinant allergens is clear for Ana o 1 and Ana o 3. Recombinant expression of a glycosylated allergen like Ana o 1 is difficult as both prokaryotic (e.g., *E. coli*) and eukaryotic (e.g., *P. Pastoris*) expression systems may not be capable of reproducing the natural glycation pattern as in cashew [32,33]. Considering Ana o 3, recombinant expression will not properly display the micro-heterogeneity of the native protein; however, recombinant proteins are required to compare micro-heterogenic variations of Ana o 3. Besides, other post-translational processing modifications such as disulfide bridges (Ana o 2, Ana o 3) might prove more difficult in *E. coli* expression systems [32].

Funding

This project was funded by Technology Foundation STW, Stichting Voedselallergie, Siemens Healthcare Diagnostics, HAL Allergy, Intersnack Nederland B.V., ALK-Abello B.V., and the Nederlands Anafylaxis Netwerk.

Acknowledgements

The authors thank Jack Yang and Yvonne Veltman for their contribution in the first steps of allergen purification. For her valuable advice for the epitope mapping of Ana o 3, we thank Dr. Lisa Steinbrueck from Pepperprint Heidelberg, Germany.

References

[1] Nwaru, B. I.; Hickstein, L.; Panesar, S. S.; Roberts, G.; Muraro, A.; Sheikh, A. Prevalence of common food allergies in Europe: a systematic review and meta-analysis. Allergy 2014, 69, 992–1007.

[2] Sicherer, S. H.; Munoz-Furlong, A.; Burks, A. W.; Sampson, H. A. Prevalence of peanut and tree nut allergy in the US determined by a random digit dial telephone survey. J. Allergy Clin. Immunol. 1999, 103, 559–62.

[3] Sicherer, S. H.; Burks, A. W.; Sampson, H. A. Clinical features of acute allergic reactions to peanut and tree nuts in children. Pediatrics 1998, 102, e6.

[4] Davoren, M.; Peake, J. Cashew nut allergy is associated with a high risk of anaphylaxis. Arch. Dis. Child. 2005, 90, 1084–1085.

[5] Hasegawa, M.; Inomata, N.; Yamazaki, H.; Morita, A.; Kirino, M.; Ikezawa, Z. Clinical features of four cases with cashew nut allergy and cross-reactivity between cashew nut and pistachio. Allergol. Int. 2009, 58, 209–15.

[6] Rance, F.; Bidat, E.; Bourrier, T.; Sabouraud, D. Cashew allergy: observations of 42 children without associated peanut allergy. Allergy 2003, 58, 1311–4.

[7] Clark, A. T.; Anagnostou, K.; Ewan, P. W. Cashew nut causes more severe reactions than peanut: case-matched comparison in 141 children. Allergy 2007, 62, 913–6.

[8] van der Valk, J. P.; Dubois, A. E. J.; Gerth van Wijk, R.; Wichers, H. J.; de Jong, N. W. Systematic review on cashew nut allergy. Allergy 2014, 69, 692–698.

[9] Wang, F.; Robotham, J. M.; Teuber, S. S.; Tawde, P.; Sathe, S. K.; Roux, K. H. Ana o 1, a cashew (Anacardium occidental) allergen of the vicilin seed storage protein family. J. Allergy Clin. Immunol. 2002, 110, 160–6.

[10] Mills, E. N. C.; Shewry, P. R. Plant food allergens; Blackwell Science Ltd.: 2004; p 219.

[11] Wang, F.; Robotham, J. M.; Teuber, S. S.; Sathe, S. K.; Roux, K. H. Ana o 2, a major cashew (Anacardium occidentale L.) nut allergen of the legumin family. Int. Arch. Allergy Immunol. 2003, 132, 27–39.

[12] Xia, L.; Willison, L. N.; Porter, L.; Robotham, J. M.; Teuber, S. S.; Sathe, S. K.; Roux, K. H. Mapping of a conformational epitope on the cashew allergen Ana o 2: a discontinuous large subunit epitope dependent upon homologous or heterologous small subunit association. Mol. Immunol. 2010, 47, 1808–16.

[13] Robotham, J. M.; Xia, L.; Willison, L. N.; Teuber, S. S.; Sathe, S. K.; Roux, K. H. Characterization of a cashew allergen, 11S globulin (Ana o 2), conformational epitope. Mol. Immunol. 2010, 47, 1830–8.

[14] Venkatachalam, M.; Monaghan, E. K.; Kshirsagar, H. H.; Robotham, J. M.; O'Donnell, S. E.; Gerber, M. S.; Roux, K. H.; Sathe, S. K. Effects of processing on immunoreactivity of cashew nut (Anacardium occidentale L.) seed flour proteins. J. Agric. Food Chem. 2008, 56, 8998–9005.

[15] Robotham, J. M.; Wang, F.; Seamon, V.; Teuber, S. S.; Sathe, S. K.; Sampson, H. A.; Beyer, K.; Seavy, M.; Roux, K. H. Ana o 3, an important cashew nut (Anacardium occidentale L.) allergen of the 2S albumin family. J. Allergy Clin. Immunol. 2005, 115, 1284–90.

[16] Moreno, F. J.; Clemente, A. 2S Albumin Storage Proteins: What Makes them Food Allergens? Open Biochem. J. 2008, 2, 16–28.

[17] Mattison, C. P.; Grimm, C. C.; Wasserman, R. L. In vitro digestion of soluble cashew proteins and characterization of surviving IgE-reactive peptides. Mol. Nutr. Food Res. 2014, 58, 884–93.

[18] Hummel, M.; Wigger, T.; Hoper, T.; Westkamp, I.; Brockmeyer, J. Simple, Rapid, and Selective Isolation of 2S Albumins from Allergenic Seeds and Nuts. J. Agric. Food Chem. 2015, 63, 6035–6040.

[19] Sathe, S. K.; SzeTao, K. W. C.; Wolf, W. J.; Hamaker, B. R. Biochemical characterization and in vitro digestibility of the major globulin in cashew nut (Anacardium occidentale). J. Agric. Food Chem. 1997, 45, 2854–2860.

[20] Teuber, S. S.; Sathe, S. K.; Peterson, W. R.; Roux, K. H. Characterization of the soluble allergenic proteins of cashew nut (Anacardium occidentale L.). J. Agric. Food Chem. 2002, 50, 6543– 6549.

[21] Shevchenko, A.; Tomas, H.; Havli, J.; Olsen, J. V.; Mann, M. Ingel digestion for mass spectrometric characterization of proteins and proteomes. Nat. Protoc. 2007, 1, 2856–2860.

[22] Vereda, A.; van Hage, M.; Ahlstedt, S.; Ibanez, M. D.; Cuesta- Herranz, J.; van Odijk, J.; Wickman, M.; Sampson, H. A. Peanut allergy: Clinical and immunologic differences among patients from 3 different geographic regions. J. Allergy Clin. Immunol. 2011, 127, 603– 607.

[23] Ballmer-Weber, B. K.; Lidholm, J.; Fernandez-Rivas, M.; Seneviratne, S.; Hanschmann, K. M.; Vogel, L.; Bures, P.; Fritsche, P.; Summers, C.; Knulst, A. C.; Le, T. M.; Reig, I.; Papadopoulos, N. G.; Sinaniotis, A.;

Belohlavkova, S.; Popov, T.; Kralimarkova, T.; de Blay, F.; Purohit, A.; Clausen, M.; Jedrzejczak-Czechowcz, M.; Kowalski, M. L.; Asero, R.; Dubakiene, R.; Barreales, L.; Clare Mills, E. N.; van Ree, R.; Vieths, S. IgE recognition patterns in peanut allergy are age dependent: perspectives of the EuroPrevall study. Allergy 2015, 70, 391−407.

[24] Mattison, C. P.; Desormeaux, W. A.; Wasserman, R. L.; Yoshioka-Tarver, M.; Condon, B.; Grimm, C. C. Decreased immunoglobulin E (IgE) binding to cashew allergens following sodium sulfite treatment and heating. J. Agric. Food Chem. 2014, 62, 6746−55.

[25] Venkatachalam, M.; Sathe, S. K. Chemical composition of selected edible nut seeds. J. Agric. Food Chem. 2006, 54, 4705−14.

[26] Bublin, M.; Kostadinova, M.; Radauer, C.; Hafner, C.; Szepfalusi, Z.; Varga, E. M.; Maleki, S. J.; Hoffmann-Sommergruber, K.; Breiteneder, H. IgE cross-reactivity between the major peanut allergen Ara h 2 and the nonhomologous allergens Ara h 1 and Ara h 3. J. Allergy Clin. Immunol. 2013, 132, 118−24.

[27] Sathe, S. K.; Venkatachalam, M.; Sharma, G. M.; Kshirsagar, H. H.; Teuber, S. S.; Roux, K. H. Solubilization and electrophoretic characterization of select edible nut seed proteins. J. Agric. Food Chem. 2009, 57, 7846−56.

[28] da Siva, J. G., Jr.; Machado, O. L.; Izumi, C.; Padovan, J. C.; Chait, B. T.; Mirza, U. A.; Greene, L. J. Amino acid sequence of a new 2S albumin from Ricinus communis which is part of a 29-kDa precursor protein. Arch. Biochem. Biophys. 1996, 336, 10−18.

[29] Moreno, F. J.; Jenkins, J. A.; Mellon, F. A.; Rigby, N. M.; Robertson, J. A.; Wellner, N.; Clare Mills, E. N. Mass spectrometry and structural characterization of 2S albumin isoforms from Brazil nuts (Bertholletia excelsa). Biochim. Biophys. Acta, Proteins Proteomics 2004, 1698, 175−86.

[30] Moreno, F. J.; Maldonado, B. M.; Wellner, N.; Mills, E. N. Thermostability and in vitro digestibility of a purified major allergen 2S albumin (Ses i 1) from white sesame seeds (Sesamum indicum L.). Biochim. Biophys. Acta, Proteins Proteomics 2005, 1752, 142−53.

[31] Sicherer, S. H.; Munoz-Furlong, A.; Godbold, J. H.; Sampson, H. A. US prevalence of self-reported peanut, tree nut, and sesame allergy: 11-year follow-up. J. Allergy Clin. Immunol. 2010, 125, 1322−6.

[32] Willison, L. N.; Sathe, S. K.; Roux, K. H. Production and analysis of recombinant tree nut allergens. Methods 2014, 66, 34−43.

[33] Lorence, A. Recombinant Gene Expression, Reviews and Protocols, 3rd ed.; Springer Science + Business Media: 2012.

Supplementary information

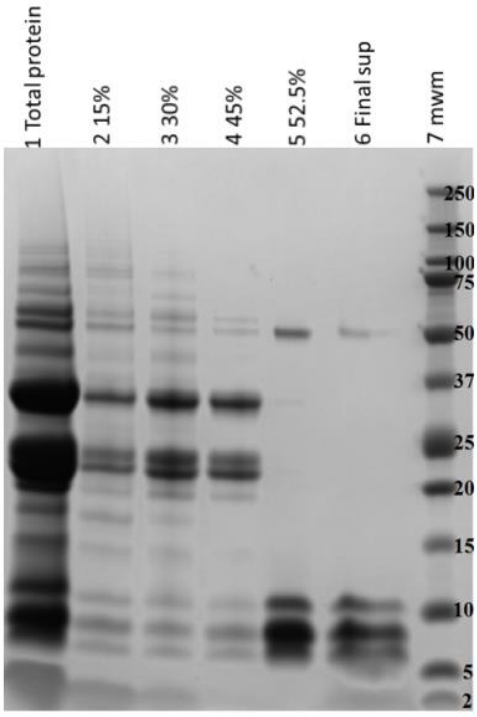


Figure 3.S1 *Reducing denaturing SDS-PAGE of ammonium sulphate precipitation. Depicted are (left to right): total protein extract from defatted cashew used for purification of Ana o 1 and Ana o 3, 15%, 30%, 45%, and 52.5% pellet of ammonium sulphate precipitation, final supernatant after 52.5% ammonium sulphate precipitation, molecular weight marker.*

100kDa protein band:

Identified as, Ana o 1, vicilin-like protein (Anacardium occidentale), GenBank AAM73729.1
14% sequence coverage, Mascot score 179

PPTKFSFSLFLVSVLVLCLGFALAKIDPELKQCKHQCKVQRQYDEQQKEQCVKECEKYYK
EKKGREREHEEEEEEWGTGGVDEPSTHEPAEKHLSQCMRQCERQEGGQQKQLCRFRCQER
YKKERGQHNYKREDDEDEDEDEAEEEDENPYVFEDEDFTTKVKTEQGKVVLLPKFTQKSK
LLHALEKYRLAVLVANPQAFVVPSHMDADSIFFVSWGRGTITKILENKRESINVRQGDIV
SISSGTPFYIANNDENEKLYLVQFLRPVNLPGHFEVFHGPGGENPESFYRAFSWEILEAA
LKTSKDTLEKLFEKQDQGTIMKASKEQVRAMSRRGEGPKIWPFTEESTGSFKLFKKDPSQ
SNKYGQLFEAERIDYPPLEKLDMVVSYANITKGGMSVPFYNSRATKIAIVVSGEGCVEIA
CPHLSSSKSSHPSYKKLRARIRKDTVFIVPAGHPFATVASGNENLEIVCFEVNAEGNIRY
TLAGKKNIIKVMEKEAKELAFKMEGEEVDKVFGKQDEEFFFQGPEWRKEKEGRADE

50kDa protein band:

Identified as Ana o 1, vicilin-like protein (Anacardium occidentale), GenBank AAM73729.1
36% sequence coverage, Mascot score 360.

PPTKFSFSLFLVSVLVLCLGFALAKIDPELKQCKHQCKVQRQYDEQQKEQCVKECEKYYK
EKKGREREHEEEEEEWGTGGVDEPSTHEPAEKHLSQCMRQCERQEGGQQKQLCRFRCQER
YKKERGQHNYKREDDEDEDEDEAEEEDENPYVFEDEDFTTKVKTEQGKVVLLPKFTQKSK
LLHALEKYRLAVLVANPQAFVVPSHMDADSIFFVSWGRGTITKILENKRESINVRQGDIV
SISSGTPFYIANNDENEKLYLVQFLRPVNLPGHFEVFHGPGGENPESFYRAFSWEILEAA
LKTSKDTLEKLFEKQDQGTIMKASKEQVRAMSRRGEGPKIWPFTEESTGSFKLFKKDPSQ
SNKYGQLFEAERIDYPPLEKLDMVVSYANITKGGMSVPFYNSRATKIAIVVSGEGCVEIA
CPHLSSSKSSHPSYKKLRARIRKDTVFIVPAGHPFATVASGNENLEIVCFEVNAEGNIRY
TLAGKKNIIKVMEKEAKELAFKMEGEEVDKVFGKQDEEFFFQGPEWRKEKEGRADE

53kDa protein band:

Identified as, Ana o 2 (Anacardium occidentale), GenBank AAN76862.1
27% sequence coverage, Mascot score 156

LSVCFLILFHGCLASRQEWQQQDECQIDRLDALEPDNRVEYEAGTVEAWDPNHEQFRCAG
VALVRHTIQPNGLLLPQYSNAPQLIYVVQGEGMTGISYPGCPETYQAPQQGRQQGQSGRF
QDRHQKIRRFRRGDIIAIPAGVAHWCYNEGNSPVVTVTLLDVSNSQNQLDRTPRKFHLAG
NPKDVFQQQQQHQSRGRNLFSGFDTELLAEAFQVDERLIKQLKSEDNRGGIVKVKDDELR
VIRPSRSQSERGSESEEESEDEKRRWGQRDNGIEETICTMRLKENINDPARADIYTPEVG
RLTTLNSLNLPILKWLQLSVEKGVLYKNALVLPHWNLNSHSIIYGCKGKGQVQVVDNFGN
RVFDGEVREGQMLVVPQNFAVVKRAREERFEWISFKTNDRAMTSPLAGRTSVLGGMPEEV
LANAFQISREDARKIKFNNQQTTLTSGESSHHMRDDA

30kDa protein band:

Identified as, Ana o 2 (Anacardium occidentale), GenBank AAN76862.1
35% sequence coverage, Mascot score 306

LSVCFLILFHGCLASRQEWQQQDECQIDRLDALEPDNRVEYEAGTVEAWDPNHEQFRCAG
VALVRHTIQPNGLLLPQYSNAPQLIYVVQGEGMTGISYPGCPETYQAPQQGRQQGQSGRF
QDRHQKIRRFRRGDIIAIPAGVAHWCYNEGNSPVVTVTLLDVSNSQNQLDRTPRKFHLAG
NPKDVFQQQQQHQSRGRNLFSGFDTELLAEAFQVDERLIKQLKSEDNRGGIVKVKDDELR
VIRPSRSQSERGSESEEESEDEKRRWGQRDNGIEETICTMRLKENINDPARADIYTPEVG
RLTTLNSLNLPILKWLQLSVEKGVLYKNALVLPHWNLNSHSIIYGCKGKGQVQVVDNFGN
RVFDGEVREGQMLVVPQNFAVVKRAREERFEWISFKTNDRAMTSPLAGRTSVLGGMPEEV
LANAFQISREDARKIKFNNQQTTLTSGESSHHMRDDA

22kDa protein band:
Identified as, 11S globulin (Pistacia vera), GenBank ABU42022.1
13% sequence coverage, Mascot score 138
MGYSSLLSFSLGFLLLFHCSFAQIEQVVNSQQRQQQQRFQTQCQIQNLNALEPKRRIESE
AGVTEFWDQNEEQLQCANVAVFRHTIQSRGLLVPSYNNAPELVYVVQGSGIHGAVFPGCP
ETFQEESQSQSRSQHSRSERSQQSGEQHQKVRHIREGDIIALPAGVAHWIYNNGQSKLVL
VALADVGNSENQLDQYLRKFVLGGSPQQEIQGGGQSWSQSRSSRKGQQSNNILSAFDEEI
LAQSLNIDTQLVKKLQREEKQRGIIVRVKEDLQVLSPQRQEKEYSDNGLEETFCTMTLKL
NINDPSRADVYNPRGGRVTSINALNLPILRFLQLSVEKGVLYQNAIMAPHWNMNAHSIVY
ITRGNGRMQIVSENGESVFDEEIREGQLVVVPQNFAVVKRASSDGFEWVSFKTNGLAKIS
QLAGRISVMRGLPLDVIQNSFDISREDAWNLKESRSEMTIFAPGSRSQRQRN

21kDa protein band:
Identified as, Ana o 2 (Anacardium occidentale), GenBank AAN76862.1
30% sequence coverage, Mascot score 400
LSVCFLILFHGCLASRQEWQQQDECQIDRLDALEPDNRVEYEAGTVEAWDPNHEQFRCAG
VALVRHTIQPNGLLLPQYSNAPQLIYVVQGEGMTGISYPGCPETYQAPQQGRQQGQSGRF
QDRHQKIRRFRRGDIIAIPAGVAHWCYNEGNSPVVTVTLLDVSNSQNQLDRTPRKFHLAG
NPKDVFQQQQQHQSRGRNLFSGFDTELLAEAFQVDERLIKQLKSEDNRGGIVKVKDDELR
VIRPSRSQSERGSESEEESEDEKRRWGQRDN**GIEETICTMRLKENINDPARADIYTPEVG**
RLTTLNSLNLPILKWLQLSVEKGVLYKNALVLPHWNLNSHSIIYGCKGKGQVQVVDNFGN
RVFDGEVREGQMLVVPQNFAVVKRAREERFEWISFKTNDRAMTSPLAGRTSVLGGMPEEV
LANAFQISREDARKIKFNNQQTTLTSGESSHHMRDDA

10kDa protein band: (LC-MS/MS analysis)
Identified as, Ana o 3, 2s Albumin (Anacardium occidentale), GenBank AAL91665.1
33% sequence coverage
MAKFLLLLSAFAVLLLVANASIYRAIVEVEEDSGREQSCQRQFEEQQRFRNCQRYVKQEV
QRGGRYNQ**RQESLRECCQELQEVDRRCRCQNLEQMVRQLQQQEQIKGEEVRELYETASEL**
PRICSISPSQGCQFQSSY

8kDa protein band:
Identified as, Ana o 3, 2s Albumin (Anacardium occidentale), GenBank AAL91665.1
54% sequence coverage, Mascot score 407
MAKFLLLLSAFAVLLLVANASIYRAIVEVEEDSGREQSCQRQFEEQQRFRNCQRYVKQEV
QRGGRYNQ**RQESLRECCQELQEVDRRCRCQNLEQMVRQLQQQEQIKGEEVRELYETASEL**
PRICSISPSQGCQFQSSY

6kDa protein band:
Identified as, Ana o 3, 2s Albumin (Anacardium occidentale), GenBank AAL91665.1
52% sequence coverage, Mascot score 336
MAKFLLLLSAFAVLLLVANASIYRAIVE**VEEDSGREQSCQRQFEEQQRFRNCQRYVKQEV**
QRGGRYNQRQESLRECCQELQEVDRRCRCQNLEQMVRQLQQQEQIKGEEVRELYETASEL
PRICSISPSQGCQFQSSY

Figure 3.S2 *Protein identification, analysis of protein bands from SDS-PAGE indicated in Figure 3.2. Results based on MALDI-MS/MS analysis unless stated otherwise. Indicated are the identified protein, GenBank number, sequence coverage, Mascot score, and the protein sequence with matched peptides underlined. For the proteins identified as Ana o 3 the large subunit (10 and 8kDa protein band) or the small subunit (6kDa protein band) as based on Figure 3.5 are indicated in bold. For the proteins identified as Ana o 2 the large subunit [11] (30kDa protein band) or the small subunit [11] (21kDa protein band) are indicated in bold.*

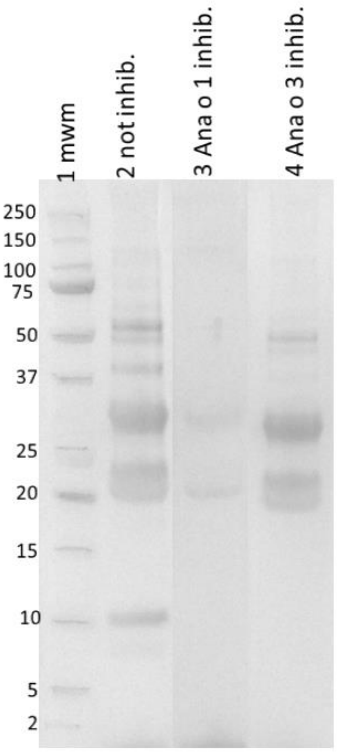
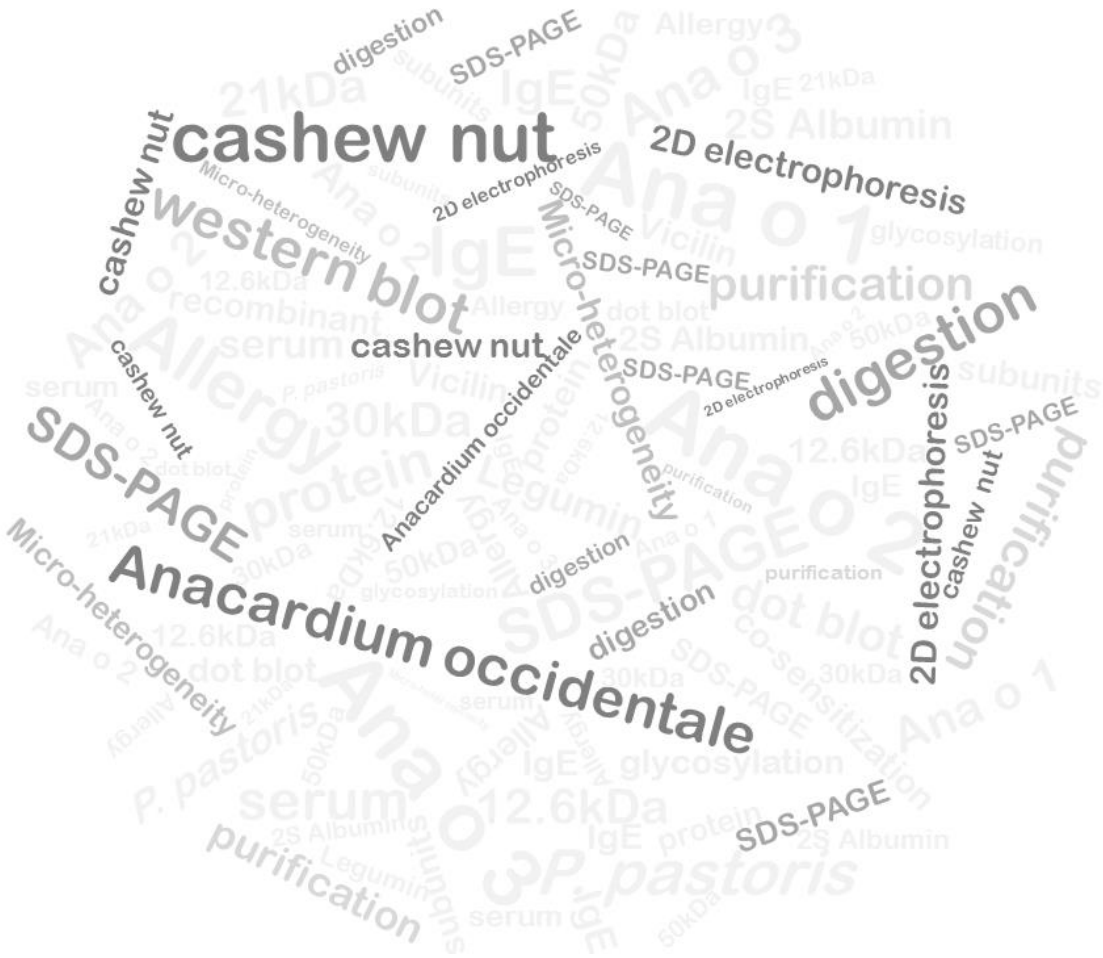


Figure 3.S3 *Inhibition western blot of Ana o 1 and Ana o 3, using plasma pool of group 2. Molecular weight marker followed by 3 lanes containing a total protein extract from defatted cashew used for purification of Ana o 1 and Ana o 3. Lane 2 is not inhibited, the plasma used in lane 3 was inhibited with 1mg/ml purified Ana o 1, the plasma used in lane 4 is inhibited with 1mg/ml purified Ana o 3.*

Chapter 4

Origin and processing methods slightly affect allergenic characteristics of cashew nuts (*Anacardium occidentale*)

Marit Reitsma, Shanna Bastiaan-Net, Lutske Sijbrandij, Evelien de Weert, Stefano Sforza, Huub F. J. Savelkoul, Nicolette W. de Jong, and Harry J. Wichers



Manuscript in preparation

Abstract

The protein content and allergen composition was studied of cashew nuts from 8 different origins (Benin, Brazil, Ghana, India, Ivory Coast, Mozambique, Tanzania, Vietnam), subjected to different in-shell heat treatments (steamed, fried, drum-roasted). On 2D electrophoresis, 9 isoforms of Ana o 1, 29 isoforms of Ana o 2 (11 of the acidic subunit, 18 of the basic subunit), and 8 isoforms of the large subunit of Ana o 3 were tentatively identified. Based on 1D and 2D electrophoresis, no difference in allergen content (Ana o 1, 2, 3) was detected between the cashew nuts of different origins ($p>0.5$), some small but significant differences were detected in allergen solubility between differently heated cashew nuts. No major differences in N- and C-terminal micro-heterogeneity of Ana o 3 were detected between cashew nuts of different origins. Between the different heat treatments, no difference was detected in glycation, pepsin digestibility, or IgE binding of the cashew nut proteins.

Introduction

Cashew nuts have been studied regarding their three major allergens [1], Ana o 1 [2], Ana o 2 [3-5] and Ana o 3 [6], as well as regarding their nutritional content [7], clinical reactivity [8-11], *in vitro* digestibility [12, 13], and their detection in food items [14, 15]. However, hardly any information is available on potential differences in allergen content between cashew nuts of different origins. Only a few studies were found describing a comparison of cashew nuts of different origins and varieties; one study counting the amount of chromosomes present in cashew nuts of two origins [16] and two other studies reporting no differences in protein content in cashew nuts of different varieties or origins [17, 18]. One paper stated in their unpublished results not to have detected differences in protein pattern on SDS-PAGE between cashew nut protein extracts from Florida and Mumbai, however no results were shown [19].

In some nuts and legumes differences in allergen composition were reported for different origins. For instance, in Balinese peanuts, very low levels of the 7S globulin Ara h 1 were detected compared to other varieties [20]. However, this lower level of Ara h 1 did not result in a decrease in immunoreactivity as determined using an RBL (rat basophilic leukemia cell line) model [21]. Other, commercial, peanut varieties do not seem to differ in major allergen content (Ara h 1, 2, 3 and 6) [22, 23], just as different pistachio nut varieties seem to be similar in protein content and IgE binding capacity [24].

With respect to the effects of thermal processing on cashew nuts, most studies have been performed on related nuts like pistachio, and legumes like peanut. For example, (steam-) roasting of peanut [25] and pistachio [24] decreased protein solubility, with steam-roasting of pistachio also inducing a decrease in patient IgE binding and a decrease in pepsin digestibility [24]. Using an RBL model, raw hazelnut appeared to be more allergenic than roasted hazelnut [26]. Roasting of total peanut protein has been shown to induce increased IgE binding in an ELISA assay [27], while no difference between raw and roasted peanut Ara h 2 and 6 was detected in an RBL assay [28]. Using purified peanut allergens, roasting of Ara h 1 (7S globulin) increased IgE cross-linking of RBL cells, while for Ara h 2/6 (2S albumins) the cross-linking capacity was decreased [29]. In a basophil activation test (BAT) model it was shown that the effect of thermal processing on peanut differed between peanut varieties and patients [30].

The effects of thermal processing on the three major allergens of cashew nuts has been studied by Venkatachalam *et al.* [31], demonstrating the stability of monoclonal antibody (mAb)-binding by these allergens. IgE binding of Ana o 1, 2, and 3 was stable (as assessed by ELISA, western blot and dot blot) between a pH range of 2 to 12. IgG-binding of Ana o 1 and

2 was heat stable, as only autoclaving, but not microwaving, roasting or y-irradiation, resulted in a decrease in mouse mAb binding as shown with western blotting [31]. Immunoreactivity of Ana o 3 was decreased by roasting and autoclaving, however this was measured by western blot using a mAb that is directed at a conformational epitope [31]. It is unclear if these heat treatments affected the multiple linear epitopes of Ana o 3. Roasting of cashew nuts was reported to increase the solubility of Ana o 3 [32] while the total cashew nut protein solubility decreased [33]. Blanching of the cashew nuts on the other hand, causes leaking of all three allergens from the cashew nut into the blanching water [31]. However, the effects of frying were not studied by Venkatachalam *et al.*

In the present study we compared cashews nuts of 8 different origins (Benin, Brazil, Ghana, India, Ivory Coast, Mozambique, Tanzania, and Vietnam) to analyze possible differences in protein and allergen content. Next, we compared the effect of different heat treatments (raw, steaming, frying, and drum roasting) on the protein and allergen content of the cashew nuts. We have studied the proteins by DUMAS method, SDS-PAGE, 2D electrophoresis, glycoprotein staining, *in vitro* digestion, and western blotting. Ana o 3 was also analyzed for N- and C-terminal micro-heterogeneity using UPLC/PDA/ESI-MS.

The cashew nuts used in this study were heat-treated (steamed, fried, drum-roasted) in the shell. Each of these three heat treatments is routinely applied in industry to cashew nuts before de-shelling. The effects of such heat treatments on the cashew nut allergens Ana o 1, 2 and 3, has not been studied before.

Materials and methods

Chemicals

Urea, thiourea, CHAPS, DTT, SDS, β-mercaptoethanol, NaCl, NH4HCO3, NaHCO3, (NH4)2CO3, Tris, glycerol, porcine pepsin, Sigmafast BCIP/NBT and alkaline phosphatase conjugated goat anti-rabbit (A3687) were obtained from Sigma-Aldrich, St. Louis, MO. KCl, KH2PO4, ammonium sulfate, glycine, Coomassie R-250 and bromophenol blue were purchased from Merck, Darmstadt, Germany. 3kDa Amicon centrifugal filter were purchased from Merck Millipore, Tullagreen, Ireland. Precision Plus Protein Dual Xtra Standard molecular weight marker was purchased from Bio-Rad, Hercules, CA. The 2-D Quant Kit, IPG strips (pH 3-11 non-linear), IPG buffer (pH 3-11), and iodoacetamide were obtained from GE Healthcare, Piscataway, NJ. NuPAGE LDS sample buffer, Mark12 unstained standard molecular weight marker, NuPAGE 1mm 4-12% Bis-Tris ZOOM gel, NuPAGE MES SDS running buffer, and SYPRO Ruby Protein Gel Stain were purchased from Invitrogen, Carlsbad, CA. CandyCane

glycoprotein molecular weight marker, MemCode Reversible Protein Stain Kit, and Pro-Q Emerald 300 Glycoprotein gel and blot stain kit were obtained from Thermo Scientific, Rockford, IL. Methanol and acetic acid were both purchased from Actu-All, Randmeer, the Netherlands. 0.45µm filters were obtained from Minisart, Satorius Stedim, Goettingen, Germany. MgCl2(H2O)6 was obtained from Boom BV, Meppel, the Netherlands. Polyclonal rabbit anti-human IgE (A0094) was obtained from Dako, Glostrup, Denmark.

Cashew nut samples

Raw, sundried cashew nuts (*Anacardium occidentale*) originating from Benin, Brazil, Ghana, India, Ivory Coast, Mozambique, Tanzania and Vietnam, were kindly provided by Intersnack B.V. (Doetinchem, the Netherlands) as raw cashew nuts in the shell. These raw cashew nuts were soaked in water for 8hr at room temperature, after which all floating cashew nuts were discarded. Heat-treatments of these raw in-shell cashew nuts were performed by Intersnack Nederland B.V.: steaming (20min 100˚C), frying (2min 180˚C), or drum roasting (8min 150˚C). Due to limited cashews from Tanzania, no Tanzania steamed sample was available. All heat-treated cashew nuts were subsequently de-shelled, and incubated for 6hr at 70˚C, after which the skin (testa) was removed. For samples that are designated as raw cashew nuts, the raw untreated cashew nuts were de-shelled and peeled directly after the 8hr soaking step. Shelled cashew nuts were stored at 4˚C, de-shelled and peeled cashew nuts were stored at -80˚C.

Protein extraction

Ten de-shelled and peeled cashew nuts were cut into small pieces, frozen by liquid nitrogen, and ground by an analytic mill (IKA A11, Staufen, Germany). Ground cashew nuts were stored at -80˚C. Protein was extracted o/n 1:25 w/v in an urea extraction buffer (7M urea, 2M thiourea, 2% (w/v) CHAPS, 50mM DTT, pH 8.8) at 4˚C while rotating. The samples were centrifuged for 20min 4˚C 10.000g and the supernatants were stored at -20˚C.

Protein content

The protein content in the extracts was determined by the 2-D Quant Kit according to the manufacturers' instructions. The protein content of the ground de-shelled and peeled cashew nuts was determined by DUMAS method. In duplicate, 10mg ground cashew nuts was analyzed in an N Analyser (Flash EA 1112, Thermo Scientific). A conversion factor of 5.3 [7] was used to calculate the percentage of protein in the cashew nuts.

In vitro gastric digestion

For a non-denaturing non-reducing protein extract, protein from Vietnam ground cashew nut was extracted 1:13 w/v in an ammonium bicarbonate buffer (0.5M NaCl, 0.1M NH4HCO3, pH 7.9) for 1hr while shaking at room temperature, followed by centrifugation for 5min at 10.000g, and 0.45µm filtering. For a denaturing reducing protein extract, protein from Vietnam ground cashew nut was extracted 1:13 w/v into the urea extraction buffer (7M urea, 2M thiourea, 2% (w/v) CHAPS, 50mM DTT, pH8.8) for 1hr while shaking at room temperature, followed by centrifugation for 5min at 10.000g, and 0.45µm filtering. Buffer exchange was performed for the urea protein extract by washing the extract five times with 0.5ml simulated gastric fluid (SGF)[34]: 6.9mM KCl, 0.9mM KH2PO4, 25mM NaHCO3, 47.2mM NaCl, 0.1mM MgCl2(H2O)6, 0.5mM (NH4)2CO3) in a 3kDa Amicon centrifugal filter. To 400µg protein, diluted in 400µl SGF buffer with pH3, 76µg porcine pepsin (equal to 61-190U) was added. After 1hr incubation at 37˚C while shaking, the reaction was stopped by adding LDS sample buffer and β-mercaptoethanol as described below for SDS-PAGE.

SDS-PAGE

15µg of sample was denatured and reduced (10% β-mercaptoethanol in NuPAGE LDS sample buffer, 5min 100˚C) and run on SDS-PAGE as described earlier [1]. The gels were Coomassie stained by incubation in fix solution (40% methanol, 10% acetic acid) for 45s 1000W in a microwave, followed by 15min incubation at room temperature. Subsequently, the gels were stained in 30% methanol, 10% acetic acid, 0.02% w/v Coomassie R-250 for 45s 1000W in a microwave, followed by 15min incubation at room temperature. The gels were de-stained in 8% acetic acid for 45sec 1000W in a microwave followed by shaking o/n at room temperature. Imaging and analysis were performed using a Universal Hood III and Image Lab 4.1 software (both Bio-Rad).

Glycoprotein stain

Samples were run on SDS-PAGE as described above, alongside a CandyCane glycoprotein molecular weight standard. Gels were stained using the Pro-Q Emerald 300 Glycoprotein gel and blot stain kit. Imaging and analysis were performed using a Universal Hood III and Image Lab 4.1 software.

2D Electrophoresis

IPG strips (pH 3-11 non-linear) were rehydrated overnight with either 40µg total cashew nut protein (urea protein extract), or 5µg purified Ana o 1, 2 or 3 [1] diluted with rehydration buffer (5.6M urea, 1.6M thiourea, 20mM DTT, 1.6% w/v chaps, a few grains of bromophenol blue, 200x diluted IPG buffer pH 3-11 NL) to a final volume of 125µl. Isoelectric focusing was

performed in a protean IEF cell (Bio- Rad): 300V 0.2kvhr linear, 1000V 0.3kVhr rapid, 5000V 4kVhr rapid, 5000V linear 2kVhr. The focused strips were reduced for 15min in 50mM Tris, 20mg/ml SDS, 10mg/ml DTT, 30% glycerol and a few grains of bromophenol blue. Subsequently, the strips were alkylated for 15min in 50mM Tris, 20mg/ml SDS, 25mg/ml iodoacetamide, 30% glycerol and a few grains of bromophenol blue. The strips were run on a NuPAGE 1mm 4-12% Bis-Tris ZOOM gel alongside a molecular weight standard (Mark12 unstained standard). Electrophoresis was performed at 160V for 55min in NuPAGE MES SDS running buffer. Gels were stained with SYPRO Ruby Protein Gel Stain according to the manufacturers' instructions.
For all gels the staining time and light exposure time for the picture were equal.

Western blotting

Western blotting of 1D and 2D SDS-PAGE was performed as described earlier [1] with the addition of a reversible protein stain step (MemCode Reversible Protein Stain Kit) directly after the protein transfer. IgE binding to the proteins was visualized by Sigmafast BCIP/NBT staining for 6 minutes.

Plasma samples

Plasma samples were obtained from PlasmaLab International (Everett, WA). Three patients with self-reported tree nut or cashew nut allergy and high cashew nut IgE titers were selected for western blotting experiments and pooled in equal volumes. Patient characteristics are presented in Table 4.1.

Table 4.1 *Clinical characteristics of cashew nut and tree nut allergic subjects. IgE levels specific for total cashew nut protein and Ana o 1, 2, 3 were measured by Immulite as described by van der Valk et al. [53]. Self-reported allergies per patient were provided by PlasmaLab International.*

Patient # plasma	Self-reported allergies	IgE level cashew nut (kU/l)	IgE level Ana o 1 (kU/l)	IgE level Ana o 2 (kU/l)	IgE level Ana o 3 (kU/l)
1	Tree nuts, hazelnut, pecan, dust mites, grass, insect venom	19.2	2.5	12.4	12.3
2	Cashew nut, peanut, walnut, tree pollen, cat, dog, horse	>100	>100	>100	>100
3	Tree nuts, banana, eggplant	61.7	23.0	58.2	>100

Analysis of Ana o 3 micro-heterogeneity

Proteins were extracted from raw cashew nuts of eight origins by an ammonium bicarbonate buffer (0.1M ammonium bicarbonate, 0.5M NaCl, pH 7.9) 1:5w/v at 4˚C o/n while rotating. Proteins in the supernatant (20min, 10.000g, 4˚C) were roughly precipitated by adding 52.5% w/w ammonium sulfate and mixing for 2hr while rotating. Subsequently, the samples were

centrifuged for 10min, 10.000g at 4˚C. Both native and reduced (1mg protein/ml in 10mM DTT, 1% SDS, 5min 100˚C) supernatants containing Ana o 3, were filtered over a 0.5ml 3kDa Amicon centrifugal filter and suspended in 200µl eluent A (0.1% trifluoracetic acid and 1% acetonitrile). The samples were analyzed by LC/PDA/ESI-MS as described before [1].

Statistical analysis

One-Way ANOVA analyses were executed with SPSS-22 (LSD). Groups were compared either by origin (comparing 8 groups of 4 samples each), or by heat treatment (comparing 4 groups of 8 samples each). Background coloring of the gels was subtracted for all analyzed spots. Significance levels were set at 0.05%. PCA plot analysis was performed in Canoco 5 version 5.04.

Results and discussion

Protein content

The average protein content for all cashew nuts, as measured by DUMAS assay, was 16.8% (±1.8%). Averages per country of origin and per heat treatment are presented in Table 4.2. On average, the highest protein concentration was measured in cashew nuts from Vietnam, and the lowest protein concentration in cashew nuts from India. Table 4.2 also shows the significant differences in protein content of cashew nuts of different origins. These differences might be related to differences in moisture or fat content of the nuts. Small differences in protein content for cashew nuts of different origins have been detected before [17]. When comparing the average protein content of raw cashew nuts, we measured a lower value (16.9±1.0%) than Rico *et al.* (18.06 ±0.8% when re-calculated with a conversion factor of 5.3) [17]. Variation in moisture content between these two results might explain this difference, as we did not dry the nuts before analysis while Rico *et al.* did.
In this study, no significant differences were detected in the total protein content of cashew nuts subjected to different heat treatments.

1D electrophoresis

Figure 4.1 shows the electrophoresis profiles of extracted proteins from cashew nuts of 8 different origins subjected to 4 different heat treatments. The allergens Ana o 1, 2 and 3 are indicated in Figure 4.1A. In Table 4.3, the percentages of Ana o 1, 2 and 3 are indicated as a percentage of all protein bands on SDS-PAGE from Figure 4.1. The extractable protein content of the cashew nuts, visualized on SDS-PAGE, demonstrates the presence of 4.1±0.9% Ana o 1, 51.5±4.9% Ana o 2, and 25.4±3.9% Ana o 3 (average of all 31 samples). There were no significant differences in allergen content ($p>0.5$, One-Way ANOVA) for cashew nuts of

Table 4.2 *Protein content (%) of cashew nuts from different origins and subjected to different heat treatments as measured by the DUMAS assay. Average (Avg.) and standard deviation (Stdev.) of duplicates are shown. Average compared: protein content of vertically depicted cashew nut origins is significantly (p≤0.05, One-Way ANOVA) lower than cashew nuts from Ghana (a), Ivory Coast (b), Mozambique (c), Tanzania (d), Vietnam (e).*

	Raw		**Steamed**		**Fried**		**Drum-roasted**		**Average**		
	Avg.	Stdev.	Avg.	Stdev.	Avg.	Stdev.	Avg.	Stdev.	Avg.	Stdev.	Avg. compared
Benin	17.3	0.2	17.7	0.5	13.7	0.5	16.1	1.6	16.2	1.8	c, e
Brazil	16.1	1.5	16.6	0.2	16.2	0.1	16.2	0.5	16.3	0.3	c, e
Ghana	17.6	0.4	17.1	1.4	15.1	0.3	17.8	0.5	16.9	1.3	e
India	16.3	0.6	14.7	0.8	15.8	0.1	14.5	0.6	15.3	0.9	a, b, c, d, e
Ivory Coast	17.4	0.2	15.7	0.4	16.2	0.3	16.7	0.1	16.5	0.8	c, e
Mozambique	17.3	1.7	19.4	0.2	16.2	0.1	18.7	0.6	17.9	1.4	
Tanzania	15.8	0.2			17.7	0.1	16.3	0.8	16.6	1.0	c, e
Vietnam	17.5	0.6	19.2	0.5	18.2	0.4	18.8	0.1	18.4	0.8	
Average	16.9	1.0	17.2	1.7	16.1	1.4	16.9	1.5	16.8	1.8	

different origins, meaning that Ana o 1, 2 and 3 were present in the same quantities in all eight cashew nut origins. In the electrophoresis profile of proteins from cashew nuts subjected to different heat treatments however, small but significant differences could be detected (One-Way ANOVA). Levels of Ana o 1 were significantly lower in steamed cashew nuts compared to raw, fried and drum-roasted cashew nuts (all p=0.00). Also, levels of Ana o 1 were significantly lower (p=0.02) in fried cashew nuts compared to drum-roasted cashew nuts. Levels of Ana o 2 were significantly lower in raw cashew nuts compared to steamed (p=0.00), fried (p=0.00) and drum-roasted (p=0.02) cashew nuts. Also levels of Ana o 2 were significantly lower in drum-roasted cashew nuts compared to steamed cashew nuts (p=0.02). Levels of Ana o 3 were significantly lower in fried cashew nuts compared to raw (p=0.04) and drum-roasted (p=0.01) cashew nuts. Table 4.3 also shows the percentage increase or decrease in detected allergen in the heat-treated cashew nut compared to the raw cashew nut.

Besides these differences in the relative quantities of Ana o 1, 2 and 3 in the protein extracts, the presence of >70kDa proteins was clearly lower in the steamed cashew nut protein extracts (Figure 4.1). It is expected that the differences in protein composition as visualized on SDS-PAGE, result from differences in protein extractability, or from heat-induced effects on the protein (e.g. aggregation). As seen before, heating causes cashew nut [32], peanut [35] and walnut [36] proteins to become insoluble. Prolonged heating (20 and especially 24min roasting at 149˚C) can induce a decrease in solubility into a borate buffered saline buffer of Ana o 1, 2 and other proteins except Ana o 3 [32, 33]. In accordance with these results, in our study, the quantity of Ana o 3 was the highest in the drum-roasted

samples. The differences in Ana o 3 solubility upon roasting, are relatively small (3% increase) compared to the effects described by Mattison *et al.* (40% increase) [32]. This can be explained by the short heat exposure (roasting 8min 150˚C), and the moderating effect of the surrounding shell, of the cashew nuts used in our study. Mattison *et al.* applied a dark roasting heat treatment of 24min at 149˚C to de-shelled cashew nuts.

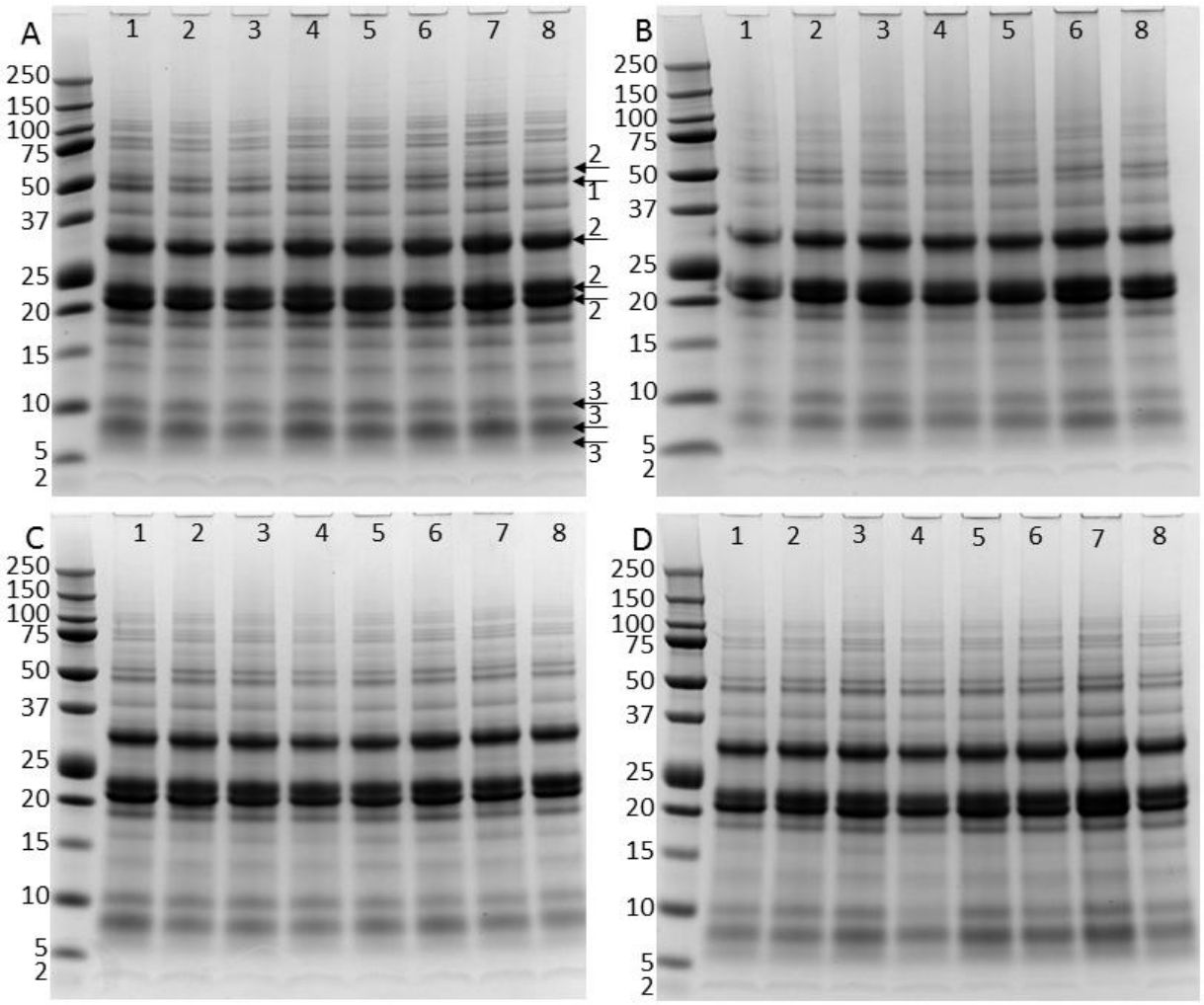


Figure 4.1 *SDS-PAGE of urea protein extracts of (A) raw, (B) steamed, (C) fried, (D) drum-roasted cashew nuts of 8 origins. All four gels, from left to right: molecular weight standard (in kDa), Benin (1), Brazil (2), Ghana (3), India (4), Ivory Coast (5), Mozambique (6), Tanzania (7, not in steamed, gel B), Vietnam (8). The position of Ana o 1, 2 and 3 are indicated by arrows in SDS-PAGE A.*

Table 4.3 *Percentage of Ana o 1, 2 and 3 in the urea protein extracts on 1D SDS-PAGE depicted in Figure 4.1. Percentages were calculated by Image Lab 4.1. Arrows indicate an increase (↑) or decrease (↓) in detected allergen in the heat-treated cashew nut compared to the raw cashew nut.*

	Benin	Brazil	Ghana	India	Ivory Coast	Mozam-bique	Tanzania	Vietnam	*Average*	*Stdev.*	compared to raw
					Raw						
Ana o 1	5.5	4.1	5.2	4.8	4.4	3.7	4.9	4.6	4.6	0.6	
Ana o 2	45.6	48.7	48.2	44.0	45.2	46.6	45.4	47.5	46.4	1.6	
Ana o 3	25.3	27.2	24.2	28.1	27.0	26.4	25.7	26.2	26.3	1.2	
					Steamed						
Ana o 1	3.1	2.9	2.9	2.1	3.7	2.2		2.9	2.8	0.5	39%↓
Ana o 2	59.8	54.8	52.5	55.1	53.8	53.6		57.6	55.3	2.5	19%↑
Ana o 3	25.8	25.6	27.8	27.1	26.6	25.2		22.3	25.8	1.8	2%↓
					Fried						
Ana o 1	4.5	4.0	4.4	3.8	4.4	3.6	4.4	3.9	4.1	0.3	11%↓
Ana o 2	48.1	53.2	59.6	62.2	51.5	51.9	51.6	54.4	54.1	4.6	17%↑
Ana o 3	28.4	23.2	14.1	15.8	26.1	24.7	23.6	23.1	22.4	4.9	15%↓
					Drum-roasted						
Ana o 1	4.9	4.7	5.0	5.5	5.2	4.1	5.0	3.9	4.8	0.5	4%↑
Ana o 2	51.6	60.2	48.5	52.0	46.8	51.5	46.2	50.1	50.9	4.4	10%↑
Ana o 3	26.0	16.1	29.4	27.9	31.0	27.4	30.1	29.1	27.1	4.7	3%↑

The IgE binding to heat-treated cashew nuts from Vietnam is shown in Figure 4.2. The control western blot showed minor non-specific binding of the antibodies to the 21kDa basic subunit of Ana o 2 as observed before [1] (data not shown).

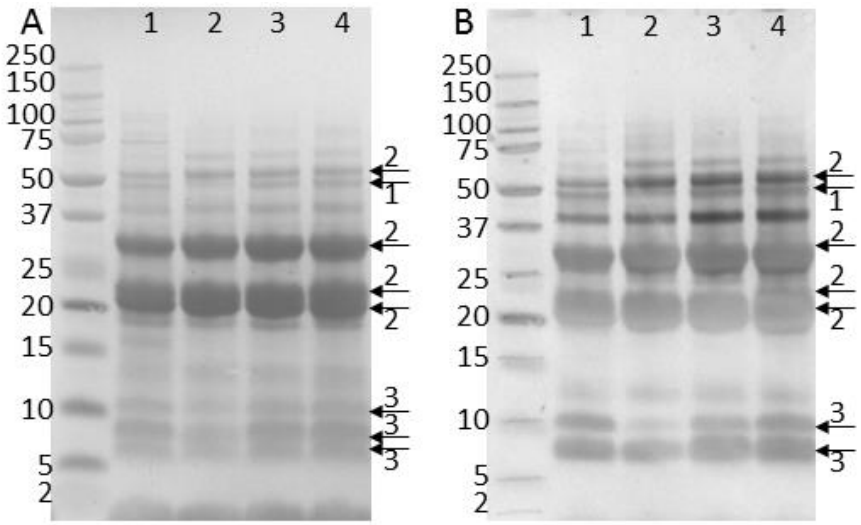


Figure 4.2 *Reversible total protein stain on nitrocellulose membrane (A) and western blot (B) of raw (1), steamed (2), fried (3) and drum-roasted (4) cashew nut proteins from Vietnam. The molecular weight standard (in kDa) is indicated on the left, arrows indicate the cashew nut allergens Ana o 1, 2 and 3.*

The western blot in Figure 4.2B shows, based on band intensity calculations of Figure 4.2A and B, the same IgE binding to the three known allergens in these four heat-treated cashew nut protein extracts. IgE binding to Ana o 1 (50kDa), Ana o 2 (53, 30, 21kDa), and Ana o 3 (10, 8kDa), as well as IgE binding to 70, 39 and 12kDa protein bands is similar between the four heat treatments. For Ana o 3, no IgE binding occurred to the 6kDa small subunit, possibly because the strongest IgE-binding epitopes are present in the large subunit for at least two out of three plasma samples [1]. IgE binding to the small subunit has been detected before [19] and seems to be more variable between persons than IgE binding to the large subunit.

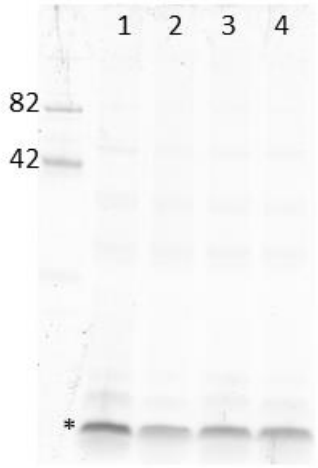


Figure 4.3 *Glycoprotein stain of raw (1), steamed (2), fried (3) and drum-roasted (4) Benin cashew nut protein extracts. The molecular weight standard (in kDa) is indicated on the left of the gels, *indicates the position of the 3kDa glycoprotein.*

Glycoprotein stain

Heating of nuts can potentially induce differences in glycation [25]. The Maillard reaction is a non-enzymatic browning reaction between proteins and non-reducing sugars which occurs faster at higher temperatures [37, 38]. Levels of glycated proteins, or Advanced Glycation End products, the products of the Maillard reaction, are therefore higher in roasted cashew nuts compared to raw cashew nuts [39].

A glycoprotein stain was performed on 1D SDS-PAGE of cashew nut proteins from the Benin origin subjected to four different heat treatments (Figure 4.3). This revealed that in all protein extracts only one protein band (3kDa) was clearly glycated/glycosylated, but not Ana o 1, 2 or 3. No differences were detected in the protein extracts of differently heat-treated cashew nuts. Glycation of heat-treated cashew nut proteins has not been described before, but was expected based on peanut studies [25]. Besides, purified cashew nut protein Ana o 1, when run at high quantity on SDS-PAGE, has shown presence of light glycosylation/glycation [1]. The relatively low levels of Ana o 1 in the protein extracts studied here, are likely too small to observe the minimal glycosylation/glycation of this protein in a total protein extract.

Based on glycoprotein staining on SDS-PAGE we do not observe differences in glycation between proteins due to the heat treatments applied to these cashew nuts. Also no

difference in the colour of the cashew nuts was observed after de-shelling and peeling. It is likely that the cashew nuts, heat-treated in the shell, were to some extent protected from the heat and the Maillard reaction by this shell. Indeed, heating in a shell decreases the level of AGEs as shown in peanut [39], an effect that is likely even stronger in cashew nuts as the cashew nut shell is rather thick and solid.

2D electrophoresis

2D electrophoresis of each of the 31 protein extracts was performed to obtain detailed information on the presence of possible isoforms of Ana o 1, 2 and 3, and to detect differences between the cashew nut origins and heat treatments. The terminology regarding isoforms is often inconsistently applied but is used here to refer to different forms of the same protein, differing in for example glycosylation or a few amino acids. The presence of multiple isoforms might be caused by allelic differences, the occurrence of alternative splicing, or the occurrence of different post-translational modifications like glycosylation, phosphorylation, deamidation, and N- or C-terminal truncation [40]. Existence of multiple isoforms for 7S and 11S globulins as well as for 2S albumins has been documented before in peanut [41].

Based on the 2D electrophoresis and subsequent western blots of purified Ana o 1, 2 and 3 proteins, and one total protein extract of drum-roasted cashew nuts from Vietnam (Figure 4.4), the allergens could be tentatively identified on all 2D electrophoresis gels (Supplementary information 1). As indicated in Figure 4.5, presumed Ana o 1 isoforms are indicated by spot number 1-9, presumed Ana o 2 isoforms are indicated by spot 10-38 with the acidic (large) subunit at 33-37kDa (spot 10-20) and the basic (small) subunit at 21-23kDa (spot 21-38), and finally presumed Ana o 3 isoforms are indicated by spot number 39-46 at a molecular weight of 10 and 8kDa.

When analyzing the results in more detail it appears that purified *Ana o 1* (Figure 4.4A) consists of four clear protein spots and five less intense protein spots. After western blotting (Figure 4.4E), IgE bound to protein spots in the same region on the blot. The individual spots can, however, not so clearly be distinguished. Besides IgE binding to these 50kDa protein spots, IgE also bound to a few minor protein spots of 30kDa, possibly indicating presence of residual Ana o 2 in this purified extract of Ana o 1. For Ana o 1, two isoforms (Ana o 1.0101 and Ana o 1.0102) are known, differing in length and displaying one amino acid substitution [2]. As these two isoforms do not differ in theoretical isoelectric point (5.6, calculated by the pI calculator, ExPASy), the differences in isoelectric points of the 9 isoforms of Ana o 1 are likely to have been caused by differences in posttranslational modifications of the isoforms.

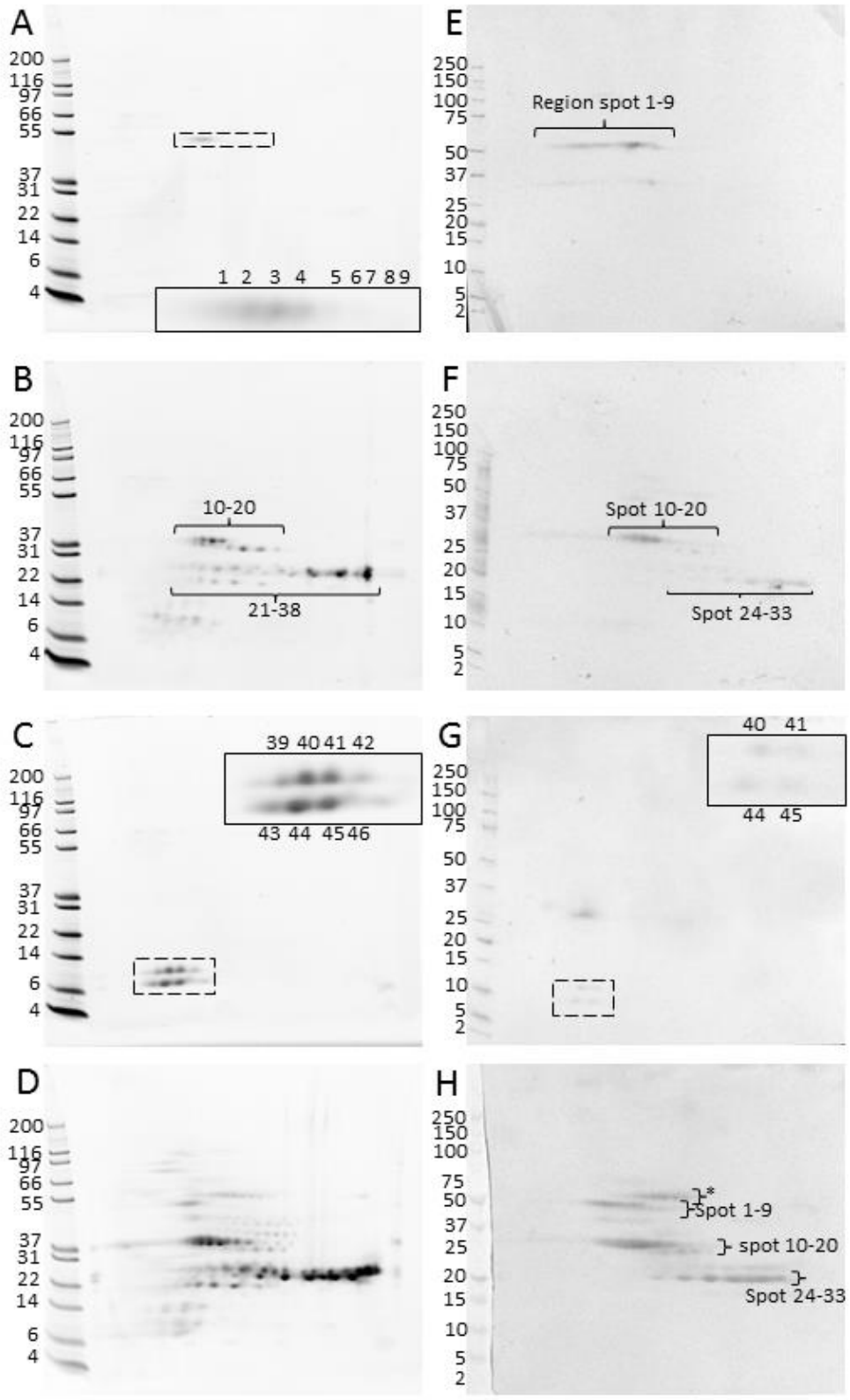


Figure 4.4 *Total protein stain (A-D) and western blot (E-H) of 2D electrophoresis of purified Ana o 1 (A, E), Ana o 2 (B, F), Ana o 3 (C, G), and a total protein extract of Vietnam drum-roasted cashew nuts (D, H). The molecular weight standard (in kDa) is indicated on the left of the gels. The 2D electrophoresis profile of Ana o 1 and Ana o 3 (dashed box) are enlarged (solid line box) in the corner of the gel (A, C) and blot (G). Protein spots and IgE binding spots and regions are indicated by numbers corresponding to Figure 4.5. *Indicates IgE binding to proteins of 55kDa.*

As the isoforms of Ana o 1 have a pI of 5.2-5.8 on 2D electrophoresis, only minor (posttranslational) differences are expected between the isoforms. Ana o 1 has 15 predicted phosphorylation sites, one N-glycosylation site, and one amidation site [42], which could cause such differences in pI between the protein isoforms. On the other hand, also the occurrence of additional alleles could explain the different isoforms of Ana o 1 observed here.

For purified *Ana o 2* (Figure 4.4B), multiple isoforms were present after 2D electrophoresis; 11 forms of the 33-37kDa acidic (large) subunit and 18 forms of the 21-23kDa basic (small) subunit. The pI range of the acidic subunit isoforms (5.2-6.3) was broader than expected (5.7). Such basic shifts on 2D electrophoresis might be caused by dethiomethyl methionine modifications [43]. The acidic subunit has 3 predicted acetylation (N-myristolyation) sites, as well as 10 phosphorylation sites. However, these posttranslational modifications would cause an acidic shift instead of the observed basic shift [44-46]. In contrast, the basic subunit of Ana o 2 showed an acidic shift on 2D electrophoresis: the protein spots migrated at a pH range of 5.2-9.3 instead of the calculated pI of 8.9. In this basic subunit of Ana o 2, 4 phosphorylation sites, as well as 1 acetylation (N-myristolyation) site were predicted [42]. Both these posttranslational modifications can induce an acidic shift in pI [44-46]. Such an acidic shift can be extensive in small basic proteins [46], such as the basic subunit of Ana o 2. Also other posttranslational modifications or a combination of different posttranslational modifications may have contributed to the large variation in protein isoelectric point. Western blotting of this 2D electrophoresis sample (Figure 4.4F) showed binding of IgE to all protein spots of the acidic subunit, but not to all protein spots of the basic subunit: IgE of the plasma pool used in this study, did not bind to the proteins in spot 21-23, and spot 34-38 while these proteins were clearly present. Possibly, these isoforms of the Ana o 2 basic subunit differ in their amino acid sequence in the epitope region. For different recombinantly produced protein isoforms from hazelnut and birch pollen (Cor a 1, Bet v 1), patient-specific variation in IgE binding patterns were reported [47, 48]. Perhaps, also in this experiment, specificity in IgE binding from different plasma samples, for specific isoforms of the protein as represented in spot 21-23, 34, and 35-38, can explain the observations. Since multiple epitopes are present on Ana o 2 [3], the difference between the protein isoforms might reflect a combination of differences in primary protein structure and variation in posttranslational processing. Alternatively, it is also possible that during the purification of Ana o 2, non-allergenic proteins of the same molecular weight were co-purified, showing up on the 2D electrophoresis gel at spot 21-23 and spot 34-38.

Ana o 3 (Figure 4.4C) appeared as 8 protein spots, 4 of 10kDa and 4 of 8kDa. The small subunit of 6kDa could not be detected on either the 2D gel nor on the western blot of purified Ana o 3 while being present in the purified extract as visualized on 1D SDS-PAGE [1].

The pI range of the large subunit of Ana o 3 is relatively small: pH 4.4-5.4 and matches with the calculated pI of 4.9. Hence, not much posttranslational or sequential diversity within this large subunit is likely. Based on the protein sequence, Ana o 3 has 5 potential phosphorylation sites, of which three are present in the protein sequence of the large subunit [42]. The small subunit of Ana o 3 has a wide diversity in isoelectric points (pI range of 6.4-10.1) based on calculation of the differently truncated isoforms [1]. This wide range of isoelectric points might cause that this subunit could not be visualized on 2D gel. On western blot (Figure 4.4G) only 4 of the 8 protein spots (spot 40, 41, 44, and 45) bound IgE, this might be due to a difference in epitopes of the Ana o 3 isoforms, specificity of the plasma samples, or impurities in the protein sample.

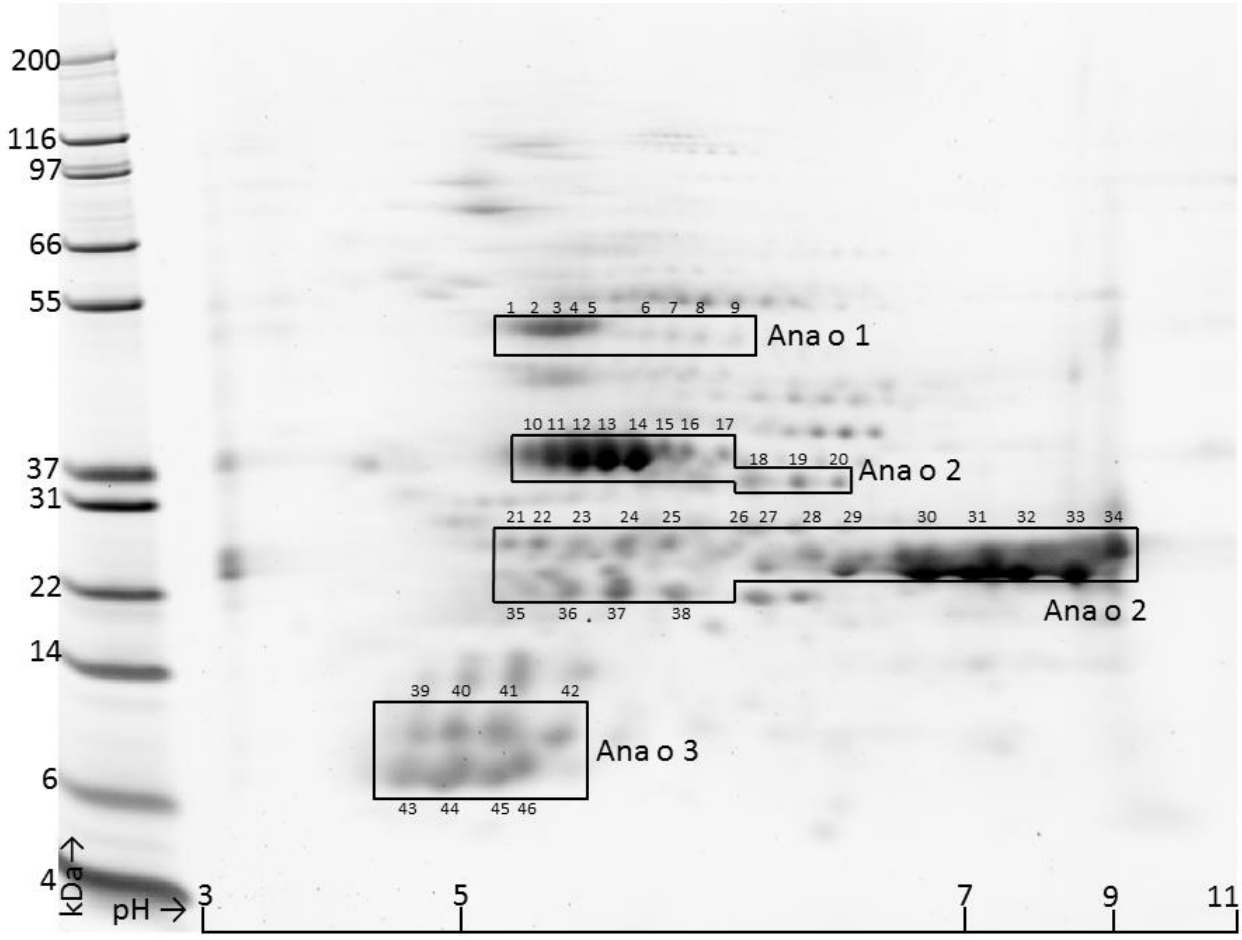


Figure 4.5 *Total protein stain of 2D electrophoresis gel of the Ghana raw cashew nut protein extract. Indicated are areas and spots with presumed isoforms of Ana o 1, 2 and 3. The molecular weight standard (in kDa) is indicated on the left of the gel.*

On the 2D electrophoresis gel in Figure 4.4D and western blot in Figure 4.4H, a total protein extract of drum-roasted cashew nut proteins from Vietnam is shown. This blot shows the same protein spots as the blots of the purified allergens, albeit that no IgE bound to Ana o 3

(Figure 4.4H), probably, as too little of this protein was present on this immunoblot and gel (Figure 4.4D, H). On 2D electrophoresis gel, drum-roasted cashew nuts from Vietnam have relatively low levels of Ana o 3 compared to drum-roasted cashew nuts of other origins, see Supplementary information 1D.

IgE clearly bound the different isoforms of Ana o 1 and Ana o 2. Besides the previously mentioned protein spots, IgE also displayed binding to multiple protein spots at ±55kDa (Figure 4.4D, H), which is likely the complete Ana o 2 monomer of 53kDa that is lost in the purification of Ana o 2 as it is not visible on either 1D [1] or 2D electrophoresis (Figure 4.4B, 3F).

When comparing all 31 2D electrophoresis gels to each other (Supplementary information 1) all tentatively identified isoforms of Ana o 1, 2 and 3 were present in each of the protein extracts, indicating no difference in presence of the isoforms of Ana o 1, 2 or 3 in cashew nuts of different origins or after exposure to different heat treatments. The spot intensity was calculated for the protein isoforms of Ana o 1, 2 and 3. In line with the results of 1D electrophoresis, no significant differences were detected between the different cashew nut origins ($p>0.5$). However, for the different heat treatments the quantities of the three allergens did differ; levels of Ana o 1 were significantly lower in steamed cashew nut protein extracts compared to fried cashew nut protein extracts ($p=0.01$). This difference in quantity of Ana o 1, that was also detected on 1D electrophoresis (Figure 4.1, Table 4.3), is due to differences in color intensity of spot 6-9. Fried cashew nut protein extracts showed higher levels of Ana o 2 than raw ($p=0.02$) and steamed ($p=0.04$) cashew nut protein extracts on 2D electrophoresis. Levels of Ana o 3 were significantly higher in fried cashew nut protein extracts compared to raw ($p=0.00$), steamed ($p=0.00$), and drum-roasted ($p=0.04$) cashew nut protein extracts. This is opposite of what was determined based on 1D electrophoresis. This discrepancy may be explained by the absence of the small subunit in the 2D electrophoresis gels while it is present on 1D electrophoresis.

Cashew nut protein profiles could be clustered per heat treatment and not per origin by means of PCA plot (Figure 4.6). Clustering of the groups was caused by the spots of Ana o 2 for the raw and the drum-roasted cashew nuts, while the combined protein spots of Ana o 1, 2 and 3 divided the steamed from the fried cashew nut group. Especially the raw cashew nut group stands out with minimal overlap from the other heat treatment groups. This PCA plot supports the statistical analysis of the 2D SDS-PAGE described above.

The differences in allergen quantity between the differently heated cashew nuts, as detected by 2D electrophoresis, can have multiple explanations. As described above, the heating steps may have caused changes in protein structure (e.g. aggregation) and protein solubility or extractability. Additionally, the tentative identification of the Ana o 1, 2 and 3 isoforms was based on the 2D electrophoresis of the purified allergens. The purification

process could have been selective for certain allergen isoforms or could still contain protein impurities that might influence the current protein isoform identifications. However, as the 2D electrophoresis patterns are highly similar between the purified allergens and the total cashew nut protein extracts on both gel and western blot, it is expected that no isoforms were missed during protein purification. Still, protein sequencing should confirm these tentative protein isoform identifications.

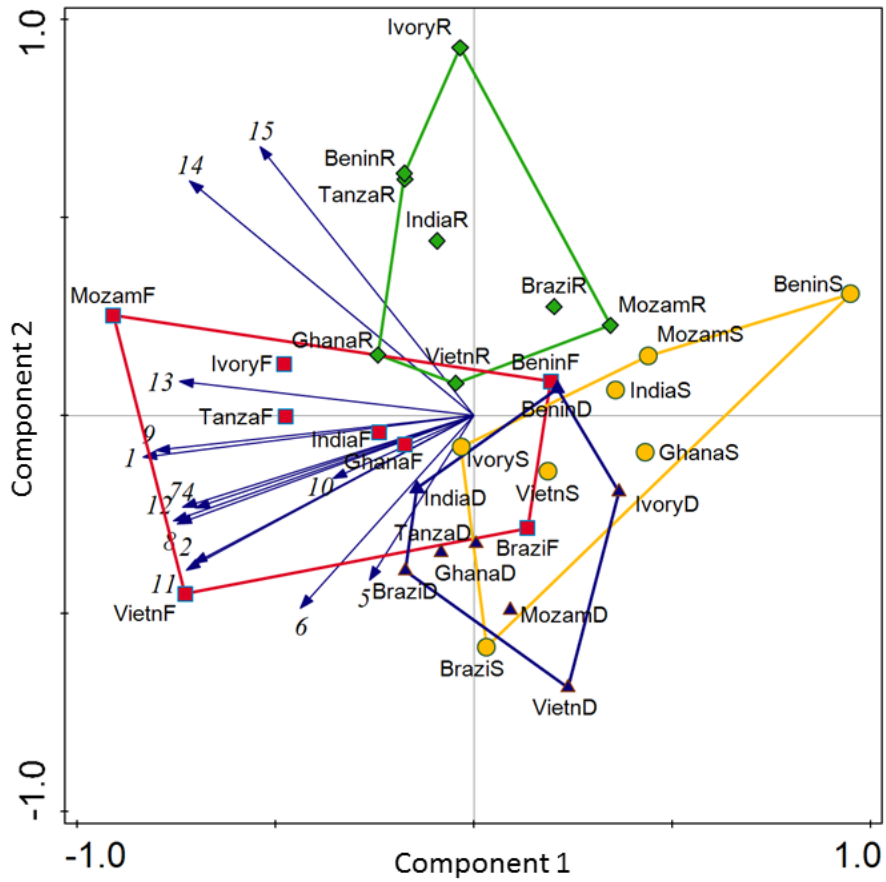


Figure 4.6 *PCA plot of spot intensities of Ana o 1, 2 and 3 on all 31 2D electrophoresis gels of cashew nut proteins from cashew nuts of diverse origins and subjected to different heat treatments. Raw (R, green diamonds), steamed (S, yellow circles), fried (F, red squares) and drum-roasted (D, blue triangles) cashew nut samples originating from Benin (Benin), Brazil (Brazi), Ghana (Ghana), India (India), Ivory Coast (Ivory), Mozambique (Mozam), Tanzania (Tanza) and Vietnam (Vietn) are shown. Grouping per heat treatment (green line raw, yellow line steamed, red line fried blue line drum-roasted cashew nut) by Canoco5. Arrows correlate to the position of the cashew nut values in the PCA plot. The arrows are derived from the data of spot 1-9 (Ana o 1, arrow 1), spot 10-20 (acidic subunit of Ana o 2, arrow 2), spot 21-38 (basic subunit of Ana o 2, arrow 5-15), and spot 39-46 (Ana o 3, arrow 4). 81.7% of the total variation can be explained by this PCA plot.*

In vitro gastric digestion

A protein extract (ammonium bicarbonate extraction buffer, prepared under non-denaturing and non-reducing conditions) from cashew nuts from Vietnam was digested by pepsin for 60 minutes as shown in Figure 4.7A. Before addition of the pepsin (Figure 4.7A, t=0, lane 1-4), Ana o 1 (50kDa), Ana o 2 (53, 30, 21kDa) and Ana o 3 (10, 8, 6kDa) were all present in the protein extracts. In the higher molecular weight range some differences were detected between the four heat treatments of the cashew nut protein extracts (Figure 4.7A lane 1-4). Multiple high molecular weight proteins (>60kDa) were present in the raw cashew nut protein extract but not in the steamed, fried or drum-roasted cashew nut protein extracts. One protein band of 72kDa seems to be absent only in the steamed cashew nut protein extract.

These differences in the protein extracts were observed repeatedly for multiple extractions and for cashew nuts of different origins. The protein electrophoresis profiles (Figure 4.7A, lane 1-4) are different from the 1D electrophoresis results of urea-extracted protein (Figure 4.1 A-D, lane 8). The increased solubility of the high molecular weight (>60kDa) proteins in the urea buffer is probably caused by the higher pH and the denaturing effect of the urea buffer [35]. The decreased solubility of high molecular weight proteins after heat treatments is similar to what has been observed before in roasted walnuts [36].
After 60 minutes of pepsin digestion (Figure 4.7A, lane 5-8), protein bands of 35kDa (pepsin), 10, 8 and 6kDa (molecular weights corresponding to Ana o 3) were visible, as well as a 5kDa band that was not detected on SDS-PAGE before digestion. This 5kDa protein likely consists of a mixture of peptides from multiple pepsin-digested proteins. Ana o 1 and 2 were degraded after 60min incubation with pepsin. This result indicates that Ana o 3 is more resistant to pepsin proteolysis than Ana o 1 and 2, as described before by Mattison *et al.* [13]. Also in other nuts and seeds the resistance of 2S albumin proteins to pepsin digestion has been shown [49-51]. The heat treatments do not seem to influence the digestibility of Ana o 3.

Besides digestion of non-reduced non-denatured heat-treated cashew nut proteins, a small experiment was performed digesting urea-extracted cashew nut proteins. These reduced denatured proteins were washed in an SGF (simulated gastric fluid) buffer to remove the urea and DTT, which interfere with pepsin activity. As shown in Figure 4.7B, Ana o 1 and 2 were again fully degraded after 60minutes digestion with pepsin, similarly as seen for the native protein extracts. In contrast to native Ana o 3, reduced and denatured Ana o 3 was largely degraded by pepsin. Only a 7kDa protein remains. As DTT reduces disulfide bridges and has been shown to thereby disrupt the alpha helical conformation of Ana o 3 [13] this

difference in digestibility between native and reduced Ana o 3 highlights the importance of the disulfide bonds and protein conformation in pepsin resistance.

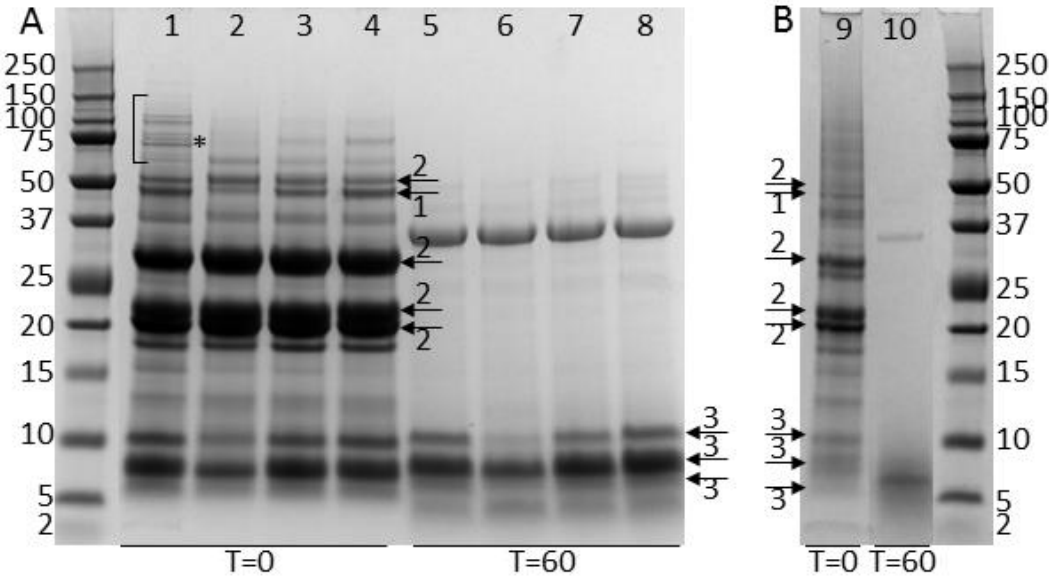


Figure 4.7 *Total protein stain on SDS-PAGE of raw (1, 5, 9, 10), steamed (2, 6), fried (3, 7) and drum-roasted (4, 8) non-reduced (A) and reduced (B) cashew nut proteins from Vietnam before (T=0, lane 1-4, lane 9) and after (t=60, lane 5-8, lane 10) 60min pepsin digestion. The molecular weight standard (in kDa) indicated on the left for SDS-PAGE A and on the right for SDS-PAGE B. [indicates the high molecular weight region (>60kDa) varying between the different heat treatments, * indicates the 72kDa protein that is absent in the protein extract of steamed Vietnam cashew nuts, arrows indicate the cashew nut allergens Ana o 1, 2 and 3.*

Micro-heterogeneity of Ana o 3

No major difference in micro-heterogeneity was detected between the raw cashew nuts of 8 different origins (Supplementary information 2). As reported earlier [1], Ana o 3 consists of a large subunit of 10 and 8kDa, and a small subunit of 6kda when visualized on reducing SDS-PAGE. The large subunit displays minimal C-terminal micro-heterogeneity while the small subunit displays both N- and C-terminal micro-heterogeneity. The highly similar pattern of Ana o 3 micro-heterogeneity in cashew nuts of 8 different origins indicates a high degree of specificity of this N- and C-terminal micro-heterogeneity. This high degree of similarity in the cashew nut genome, post-translational modification process or shift in cleavage site matches with the overall high degree of similarity between the different origins of the cashew nuts studied here.

In the present study, we have compared cashew nuts of 8 origins, and subjected (in-shell) to three different heat treatments. This is the first study on the electrophoresis profile of cashew nuts that have been heat-treated in-shell, a treatment that is always applied industrially to cashew nuts prior to de-shelling. A summary of the presented data regarding Ana o 1, 2 and 3 is shown in Table 4.4. This table shows the protein characteristics of Ana o

1, 2 and 3 after 1D and 2D electrophoresis of protein extracts from cashew nuts from different origins and subjected to different heat treatments. Minimal differences (maximum 3.1%) were detected in total protein content between cashews from different origins. Based on 1D and 2D electrophoresis no difference in the Ana o 1, 2 and 3 content of cashew nuts from these 8 different origins could be detected. Even a very specific protein characteristic, namely the micro-heterogeneity of Ana o 3, was highly similar between the cashew nuts of different origins. Cashew nut trees were distributed from Brazil to India and Mozambique in the 16th century [52]. In this short time period of 500 years limited to no variation in cashew nut allergens has occurred between these populations of cashew nut trees, as based on SDS-PAGE and 2D electrophoresis. Apparently, possible differences in soil or climate also did not strongly affect the quantities of Ana o 1, 2 and 3 in these cashew nuts.

Table 4.4 *Overview of protein and isoform characteristics of Ana o 1, 2 and 3 as based on 1D and 2D electrophoresis. Significant effects (*p<0.05, One-Way ANOVA) of heat treatments are indicated for raw (R), steamed (S), fried (F), and drum-roasted (D) cashew nut protein extracts. The small subunit of Ana o 3 is not detectable on 2D electrophoresis gel (-).*

	Characteristics	Ana o 1	Ana o 2		Ana o 3	
			Acidic subunit	Basic subunit	Large subunit	Small subunit
1D electrophoresis	Molecular weight, reduced (kDa)	50	30	21	10, 8	6
	Percentage present in urea protein extract	4.1±0.9%	51.5±4.9%		25.4±3.9%	
	Present after heat treatment*	S<R, F, D F<D	R<S, F, D D<S		F<R, D	
	Glycated/glycosylated	Only visible when purified [1]	No		No	
	Pepsin digestion resistance, 60min, 400µg native protein + 76µg pepsin	Degraded	Degraded		Resistant	
	Pepsin digestion resistance, 60min, 400µg denatured reduced protein + 76µg pepsin	Degraded	Degraded		Degraded	
2D electrophoresis	pI calculated	5.6	5.7	8.9	4.9	6.4-10.1
	pI based on 2D gel electrophoresis	5.2-5.8	5.2-6.3	5.2-9.3	4.4-5.4	-
	Number of isoforms	9	11	18	8	-
	Number of IgE reactive isoforms	9	11	11	4	-
	Present after heat treatment*	S<F	F>R, S		F>R, S, D	-

Only small differences were detected between cashew nut proteins of cashew nuts subjected to different heat treatments. On both 1D and 2D electrophoresis, the quantity of extracted Ana o 1 was lower in steamed cashew nuts compared to fried cashew nuts. The quantity of extracted Ana o 3 was, for both 1D and 2D electrophoresis, lower in fried cashew nuts compared to raw and drum-roasted cashew nuts. No difference in glycation,

digestibility or IgE binding was detected between cashew nuts subjected to different heat treatments. As the cashew nuts studied here were heat-treated within the shell, a step that occurs prior to the final roasting/frying step, most likely more extreme differences can be obtained when focusing on these secondary heat treatment steps as shown by Mattison *et al*. [32]. In further research, cellular studies such as a basophil activation test or RBL cell assay might be useful in order to detect possible effects of heat treatments on cashew nut proteins on cross-linking of receptor-bound IgE.

Acknowledgements

The authors thank Rebekah Sayers for her help with 2D electrophoresis. The authors would also like to thank H. Overkamp from Intersnack Nederland B.V. for providing the cashew nuts and applying the heat treatments to these cashew nuts.
This article has been supported by COST Action FA 1402.

Funding

This project was funded by Technology Foundation STW, Stichting Voedselallergie, Siemens Healthcare Diagnostics, HAL Allergy, Intersnack Nederland B.V., ALK-Abello B.V., and the Nederlands Anafylaxis Netwerk.

Conflict of interest

The authors declare no competing financial interest.

References

[1] Reitsma, M., et al., Purification and Characterization of Anacardium occidentale (Cashew) Allergens Ana o 1, Ana o 2, and Ana o 3. J Agric Food Chem, 2016. 64(5): p. 1191-201.

[2] Wang, F., et al., Ana o 1, a cashew (Anacardium occidental) allergen of the vicilin seed storage protein family. J Allergy Clin Immunol, 2002. 110(1): p. 160-6.

[3] Wang, F., et al., Ana o 2, a major cashew (Anacardium occidentale L.) nut allergen of the legumin family. Int Arch Allergy Immunol, 2003. 132(1): p. 27-39.

[4] Xia, L., et al., Mapping of a conformational epitope on the cashew allergen Ana o 2: a discontinuous large subunit epitope dependent upon homologous or heterologous small subunit association. Mol Immunol, 2010. 47(9): p. 1808-16.

[5] Robotham, J.M., et al., Characterization of a cashew allergen, 11S globulin (Ana o 2), conformational epitope. Mol Immunol, 2010. 47(9): p. 1830-8.

[6] Robotham, J.M., et al., Ana o 3, an important cashew nut (Anacardium occidentale L.) allergen of the 2S albumin family. J Allergy Clin Immunol, 2005. 115(6): p. 1284-90.

[7] Venkatachalam, M. and S.K. Sathe, Chemical composition of selected edible nut seeds. J Agric Food Chem, 2006. 54(13): p. 4705-14.

[8] Davoren, M. and J. Peake, Cashew nut allergy is associated with a high risk of anaphylaxis. Arch Dis Child, 2005. 90(10): p. 1084-5.

[9] Clark, A.T., K. Anagnostou, and P.W. Ewan, Cashew nut causes more severe reactions than peanut: case-matched comparison in 141 children. Allergy, 2007. 62(8): p. 913-6.

[10] Hourihane, J.O., et al., Clinical features of cashew allergy. Allergy, 2001. 56(3): p. 252-3.

[11] van der Valk, J.P., et al., Multicentre Double-Blind Placebo-Controlled Food Challenge Study in Children Sensitised to Cashew Nut. PLoS One, 2016. 11(3): p. e0151055.

[12] Sathe, S.K., et al., Biochemical characterization and in vitro digestibility of the major globulin in cashew nut (Anacardium occidentale). Journal of Agricultural and Food Chemistry, 1997. 45(8): p. 2854-2860.

[13] Mattison, C.P., C.C. Grimm, and R.L. Wasserman, In vitro digestion of soluble cashew proteins and characterization of surviving IgE-reactive peptides. Mol Nutr Food Res, 2014. 58(4): p. 884-93.

[14] Cheng, F., et al., Development and inter-laboratory transfer of a decaplex polymerase chain reaction assay combined with capillary electrophoresis for the simultaneous detection of ten food allergens. Food Chem, 2016. 199: p. 799-808.

[15] Gaskin, F.E. and S.L. Taylor, Sandwich enzyme-linked immunosorbent assay (ELISA) for detection of cashew nut in foods. J Food Sci, 2011. 76(9): p. T218-26.

[16] Aliyu, O.M. and J.A. Awopetu, Chromosome studies in cashew (Anacardium occidentale L.). African Journal of Biotechnology, 2007. 6(2): p. 131-136.

[17] Rico, R., M. Bulló, and J. Salas-Salvadó Nutritional composition of raw fresh cashew (Anacardium occidentale L.) kernels from different origin Food Science & Nutrition Early View. Food Science & Nutrition, 2015. n/a.

[18] Nagaraja, K.V. and V.M.K. Nampoothiri, Chemical Characterization of High-Yielding Varieties of Cashew (Anacardium-Occidentale). Qualitas Plantarum-Plant Foods for Human Nutrition, 1986. 36(3): p. 201-206.

[19] Teuber, S.S., et al., Characterization of the soluble allergenic proteins of cashew nut (Anacardium occidentale L.). Journal of Agricultural and Food Chemistry, 2002. 50(22): p. 6543-6549.

[20] Schmidt,H. C.G., T. Latendorf, M. Nebendahl, A. Petersen, S. Krause, M. Leippe, W. Becker, O. Janssen, 2-D DIGE analysis of the proteome of extracts from peanut variants reveals striking differences in major allergen contents. Proteomics, 2009. 9: p. 3507-3521.

[21] Krause, S., et al., Peanut varieties with reduced Ara h 1 content indicating no reduced allergenicity. Mol Nutr Food Res, 2010. 54(3): p. 381-7.

[22] Koppelman, S.J., et al., 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, 2001. 56(2): p. 132-7.

[23] Koppelman, S.J., et al., Allergenicity Attributes of Different Peanut Market Types. Food Chem Toxicol, 2016.

[24] Noorbakhsh, R., et al., Influence of processing on the allergenic properties of pistachio nut assessed in vitro. J Agric Food Chem, 2010. 58(18): p. 10231-5.

[25] Hebling, C.M., et al., Global proteomic screening of protein allergens and advanced glycation endproducts in thermally processed peanuts. J Agric Food Chem, 2013. 61(24): p. 5638-48.

[26] Vogel, L. T.H., S. Vieths, Development of a biological assay to determine the allergenic potential of foods. Journal für verbraucherschutz und Lebensmittelsicherheit, 2006. 1: p. 317-324.
[27] Maleki, S.J., et al., The effects of roasting on the allergenic properties of peanut proteins. Journal of Allergy and Clinical Immunology, 2000. 106(4): p. 763-768.
[28] Vissers, Y.M., et al., Effect of heating and glycation on the allergenicity of 2S albumins (Ara h 2/6) from peanut. PLoS One, 2011. 6(8): p. e23998.
[29] Vissers, Y.M., et al., Effect of roasting on the allergenicity of major peanut allergens Ara h 1 and Ara h 2/6: the necessity of degranulation assays. Clin Exp Allergy, 2011. 41(11): p. 1631-42.
[30] Sabato, V., et al., Basophil activation reveals divergent patient-specific responses to thermally processed peanuts. J Investig Allergol Clin Immunol, 2011. 21(7): p. 527-31.
[31] Venkatachalam, M., et al., Effects of processing on immunoreactivity of cashew nut (Anacardium occidentale L.) seed flour proteins. J Agric Food Chem, 2008. 56(19): p. 8998-9005.
[32] Mattison, C.P., B. Vant-Hullc, A. M. Vargasa, R. L. Wassermand, C. C. Grimm, Heat-induced alterations in cashew allergen solubility and IgE binding. Toxicology Reports, 2016(3): p. 244-251.
[33] Sealey-Voyksner, J., J. Zweigenbaum, and R. Voyksner, Discovery of highly conserved unique peanut and tree nut peptides by LC-MS/MS for multi-allergen detection. Food Chem, 2016. 194: p. 201-11.
[34] Minekus, M., et al., A standardised static in vitro digestion method suitable for food - an international consensus. Food Funct, 2014. 5(6): p. 1113-24.
[35] Poms, R.E., C. Capelletti, and E. Anklam, Effect of roasting history and buffer composition on peanut protein extraction efficiency. Mol Nutr Food Res, 2004. 48(6): p. 459-64.
[36] Downs, M.L., et al., Insoluble and soluble roasted walnut proteins retain antibody reactivity. Food Chem, 2016. 194: p. 1013-21.
[37] Chung, S.Y. and E.T. Champagne, Allergenicity of Maillard reaction products from peanut proteins. J Agric Food Chem, 1999. 47(12): p. 5227-31.
[38] Teodorowicz, M., et al., Effect of Maillard reaction on biochemical properties of peanut 7S globulin (Ara h 1) and its interaction with a human colon cancer cell line (Caco-2). Eur J Nutr, 2013. 52(8): p. 1927- 38.
[39] Uribarri, J., et al., Advanced glycation end products in foods and a practical guide to their reduction in the diet. J Am Diet Assoc, 2010. 110(6): p. 911-16 e12.
[40] Agrawal, P., S. Kumar, and H.R. Das, Mass spectrometric characterization of isoform variants of peanut (Arachis hypogaea) stem lectin (SL-I). Journal of Proteomics, 2010. 73(8): p. 1573-1586.
[41] Chassaigne, H., et al., Resolution and identification of major peanut allergens using a combination of fluorescence two-dimensional differential gel electrophoresis, Western blotting and Q-TOF mass spectrometry. J Proteomics, 2009. 72(3): p. 511-26.
[42] de Castro, E., et al., ScanProsite: detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins. Nucleic Acids Res, 2006. 34(Web Server issue): p. W362-5.
[43] Lengqvist, J., et al., Observed peptide pI and retention time shifts as a result of post-translational modifications in multidimensional separations using narrow-range IPG-IEF. Amino Acids, 2011. 40(2): p. 697-711.
[44] Khaldi, N. and D.C. Shields, Shift in the isoelectric-point of milk proteins as a consequence of adaptive divergence between the milks of mammalian species. Biol Direct, 2011. 6: p. 40.
[45] Zhu, K., et al., Protein pI shifts due to posttranslational modifications in the separation and characterization of proteins. Anal Chem, 2005. 77(9): p. 2745-55.
[46] Kumar, Y., et al., ProteoMod: A new tool to quantitate protein post-translational modifications. Proteomics, 2004. 4(6): p. 1672-83.
[47] Breiteneder, H., et al., Four recombinant isoforms of Cor a I, the major allergen of hazel pollen, show different IgE-binding properties. Eur J Biochem, 1993. 212(2): p. 355-62.
[48] Ferreira, F., et al., Dissection of immunoglobulin E and T lymphocyte reactivity of isoforms of the major birch pollen allergen Bet v 1: potential use of hypoallergenic isoforms for immunotherapy. J Exp Med, 1996. 183(2): p. 599-609.
[49] Pfeifer, S., et al., Cor a 14, the allergenic 2S albumin from hazelnut, is highly thermostable and resistant to gastrointestinal digestion. Mol Nutr Food Res, 2015. 59(10): p. 2077-86.
[50] Moreno, F.J., et al., Stability of the major allergen Brazil nut 2S albumin (Ber e 1) to physiologically relevant in vitro gastrointestinal digestion. FEBS J, 2005. 272(2): p. 341-52.

[51] Moreno, F.J., et al., Thermostability and in vitro digestibility of a purified major allergen 2S albumin (Ses i 1) from white sesame seeds (Sesamum indicum L.). Biochim Biophys Acta, 2005. 1752(2): p. 142-53.
[52] Dendena, B. and S. Corsi, Cashew, from seed to market: a review. Agronomy for Sustainable Development, 2014. 34(4): p. 753-772.
[53] van der Valk, J.P., et al., sIgE Ana o 1, 2 and 3 accurately distinguish tolerant from allergic children sensitised to cashew nuts. Clin Exp Allergy, 2016.

Supplementary information

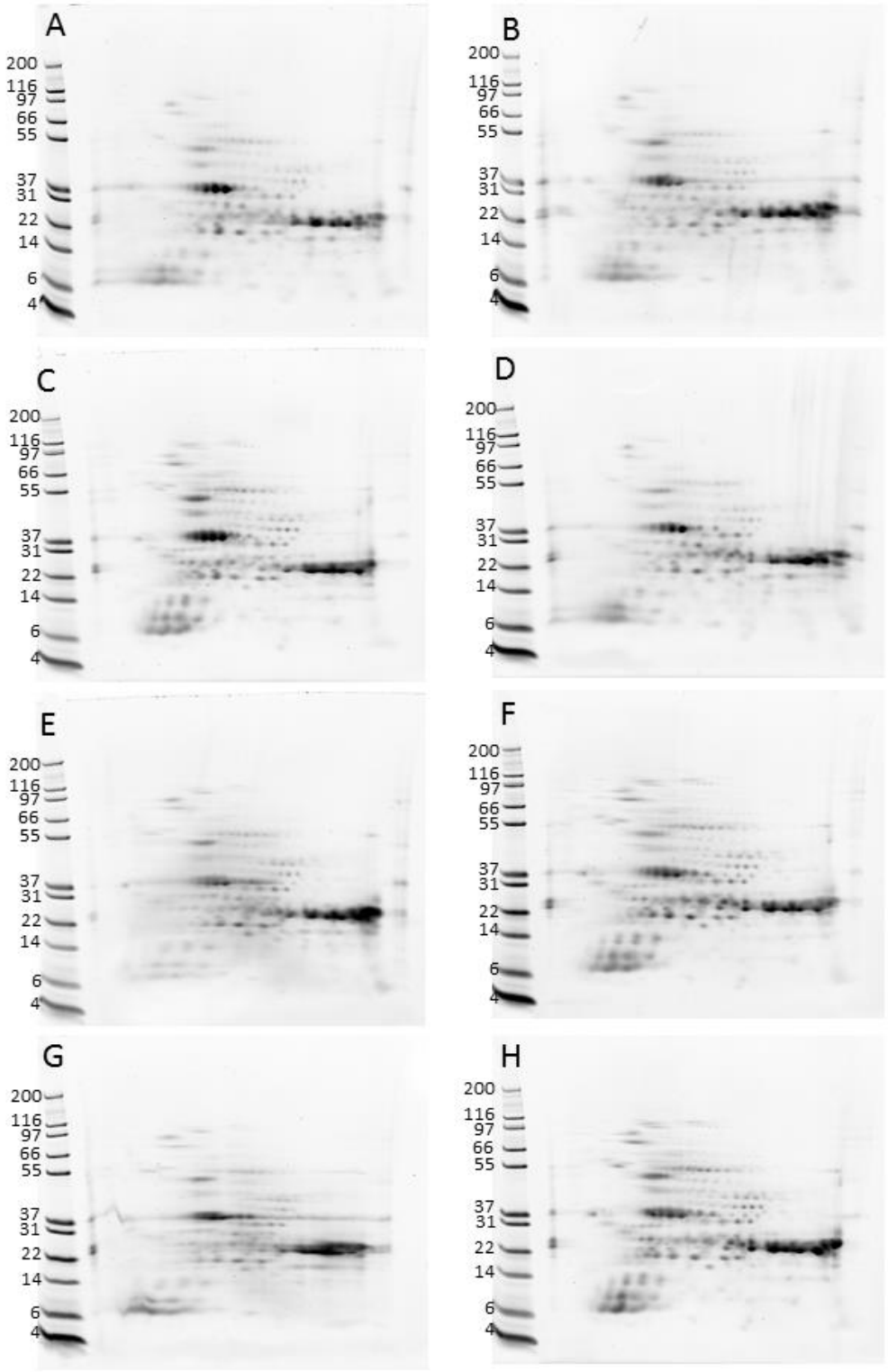


Figure 4.S1A *Total protein stain of 2D electrophoresis of raw cashew nut protein extracts from Benin (A), Brazil (B), Ghana (C), India (D), Ivory Coast (E), Mozambique (F), Tanzania (G), Vietnam (H). The molecular weight standard (in kDa) is indicated on the left of the gels.*

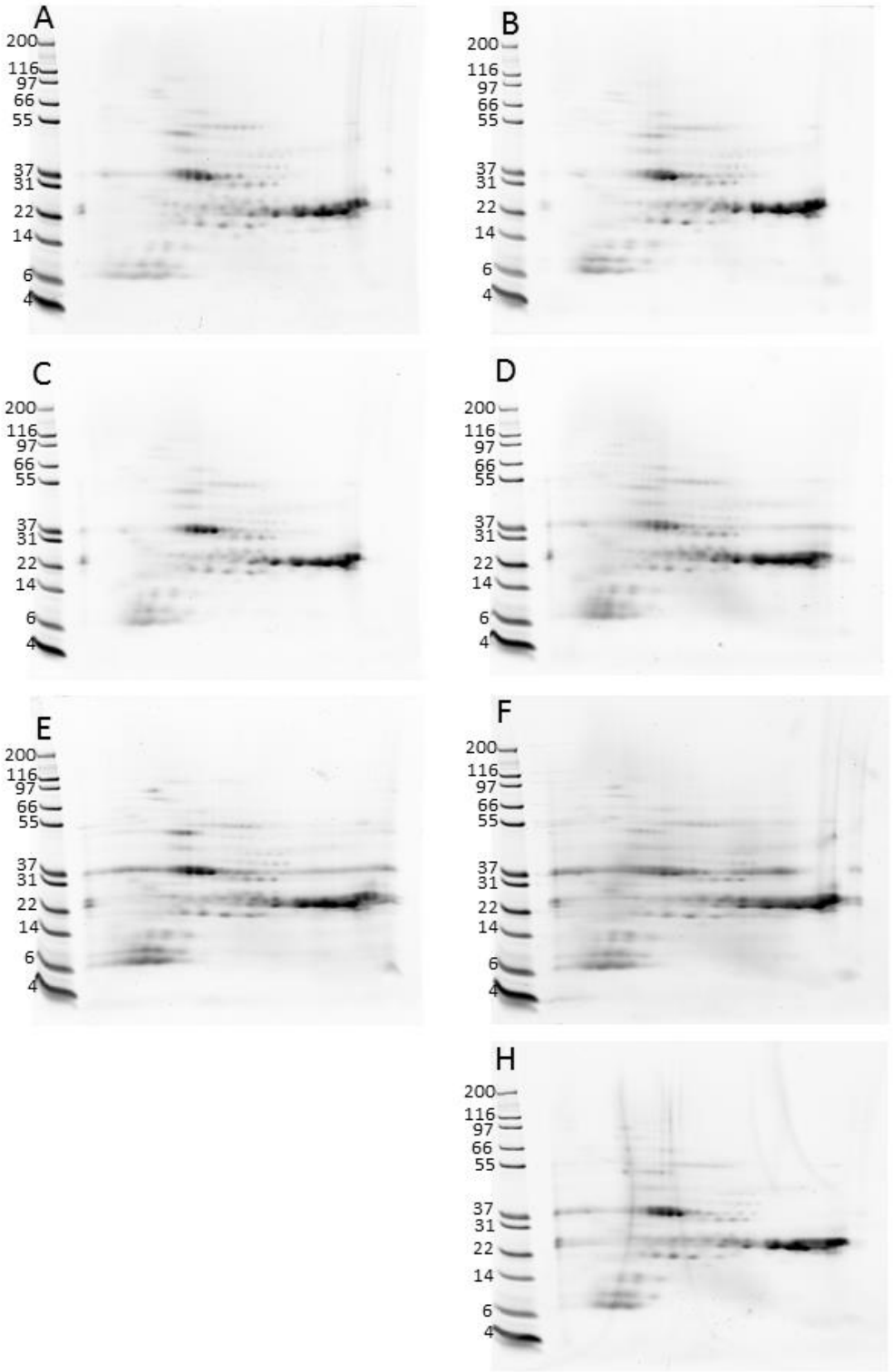


Figure 4.S1B *Total protein stain of 2D electrophoresis of steamed cashew nut protein extracts from Benin (A), Brazil (B), Ghana (C), India (D), Ivory Coast (E), Mozambique (F), Vietnam (H). The molecular weight standard (in kDa) is indicated on the left of the gels.*

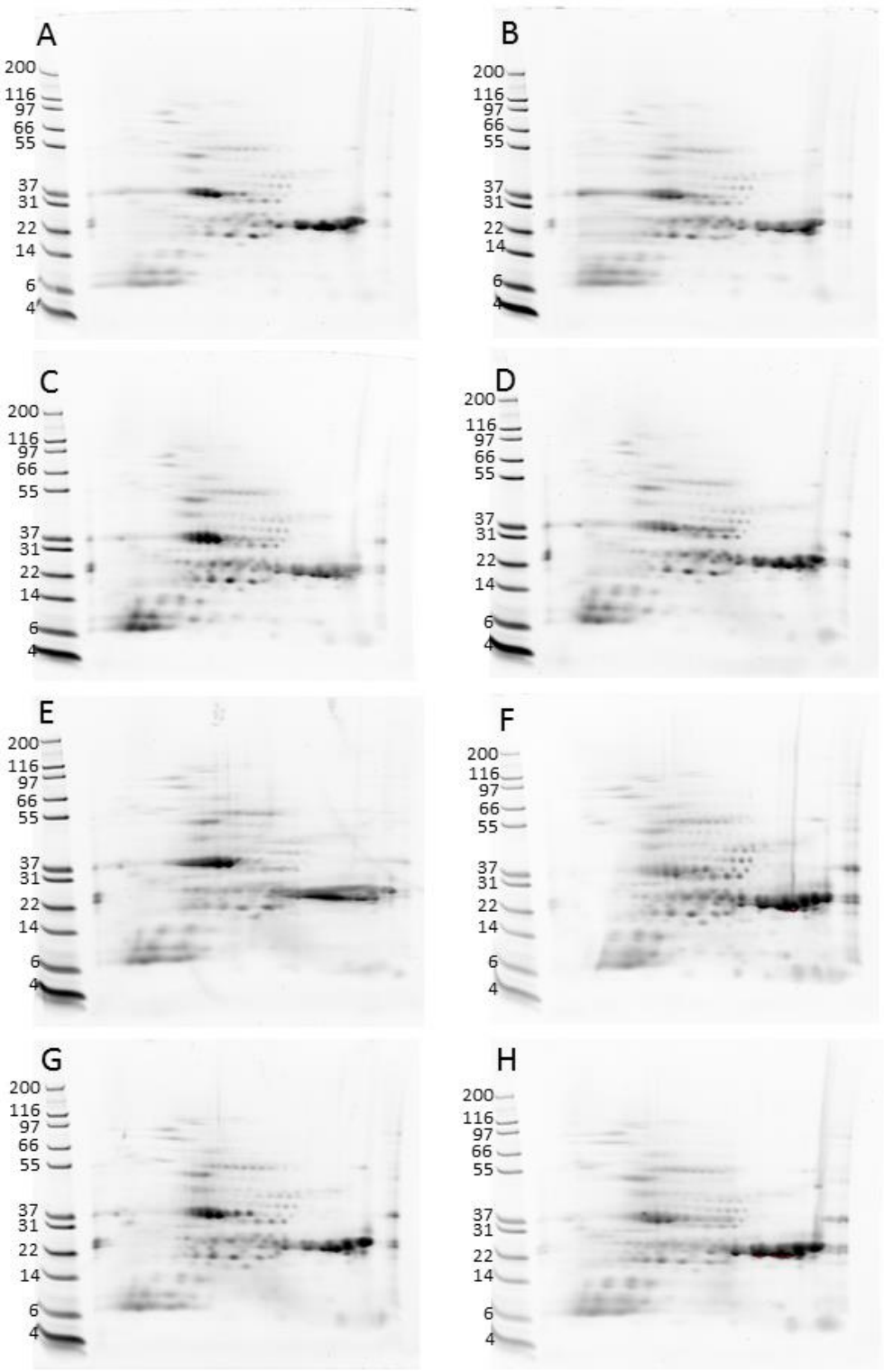


Figure 4.S1C *Total protein stain of 2D electrophoresis of fried cashew nut protein extracts from Benin (A), Brazil (B), Ghana (C), India (D), Ivory Coast (E), Mozambique (F), Tanzania (G), Vietnam (H). The molecular weight standard (in kDa) is indicated on the left of the gels.*

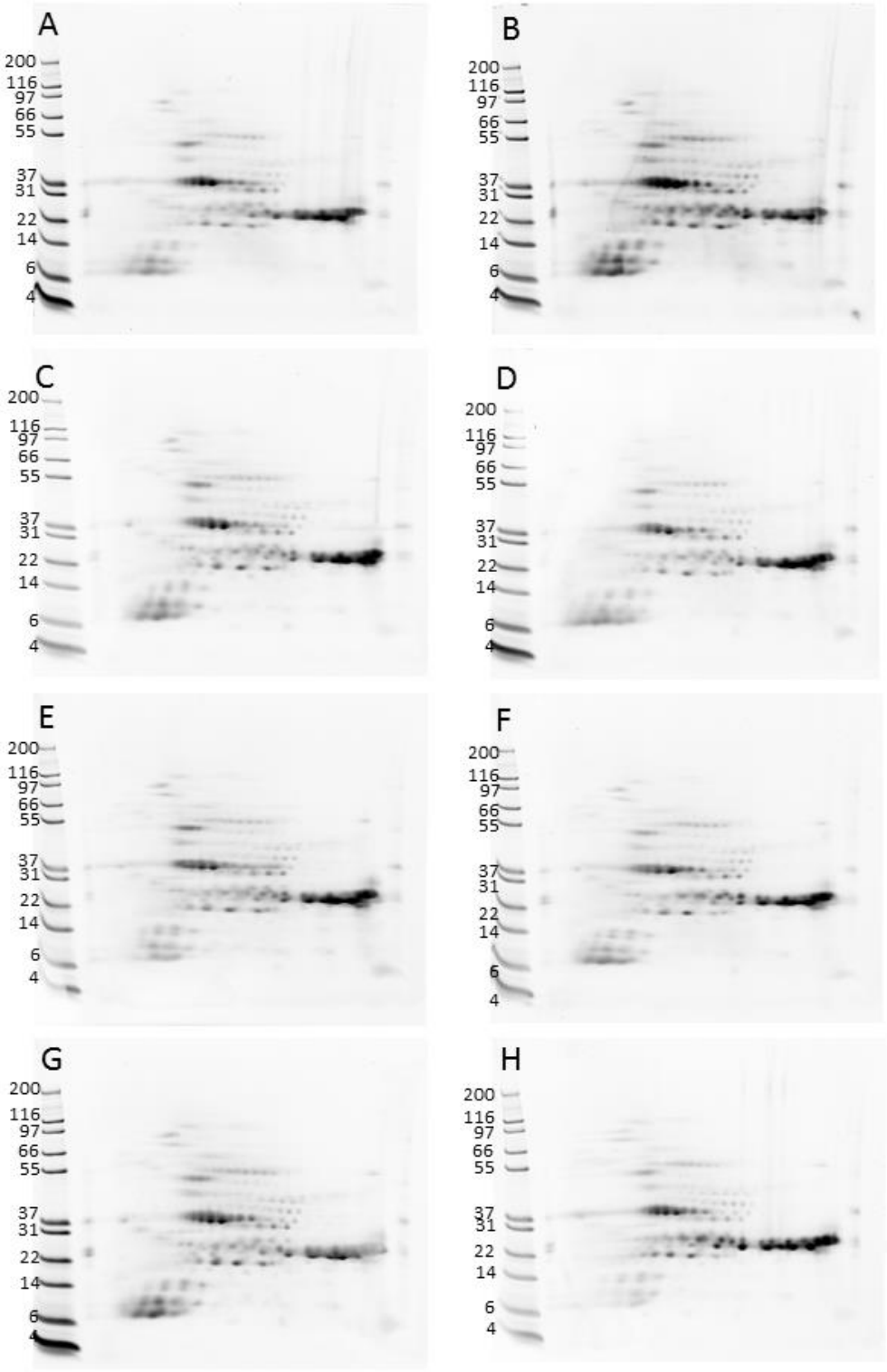


Figure 4.S1D *Total protein stain of 2D electrophoresis of drum-roasted cashew nut protein extracts from Benin (A), Brazil (B), Ghana (C), India (D), Ivory Coast (E), Mozambique (F), Tanzania (G), Vietnam (H). The molecular weight standard (in kDa) is indicated on the left of the gels.*

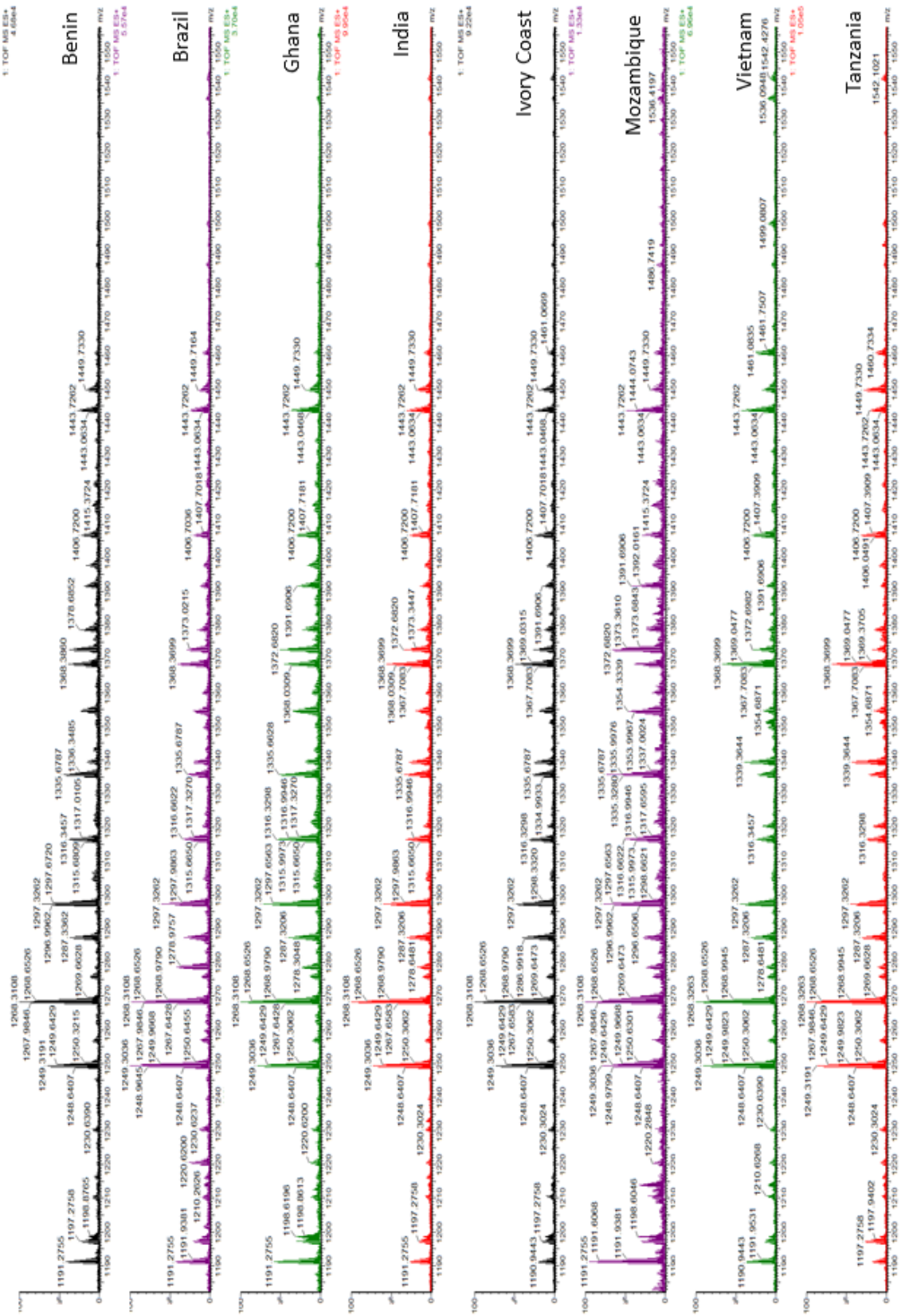
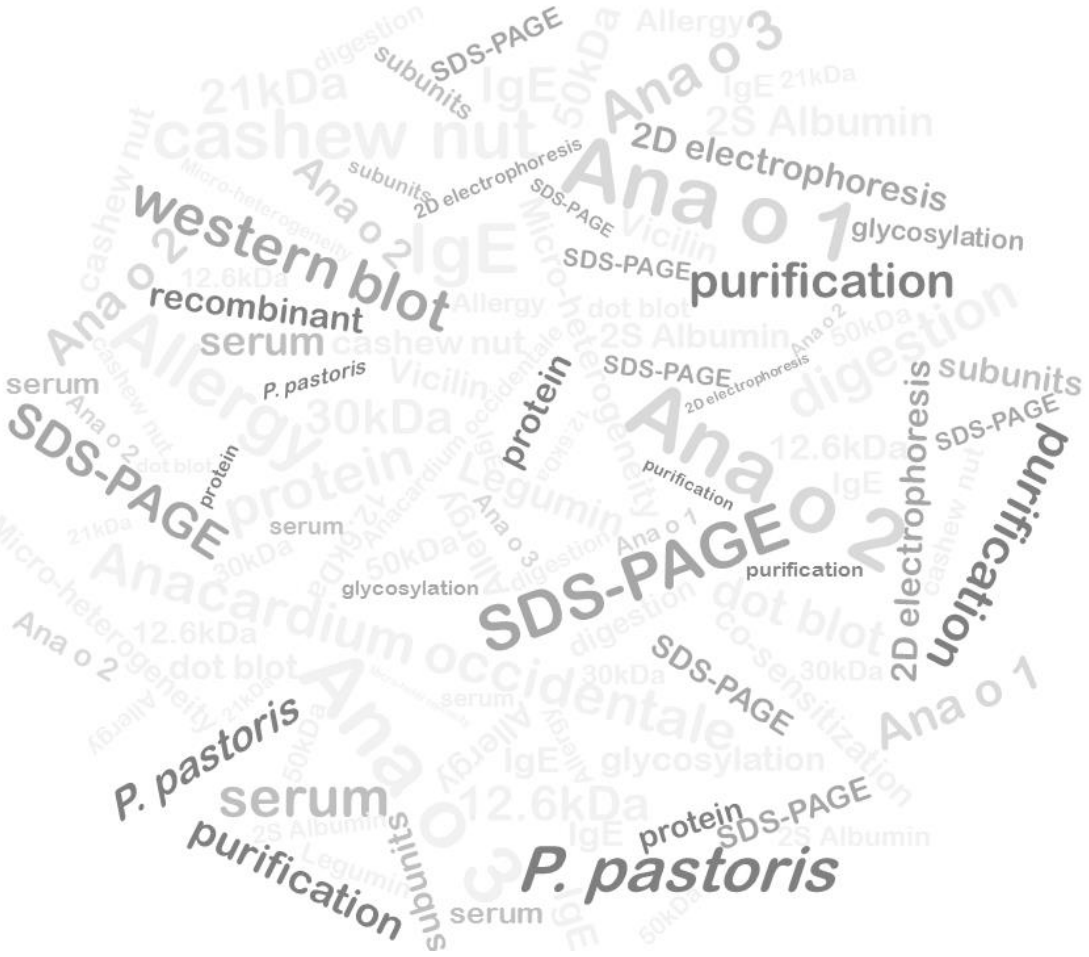


Figure 4.S-2 *Chromatographic analysis of precipitated and denatured Ana o 3 (small subunit) from raw cashew nuts of different origin.*

Chapter 5

Pichia pastoris recombinantly expressed *Anacardium occidentale* (cashew nut) allergens Ana o 1, Ana o 2 and Ana o 3

Marit Reitsma, Shanna Bastiaan-Net, Johanna P. M. van der Valk, Huub F. J. Savelkoul, Nicolette W. de Jong, Harry J. Wichers



Manuscript in preparation

Abstract

This paper describes the recombinant expression of rAna o 1, 2 and 3 by *P. pastoris*. Like the native allergens, rAna o 1 is glycosylated, rAna o 2 and 3 are not. All three recombinant proteins are produced as a mixture of complete and truncated or post-translationally cleaved polypeptides, of which some N-terminal polypeptides are lost during the purification process. Despite the truncation of some of the proteins, all three recombinant allergens were IgE reactive as shown by inhibition western blot. 2D electrophoresis showed that rAna o 1 and 3 are subjected to only minor posttranslational modifications, while rAna o 2 is more extensively modified by *P. pastoris*.

Introduction

Native, and recombinantly produced allergens have been proven valuable tools in allergy diagnostics. The usefulness of pure allergens has been well recognised in studies focussed on cross-allergenicity of related allergens [1], patient population characterization via IgE binding [2, 3], protein biochemical characterization (glycation [4], heat stability [5], digestibility[6]), and the study of (conformational) epitopes [7]. Allergens can either be purified directly from the food source or recombinantly produced in expression hosts like bacterial, yeast, insect or human cells. Recently, we have described the purification of the three major cashew allergens Ana o 1, 2 and 3 [8]. The purity of these allergens was high (96%, 93%, 99% respectively as based on SDS-PAGE analysis), however, impurities can contain other allergens (e.g. contamination of Ana o 1 with Ana o 2) or other cashew nut proteins. Using a recombinant production system this potential contamination is avoided.

Ana o 1, 2, and 3 have been recombinantly produced in BL21 (DE3) *E. coli* by Wang and Robotham *et al.* [9-11]. The resulting 55 and 65kDa rAna o 1 (obtained from 2 colonies, transformants differing in start site), 52kDa (and 120kDa dimer) rAna o 2, and 14kDa rAna o 3 were used to identify the native cashew allergens. The choice for recombinant allergen production in *E. coli* is common in the tree nut allergen field. As described by Willison *et al.* [12], most (25/27) tree nut allergens are recombinantly produced by *E. coli*, while only two allergens (Brazil nut Ber e 1, and hazelnut Cor a 8) were produced in *Pichia pastoris*. However, *E. coli* is known for its inability to perform post-translational modifications like disulphide bond formation and glycosylation, except when specific strains are used [13, 14]. *P. pastoris* on the other hand is capable of disulphide bond formation and glycosylation, albeit that the glycosylation pattern can differ from that of the native protein [13, 15]. As the major allergens from cashew nuts undergo post-translational modifications such as glycosylation (Ana o 1 [8]), and disulphide bond formation (Ana o 2 [6], Ana o 3 [8]), we chose to produce these three allergens in a *P. pastoris* expression system.

In this study we describe the protein characterization of *P. pastoris* expressed Ana o 1, 2, and 3. We compare the recombinant allergens to the purified native allergens. The three allergens were studied on reducing and non-reducing SDS-PAGE, (inhibition) western blot, and by 2D electrophoresis.

Materials and methods

Chemicals

The AccuPrime pfx polymerase, TOP 10 *E. coli* cells, pPICzα-A vector, the Fermentas enzymes EcoRI, XbaI, Sac I, and T4 DNA ligase, Pichia EasyCompTM kit, Pichia Transformation kit, biotin, HisPur Cobalt Spin Columns, MemCodeTM Reversible Protein Stain Kit, anti-c-Myc

antibody (R950-25), Mark12TM Unstained Standard, CandyCaneTM glycoprotein molecular weight standard, NuPAGE LDS sample buffer, NuPAGE MES SDS running buffer, 10% Bis-Tris mini gels, NuPAGE 1mm 4-12% Bis-Tris ZOOM gels, Sypro Ruby Protein Gel Stain, Pro-Q Emerald 300 Glycoprotein stain, Pierce™ ECL western blotting substrate were obtained from Thermo Scientific, Rockford, IL. The QIAquick PCR purification kit, and the Plasmid Mini and Midi kit were purchased from Qiagen, Hilden, Germany. Yeast nitrogen base, potassium phosphate, glycerol, sodium phosphate, NaCl, imidazole, β-mercaptoethanol, urea, thiourea, DTT, SDS, chaps, Tris, Tween-20, BSA, alkaline phosphatase conjugated goat anti-rabbit antibody (A3687), and Sigmafast BCIP/NBT were obtained from Sigma-Aldrich, St. Louis, MO. YPD, micro agar, yeast extract, peptone, and sorbitol were obtained from Duchefa Biochemie b.v., Haarlem, the Netherlands. The IPG buffer pH 3-11 NL, IPG strips (pH 3-11 nonlinear), iodoacetamide, and the 2-D Quant Kit were obtained from GE Healthcare, Piscataway, NJ. The Precision Plus Protein Dual Xtra Standard and the nitrocellulose membrane (0.2µm) were obtained from Bio-Rad, Hercules, CA. Bromophenol blue and glycine were purchased from Merck, Darmstadt, Germany. Methanol was purchased from Actu-All, Randmeer, the Netherlands, and 30kDa and 3kDa Amicon centrifugal filters were obtained from Merck Millipore, Tullagreen, Ireland. Goat anti-mouse HRP (P0447) and polyclonal rabbit anti-human IgE (A0094) antibodies were obtained from Dako, Glostrup, Denmark.

Gene construction

The nucleotide sequence of the three allergens were codon optimised for expression in *P. pastoris* and synthesised by BaseClear B.V. (Leiden the Netherlands), based on the sequences identified by Wang and Robotham *et al.*; GenBank number Ana o 1.0101 AAM73730.2 ; Ana o 2.0101, AAN76862.1; and Ana o 3.0101, AAL91665.1 [9-11]. Start and end of the sequences were based on Wang and Robotham *et al.* [9-11]. For Ana o 1 and 3, restriction sites were added, and leader peptides [9-11] were removed by PCR, using the primers mentioned below. Ana o 2 cDNA was synthesized without the predicted leader sequence. Figure 5.1 and Supplementary information 1 show the synthesised nucleotide sequences, the nucleotide sequences obtained by PCR and inserted into a piczα-A vector, and the deduced amino acid sequences as expressed by *P. pastoris*.

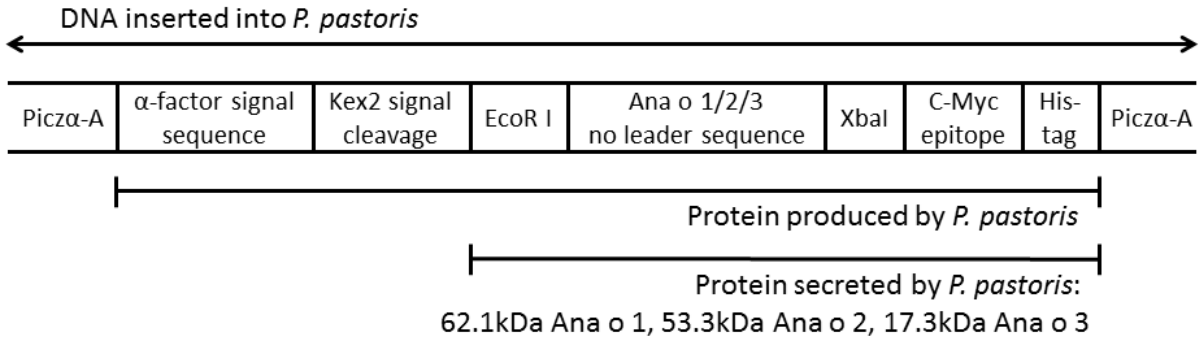


Figure 5.1. *Graphic representation of the Piczα-A vector with inserted DNA, and the recombinant proteins as produced and secreted by P. pastoris.*

PCR

For Ana o 1 and 3, synthesised nucleotide sequences were amplified and leader peptides were removed by PCR. The following primers were used (Biolegio, Nijmegen, the Netherlands): Ana o 1 forward 5'GGGGGAATTCAAGATTGACCCGGAGTTG-3', Ana o 1 reverse 5'GGGGTCTAGAAATTCATCAGCACGCC-3', Ana o 3 forward 5'GGGGGAATTCTCCATTTACCGAGCCATTG-3', Ana o 3 reverse 5'GGGGTCTAGAAAATAAGATGACTGAAACTGACAGC-3' with EcoRI and XbaI restriction sites underlined and GGGG inserted for primer stability. Using an AccuPrime pfx polymerase, cDNA was amplified in a Primus 96 plus Thermal Cycler (MWG AG Biotech, Ebersberg, Germany), for 2min 95˚C, followed by 30 cycles of 15sec 95˚C, 30sec 57.5˚C, and 2min 68˚C for Ana o 1; 2min 95˚C, followed by 5 cycles of 15sec 95˚C, 30sec 56˚C, 1min 68˚C, followed by 30 cycles of 15sec 95˚C and 90sec 68˚C for Ana o 3. Amplified cDNA was purified from a 1% agarose gel by the QIAquick PCR purification kit.

Transformation of *Pichia pastoris*

The nucleotide sequence of Ana o 2 was synthesised in a pUC57 vector by BaseClear B.V. and by heat-shock inserted into TOP 10 *E. coli* cells. After overnight growth of the Ana o 2-pUC57 transformed *E. coli,* cDNA was isolated using a Plasmid Mini Kit.

cDNA of Ana o 1, 2 and 3 was cut (EcoRI and XbaI) and ligated (T4 DNA ligase) into a pPICzα-A vector. TOP 10 *E. coli* cells were transformed with the plasmid by heat-shock, and zeocin resistant transformants were verified by sequencing (BaseClear B.V.). In accordance to the protocol of the Pichia EasyCompTM Kit, midiprepped plasmid (Plasmid Midi Kit) was treated with restriction enzyme SacI, and, after a heat inactivation step of 20min 65˚C, used for transformation into competent *P. pastoris* cells of strain X33 by use of the EasyCompTM *Pichia* Transformation kit. Transformed cells were plated onto YPDS plates (4% YPD, 2% agar, 16% sorbitol) containing increasing Zeocin concentrations (100, 250, 500, 750, and

1000ug/ml) and three colonies from the 500 and 750ug/ml plates were verified by colony PCR.

Expression and purification of recombinant proteins

The colony expressing most protein out of three colonies was picked for Ana o 1, Ana o 2 and Ana o 3 and expression was optimized for time and methanol induction. Single colonies were grown in 25ml BMGY (BMY (1% yeast extract, 2% peptone, 1.34% yeast nitrogen base, 4 × 10-5%biotin, 100mM potassium phosphate pH 6) with 1% glycerol) with Zeocin selection, at 30˚C for 24hr 230rpm. Subsequently, the cultures were centrifuged (5min 3000g, room temperature) and the pellet was re-suspended in 100ml BMMY at an OD600 of 1.0. The cultures were grown in BMMY (BMY supplemented with methanol) containing 1% methanol for expression during 96hr (rAna o 1, rAna o 2) or 24hr (rAna o 3) at 30˚C 230rpm. During the expression period methanol was added to the desirable concentration every 24hr. A control colony of empty *P. pastoris* strain X33 was grown for either 24hr in BMGY+24hr in BMMY or 24hr in BMGY+96hr in BMMY (the BMMY contained 0.5% methanol as carbon source, higher levels than 0.5% was toxic for empty *P. pastoris*), at 30˚C 230rpm. Supernatant was collected by centrifugation at 5000g for 30min at 4˚C, and concentrated in a 50ml 30kDa (rAna o 1, rAna o 2) or 3kDa (rAna o 3) Amicon centrifugal filter.

Recombinantly expressed proteins were purified in 3ml HisPur Cobalt Spin Columns according to the manufacturers' instructions. In short maximally 20mg protein was applied onto an equilibrated column (50mM sodium phosphate, 300mM NaCl, 10mM imidazole, pH 7.4), followed by 30min mixing, washing with equilibration buffer, and elution of the His-tagged proteins (50mM sodium phosphate, 300mM NaCl, 150mM imidazole, pH 7.4). All steps were performed by gravity flow. Eluted His-affinity-purified proteins were concentrated in a 0.5ml 3kDa Amicon centrifugal filter.

Protein quantification

The protein content of the *P. pastoris* supernatant and the purified recombinant proteins was determined by the 2-D Quant Kit according to the manufacturers' instructions.

1D SDS-PAGE

Denatured samples containing 5µg protein (10% β-mercaptoethanol in NuPAGE LDS sample buffer, 5min 100˚C) were run on SDS-PAGE as described earlier [8]. Mark12TM Unstained Standard was used for Sypro Ruby stained gels, CandyCaneTM glycoprotein molecular weight standard for Pro-Q Emerald stained gels, and the Precision Plus Protein Dual Xtra Standard was used for western blots. Gels were stained o/n using a Sypro Ruby Protein Gel Stain for a total protein stain or a Pro-Q Emerald 300 Glycoprotein stain to stain

glycoproteins. Imaging and analysis were performed using a Universal Hood III and Image Lab 4.1 software (both Bio-Rad).
Non-Reducing SDS-PAGE was performed as described above, but during the sample preparation no β- mercaptoethanol was added.

2D electrophoresis

IPG strips (pH 3-11 nonlinear) were rehydrated overnight with 8µg protein mixed with 125µl rehydration buffer (5.6M urea, 1.6M thiourea, 20mM DTT, 1.6% chaps w/v, a few grains of bromophenol blue, and 200x diluted IPG buffer pH 3-11 NL). Isoelectric focussing was performed on a protean IEF cell (Bio- Rad): 300V 0.2kvhr linear, 1000V 0.3kVhr rapid, 5000V 4kVhr rapid, 5000V linear 2kVhr. The focussed strips were reduced for 15min in 50mM Tris, 20mg/ml SDS, 10mg/ml DTT, 30% glycerol, and a few grains of bromophenol blue. Subsequently, the strips were alkylated for 15min in 50mM Tris, 20mg/ml SDS, 25mg/ml iodoacetamide, 30% glycerol, and a few grains of bromophenol blue. The strips were run on a NuPAGE 1mm 4-12% Bis-Tris ZOOM gel alongside a molecular weight marker (Mark12TM Unstained Standard). Electrophoresis was performed at 160V for 55min in NuPAGE MES SDS running buffer. Gels were stained o/n with Sypro Ruby Protein Gel Stain according to the manufacturers' instructions.

Western blotting

Transfer of proteins from 1D and 2D SDS-PAGE gel to a nitrocellulose membrane (0.2µm), was performed using a criterion blotter (Bio-Rad) in a cold Tris-glycine buffer (25mM Tris, 190mM glycine, 0.1% SDS, 20% methanol) for 36min at 70V. Transferred proteins were visualised and destained using the MemCodeTM Reversible Protein Stain Kit according to the manufacturers' instructions. Membranes were washed with TBS-T (50mM Tris, 150mM NaCl, 0.1% Tween-20, pH 7.4), and blocked with 3% BSA in TBS (TBS-T without Tween-20) for 1hr at 4 degrees.
For the c-Myc western blots, after washing with TBST and TBS, the blots were incubated with mouse anti-c-Myc antibody, 1:5000 in TBS o/n. Subsequently, the blots were washed with TBST and TBS and incubated for 1hr with goat anti-mouse HRP antibody diluted 1:1000 in TBS. Next, the blots were washed with TBST and TBS, covered with ECL western blotting substrate and images were taken by a Universal Hood III. The antibody control blot was performed as described above but with omission of the first antibody, instead incubation was performed with TBS.
For the serum western blots, the blocked membranes were incubated o/n using a serum pool diluted 1:5 in TBS, followed by washing in TBST and TBS, and incubating for 1hr using polyclonal rabbit anti-human IgE antibody diluted 1:15000. This was followed by washing in TBST and TBS and incubating for 1hr using alkaline phosphatase conjugated goat anti-rabbit

antibody diluted 1:20.000 in TBS, washing in TBST and TBS, and staining by Sigmafast BCIP/NBT for 6min.

Inhibition western blotting was performed as stated above for serum western blots, however, the plasma pool was pre-incubated with 40µg (40µl) His-affinity-purified Ana o 1, 40µg (100µl) His-affinity-purified Ana o 2, 25µg (50µl) His-affinity-purified Ana o 3, or 40µl (equal to the volume of 40µg rAna o 1) His affinity-purified X33 control. Pre-incubation of a 100µl serum pool with 25-40µg recombinant protein, was performed for 2hr shaking at room temperature.

Analysis of all blots was performed using the Image Lab 4.1 software.

Serum

The serum pool used for western blotting was composed of 13 sera from cashew-allergic patients as diagnosed by a double-blind placebo-controlled food challenge, as described in the multicentre prospective study "Improvement of diagnostic methods for allergy assessment" with cashew allergy in children as a showcase (trial number NTR3572, medical ethical approval number 2012-125, Erasmus Medical Centre Rotterdam) [16]. The plasma pool was composed from 3 patients with self-reported cashew nut or tree nut allergy and high cashew nut IgE titers. The plasma samples were obtained from PlasmaLab International (Everett, WA). Both the serum pool and the plasma pool were obtained by pooling the samples in equal volume. For patient characteristics, see Table 5.1.

Protein identification

Protein identification by MALDI-MS/MS was performed by Alphalyse Denmark as described earlier [8].

Table 5.1 *Clinical characteristics of the serum and plasma pool used for western blotting. Serum samples are from cashew nut allergic subjects diagnosed by double blind placebo controlled food challenges [3]. Plasma samples of cashew[*] and tree nut[**] allergic subjects obtained from PlasmaLab International.*

Serum pool		**Plasma pool**	
patient #	sIgE cashew (kU/L)	patient #	sIgE cashew (kU/L)
1	58.3	1[**]	19.2
2	100	2[*]	>100
3	10	3[**]	61.7
4	32		
5	100		
6	32.6		
7	22.5		
8	21.1		
9	11.2		
10	7.9		
11	5.83		

Results and discussion

Expression and purification of recombinant Ana o 1, 2, 3

Recombinant Ana o 1, 2 and 3 (rAna o 1, 2, 3) were produced without their leader peptide, following the example of Wang and Robotham *et al.* [9-11]. This was done as previous attempts of cloning and protein expression of rAna o 1 and rAna o 3 by *P. pastoris* failed when the leader peptide was present (data not shown). The inability of *P. pastoris* to produce recombinant proteins when the leader peptide is present has been shown before [17], and might be explained by hindrance of transcription/translation or instability of the protein after production.

rAna o 1, 2 and 3 were produced in *Pichia pastoris* and concentrated using a 30kDa (rAna o 1, 2) or 3kDa (rAna o 3) molecular weight cut-off. The cut-off of 30kDa was chosen for rAna o 1 and 2 in order to speed up the concentration process and to already remove small proteins originating from *P. pastoris*. The recombinant proteins were purified from the concentrated *P. pastoris* supernatant by His-affinity-purification. The purified proteins were studied on reducing and non-reducing SDS-PAGE, (inhibition) western blot, and 2D electrophoresis. The expressed proteins will be discussed one by one. An overview of all the results is shown in Table 5.2.

Controls

The *P. pastoris* control colony, not containing any inserted cashew DNA, was used in order to check the purification procedure, and was used as a negative control for the different western blots. Figure 5.2A shows the presence of multiple protein bands of 14kDa to >200kDa in the *P. pastoris* supernatant before His-affinity purification. His-affinity purification removes all proteins from this extract (lane ctrl). Therefore it is expected that for the rAna o 1, 2 and 3 colonies no significant amounts of *P. pastoris*-specific proteins will elute along with the His-tagged rAna o 1, 2 and 3 proteins. As expected [18], the larger molecular weight proteins from the *P. pastoris* supernatant, 50kDa to >180kDa proteins, are glycosylated (Figure 5.2B). Figure 5.2C shows no binding of the c-Myc antibody to the control samples as no recombinant His- and c-Myc-tagged proteins are produced by this control colony. Serum IgE (Figure 5.2D) does bind to a 230kDa *P. pastoris* protein, however, this only occurs in the supernatant and not in the His-affinity purified samples, confirming the specificity of the His-affinity purification. As a positive control for the western blots of rAna o 1, 2 and 3, a total cashew nut protein extract was applied on western blot (Figure 5.2D). In this western blot, serum IgE binding to native Ana o 1, 2 and 3 is confirmed: 53kDa monomeric Ana o 2, 50kDa Ana o 1, the 20-22kDa basic (small) subunit of Ana o 2, the 33kDa acidic (large) subunit of Ana o 2, and the 10 and 8kDa large subunit of Ana o 3 [8]. Finally, in

Figure 5.2E an inhibition western blot is shown. A positive total cashew nut protein control is included, indicating binding of the plasma IgE to native Ana o 1, 2 and 3. Pre-incubation of the plasma IgE with the His-affinity-purified control sample results in a decreased, but not completely inhibited, IgE binding to 53kDa Ana o 2 (the complete monomer), 50kDa Ana o 1, and a 39kDa protein. No decrease in IgE binding was detected to the acidic and basic subunits of Ana o 2 or to Ana o 3. Overall, Figure 5.2 shows by SDS-PAGE and western blot, that during the His-affinity purification no significant amount of *P. pastoris* proteins contaminate the His-affinity-purified proteins, and that although some *P. pastoris* proteins (230kDa) bind IgE, these do not contaminate the His-affinity-purified samples.

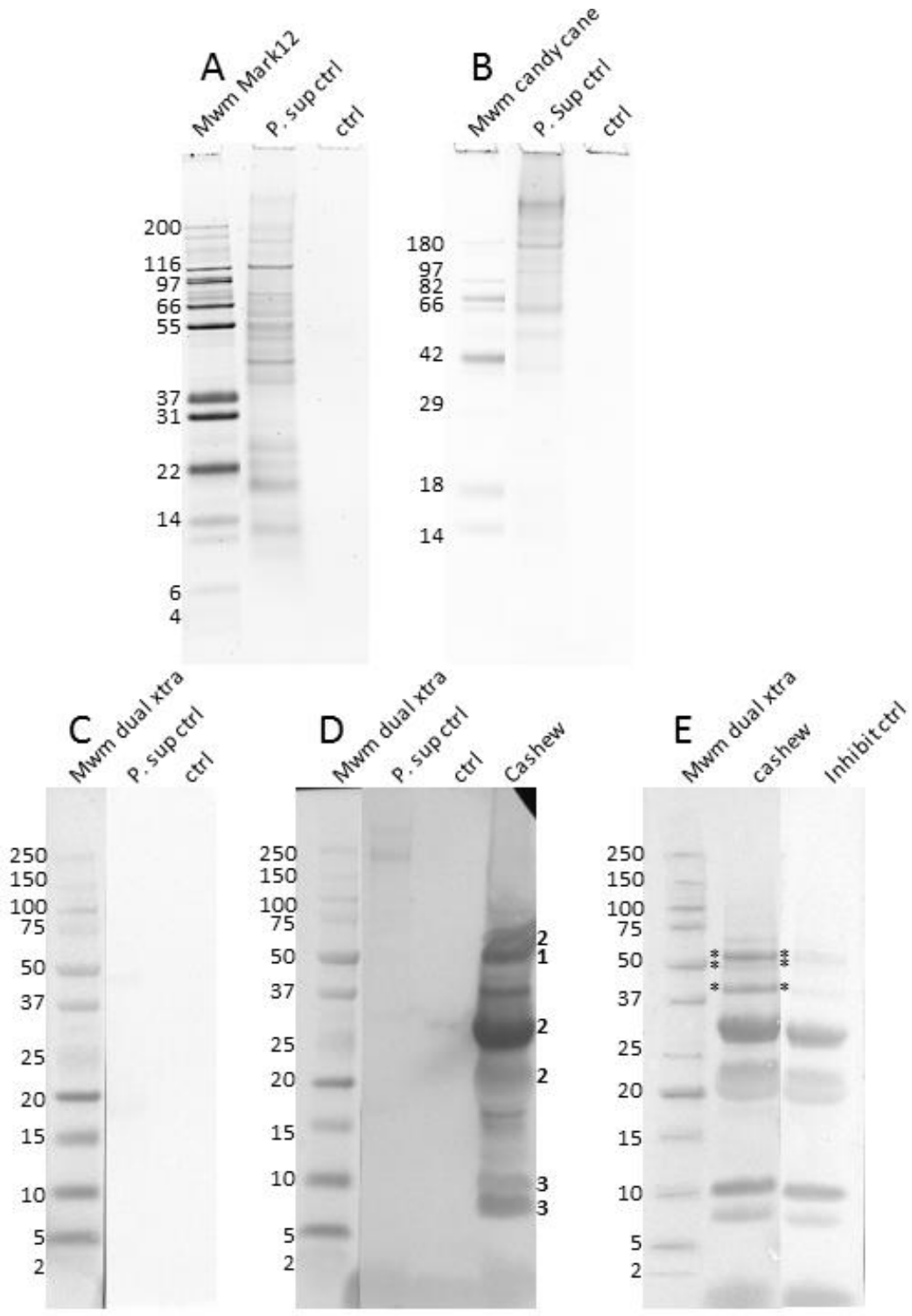


Figure 5.2 *Sypro Ruby total protein stain (A), Emerald glycoprotein stain (B), c-Myc western blot (C), and serum western blot (D) of 3kDa concentrated P. pastoris supernatant of control colony (P. sup ctrl), and His-affinity-purified control (ctrl). Inhibition western blot (F) with IgE binding to a total cashew nut protein extract either not inhibited (cashew lane), or inhibited by pre-incubation of the plasma pool with an equal volume of His-affinity-purified control as used for inhibition with rAna o 1 and 2 (Inhibit ctrl lane). The molecular weight in kDa is indicated on the left of the figures, the bold numbers on the right of Figure D indicates the protein bands of native Ana o 1, 2 and 3, * in Figure E indicates the inhibited proteins.*

Ana o 1

Native Ana o 1 is a 7S globulin of 50kDa [8], forming trimers in non-reduced state [19]. Recombinantly produced Ana o 1 from *E. coli* has a molecular weight of 65 and 55kDa, depending on the location of the start site [9]. The *P. pastoris* recombinantly produced Ana o 1 described in this paper is produced from the same nucleotide sequence as the 65kDa Ana o 1 that was expressed in *E. coli*. When calculated based on the amino acid sequence of the expected protein, rAna o 1 is expected to be 62kDa including c- Myc epitope and Histidine tag.

Figure 5.3A shows on reducing SDS-PAGE, that *P. pastoris* colony rAna o 1 produces a protein of 51kDa besides *P. pastoris* proteins of 14kDa to >200kDa. His-affinity-purified rAna o 1 contains the same 51kDa protein, as well as several less prevalent proteins of 105 and 30kDa. MALDI-MS/MS protein identification confirmed the 51kDa protein as Ana o 1, matching amino acids from E111 up to the C-terminus, including the Histidine-tag. The 30kDa protein was also identified as Ana o 1, however, only C-terminal peptides (I202-H537) were identified. A 17kDa protein from the *P. pastoris* supernatant was also identified as Ana o 1, however, only N-terminal peptides were identified (Q19-R98). The protein identification results are shown in Supplementary information 2. The 51kDa protein was identified as a glycoprotein as determined by a glycoprotein stain (Figure 5.3B). Native Ana o 1 is also glycosylated [8], however, *P. pastoris* might glycosylate this protein in a different manner [20].

Non-reduced (Figure 5.3C) rAna o 1 clearly showed higher quantities of the 105kDa protein in addition to the 51kDa protein. This 105kDa protein is expected to be a dimeric form of rAna o 1, aggregated under non-reducing conditions. Next to these proteins, smaller proteins of 25-30kDa are visible, indicating that the 30kDa C-terminal rAna o 1 polypeptide is not covalently attached by disulphide bridges, to an N-terminus in non-reducing state.

On western blot (Figure 5.3D) the anti c-Myc antibody binds to the 105kDa rAna o 1 dimer, the 51kDa rAna o 1, and the 30kDa C-terminal rAna o 1, confirming these are either full-length rAna o 1, or C-terminal rAna o 1. The 17kDa protein does not bind to the anti c-Myc antibody, confirming this protein does not contain the C-terminus of rAna o 1.

On western blot (Figure 5.3E), serum IgE binds to the 51 and 17kDa proteins, confirming the immunoreactivity of these proteins. The 105kDa rAna o 1 dimer on the other hand does not appear on the serum western blot, likely due to the relatively low abundance of this protein, combined with the lower sensitivity of the serum western blot compared to the c-Myc western blot. Also the 30kDa C-terminal rAna o 1 is not visible on this serum western blot. The relatively weak binding of IgE to the 30kDa C-terminal protein, compared to the stronger IgE binding to the 17kDa protein, while both proteins are more or less equally present on SDS-PAGE, indicates that, for the sera used, most IgE epitopes are present on the N-terminus of rAna o 1.

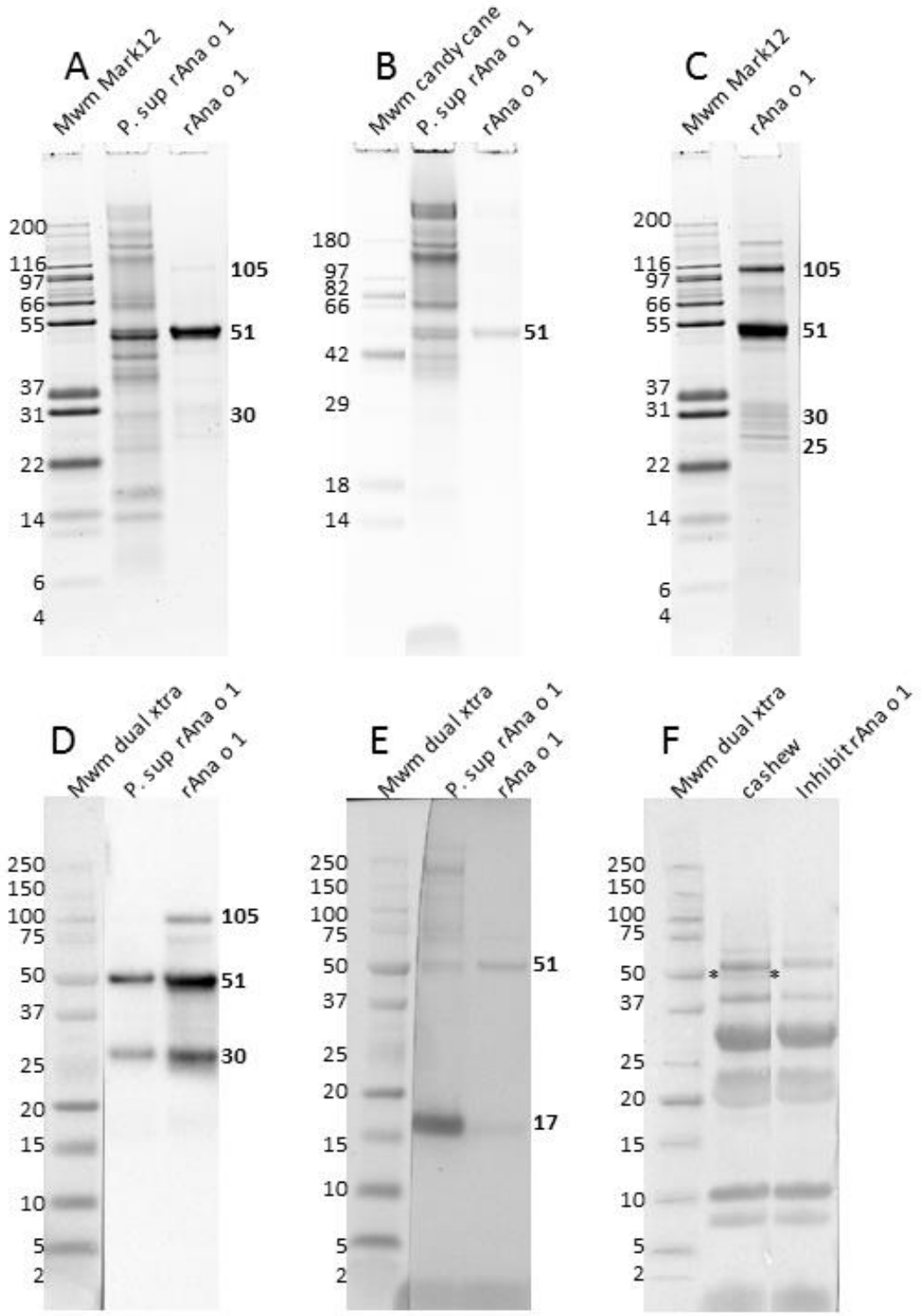


Figure 5.3 *Sypro Ruby total protein stain (A), Emerald glycoprotein stain (B), non-reducing Sypro Ruby total protein stain (C), c-Myc western blot (D), and serum western blot (E) of 30kDa concentrated P. pastoris supernatant of colony rAna o 1 (P. sup rAna o 1), and His-affinity-purified rAna o 1 (rAna o 1). Inhibition western blot (F) with IgE binding to a total cashew nut protein extract either not inhibited (cashew lane), or inhibited by pre-incubation of the plasma pool with 40µg His-affinity-purified rAna o 1 (Inhibit rAna o 1 lane). The molecular weight in kDa is indicated on the left of the figures, the bold numbers on the right of the figures indicate the protein bands described in Table 5.2, * indicates the inhibited protein binding after pre-incubation of the plasma with rAna o 1.*

As a last experiment the immunoreactivity of rAna o 1 and native Ana o 1 were compared on a plasma inhibition western blot, where plasma IgE was inhibited with rAna o 1. Figure 5.3F shows that IgE binding to native Ana o 1 (50kDa) is completely inhibited by rAna o 1. This inhibition western blot confirms the immunoreactivity of the linear epitopes of rAna o 1, matching that of native Ana o 1.
We have shown that rAna o 1 is glycosylated, and is, on western blot, comparable in immunoreactivity to native Ana o 1. However, these results do raise some questions. First of all, rAna o 1 is expressed as a 51kDa protein instead of the expected 62kDa. MALDI-MS/MS analysis of the 51kDa protein detected mostly C-terminal peptides: all matched peptides were found between amino acid E111 and H537, covering 74% of the sequence. Such a protein sequence would result in a 48.5kDa protein, very close to the 51kDa detected on SDS-PAGE. Protein identification of the 17kDa protein resulted in a matched protein sequence (Q19-R98) of only 9.6kDa. Hence, it is likely that not all the peptides were identified. This N-terminal 17kDa protein detected in the *P. pastoris* supernatant might be the N-terminus belonging to the C-terminal 51kDa protein. A fraction of these two subunits are bound to one another by *P. pastoris*, as the 17kDa would otherwise be discarded during the 30kDa concentration step, as well as during the His-affinity purification of rAna o 1.

It can be seen on reducing SDS-PAGE (Figure 5.3A) and serum western blot (Figure 5.3E), that a large amount of the 17kDa protein is lost during purification, indicating that the principal fraction of the 51 and 17kDa proteins were not attached to each other during His-affinity purification. Like other 7S globulins, native Ana o 1 is not processed into subunits, but is produced as one 50kDa protein. *E. coli*-produced Cor a 11, a 7S vicilin from hazelnut was produced as one full-length protein [21].

At this point it is still unclear where the 30kDa C-terminal polypeptide fits in. Possibly, *P. pastoris* does not only cleave the 62kDa protein into two subunits of 51 and 17kDa, but also cleaves the 51kDa C-terminal protein at another site, resulting in a 30 and 25kDa protein. It is also possible that these proteins are the product of proteolytic cleavage during the production or purification process.
Overall it can be stated that the production of rAna o 1 in *P. pastoris* leads to a different product, when compared molecularly, as can be obtained from cashew nuts. No comparison can be made with the *E. coli*-produced rAna o 1 as no SDS-PAGE results were shown in the paper by Wang *et al.*, only a total cashew nut protein western blot in which the serum was pre-incubated (inhibited) with rAna o 1 [9]. Truncated, or proteolytically cleaved expression of 7S globulins (*E.coli*-produced rAha h 1), has been described before [22, 23] and was suggested to have been caused by inefficient translation of the protein due to rare codons, or due to numerous cysteine residues [23].

A graphic representation of the various recombinantly produced protein polypeptides, as well as of the native proteins, are presented in Figure 5.6. As based on SDS-PAGE results, in *P. pastoris,* rAna o 1, is produced as a two connected 51 + 17kDa subunits, from which the 51kDa subunit is partly cleaved into 30 + 25kDa polypeptides. The occurrence of this 7S globulin in the form of subunits is different from native Ana o 1. However, despite these differences between the cashew native and the *P. pastoris* recombinant Ana o 1 proteins, IgE binding did occur to both proteins. It should be kept in mind that the 17kDa protein did bind the serum IgE very strongly, and that a considerable amount of this specific polypeptide is lost during either the concentration or the purification step.

Ana o 2

Ana o 2, the 11S globulin from cashew nut, is a 53kDa protein that is composed of a small (basic) subunit of about 21kDa, and a large (acidic) subunit of about 33kDa [8, 10]. In native state this protein forms hexameric structures [19]. Cloning of Ana o 2 cDNA into an *E. coli* expression system yielded a 52kDa protein, although no SDS-PAGE data was shown to confirm the protein size [10]. Here, recombinant Ana o 2 was produced in a *P. pastoris* expression system.

In our experiments, the expected molecular weight of the recombinant Ana o 2 protein is 53kDa, with possible 30 and 24kDa subunits (calculated from [10]). Figure 5.4A shows the proteins produced by *P. pastoris*, and the purified fraction after His-affinity purification. Again it can be seen that *P. pastoris* produced proteins of 14kDa to >200kDa, with the 51-55kDa protein bands being most pronounced. His-affinity purification selectively purifies these 51-53kDa proteins, as well as proteins of 41, 39, 36, and 25-30kDa, from the *P. pastoris* supernatant. The 55, 51, and 36kDa proteins were analysed by MALDI-MS/MS. Both the 55 and 51kDa proteins were identified as rAna o 2, matching peptides from D171- R416. The 36kDa protein was also identified as rAna o 2, matching peptides from C45-R416. However, this 371 amino acid-long polypeptide sequence would result in a 42kDa protein (C45-R416), or a 48kDa protein including the complete N-terminus (C45-H467). As most sequence matches were found in the C-terminus of the protein sequence, it is expected that this 36kDa protein is a C-terminal polypeptide of rAna o 2.

Native Ana o 2 is not glycosylated as shown by glycoprotein staining [8]. *P. pastoris* recombinant Ana o 2 is also not glycosylated as shown in Figure 5.4B. The glycoprotein stain does show presence of 150 and >180kDa glycosylated proteins, in both the *P. pastoris* supernatant and the His-affinity-purified sample. As these proteins are present in low quantity on the SDS-PAGE of Figure 5.4A, these proteins must be highly glycosylated to be stained so dark in Figure 5.4B. These proteins are likely impurities in the His-affinity-purified rAna o 2 sample. The 55, 51, 36, and 25-30kDa proteins are not glycosylated.

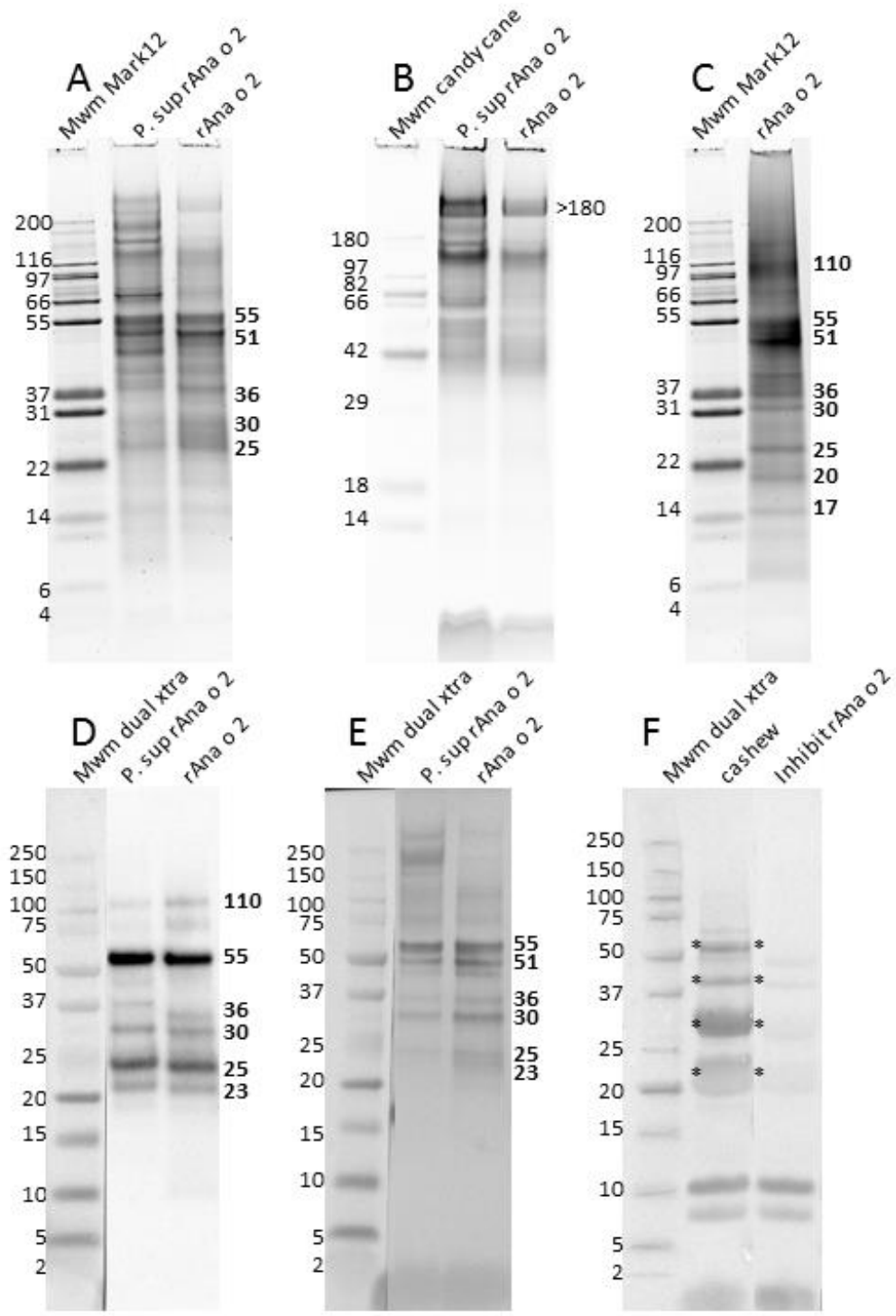


Figure 5.4 *Sypro Ruby total protein stain (A), Emerald glycoprotein stain (B), non-reducing Sypro Ruby total protein stain (C), c-Myc western blot (D), and serum western blot (E) of 30kDa concentrated P. pastoris supernatant of colony rAna o 2 (P. sup rAna o 2), and His-affinity-purified rAna o 2 (rAna o 2). Inhibition western blot (F) with IgE binding to a total cashew nut protein extract either not inhibited (cashew lane), or inhibited by pre-incubation of the plasma pool with 40µg His-affinity-purified rAna o 2 (Inhibit rAna o 2 lane). The molecular weight in kDa is indicated on the left of the figures, the bold numbers on the right of the figures indicate the protein bands described in Table 5.2. >180 indicates the >180kDa impurities in the His-affinity-purified rAna o 2 sample * indicates the inhibited protein binding after pre-incubation of the plasma with rAna o 2.*

Non-reduced (Figure 5.4C) rAna o 2 shows a higher abundance of a 110kDa protein, which is most likely an aggregated form (dimer) of rAna o 2. Beside this possible dimer, the 51-55kDa proteins are still clearly visible, as well as proteins of 35-38kDa, 25, 20, and 17kDa.

Several western blots were performed with rAna o 2. First of all a western blot using an anti c-Myc antibody was done (Figure 5.4D). In both the *P. pastoris* supernatant and the His-affinity-purified rAna o 2 sample, the same protein bands are bound by the anti c-Myc antibody: 110, 55, 36, 30, 25, 23kDa proteins. Subsequently, a serum western blot (Figure 5.4E) shows IgE binding to a 230kDa *P. pastoris* protein, as already shown for the control sample of Figure 5.2D. In the His-affinity-purified sample this 230kDa protein is not present and therefore does not bind the serum IgE. The serum IgE does bind to the 55 and 51kDa rAna o 2 proteins, as well as to the 36, 30, 25 and 23kDa proteins that were also detected on the c-Myc western blot. This western blot confirms the IgE reactivity of the recombinantly produced Ana o 2.

Lastly, in Figure 5.4F an inhibition western blot is shown, in which a plasma pool was pre-incubated with rAna o 2 before western blotting with a total cashew nut protein extract. Compared to the non-inhibited cashew nut lane, the rAna o 2-inhibited lane shows strongly reduced IgE binding to the 21 and 33kDa Ana o 2 subunits, as well as complete inhibition of IgE binding to the 53kDa monomeric Ana o 2. Also decreased IgE binding to a 39kDa protein and to Ana o 1 was detected, similar as observed in the control sample of Figure 5.2E, and therefore likely caused by the impurities present in the rAna o 2 sample. The inhibition western blot indicates that the IgE reactivity of rAna o 2 and native Ana o 2 are highly comparable on western blot.

The characterisation of rAna o 2 from *P. pastoris* brings about some confusion due to the proteins of different polypeptides produced by *P. pastoris*. First of all it is clear that the 55kDa protein is the complete rAna o 2 protein. Besides this complete protein several truncated rAna o 2 proteins are produced by *P. pastoris* (Figure 5.6B): 51, 36, 30, 25, and 23kDa proteins, of which the 36kDa protein was identified as a C-terminal polypeptide of rAna o 2. All these proteins contain the C-terminus of rAna o 2, as they are purified via their Histidine-tag, and contain the c-Myc epitope as shown by western blotting. In cashew nut, the 11S globulin Ana o 2 is produced as a single polypeptide, which is subsequently linked by disulphide bridges, and cleaved by asparaginyl endopeptidase [24-26]. The cleavage site of this enzyme in native Ana o 2 is after the amino acids D257, N258. In the protein sequence of rAna o 2 four sites are present in which an asparagine is preceded by an aspartic acid. Possibly, proteolytic enzymes from *P. pastoris* not only cleave the protein at the designated site (N258), but also at these other sites (N24, N214, N344). The C-terminal protein sizes of such possibly asparaginyl endopeptidase truncated proteins would be 51, 29, 24 and 14kDa, matching with the detected C-terminal polypeptides of 51, 30, and 25/23kDa. Asparaginyl endopeptidase cleaves specifically after an asparagine but is not specific towards the amino

acid following this asparagine amino acid [27]. Therefore the 36kDa protein might originate from a cleavage site at N154 (36kDa C-terminal polypeptide). Similar truncation of a recombinantly produced 11S globulin, walnut Jug r 4 in *E. coli*, has been described before [28]. It is clear that the C-terminal polypeptides have not been connected to their N-terminal counterpart by disulphide bridges, as occurs in native Ana o 2. This can be concluded as these N-terminal polypeptides were not detected on the serum western blot, nor on (non-reducing) SDS-PAGE. It is unclear why *P. pastoris* does not connect the N- and C-terminal polypeptides, while it is capable of doing so [13, 15].

All the C-terminal polypeptides are bound by serum IgE on western blot, indicating presence of epitopes on these truncated C-terminal rAna o 2 polypeptides. However, only the 55kDa protein represents the complete rAna o 2 protein, and as several different N-terminal polypeptides are missed during the His-affinity purification of the *P. pastoris* supernatant, the N-terminal epitopes are underrepresented. Next to this, also the conformational folding of the proteins is expected to differ from native Ana o 2, thereby possibly affecting the conformational epitopes. While native Ana o 2 is composed of two subunits, linked together by disulphide bonds [29], a part of the native protein is produced as a 53kDa monomeric protein. However, rAna o 2 is produced mostly in monomeric form, and only a part is processed into subunits. This difference in subunit structure will change the conformational folding of this protein, potentially affecting the conformational epitope present on both native and *E. coli* or *P. pastoris* recombinantly produced Ana o 2 [30].

Ana o 3

As shown in Figure 5.5A, the recombinant expression of Ana o 3 in *P. pastoris* results in the production of proteins of 12kDa to >200kDa. Of these proteins, only an 18kDa protein and a 16kDa protein are purified by His-affinity purification. The expected molecular weight of rAna o 3 is 17.3kDa, with possible subunits of 6 and 11kDa (based on [8], and calculated from the protein sequence cloned into *P. pastoris* as described in Supplementary information 1). Both the 18kDa protein and the 16kDa protein were identified as Ana o 3 by MALDI-MS/MS (A7-R104), see Supplementary information 2.

The high molecular weight proteins (>50kDa) from *P. pastoris* are glycosylated (Figure 5.5B). The 18 and 16kDa recombinant proteins are not glycosylated. The nucleotide sequence of Ana o 3 does contain one putative N-glycosylation motif [8], however, this N-glycosylation motif is present in the leader peptide, which is not included in to the recombinant protein. Native Ana o 3 is also not glycosylated [8].

In order to verify the presence of disulphide bridges in the recombinant Ana o 3 protein, a non-reducing SDS-PAGE was performed (Figure 5.5C). Native Ana o 3 has a large and a small subunit, bound together by disulphide bridges to form a protein of 12.6kDa. The molecular weight of recombinant Ana o 3 does not differ between a reducing and non-reducing SDS-

PAGE, indicating that rAna o 3 is not proteolytically cleaved into a small and a large subunit, and that no disulphide bonds are present in rAna o 3.

Figure 5.5D shows a western blot using a c-Myc antibody, targeting the c-Myc epitope located C-terminally in the recombinant proteins. In the *P. pastoris* supernatant, the c-Myc antibody only binds the 18 and 16kDa rAna o 3 proteins, indicating that no recombinant proteins containing a c-Myc epitope are discarded during the purification step. In the His-affinity-purified rAna o 3 sample, the c-Myc antibody does not only bind to the expected 18 and 16kDa proteins, but also to a 12kDa protein. This 12kDa protein did not appear before on SDS-PAGE, however, the c-Myc western blot is more sensitive than SDS-PAGE. It is possible this C-terminal protein is produced by *P. pastoris* at relatively low levels. This protein might be produced without (part of) the N-terminus like described for rAna o 2, or it might be proteolytically cleaved after production of the complete protein. However, due to the low quantity on SDS-PAGE, the 12kDa protein it could not be analysed by MALDI-MS/MS. Also the relative amount of this protein, compared to the amount of 18 or 16kDa proteins, could not be calculated as it is not visible on SDS-PAGE. Further analysis by mass spectrometry could give more information.

Figure 5.5E shows IgE binding of a serum pool of cashew allergic patients to both the 18 and the 16kDa protein of rAna o 3, confirming the immunoreactivity of these recombinant Ana o 3 proteins. The 12kDa protein that is visible in the c-Myc western blot, is not visible in the serum western blot. This protein either does not bind IgE, or is present in too small quantity to be visualised on this western blot. Besides binding of IgE to recombinant Ana o 3 in a standard serum western blot, also an inhibition western blot was performed. For this inhibition western blot, a plasma pool from cashew nut and tree nut allergic patients was pre-incubated with His-affinity-purified rAna o 3, in order to analyse if rAna o 3 could bind all IgE in this serum pool, and therefore shows the same immunoreactivity on western blot as native Ana o 3. Indeed Figure 5.5F shows that, after inhibition, no IgE binds to native Ana o 3 (10, 8kDa). This indicates that not only does rAna o 3 bind IgE, but also that rAna o 3 contains the same linear epitopes and has comparable immunoreactivity on western blot as native Ana o 3.

It can be concluded that *P. pastoris* expresses rAna o 3 as 18 and 16kDa proteins (Figure 5.6C), that can be His-affinity purified, and are, on western blot, immunologically comparable to native Ana o 3. It is unclear why rAna o 3 is expressed as both an 18kDa protein and as a 16kDa protein. Possibly two different start sites for DNA transcription are used by *P. pastoris*. This has been reported for the expression of the 2S albumin from brazil nut, Ber e 1, in *P. pastoris* [31]. For Ber e 1, approximately 20% of the recombinant protein produced was a ±4kDa smaller protein than the expected 14kDa protein [31]. For rAna o 3,

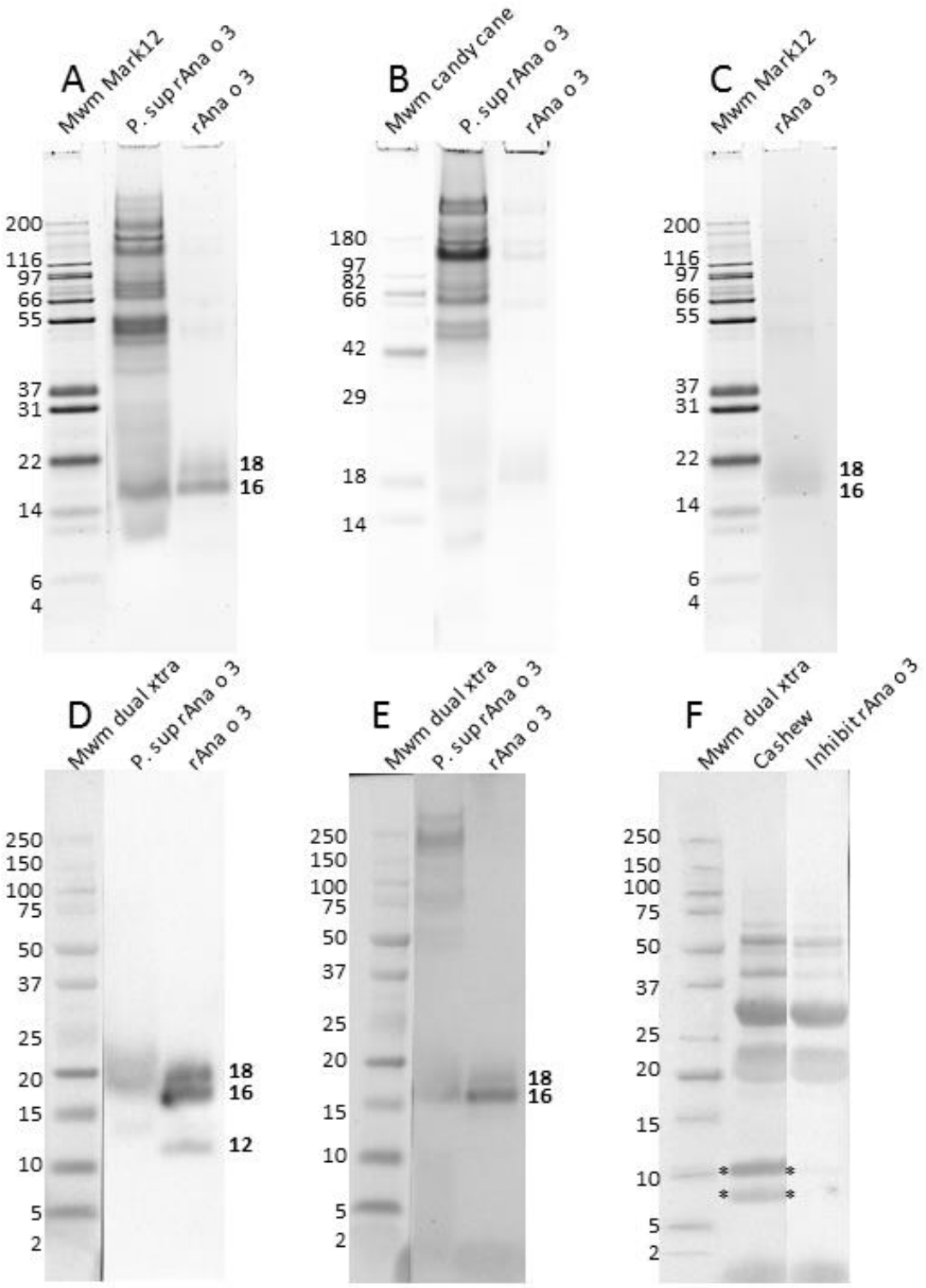


Figure 5.5 *Sypro Ruby total protein stain (A), Emerald glycoprotein stain (B), non-reducing Sypro Ruby total protein stain (C), c-Myc western blot (D), and serum western blot (E) of 3kDa concentrated P. pastoris supernatant of colony rAna o 3 (P. sup rAna o 3), and His-affinity-purified rAna o 3 (rAna o 3). Inhibition western blot (F) with IgE binding to a total cashew nut protein extract either not inhibited (cashew lane), or inhibited by pre-incubation of the plasma pool with 25µg His-affinity-purified rAna o 3 (Inhibit rAna o 3 lane). The molecular weight in kDa is indicated on the left of the figures, the bold numbers on the right of the figures indicate the protein bands described in Table 5.2, * indicates the inhibited protein binding after pre-incubation of the plasma with rAna o 3.*

P. pastoris produces ±15% 18kDa protein and ±85% of the 16kDa protein. Confirmation of a difference in start site between the two proteins could be provided by N-terminal sequencing of the proteins. No comparison can be made to the *E.coli*-produced rAna o 3 as in this study by Robotham *et al*. no SDS-PAGE or western blot data are shown besides an inhibition western blot [11]. Native Ana o 3 is produced as a large and a small subunit, linked together by disulphide bonds [8, 11]. Here, *P. pastoris*-produces rAna o 3 as the complete protein, and as an N-terminally truncated protein. rAna o 3 is apparently not processed into subunits as it would be by the cashew tree. Production of a complete 2S albumin, not processed into subunits, has been shown before for walnut Jug r 1 and a hazelnut 2S albumin, produced in *E. coli* [32, 33]. It is unclear if *E.coli*-produced rAna o 3 was cleaved into subunits or not [11].

Table 5.2 *Overview of protein characteristics of recombinant Ana o 1, 2, and 3 in P. pastoris supernatant and His affinity-purified samples. Molecular weight indicated in kDa, - indicates that the method was not applied to that sample, (..) indicates the number of spots detected on 2D gel.*

Methods	**rAna o 1**		**rAna o 2**		**rAna o 3**	
	Supernatant	His-purified	Supernatant	His-purified	Supernatant	His-purified
Calculated	62		53 (30+24kDa subunits)		17 (6+11Da subunits)	
Reducing SDS-PAGE	51, 14->200	105, 51, 30	51-53, 14->200	51-53, 41, 39, 36, 25-30	16, 10, 14->200	18, 16
Glycosylation	50->180	51	50->180	Impurities, not 51-53	50->180	No bands
Non-reducing SDS-PAGE	-	105, 51, 25-30	-	110, 51-53, 35-38, 25, 20, 17	-	16-18
c-Myc western	51, 30	105, 51, 30	110, 53, 36, 25, 23	110, 53, 36, 30, 25, 23	18, 16	18, 16, 12
Serum western	230, 51, 17	51, 17	230, 53, 51, 36, 30, 25, 23	53, 51, 36, 30, 25, 23	230, 18, 16	18, 16
2D gel	-	51 (1), 50 (3), 30 (5), 25-30 (5), 17 (1)	ND	54 (5), 53 (7), 51 (7), 50 (4), 39 (2), 36 (4), 25 (5), 23 (6), 20 (1), 15 (1), 13 (1), 7 (1)	-	18 (2), 16 (2)
2D serum western	-	-	-	54 (5), 53 (7), 51 (7), 50 (4), 39 (2), 36 (4), 25 (5), 24 (4)	-	-
Protein identification	17 (N-terminal)	51, 30 (C-terminal)	-	53, 51, 36 (C-terminal)	-	18, 16

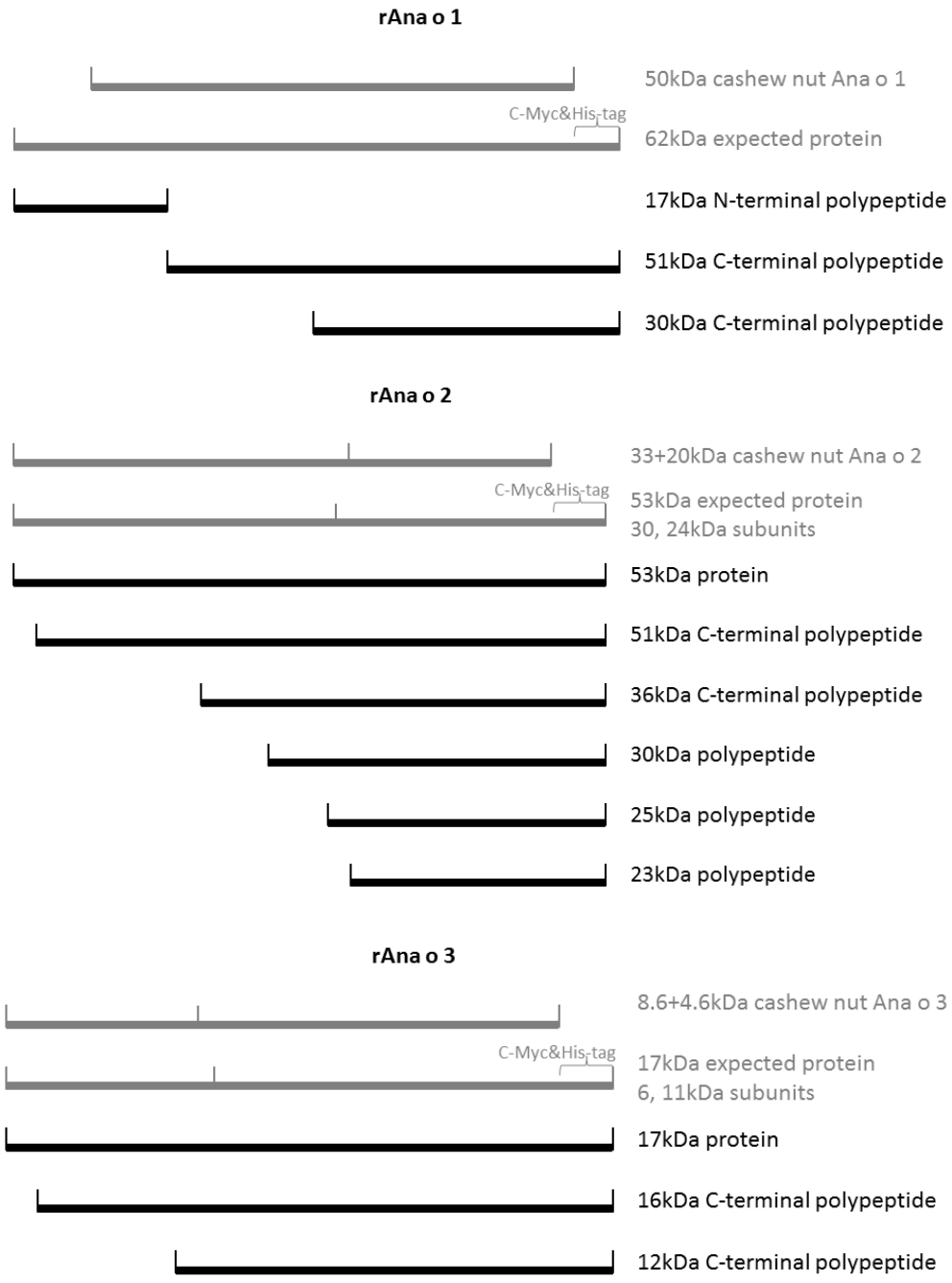


Figure 5.6 *Representation of rAna o 1, 2 and 3 as produced by P. pastoris. The top grey bar is the native cashew nut protein, the second grey bar is the expected P. pastoris-produced protein with the c-Myc and His-tag indicated. The black bars below are the recombinant proteins produced by P. pastoris and detected on SDS-PAGE and western blot.*

2D electrophoresis

As 2D electrophoresis of the native cashew nut allergens revealed presence of multiple isoforms for both Ana o 1, 2 and 3 (chapter 4 of this thesis), 2D electrophoresis of His-affinity-purified rAna o 1, 2 and 3 was performed. 2D electrophoresis of rAna o 1 (Figure 5.7A) shows an abundant spot at 51kDa as well as a few minor spots at 50, 30, 25 and 17kDa.

The isoelectric point (pI) of the abundant 51kDa protein spot is 5.4, in line with the expected pI of 5.6 [34]. The occurrence of one protein spot for the 51 protein, suggests that *P. pastoris* does not produce many different isoforms of the rAna o 1 protein. The multiple spots for the low abundance 30 and 25kDa proteins indicates that these proteins are either different forms of incorrectly produced rAna o 1, or that they are post-translationally processed forms, which are e.g. glycosylated or phosphorylated.

rAna o 2 has by far the most complicated spot pattern on 2D electrophoresis of the three recombinant proteins, see Figure 5.7B. At a molecular weight of 55kDa 7 spots are present (pI 5.4-6.1), 5 spots at 54kDa (pI 5.5-6.0), 7 spots at 51kDa (pI 5.4-6.1), 4 spots at 50kDa (pI 5.4-5.8), 2 spots at 39kDa (pI 5.0-5.2), 4 pots at 36kDa (pI 5.2-5.4), 5 spots at 25kDa (pI 5.2-6.1), 6 spots at 23kDa (pI 5.2-6.2), and several separate spots are present at 20 (pI 6.0, 6.2, 6.8), 15 (pI 4.3), 13 (pI 4.0) and 7kDa (pI 10.5). The spots of 50-54kDa represent most likely the same proteins as the 55 and 51kDa bands on SDS-PAGE, differing slightly in electrophoresis profile due to the difference in polyacrylamide concentrations between the 1D and 2D gels. Apparently the full length rAna o 2 protein is subjected to several post-translational modifications, leading to a change in isoelectric point. On 2D electrophoresis of native Ana o 2, 11 spots can be detected for the large acidic subunit (pI 5.4-6.3), and 18 spots can be detected for the small basic subunit (pI 5.4-9.3, chapter 4 of this thesis). It is expected that native Ana o 2 is subjected to several post-translational modifications, and this is likely to be the case as well for rAna o 2. However, the pI range of the native protein isoforms is much broader (5.4-9.3) compared to the pI range of the most abundant protein spots of rAna o 2 (pI 5.0-6.2). This might be related to a difference in protein translation, post-translational modifications, or due to the truncation of the recombinant protein. For the 2D electrophoresis gel of rAna o 2 also a western blot was performed, in order to determine if the isoforms are immuno-reactive (Figure 5.7D). Serum IgE binds to all recombinant Ana o 2 protein isoforms mentioned above, except the single protein spots of 20, 15, 13 and 7kDa, suggesting these single protein spots might be impurities in the His-affinity-purified protein sample.

Finally, also 2D electrophoresis was performed for rAna o 3. 2D electrophoresis of native Ana o 3 shows that the large subunit (10 and 8kDa on reducing SDS-PAGE) has 8 isoforms, 4 of 10kDa and 4 of 8kDa (Chapter 4 of this thesis). The small subunit of 6kDa was not detected on 2D SDS-PAGE. 2D electrophoresis of recombinant Ana o 3 (Figure 5.7C) shows four distinct protein spots: two of 18kDa (pI 5.5 and 5.7) and two of 16kDa (pI 5.5 and 5.7). The isoelectric points of 5.5 and 5.7 are in line with the expected pI of 5.4 [34]. It is likely that rAna o 3 is post-translationally processed by *P. pastoris*, for example by phosphorylation as possible phosphorylation sites are present in the amino acid sequence [35], and *P. pastoris* is capable of phosphorylation [36].

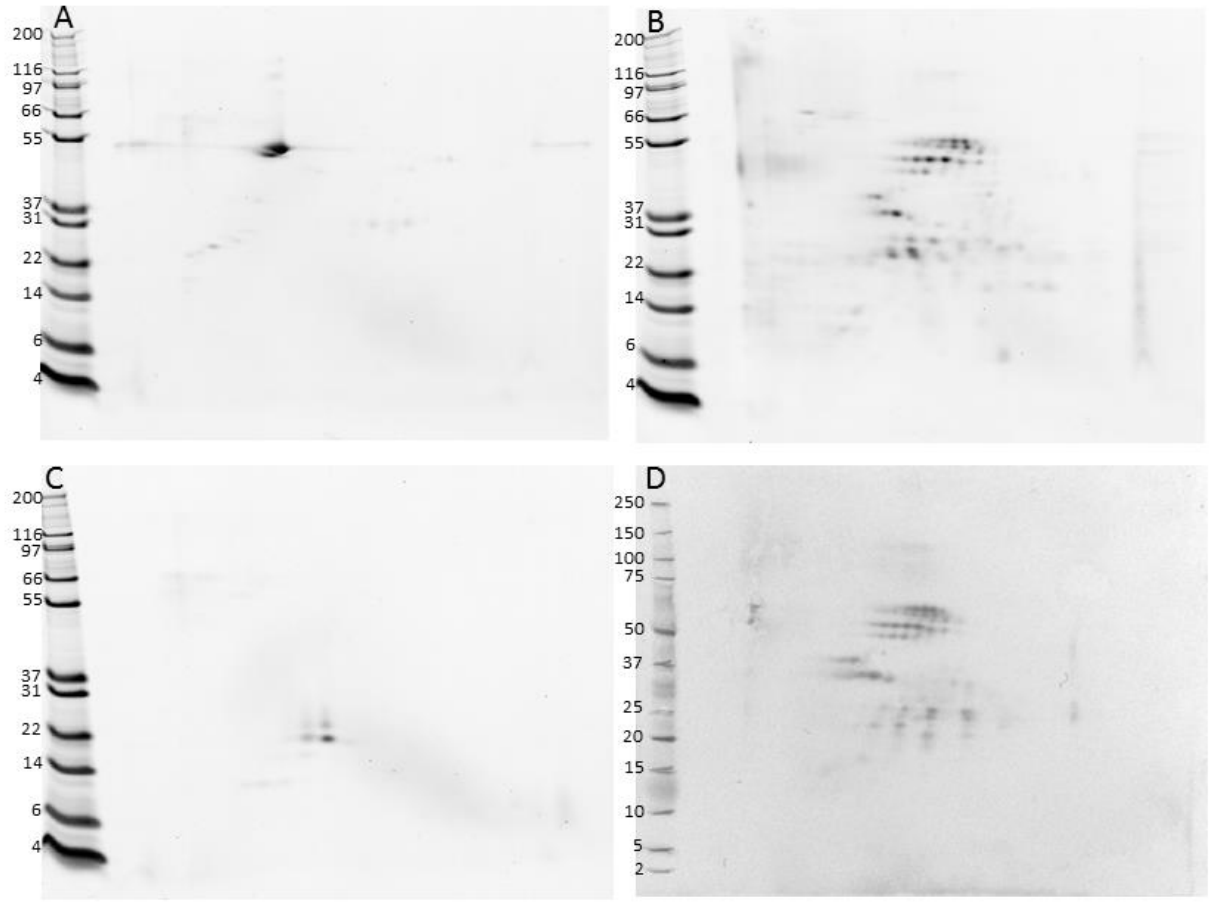


Figure 5.7 *2D electrophoresis of rAna o 1 (A), rAna o 2 (B), rAna o 3 (C), and a 2D electrophoresis western blot of rAna o 2 (D) using a serum pool. The molecular weight in kDa is indicated on the left of the figures.*

Protein yield

The protein yield of the three recombinant allergens was calculated as mg His-affinity-purified protein per litre of *P. pastoris* supernatant: 10.5mg/L (±1.9mg/L) for rAna o 1, 4.8mg/L (±3.2mg/L) for rAna o 2 and 11.1mg/L (±7.6mg/L) for rAna o 3. The low protein yield of rAna o 2 is probably related to the lower total protein levels present in the *P. pastoris* supernatant of 194.5mg/L (±34.6mg/L) for rAna o 2 compared to 282.5mg/L (±27.6mg/L) for Ana o 1 and 267.5mg/L (±34.6mg/L) for rAna o 3. Furthermore, as many truncated C-terminal rAna o 2 polypeptides were produced by *P. pastoris*, these proteins lack the N-terminal α-signal peptide that induces secretion of the recombinant proteins into the *P. pastoris* supernatant. Furthermore, the protein yield might be low as several N-terminal polypeptides are discarded during the protein purification step.

Overall the three recombinant proteins are produced at relatively low quantity when compared to the 4- 12g recombinant protein per litre *P. pastoris* supernatant [15] that *P. pastoris* is capable of. Other tree nut proteins (2S albumin Ber e 1) have been produced at 200mg/L in *E. coli* [31].The relatively low yield of all three allergens might be improved by

the addition of a Maltose-binding fusion tag to increase protein solubility [12]. Also, the recombinant protein production process was optimised for expression time and methanol induction, but not for temperature, pH of the growth medium, time of the glycerol growth phase, or production system (shake-flask VS bioreactor) [37, 38].

Summary and recommendations

It can be stated that *P. pastoris* does produce immunologically reactive cashew nut allergens, however, all three proteins are cleaved into polypeptides in a different manner than the cashew tree does. As is shown in Figure 5.6A, the primary translation product of rAna o 1 is a 62kDa protein. The protein is however, produced as, or cleaved into, two polypeptides of 51 and 17kDa. A small fraction of these two polypeptides are connected by disulphide bridges, but most are not. Besides this unexpected cleavage, also the 51kDa protein is for a small fraction of the total protein production, cleaved into polypeptides of 21 and 30kDa.

The production of rAna o 2 is shown in Figure 5.6B. The coding sequence of rAna o 2 encodes a 53kDa protein, that is potentially split into two subunits of 30 and 24kDa, linked via disulphide bridges. However, only a part of the recombinant protein is produced as a 55kDa protein, which is likely this 53kDa complete protein. rAna o 2 is post-translationally cleaved into C-terminal polypeptides of 51, 36, 30, 25 and 23kDa. These C-terminal polypeptides are not connected to their N-terminal counterparts, causing these N-terminal polypeptides to be discarded during the purification process.

Last but not least, the peptide structure of rAna o 3 is shown in Figure 5.6C. rAna o 3 encodes a 17kDa protein, that is potentially split into two subunits of 6 and 11kDa. rAna o 3 however, is produced as an 18 and 16kDa protein, and minor amount is produced as a 12kDa C-terminal polypeptide. It is unclear if the N-terminal polypeptides are attached to their C-terminal polypeptides, as they are not visible on SDS-PAGE or western blot. However, the N-terminal polypeptides are theoretically rather small (2-6kDa), and can therefore be difficult to detect on SDS-PAGE. It is unclear if there are epitopes present on these N-terminal polypeptides.

Similar truncation or cleavage by *P. pastoris* as observed here for rAna o 1, 2 and 3, has been reported before [22, 23, 28, 31]. It is possible that the polypeptides are formed due to the occurrence of different start sites for protein transcription/translation, or due to post translational processing. The recombinant allergens are IgE reactive, despite the differences between with the native allergens.

If these *P. pastoris*-produced recombinant proteins will be used for further experiments, the exact transcription/translation start sites or post-translational cleavage sites should be studied by N-terminal amino acid sequencing and mass spectrometry. Also, the protein production should be optimized. Furthermore a more thorough inhibition blotting comparison should be made. Western blotting with native purified protein inhibition should

be compared to western blotting with the recombinant proteins. The native and recombinant proteins should inhibit IgE binding comparably at similar protein concentrations. Also cell experiments could be performed, using for example the rat basophil leukemia (RBL) cell line, to compare cross-linking activity of the native proteins to the recombinant proteins.

Acknowledgements

The authors would like to thank Ying Deng, Renata Ariens, Gracian Camps Ramon, and Eveline de Jong for their help with the cloning of rAna o 1, 2 3.

Funding

This work was supported by Technology Foundation STW, Stichting Voedselallergie, Siemens Healthcare Diagnostics, HAL Allergy, Intersnack Nederland B.V., ALK-Abello B.V., and the Nederlands Anafylaxis Netwerk.

References

[1] Willison, L. N., Tawde, P., Robotham, J. M., Penney, R. M. t., et al., Pistachio vicilin, Pis v 3, is immunoglobulin E-reactive and cross-reacts with the homologous cashew allergen, Ana o 1. Clinical and experimental allergy : journal of the British Society for Allergy and Clinical Immunology 2008, 38, 1229-1238.

[2] Savvatianos, S., Konstantinopoulos, A. P., Borga, A., Stavroulakis, G., et al., Sensitization to cashew nut 2S albumin, Ana o 3, is highly predictive of cashew and pistachio allergy in Greek children. The Journal of allergy and clinical immunology 2015, 136, 192-194.

[3] van der Valk, J. P., Gerth van Wijk, R., Vergouwe, Y., Steyerberg, E. W., et al., sIgE Ana o 1, 2 and 3 accurately distinguish tolerant from allergic children sensitised to cashew nuts. Clinical and experimental allergy : journal of the British Society for Allergy and Clinical Immunology 2016.

[4] Teodorowicz, M., Fiedorowicz, E., Kostyra, H., Wichers, H., Kostyra, E., Effect of Maillard reaction on biochemical properties of peanut 7S globulin (Ara h 1) and its interaction with a human colon cancer cell line (Caco-2). European journal of nutrition 2013, 52, 1927-1938.

[5] Blanc, F., Vissers, Y. M., Adel-Patient, K., Rigby, N. M., et al., Boiling peanut Ara h 1 results in the formation of aggregates with reduced allergenicity. Molecular nutrition & food research 2011, 55, 1887-1894.

[6] Mattison, C. P., Grimm, C. C., Wasserman, R. L., In vitro digestion of soluble cashew proteins and characterization of surviving IgE-reactive peptides. Molecular nutrition & food research 2014, 58, 884-893.

[7] Guan, X., Noble, K. A., Tao, Y., Roux, K. H., Sathe, S. K., Young, N. L., Marshall, A. G., Epitope mapping of 7S cashew antigen in complex with antibody by solution-phase H/D exchange monitored by FT-ICR mass spectrometry. Journal of Mass Spectrometry 2015, 50, 812-819.

[8] Reitsma, M., Bastiaan-Net, S., Sforza, S., van der Valk, J. P., et al., Purification and Characterization of Anacardium occidentale (Cashew) Allergens Ana o 1, Ana o 2, and Ana o 3. Journal of agricultural and food chemistry 2016, 64, 1191-1201.

[9] Wang, F., Robotham, J. M., Teuber, S. S., Tawde, P., et al., Ana o 1, a cashew (Anacardium occidental) allergen of the vicilin seed storage protein family. The Journal of allergy and clinical immunology 2002, 110, 160-166.

[10] Wang, F., Robotham, J. M., Teuber, S. S., Sathe, S. K., Roux, K. H., Ana o 2, a major cashew (Anacardium occidentale L.) nut allergen of the legumin family. International archives of allergy and immunology 2003, 132, 27-39.

[11] Robotham, J. M., Wang, F., Seamon, V., Teuber, S. S., et al., Ana o 3, an important cashew nut (Anacardium occidentale L.) allergen of the 2S albumin family. The Journal of allergy and clinical immunology 2005, 115, 1284-1290.

[12] Willison, L. N., Sathe, S. K., Roux, K. H., Production and analysis of recombinant tree nut allergens. Methods 2014, 66, 34-43.

[13] Mattanovich, D., Branduardi, P., Dato, L., Gasser, B., et al., Recombinant protein production in yeasts. Methods in molecular biology 2012, 824, 329-358.

[14] Makino, T., Skretas, G., Georgiou, G., Strain engineering for improved expression of recombinant proteins in bacteria. Microbial cell factories 2011, 10, 32.

[15] Demain, A. L., Vaishnav, P., Production of recombinant proteins by microbes and higher organisms. Biotechnology advances 2009, 27, 297-306.

[16] van der Valk, J. P., Gerth van Wijk, R., Dubois, A. E., de Groot, H., et al., Multicentre Double-Blind Placebo-Controlled Food Challenge Study in Children Sensitised to Cashew Nut. PloS one 2016, 11, e0151055.

[17] Brocca, S., Schmidt-Dannert, C., Lotti, M., Alberghina, L., Schmid, R. D., Design, total synthesis, and functional overexpression of the Candida rugosa lip1 gene coding for a major industrial lipase. Protein science : a publication of the Protein Society 1998, 7, 1415-1422.

[18] Huang, C. J., Damasceno, L. M., Anderson, K. A., Zhang, S., et al., A proteomic analysis of the Pichia pastoris secretome in methanol-induced cultures. Applied microbiology and biotechnology 2011, 90, 235-247.

[19] Mills, E. N. C., Shewry, P. R., Plant food allergens, Blackwell science Ltd 2004.

[20] Cregg, J. M., Introduction: distinctions between Pichia pastoris and other expression systems. Methods in molecular biology 2007, 389, 1-10.

[21] Lauer, I., Foetisch, K., Kolarich, D., Ballmer-Weber, B. K., et al., Hazelnut (Corylus avellana) vicilin Cor a 11: molecular characterization of a glycoprotein and its allergenic activity. The Biochemical journal 2004, 383, 327-334.
[22] Kleber-Janke, T., Becker, W. M., Use of modified BL21(DE3) Escherichia coli cells for high-level expression of recombinant peanut allergens affected by poor codon usage. Protein expression and purification 2000, 19, 419-424.
[23] Burks, A. W., Cockrell, G., Stanley, J. S., Helm, R. M., Bannon, G. A., Recombinant peanut allergen Ara h I expression and IgE binding in patients with peanut hypersensitivity. The Journal of clinical investigation 1995, 96, 1715-1721.
[24] Tai, S. S., Wu, L. S., Chen, E. C., Tzen, J. T., Molecular cloning of 11S globulin and 2S albumin, the two major seed storage proteins in sesame. Journal of agricultural and food chemistry 1999, 47, 4932-4938.
[25] Beyer, K., Grishina, G., Bardina, L., Grishin, A., Sampson, H. A., Identification of an 11S globulin as a major hazelnut food allergen in hazelnut-induced systemic reactions. The Journal of allergy and clinical immunology 2002, 110, 517-523.
[26] Chua, A. C., Hsiao, E. S., Yang, Y. C., Lin, L. J., et al., Gene families encoding 11S globulin and 2S albumin isoforms of jelly fig (Ficus awkeotsang) Achenes. Bioscience, biotechnology, and biochemistry 2008, 72, 506- 513.
[27] Nielsen, N. C., in: Vasil, B. A. L. I. K. (Ed.), Springer Science 1997.
[28] Wallowitz, M., Peterson, W. R., Uratsu, S., Comstock, S. S., et al., Jug r 4, a legumin group food allergen from walnut (Juglans regia Cv. Chandler). Journal of agricultural and food chemistry 2006, 54, 8369-8375.
[29] Xia, L., Willison, L. N., Porter, L., Robotham, J. M., et al., Mapping of a conformational epitope on the cashew allergen Ana o 2: a discontinuous large subunit epitope dependent upon homologous or heterologous small subunit association. Mol Immunol 2010, 47, 1808-1816.
[30] Robotham, J. M., Xia, L., Willison, L. N., Teuber, S. S., et al., Characterization of a cashew allergen, 11S globulin (Ana o 2), conformational epitope. Mol Immunol 2010, 47, 1830-1838.
[31] Alcocer, M. J., Murtagh, G. J., Bailey, K., Dumoulin, M., et al., The disulphide mapping, folding and characterisation of recombinant Ber e 1, an allergenic protein, and SFA8, two sulphur-rich 2S plant albumins. Journal of molecular biology 2002, 324, 165-175.
[32] Sordet, C., Culerrier, R., Granier, C., Rance, F., et al., Expression of Jug r 1, the 2S albumin allergen from walnut (Juglans regia), as a correctly folded and functional recombinant protein. Peptides 2009, 30, 1213-1221.
[33] Garino, C., Zuidmeer, L., Marsh, J., Lovegrove, A., et al., Isolation, cloning, and characterization of the 2S albumin: a new allergen from hazelnut. Molecular nutrition & food research 2010, 54, 1257-1265.
[34] http://web.expasy.org/compute_pi/
[35] Blom, N., Gammeltoft, S., Brunak, S., Sequence and structure-based prediction of eukaryotic protein phosphorylation sites. Journal of molecular biology 1999, 294, 1351-1362.
[36] Montesino, R., Nimtz, M., Quintero, O., Garcia, R., et al., Characterization of the oligosaccharides assembled on the Pichia pastoris-expressed recombinant aspartic protease. Glycobiology 1999, 9, 1037-1043.
[37] Macauley-Patrick, S., Fazenda, M. L., McNeil, B., Harvey, L. M., Heterologous protein production using the Pichia pastoris expression system. Yeast 2005, 22, 249-270.
[38] Anasontzis, G. E., Salazar Pena, M., Spadiut, O., Brumer, H., Olsson, L., Effects of temperature and glycerol and methanol-feeding profiles on the production of recombinant galactose oxidase in Pichia pastoris. Biotechnology progress 2014, 30, 728-735.

Supplementary information

Synthesized Ana o 1 nucleotide sequence:

CTC~~GAG~~~~AAAAGAGAGGCTGAAGCTGGTCCTCCTACAAAGTTCTCTTTTTCTCTGTTTTTAGTTTCTGTTTTGGTCCTGTGTTTAGGTTTTGCTTTGGGCT~~AAGATTGACCCGGAGTTGAAACAGTGCAAGCACCAGTGTAAAGTGCAGAGGCAGTATGATGAGC
AACAGAAGGAGCAATGTGTGAAAGAGTGTGAAAAGTACTACAAAGAGAAGAAGGGACGGGAACGAGAGCATGAAGAG
GAAGAAGAGGAATGGGGTACTGGTGGCGTTGATGAACCCAGCACTCATGAGCCAGCTGAAAAGCATCTCTCCCAGTGTAT
GAGACAGTGCGAGAGACAAGAAGGAGGTCAACAAAAGCAATTATGTCGTTTTAGGTGTCAAGAGCGTTATAAGAAAGAG
AGAGGACAACATAATTACAAGAGAGAAGACGATGAAGACGAGGACGAAGATGAAGCCGAGGAGGAAGATGAGAATCCC
TATGTATTCGAGGATGAAGATTTCACCACCAAAGTCAAGACTGAGCAAGGTAAAGTAGTTTTGTTGCCCAAGTTTACTCAA
AAATCGAAGCTTCTTCACGCCCTGGAGAAATACCGTCTAGCCGTTTTGGTTGCAAATCCTCAGGCTTTTGTAGTTCCAAGTC
ACATGGATGCTGACAGTATTTTCTTTGTTTCTTGGGGACGCGGAACTATCACCAAGATCTTAGAGAACAAACGAGAGAGCA
TTAATGTCAGACAGGGAGATATCGTCAGCATTAGTTCTGGTACTCCTTTTTATATCGCCAATAACGATGAAAACGAGAAGC
TTTACCTCGTCCAATTCCTAAGACCTGTAAATTTACCAGGTCATTTCGAAGTGTTTCATGGTCCAGGCGGTGAAAATCCAGA
GTCTTTCTATAGAGCTTTCTCGTGGGAAATATTGGAAGCCGCTCTGAAGACCTCAAAAGACACACTTGAGAAACTTTTCGA
GAAACAGGACCAAGGTACTATCATGAAAGCCTCCAAAGAACAAATTAGAGCTATGTCCAGGAGAGGCGAAGGTCCTAAA
ATTTGGCCATTTACAGAGGAATCAACGGGTTCATTCAAACTGTTCAAAAAGGATCCCTCTCAATCCAATAAATACGGCCAAT
TGTTTGAAGCTGAACGTATTGATTATCCGCCACTTGAAAAGTTGGACATGGTTGTCTCCTATGCGAACATCACCAAGGGAG
GAATGTCTGTTCCTTTCTACAACTCAAGAGCAACGAAAATAGCCATTGTGGTTTCTGGAGAAGGTTGCGTTGAGATAGCAT
GTCCTCATCTATCCTCTTCGAAAAGTTCACACCCAAGTTACAAGAAATTGAGGGCAAGAATTAGAAAGGACACAGTGTTTA
TTGTCCCAGCGGGTCACCCTTTCGCAACTGTTGCTTCAGGAAATGAAAACTTGGAAATCGTGTGCTTTGAAGTAAACGCAG
AAGGCAACATAAGGTACACATTAGCCGGGAAGAAGAACATTATTAAGGTCATGGAGAAGGAAGCTAAAGAGTTGGCATT
CAAGATGGAAGGAGAAGAAGTGGACAAAGTGTTTGGAAAACAAGATGAAGAGTTTTTCTTCCAGGGGCCAGAATGGCGA
AAGGAAAAAGAAGGGCGTGCTGATGAATT**TCTAGA**

Ana o 1 nucleotide sequence as expressed in *P. pastoris*:

~~ATGAGATTTCCTTCAATTTTTACTGCTGTTTTATTCGCAGCATCCTCCGCATTAGCTGCTCCAGTCAACACTACAACAGAAGA~~
~~TGAAACGGCACAAATTCCGGCTGAAGCTGTCATCGGTTACTCAGATTTAGAAGGGGATTTCGATGTTGCTGTTTTGCCATT~~
~~TTCCAACAGCACAAATAACGGGTTATTGTTTATAAATACTACTATTGCCAGCATTGCTGCTAAAGAAGAAGGGGTATCTCTC~~
~~GAGAAAAGAGAGGCTGAAGCT~~**GAATTC**AAGATTGACCCGGAGTTGAAACAGTGCAAGCACCAGTGTAAAGTGCAGAGG
CAGTATGATGAGCAACAGAAGGAGCAATGTGTGAAAGAGTGTGAAAAGTACTACAAAGAGAAGAAGGGACGGGAACGA
GAGCATGAAGAGGAAGAAGAGGAATGGGGTACTGGTGGCGTTGATGAACCCAGCACTCATGAGCCAGCTGAAAAGCAT
CTCTCCCAGTGTATGAGACAGTGCGAGAGACAAGAAGGAGGTCAACAAAAGCAATTATGTCGTTTTAGGTGTCAAGAGCG
TTATAAGAAAGAGAGAGGACAACATAATTACAAGAGAGAAGACGATGAAGACGAGGACGAAGATGAAGCCGAGGAGGA
AGATGAGAATCCCTATGTATTCGAGGATGAAGATTTCACCACCAAAGTCAAGACTGAGCAAGGTAAAGTAGTTTTGTTGCC
CAAGTTTACTCAAAAATCGAAGCTTCTTCACGCCCTGGAGAAATACCGTCTAGCCGTTTTGGTTGCAAATCCTCAGGCTTTT
GTAGTTCCAAGTCACATGGATGCTGACAGTATTTTCTTTGTTTCTTGGGGACGCGGAACTATCACCAAGATCTTAGAGAAC
AAACGAGAGAGCATTAATGTCAGACAGGGAGATATCGTCAGCATTAGTTCTGGTACTCCTTTTTATATCGCCAATAACGAT
GAAAACGAGAAGCTTTACCTCGTCCAATTCCTAAGACCTGTAAATTTACCAGGTCATTTCGAAGTGTTTCATGGTCCAGGC
GGTGAAAATCCAGAGTCTTTCTATAGAGCTTTCTCGTGGGAAATATTGGAAGCCGCTCTGAAGACCTCAAAAGACACACTT
GAGAAACTTTTCGAGAAACAGGACCAAGGTACTATCATGAAAGCCTCCAAAGAACAAATTAGAGCTATGTCCAGGAGAGG
CGAAGGTCCTAAAATTTGGCCATTTACAGAGGAATCAACGGGTTCATTCAAACTGTTCAAAAAGGATCCCTCTCAATCCAA
TAAATACGGCCAATTGTTTGAAGCTGAACGTATTGATTATCCGCCACTTGAAAAGTTGGACATGGTTGTCTCCTATGCGAA
CATCACCAAGGGAGGAATGTCTGTTCCTTTCTACAACTCAAGAGCAACGAAAATAGCCATTGTGGTTTCTGGAGAAGGTTG
CGTTGAGATAGCATGTCCTCATCTATCCTCTTCGAAAAGTTCACACCCAAGTTACAAGAAATTGAGGGCAAGAATTAGAAA
GGACACAGTGTTTATTGTCCCAGCGGGTCACCCTTTCGCAACTGTTGCTTCAGGAAATGAAAACTTGGAAATCGTGTGCTT
TGAAGTAAACGCAGAAGGCAACATAAGGTACACATTAGCCGGGAAGAAGAACATTATTAAGGTCATGGAGAAGGAAGCT
AAAGAGTTGGCATTCAAGATGGAAGGAGAAGAAGTGGACAAAGTGTTTGGAAAACAAGATGAAGAGTTTTTCTTCCAGG
GGCCAGAATGGCGAAAGGAAAAAGAAGGGCGTGCTGATGAATT**TCTAGA**ACAAAAACTCATCTCAGAAGAGGATCTGAA
TAGCGCCGTCGACCATCATCATCATCATCAT

Ana o 1 amino acid sequence as obtained from *P. pastoris:* 62.1kDa

EFKIDPELKQCKHQCKVQRQYDEQQKEQCVKECEKYYKEKKGREREHEEEEEEWGTGGVDEPSTHEPAEKHLSQCMRQCERQ
EGGQQKQLCRFRCQERYKKERGQHNYKREDDEDEDEDEAEEEDENPYVFEDEDFTTKVKTEQGKVVLLPKFTQKSKLLHALEKY
RLAVLVANPQAFVVPSHMDADSIFFVSWGRGTITKILENKRESINVRQGDIVSISSGTPFYIANNDENEKLYLVQFLRPVNLPGHF
EVFHGPGGENPESFYRAFSWEILEAALKTSKDTLEKLFEKQDQGTIMKASKEQIRAMSRRGEGPKIWPFTEESTGSFKLFKKDPS

QSNKYGQLFEAERIDYPPLEKLDMVVSYANITKGGMSVPFYNSRATKIAIVVSGEGCVEIACPHLSSSKSSHPSYKKLRARIRKDTV
FIVPAGHPFATVASGNENLEIVCFEVNAEGNIRYTLAGKKNIIKVMEKEAKELAFKMEGEEVDKVFGKQDEEFFFQGPEWRKEKE
GRADE**FL**EQKLISEEDLNSAVDHHHHHH

Synthesized Ana o 2 nucleotide sequence:

GAATTCCGTCAGGAATGGCAACAACAAGATGAGTGCCAAATCGATAGGCTGGATGCCTTGGAACCCGATAACCGAGTTG
AGTATGAAGCCGGTACGGTTGAAGCCTGGGACCCTAACCATGAGCAATTCCGATGTGCTGGTGTTGCATTGGTTAGGCAT
ACCATACAACCTAATGGCCTGTTGTTGCCTCAATATTCTAATGCTCCTCAATTGATTTACGTTGTCCAGGGTGAGGGTATGA
CAGGTATTTCATATCCAGGATGTCCAGAAACTTACCAAGCACCCCAACAGGGAAGACAGCAGGGACAGTCTGGTAGATTC
CAGGACCGGCATCAAAAGATCCGAAGATTTAGACGAGGCGATATTATCGCAATCCCCGCCGGAGTAGCACACTGGTGTTA
CAACGAGGGCAATTCCCCGGTCGTCACTGTTACTCTTCTGGACGTCTCAAACTCGCAAAATCAGCTTGATAGGACGCCTAG
AAAATTTCATCTGGCTGGTAACCCTAAAGATGTATTTCAGCAACAGCAACAACACCAATCCAGAGGGAGAAACCTTTTTTCT
GGTTTTGATACAGAGTTATTGGCTGAGGCTTTCCAAGTGGACGAACGTCTGATAAAGCAGTTGAAAAGCGAGGACAACAG
GGGTGGAATTGTTAAGGTTAAGGATGACGAACTTAGAGTGATCAGACCATCAAGATCCCAATCCGAGCGTGGAAGTGAA
AGTGAAGAGGAAAGTGAGGATGAAAAAAGAAGATGGGGACAACGTGACAATGGGATTGAAGAAACCATTTGCACTATG
AGATTAAAAGAGAATATCAATGATCCTGCTCGCGCTGACATTTACACCCCAGAAGTTGGTCGTCTTACTACATTAAACTCCC
TCAACCTCCCAATTTTGAAATGGCTTCAACTCAGTGTTGAAAAGGGTGTGCTATACAAAAATGCTCTAGTCCTGCCACACTG
GAACTTGAACTCGCATTCCATAATATATGGATGTAAGGGTAAAGGTCAGGTTCAAGTAGTAGACAACTTCGGCAACAGAG
TGTTCGACGGCGAAGTCAGAGAGGGACAGATGTTGGTCGTGCCACAAAACTTTGCAGTAGTTAAGCGTGCAAGAGAGGA
AAGATTCGAATGGATTTCTTTCAAGACCAATGATCGGGCCATGACTTCTCCGTTAGCTGGACGCACCTCTGTTTTAGGTGG
GATGCCAGAGGAAGTGTTAGCCAATGCGTTCCAGATCTCAAGAGAAGATGCTAGGAAGATCAAGTTTAATAATCAGCAGA
CAACTTTGACATCTGGAGAGTCAAGCCACCATATGAGGGATGATGCTGG**TCTAGA**

Ana o 2 nucleotide sequence as expressed in *P. pastoris*:

~~ATGAGATTTCCTTCAATTTTTACTGCTGTTTTATTCGCAGCATCCTCCGCATTAGCTGCTCCAGTCAACACTACAACAGAAGA~~
~~TGAAACGGCACAAATTCCGGCTGAAGCTGTCATCGGTTACTCAGATTTAGAAGGGGATTTCGATGTTGCTGTTTTGCCATT~~
~~TTCCAACAGCACAAATAACGGGTTATTGTTTATAAATACTACTATTGCCAGCATTGCTGCTAAAGAAGAAGGGGTATCTCTC~~
~~GAGAAAAGAGAGGCTGAAGCT~~**GAATTC**CGTCAGGAATGGCAACAACAAGATGAGTGCCAAATCGATAGGCTGGATGCCT
TGGAACCCGATAACCGAGTTGAGTATGAAGCCGGTACGGTTGAAGCCTGGGACCCTAACCATGAGCAATTCCGATGTGCT
GGTGTTGCATTGGTTAGGCATACCATACAACCTAATGGCCTGTTGTTGCCTCAATATTCTAATGCTCCTCAATTGATTTACGT
TGTCCAGGGTGAGGGTATGACAGGTATTTCATATCCAGGATGTCCAGAAACTTACCAAGCACCCCAACAGGGAAGACAGC
AGGGACAGTCTGGTAGATTCCAGGACCGGCATCAAAAGATCCGAAGATTTAGACGAGGCGATATTATCGCAATCCCCGCC
GGAGTAGCACACTGGTGTTACAACGAGGGCAATTCCCCGGTCGTCACTGTTACTCTTCTGGACGTCTCAAACTCGCAAAAT
CAGCTTGATAGGACGCCTAGAAAATTTCATCTGGCTGGTAACCCTAAAGATGTATTTCAGCAACAGCAACAACACCAATCC
AGAGGGAGAAACCTTTTTTCTGGTTTTGATACAGAGTTATTGGCTGAGGCTTTCCAAGTGGACGAACGTCTGATAAAGCAG
TTGAAAAGCGAGGACAACAGGGGTGGAATTGTTAAGGTTAAGGATGACGAACTTAGAGTGATCAGACCATCAAGATCCC
AATCCGAGCGTGGAAGTGAAAGTGAAGAGGAAAGTGAGGATGAAAAAAGAAGATGGGGACAACGTGACAATGGGATT
GAAGAAACCATTTGCACTATGAGATTAAAAGAGAATATCAATGATCCTGCTCGCGCTGACATTTACACCCCAGAAGTTGGT
CGTCTTACTACATTAAACTCCCTCAACCTCCCAATTTTGAAATGGCTTCAACTCAGTGTTGAAAAGGGTGTGCTATACAAAA
ATGCTCTAGTCCTGCCACACTGGAACTTGAACTCGCATTCCATAATATATGGATGTAAGGGTAAAGGTCAGGTTCAAGTAG
TAGACAACTTCGGCAACAGAGTGTTCGACGGCGAAGTCAGAGAGGGACAGATGTTGGTCGTGCCACAAAACTTTGCAGTA
GTTAAGCGTGCAAGAGAGGAAAGATTCGAATGGATTTCTTTCAAGACCAATGATCGGGCCATGACTTCTCCGTTAGCTGG
ACGCACCTCTGTTTTAGGTGGGATGCCAGAGGAAGTGTTAGCCAATGCGTTCCAGATCTCAAGAGAAGATGCTAGGAAGA
TCAAGTTTAATAATCAGCAGACAACTTTGACATCTGGAGAGTCAAGCCACCATATGAGGGATGATGCTGG**TCTAGA**ACAA
AAACTCATCTCAGAAGAGGATCTGAATAGCGCCGTCGACCATCATCATCATCATCAT

Ana o 2 amino acid sequence as obtained from *P. pastoris:* 53.3kDa

EFRQEWQQQDECQIDRLDALEPDNRVEYEAGTVEAWDPNHEQFRCAGVALVRHTIQPNGLLLPQYSNAPQLIYVVQGEGMT
GISYPGCPETYQAPQQGRQQGQSGRFQDRHQKIRRFRRGDIIAIPAGVAHWCYNEGNSPVVTVTLLDVSNSQNQLDRTPRKF
HLAGNPKDVFQQQQQHQSRGRNLFSGFDTELLAEAFQVDERLIKQLKSEDNRGGIVKVKDDELRVIRPSRSQSERGSESEEESE
DEKRRWGQRDNGIEETICTMRLKENINDPARADIYTPEVGRLTTLNSLNLPILKWLQLSVEKGVLYKNALVLPHWNLNSHSIIYG
CKGKGQVQVVDNFGNRVFDGEVREGQMLVVPQNFAVVKRAREERFEWISFKTNDRAMTSPLAGRTSVLGGMPEEVLANAF
QISREDARKIKFNNQQTTLTSGESSHHMRDDA**GL**EQKLISEEDLNSAVDHHHHHH

Synthesized Ana o 3 nucleotide sequence:
~~CTCGAGAAAAGAGAGGCTGAAGCTGCAAAGTTCTTACTCCTCCTATCTGCCTTCGGCAGTCCTCCTCCTGGTGGCTAACGCC~~
GAATTCTCCATTTACCGAGCCATTGTGGAGGTTGAAGAAGACTCGGGCCGTGAGCAGAGTTGCCAACGGCAGTTCGAAGA
GCAGCAGCGATTCCGGAACTGTCAAAGGTACGTGAAGCAGGAGGTCCAGAGGGGAGGACGCTATAACCAGAGACAAGA
AAGCTTGAGGGAATGCTGCCAGGAGTTGCAGGAAGTAGACAGAAGGTGCCGCTGCCAGAACCTAGAGCAAATGGTGAG
GCAGCTGCAGCAACAGGAACAAATAAAGGGTGAGGAGGTTCGGGAACTTTATGAAACAGCCAGTGAATTGCCTCGCATTT
GCAGTATTTCACCCAGCCAGGGCTGTCAGTTTCAGTCATCTTATTT**TCTAGA**

Ana o 3 nucleotide sequence as expressed in *P. pastoris*:
~~ATGAGATTTCCTTCAATTTTTACTGCTGTTTTATTCGCAGCATCCTCCGCATTAGCTGCTCCAGTCAACACTACAACAGAAGA~~
~~TGAAACGGCACAAATTCCGGCTGAAGCTGTCATCGGTTACTCAGATTTAGAAGGGGATTTCGATGTTGCTGTTTTGCCATT~~
~~TTCCAACAGCACAAATAACGGGTTATTGTTTATAAATACTACTATTGCCAGCATTGCTGCTAAAGAAGAAGGGGTATCTCTC~~
~~GAGAAAAGAGAGGCTGAAGCT~~**GAATTC**TCCATTTACCGAGCCATTGTGGAGGTTGAAGAAGACTCGGGCCGTGAGCAGA
GTTGCCAACGGCAGTTCGAAGAGCAGCAGCGATTCCGGAACTGTCAAAGGTACGTGAAGCAGGAGGTCCAGAGGGGAG
GACGCTATAACCAGAGACAAGAAAGCTTGAGGGAATGCTGCCAGGAGTTGCAGGAAGTAGACAGAAGGTGCCGCTGCCA
GAACCTAGAGCAAATGGTGAGGCAGCTGCAGCAACAGGAACAAATAAAGGGTGAGGAGGTTCGGGAACTTTATGAAAC
AGCCAGTGAATTGCCTCGCATTTGCAGTATTTCACCCAGCCAGGGCTGTCAGTTTCAGTCATCTTATTT**TCTAGA**ACAAAAA
CTCATCTCAGAAGAGGATCTGAATAGCGCCGTCGACCATCATCATCATCATCAT

Ana o 3 amino acid sequence as obtained from *P. pastoris*: 17.3kDa
EFSIYRAIVEVEEDSGREQSCQRQFEEQQRFRNCQRYVKQEVQRGGRYNQRQESLRECCQELQEVDRRCRCQNLEQMVRQLQ
QQEQIKGEEVRELYETASELPRICSISPSQGCQFQSSY**FL**EQKLISEEDLNSAVDHHHHHH

Figure 5.S1 *Nucleotide and protein sequences of rAna o 1, rAna o 2 and rAna o 3. Restriction sites are indicated in bold, leader peptides are double struck through, a-factor signal sequence and kex2 signal cleavage sequence are single struck through, c-Myc epitope and His-tag are underlined.*

rAna o 1: 51kDa protein band, 58% sequence coverage, Mascot score 855. Amino acid E111 and H537 are indicated in bold.

EFKIDPELKQCKHQCKVQRQYDEQQKEQCVKECEKYYKEKKGREREHEEEEEEWGTGGVDEPSTHEPAEKHLSQCMRQCERQ
EGGQQKQLCRFRCQERYKKERGQHNYKR**E**DDEDEDEDEAEEEDENPYVFEDEDFTTKVKTEQGKVVLLPKFTQKSKLLHALEKY
RLAVLVANPQAFVVPSHMDADSIFFVSWGRGTITKILENKRESINVRQQGDIVSISSGTPFYIANNDENEKLYLVQFLRPVNLPGHF
EVFHGPGGENPESFYRAFSWEILEAALKTSKDTLEKLFEKQDQGTIMKASKEQIRAMSRRGEGPKIWPFTEESTGSFKLFKKDPS
QSNKYGQLFEAERIDYPPLEKLDMVVSYANITKGGMSVPFYNSRATKIAIVVSGEGCVEIACPHLSSSKSSHPSYKKLRARIRKDTV
FIVPAGHPFATVASGNENLEIVCFEVNAEGNIRYTLAGKKNIIKVMEKEAKELAFKMEGEEVDKVFGKQDEEFFFQGPEWRKEKE
GRADEFLEQKLISEEDLNSAVDHHHHHH**H**

rAna o 1 30kDa protein band, 30% sequence coverage, Mascot score 636. Amino acid I202 and H537 are indicated in bold.

EFKIDPELKQCKHQCKVQRQYDEQQKEQCVKECEKYYKEKKGREREHEEEEEEWGTGGVDEPSTHEPAEKHLSQCMRQCERQ
EGGQQKQLCRFRCQERYKKERGQHNYKREDDEDEDEDEAEEEDENPYVFEDEDFTTKVKTEQGKVVLLPKFTQKSKLLHALEKY
RLAVLVANPQAFVVPSHMDADSIFFVSWGRGTITK**I**LENKRESINVRQGDIVSISSGTPFYIANNDENEKLYLVQFLRPVNLPGHF
EVFHGPGGENPESFYRAFSWEILEAALKTSKDTLEKLFEKQDQGTIMKASKEQIRAMSRRGEGPKIWPFTEESTGSFKLFKKDPS
QSNKYGQLFEAERIDYPPLEKLDMVVSYANITKGGMSVPFYNSRATKIAIVVSGEGCVEIACPHLSSSKSSHPSYKKLRARIRKDTV
FIVPAGHPFATVASGNENLEIVCFEVNAEGNIRYTLAGKKNIIKVMEKEAKELAFKMEGEEVDKVFGKQDEEFFFQGPEWRKEKE
GRADEFLEQKLISEEDLNSAVDHHHHHH**H**

rAna o 1 17kDa protein band, 13% sequence coverage, Mascot score 113. Amino acid Q19 and R98 are indicated in bold.

EFKIDPELKQCKHQCKVQR**Q**YDEQQKEQCVKECEKYYKEKKGREREHEEEEEEWGTGGVDEPSTHEPAEKHLSQCMRQCERQ
EGGQQKQLCRFRCQE**R**YKKERGQHNYKREDDEDEDEDEAEEEDENPYVFEDEDFTTKVKTEQGKVVLLPKFTQKSKLLHALEKY
RLAVLVANPQAFVVPSHMDADSIFFVSWGRGTITKILENKRESINVRQGDIVSISSGTPFYIANNDENEKLYLVQFLRPVNLPGHF
EVFHGPGGENPESFYRAFSWEILEAALKTSKDTLEKLFEKQDQGTIMKASKEQIRAMSRRGEGPKIWPFTEESTGSFKLFKKDPS
QSNKYGQLFEAERIDYPPLEKLDMVVSYANITKGGMSVPFYNSRATKIAIVVSGEGCVEIACPHLSSSKSSHPSYKKLRARIRKDTV
FIVPAGHPFATVASGNENLEIVCFEVNAEGNIRYTLAGKKNIIKVMEKEAKELAFKMEGEEVDKVFGKQDEEFFFQGPEWRKEKE
GRADEFLEQKLISEEDLNSAVDHHHHHH

rAna o 2: 53kDa protein band, 24% sequence coverage, Mascot score 241. Amino acid D171 and R416 are indicated in bold.

EFRQEWQQQDECQIDRLDALEPDNRVEYEAGTVEAWDPNHEQFRCAGVALVRHTIQPNGLLLPQYSNAPQLIYVVQGEGMT
GISYPGCPETYQAPQQGRQQGQSGRFQDRHQKIRRFRRGDIIAIPAGVAHWCYNEGNSPVVTVTLLDVSNSQNQLDRTPRKF
HLAGNPK**D**VFQQQQQHQSRGRNLFSGFDTELLAEAFQVDERLIKQLKSEDNRGGIVKVKDDELRVIRPSRSQSERGSESEEESE
DEKRRWGQRDNGIEETICTMRLKENINDPARADIYTPEVGRLTTLNSLNLPILKWLQLSVEKGVLYKNALVLPHWNLNSHSIIYG
CKGKGQVQVVDNFGNRVFDGEVREGQMLVVPQNFAVVKRAREERFEWISFKTNDRAMTSPLAGRTSVLGGMPEEVLANAF
QIS**R**EDARKIKFNNQQTTLTSGESSHHMRDDAGLEQKLISEEDLNSAVDHHHHHH

rAna o 2: 51kDa protein band, 32% sequence coverage, Mascot score 263. Amino acid D171 and R416 are indicated in bold.

EFRQEWQQQDECQIDRLDALEPDNRVEYEAGTVEAWDPNHEQFRCAGVALVRHTIQPNGLLLPQYSNAPQLIYVVQGEGMT
GISYPGCPETYQAPQQGRQQGQSGRFQDRHQKIRRFRRGDIIAIPAGVAHWCYNEGNSPVVTVTLLDVSNSQNQLDRTPRKF
HLAGNPK**D**VFQQQQQHQSRGRNLFSGFDTELLAEAFQVDERLIKQLKSEDNRGGIVKVKDDELRVIRPSRSQSERGSESEEESE
DEKRRWGQRDNGIEETICTMRLKENINDPARADIYTPEVGRLTTLNSLNLPILKWLQLSVEKGVLYKNALVLPHWNLNSHSIIYG
CKGKGQVQVVDNFGNRVFDGEVREGQMLVVPQNFAVVKRAREERFEWISFKTNDRAMTSPLAGRTSVLGGMPEEVLANAF
QIS**R**EDARKIKFNNQQTTLTSGESSHHMRDDAGLEQKLISEEDLNSAVDHHHHHH

rAna o 2: 36kDa protein band, 45% sequence coverage, Mascot score 423. Amino acid C45 and R416 are indicated in bold.

EFRQEWQQQDECQIDRLDALEPDNRVEYEAGTVEAWDPNHEQFR**C**AGVALVRHTIQPNGLLLPQYSNAPQLIYVVQGEGMT GISYPGCPETYQAPQQGRQQGQSGRFQDRHQKIRRFRRGDIIAIPAGVAHWCYNEGNSPVVTVTLLDVSNSQNQLDRTPRKF HLAGNPKDVFQQQQQHQSRGRNLFSGFDTELLAEAFQVDERLIKQLKSEDNRGGIVKVKDDELRVIRPSRSQSERGSESEEESE DEKRRWGQRDNGIEETICTMRLKENINDPARADIYTPEVGRLTTLNSLNLPILKWLQLSVEKGVLYKNALVLPHWNLNSHSIIYG CKGKGQVQVVDNFGNRVFDGEVREGQMLVVPQNFAVVKRAREERFEWISFKTNDRAMTSPLAGRTSVLGGMPEEVLANAF QIS**R**EDARKIKFNNQQTTLTSGESSHHMRDDAGLEQKLISEEDLNSAVDHHHHHH

rAna o 3: 18kDa protein band, 69% sequence coverage, Mascot score 452. Amino acid A7 and R104 are indicated in bold.

EFSIYR**A**IVEVEEDSGREQSCQRQFEEQQRFRNCQRYVKQEVQRGGRYNQRQESLRECCQELQEVDRRCRCQNLEQMVRQLQ QQEQIKGEEVRELYETASELP**R**ICSISPSQGCQFQSSYFLEQKLISEEDLNSAVDHHHHHH

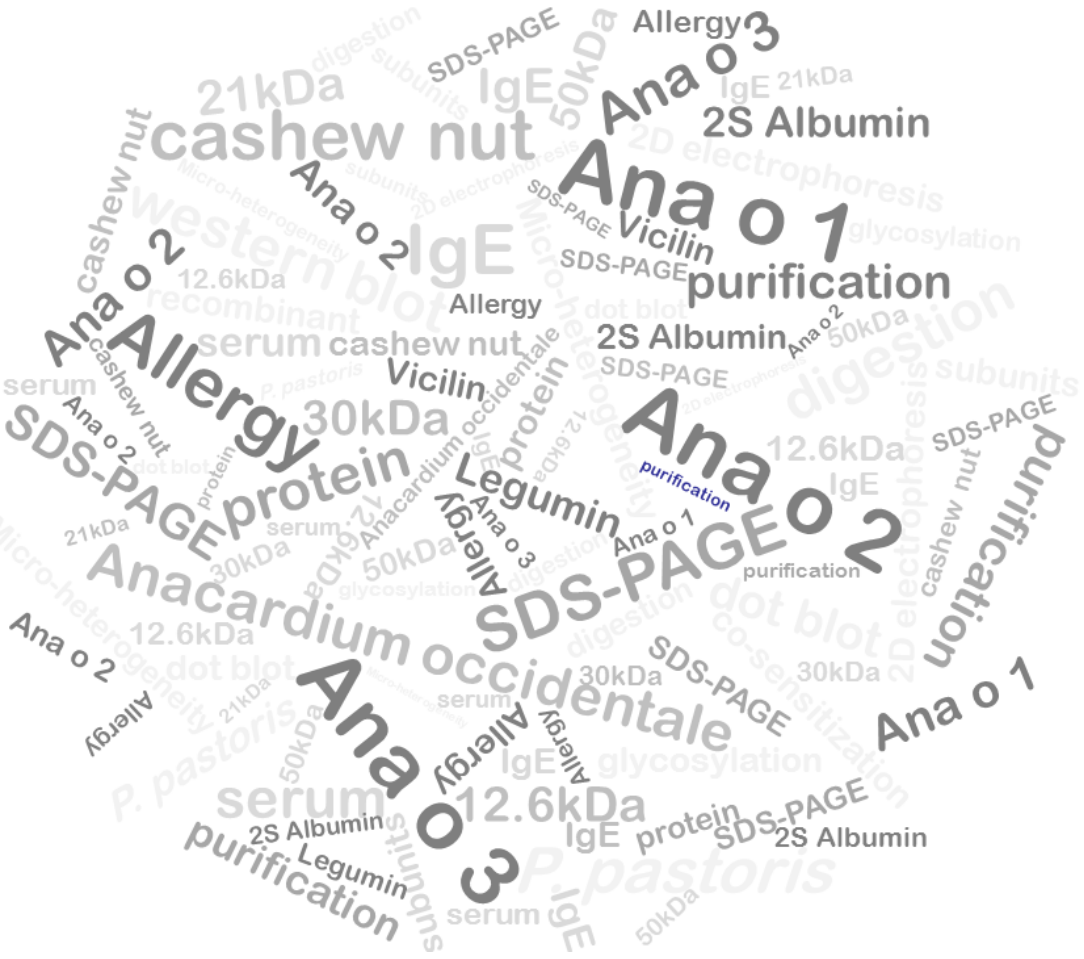
rAna o 3: 16Da protein band, 69% sequence coverage, Mascot score 513. Amino acid A7 and R104 are indicated in bold.

EFSIYR**A**IVEVEEDSGREQSCQRQFEEQQRFRNCQRYVKQEVQRGGRYNQRQESLRECCQELQEVDRRCRCQNLEQMVRQLQ QQEQIKGEEVRELYETASELP**R**ICSISPSQGCQFQSSYFLEQKLISEEDLNSAVDHHHHHH

Figure 5.S2 *Protein identification of P. pastoris recombinantly produced proteins, matched peptides are underlined.*

Chapter 6

General discussion



This thesis focussed on the cashew nut (*Anacardium occidentale*) allergens, Ana o 1, Ana o 2 and Ana o 3. This research was carried out within a larger project, the IDEAL project (Improvement of Diagnostic mEthods for ALlergy assessment, cashew allergy in children as a showcase), a collaborative project integrating molecular and clinical aspects of cashew nut allergy. While this thesis is focussed mostly on cashew nut allergen characterisation, the clinical research in Rotterdam, Delft and Groningen focused on the diagnosis of cashew nut allergy in children using double blind placebo controlled food challenges (DBPCFC), skin prick tests (SPT), serum IgE (sIgE) measurements, and quality of life assessments in cashew nut sensitized children. Merging different disciplines in this way, led to a collaboration in which molecular and clinical data could be used to deepen our understanding of this type of allergy, and thoroughly characterised allergen samples could be biologically characterised making use of sera of patients included in this study. This collaboration led to new insights into cashew nut allergy.

Cashew nut allergy

In the IDEAL project, the cashew nut proteins were not only studied at laboratory scale, as will be discussed later in this chapter, the cashew nut proteins were also studied regarding their clinical profile [1-4]. The study was initiated because at the children's hospital "Kinderhaven", in Rotterdam, a high number of children were reported having possible cashew nut allergy (290 children). Of these 290 children, 90 were tested by SPT, of which 69 (73%) were sensitized to cashew nuts. In the IDEAL study, cashew nut sensitized children were diagnosed for cashew nut allergy by a double blind placebo controlled food challenge (DBPCFC) [1]. In 137 (76.5%) of the 179 tested children, an IgE mediated cashew nut allergy could be confirmed by a positive DBPCFC. Symptoms experienced during these positive challenges were gastro-intestinal (64% oral allergy symptoms, 72% nausea/stomach pain/vomiting/diarrhoea), symptoms of the skin (21% urticarial, 28% redness and itchiness, 27% angioedema), 19% symptoms of the eye, 15% symptoms of the upper airways, and 7% symptoms of the lower airways. No cardiovascular symptoms were observed [1]. The remarkably low dose of cashew nut protein required for the induction of an allergic reaction was already specified by Blom *et al.* (7.4mg cashew nut causing an allergic reaction in 5% of the allergic population) [5] and by Davoren *et al.* (anaphylaxis after contact via skin or oral mucosa) [6]. This was confirmed by the study of van der Valk *et al.*, as 17% of the children experienced anaphylaxis upon ingestion of only 1mg cashew nut protein [1]. This lowest observed adverse effect level (LOAEL) of cashew nut protein is comparable to the LOAEL of peanut protein. Ingestion of 100µg peanut protein can already result in subjective symptoms in peanut allergic patients. Objective symptoms have been observed from 5mg peanut protein onwards [7].

The age of the cashew nut allergic patients studied here was between 2 and 17 years old, with an average of 9 years old [1]. It is known that food allergy is more prevalent in children than in adults [8]. This might be caused by the low gastric acid and pepsin secretion early in life [9, 10], the overall immaturity of the gastro-intestinal tract in new-borns [11], and possible sensitization to allergens via the umbilical cord during pregnancy [12], or via breast milk [13]. Of the cashew nut sensitized children, 38% reported not having consumed cashew nut [2]. Sensitization in-utero, or via the breast milk, might explain this phenomenon. Of course it is also possible that children or their parents are unaware of previous consumption of, or skin contact with, (traces of) cashew nuts.

Based on the review presented in **Chapter 2**, intestinal exposure to soluble monomeric Ana o 1, 2, and 3 should induce tolerance. However, in cashew nut also trimeric (Ana o 1) and hexameric (Ana o 2) allergens are present, which have a higher probability of transport by M-cells, inducing a local immune response.

No improvement in quality of life was gained by the patients participating in the IDEAL study after the food challenge or diagnosis (either positive or negative) [4]. This was explained by the relative ease of avoiding cashew nuts in the diet, which is not as burdensome as avoiding for example dairy- or egg-based products [4]. A negative food challenge also did not often result in the introduction of cashew nuts into the diet on regular basis. Only 43% of the negatively diagnosed children successfully introduced cashew nuts into their diet [3].

Allergy diagnostics

At the moment, the golden standard of food allergy diagnostics is the DBPCFC [14]. However, a DBPCFC is costly and time consuming, as well as invasive for the patients. Other diagnostic tools are the patient history, SPT, mediator release assays, specific IgE measurements, and elimination diets [15]. Many of these tests require well-characterised allergen extracts or purified allergens. In component resolved diagnostics (CRD), purified natural or recombinant allergens are used for serologic patient characterisation; enhancing the knowledge regarding clinical diagnostics and possible cross-reactivity towards other allergens, and allowing the classification of patients into different clinical phenotypes [16, 17]. In the IDEAL study the clinical centres have performed the DBPCFC, while we have purified native allergens and produced recombinant allergens for specific IgE measurements. This was done in order to possibly add more diagnostic tools for the diagnosis of cashew nut allergy, and to avoid the invasive and costly DBPCFC wherever possible.

The purification process for native Ana o 1, 2 and 3, has been described in **Chapter 3**. The purified allergens have subsequently been applied for clinical diagnostics: serum IgE

measurements of Ana o 1, 2 and 3 [2]. Cashew allergic children had (median values) sIgE levels of 2.0, 6.3, and 13.0 kU/l for Ana o 1, 2 and 3 respectively. Sensitized, but not allergic, children had median sIgE levels of 0.2, 1.2, and 0.6 kU/l for Ana o 1, 2 and 3 respectively [2]. The serum levels of Ana o 1, 2 and 3 could better distinct between allergic and sensitized-only patients, than sIgE or SPT of total cashew nut protein. However, based on allergen specific sIgE levels, no distinction could be made between mild and severe allergic reactions [2].
Based on the sIgE data, a predictive model was developed by van der Valk *et al.* that should help to decide whether or not a DBPCFC is required for the diagnosis of a patient [25]. This model includes gender, SPT results, and Ana o 3 sIgE levels.

Besides the use of native purified allergens, also recombinant allergens can be used in sIgE testing. As mentioned in **Chapter 1**, recombinant protein production can give higher yields, is easier to standardise, and the produced allergen will not be contaminated with the presence of other allergens. Purification of allergens directly from the food, as the allergens described in **Chapter 3**, yields allergens with correct post-translational modifications, and the possibility to purify multiple isoforms at once. In **Chapter 5** we described the recombinant production of Ana o 1, 2 and 3 in a yeast (*P. pastoris*) system, and compared them to the native allergens. As described in **Chapter 5**, *P. pastoris*-produced Ana o 1, 2 and 3 were IgE reactive, but differed from the native allergens. Like the native allergen, rAna o 1 was glycosylated, but it was cleaved into two subunits that were mostly not bound to one another, and therefore significant amounts of the N-terminal rAna o 1 polypeptide was lost during purification. Ana o 2 was cleaved into at least 5 subunits, probably by endo peptidases that should cleave Ana o 2 only once to obtain the acidic and basic subunit. Also for Ana o 2 the N-terminal polypeptides were removed during purification as they were not bound to their C-terminal counterparts. Finally, Ana o 3 was also produced differently than it should have been, with the production of two proteins of which one misses its N-terminal polypeptide. Furthermore, Ana o 3 was not cleaved into two subunits as occurs in the cashew nut. All in all the recombinantly-produced cashew nut allergens either need to be further studied to better understand why the proteins were produced as they were, or cloning and production of these allergens should be done again. However, it is difficult to decide whether or not our recombinant allergens are exceptional, or if such differences between native and recombinant allergens occur more often. For example the *E.coli*-produced allergens from Wang and Robotham *et al.* [18-20] were never shown on SDS-PAGE, so if truncation or endopeptidase cleavage of the proteins has occurred remains unclear to the reader. Furthermore, papers that do note such discrepancies between native and recombinant allergens [21-24], only mention such observations briefly, and do not relate to the effect this might have on subsequent applications of such a recombinant allergen.

The cashew nut allergens

Ana o 1, 2 and 3 are the only characterised allergens from cashew nut so far. These three major allergens have been studied regarding their nucleotide sequence [18-20], electrophoretic profile on 1D [18-20, 26], state of glycosylation [26], heat and pepsin/trypsin digestibility [27-29], the identification of linear [18-20] and conformational epitopes (Ana o 2 only) [30, 31], and they have been recombinantly expressed in *E.coli* [18-20]. In this thesis we have characterised purified native allergens further by 2D-electrophoresis (**Chapter 4**) and have expressed recombinant proteins also in the *P. pastoris expression* system (**Chapter 5**). The protein characteristics of these three allergens are summarized in Table 6.1.

As this thesis focusses mostly on the allergens Ana o 1, 2 and 3, a visual representation is provided in Figure 6.1. This Figure shows an SDS-PAGE of a total cashew nut protein extract and the purified allergens Ana o 1, 2 and 3 which are described in **Chapter 3** of this thesis. Ana o 1 is a 50kDa vicilin (7S globulin), Ana o 2 is a 53kDa 11S globulin with acidic (30kDa) and basic (21kDa) subunits, Ana o 3 is a 2S albumin with large subunits of 8.1-8.4kDa (molecular weight on SDS-PAGE 10 and 8kDa) and a small subunit of 3.7-4.5kDa (molecular weight on SDS-PAGE 6kDa). A 53kDa band is indicated containing Ana o 2 protein [26].

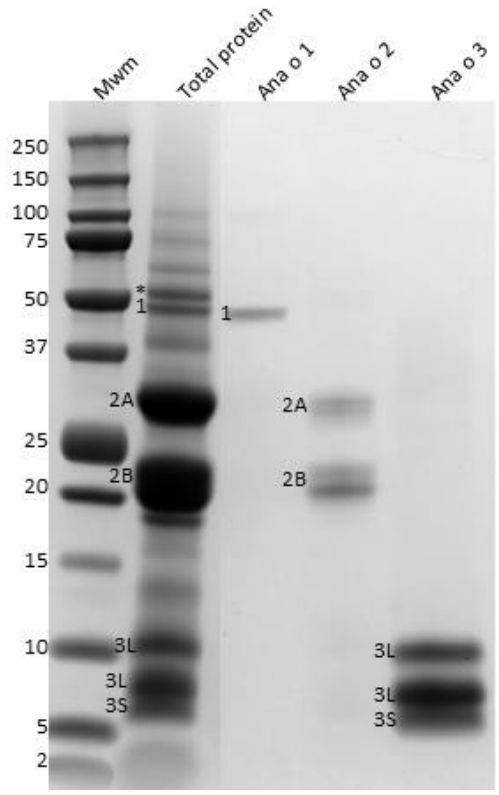


Figure 6.1 *SDS-PAGE of a total cashew nut protein extract, purified Ana o 1, purified Ana o 2, and purified Ana o 3. The molecular weights are indicated on the left. In the figure Ana o 1 is indicated (1, 50kDa), Ana o 2 acidic subunit (2A, ±30kDa), Ana o 2 basic subunit (2B, ±21kDa), Ana o 3 large subunits (3L, 10 and 8kDa), Ana o 3 small subunit (3S, 6kDa). *indicates a 53kDa protein band containing Ana o 2.*

Table 6.1 *Protein characteristics of cashew nut allergens Ana o 1, Ana o 2 and Ana o 3. Data obtained from* ***Chapter 3*** *[26],* ***Chapter 4*** *[27], and* ***Chapter 5*** *[32] from this thesis, and from literature.*

Protein characteristics		**Ana o 1**	**Ana o 2**	**Ana o 3**
General protein characteristics	Genbank number	Ana o 1.0101 AAM73730.2, Ana o 1.0102 AAM73729.1	AAN76862.1	AAL91665.1
	Protein family	Vicilin, 7S globulin [20]	Legumin, 11S globulin [19]	2S albumin [18]
	Known by other names	-	Cashew major protein [29], Anacardein [31]	-
	Leader peptide	Yes, M3-A28 [20]	Yes, L1-A14 [19]	Yes, M1-A20 [18]
	Recombinant expression *E.coli*	Yes [20]	Yes [19]	Yes [18]
	Recombinant expression *P. pastoris*	Yes [32]	Yes [32]	Yes [32]
	Protocol for purification available	Yes [26]	Yes [26]	Yes [26, 33, 34]
1D electrophoresis	Molecular weight reduced	50kDa [20, 26]	21, 30kDa [19, 26, 30]	10, 8, 6kDa [18, 26]
	Molecular weight native	50kDa [20, 26]	53kDa [19, 26, 30]	12.6kDa [18]
	Subunits on SDS-PAGE	No [20, 26]	Yes, 30kDa acidic subunit, 21kDa basic subunit [19, 26, 30]	Yes, 8-10kDa large subunit, 6kDa small subunit [26]
	Quaternary structure	Homo-trimer, 150kDa [35]	Homo-hexamer, ±360kDa [35]	Homo-monomer, 12.6kDa [35]
	Disulphide bridges	No	Yes, 1 [33, 35]	Yes, 4 [35, 36]
	Predicted N-glycosylation	Yes, N391 [26]	No [26]	Yes, N19 [26]
	Glycosylation (native)	Yes [26]	No [26]	No [26]
Mass spectrometry	N-terminal micro-heterogeneity	Unknown, not expected	Unknown, not expected	Yes, large and small subunit [26]
	C-terminal micro-heterogeneity	Unknown, not expected	Unknown, not expected	Yes, small subunit [26]
	Size of subunits	No subunits	Not done	Large subunit 8.1-8.4kDa, small subunit 3.7-4.5kDa [26]
2D electrophoresis	Isoelectric point calculated	5.6 [27]	5.7 (acidic subunit), 8.9 (basic subunit) [27]	4.9 (large subunit) 6.4-10.1 (small subunit) [27]
	Isoelectric point	5.4-5.8 [27]	5.4-6.3 (acidic subunit), 5.4-9.3 (basic subunit) [27]	4.4-5.4 (large subunit), small subunit not visible [27]
	Number of isoforms (spots)	9 [27]	11 (acidic subunit) 18 (basic subunit) [27]	8 (large subunit) [27]
Stability	Pepsin digestibility (400µg protein, 126U	Degraded [27]	Degraded [27]	Resistant [27], degraded after

	pepsin, 60min)			60min when reduced [27]
	Trypsin digestibility (25µg protein, 20U trypsin, 5min)	Degraded [33]	Resistant [33]	Resistant [33]
	Pepsin (25µg protein, 0.8U pepsin, 30min) + trypsin (2U trypsin, 30min)	Degraded [33]	Degraded [33]	Partially degraded [33]
	Heat stability in-shell: steaming, frying, drum roasting	Decreased solubility after steaming [27]	Increased solubility after frying [27]	Decreased solubility after frying [27]
	Heat stability de-shelled: roasting	Decreased solubility [28]	Decreased solubility [28]	Increased solubility [28]
	Heat stability de-shelled: Blanching, γ - irradiation, microwaving, dry roasting	No effect on linear epitopes (antibody binding) [29]	No effect on linear epitopes (antibody binding) [29]	Decreased IgG binding to a conformational epitope [29]
	Heat stability: autoclaving	Decreased mAb binding to linear epitopes [29]	Decreased mAb binding to linear epitopes [29]	Decreased IgG binding to a conformational epitope [29]
	pH stability	Stable except pH 1, 13 [29]	Stable except pH 1, 13 [29]	Stable except pH 1, 12, 13 [29]
Difference between multiple cashew nut origins	1D electrophoresis	No significant differences [27]	No significant differences [27]	No significant differences [27]
	2D electrophoresis	No significant differences [27]	No significant differences [27]	No significant differences [27]
IgE binding characteristics	Number of linear epitopes	11 [20]	22 [19]	16 [18], 12 [26]
	Conformational epitopes	Unknown	1 [30, 31]	Unknown
	IgE reactive isoforms 2D electrophoresis	9/9 [27]	11/11 (acidic subunit) 11/18 (basic subunit) [27]	4/8 [27]
Clinical characteristics	sIgE cashew nut allergic children	2.0 kU/l [2]	6.3 kU/l [2]	13.0 kU/l [2]
	sIgE cashew nut sensitized children (not allergic)	0.2 kU/l [2]	1.2 kU/l [2]	0.6 kU/l [2]
	Predictive value for cashew nut allergy	Predictive [2]	Predictive [2]	Predictive [2, 37, 38]
Cross-reactivity	Homology to other proteins	80% identical to Pis v 3 [39], 27% identical to Ara h 1 [20]	48% identical to Pis v 2 [40], 43% identical to Ara h 3, 42% identical to Ara h 4 [19]	64% identical to Pis v 1 [40]
	In vitro cross-reactivity	Pis v 3 [39]	Pis v 2 [40]	Pis v 1 [40]
	Clinical cross-reactivity	Pistachio [38-43], mango [44-46]		

The development of allergies towards cashew nut proteins is not unexpected since vicilin proteins (Ana o 1), 11S globulins (Ana o 2), and 2S albumins (Ana o 3) have been identified as allergens in a multitude of plants, including most tree nuts. Vicilin proteins, like Ana o 1, have been identified in, among others, chickpea (Cic a 1), coconut (Coc n 1), hazelnut (Cor a 11), lentil (Len c 1), peanut (Ara h 1), pea (Pis s 1, Pis s 2), pecan (Car i 2), pistachio (Pis v 3), sesame (Ses i 3), soy (Gly m 5), and walnut (Jug r 2) [48]. 11S Globulins, such as Ana o 2, have been identified in almond (Pru du 6), Brazil nut (Ber e 2), chickpea (Cic a 6), coconut (Coc n 4), hazelnut (Cor a 9), kiwi (Act d 12), peanut (Ara h 3, Ara h 4), pecan (Car i 4), pistachio (Pis v 2, Pis v 5), sesame (Ses i 6, Ses i 7), soy (Gly m 6), walnut (Jug r 4), and white mustard (Sin a 2) [48]. 2S Albumins, like Ana o 3, have been identified in Brazil nut (Ber e 1), butternut (Jug ca 1), castor bean (Ric c 1, Ric c 3), hazelnut (Cor a 14), kiwi (Act d 13), oriental mustard (Bra j 1), peanut (Ara h 2, Ara h 6, Ara h 7), pecan (Car i 1), pistachio (Pis v 1), rapeseed (Bra n 1), sesame (Ses i 1, Ses i 2), soy (Gly m 8), walnut (Jug n 1, Jug r 1), and white mustard (Sin a 1) [48]. Not only do these three allergens have much in common with other allergenic proteins, but also many other common characteristics of allergenic proteins can be noticed for these three allergens (see Table 6.2):

1. First of all, a crucial requirement for proteins to act as allergens is to contain at least two IgE-binding epitopes. Without these two epitopes, the required cross-linking of two IgE molecules cannot occur, and no allergic reaction will take place. Ana o 1 contains 11 linear epitopes [20], Ana o 2 contains 22 linear epitopes [19], as well as at least one conformational epitope [30, 31], and Ana o 3 contains at least 12 [26] or 16 [18] epitopes. The difference in number of epitopes detected in our study (12 epitopes, **Chapter 3**) and in the paper of Robotham *et al.* (16 epitopes [18]) might have been caused by the different patient serum used (derived from American adults with self-reported cashew nut allergy), different method of peptide production, different manner of IgE detection on blot, or it could be caused by a difference in cut-off point of positive IgE binding. In our study, not only did we detect 9 epitopes within the large and small subunit of Ana o 3, we also detected 3 mildly IgE-reactive epitopes in the leader peptide of Ana o 3. This leader peptide was not included in the epitope analysis of Robotham *et al.* A leader peptide (signal peptide) directs the transport of the protein towards the endoplasmic reticulum, and is cleaved from the pre-protein, after which the leader peptide is expected to be degraded [49]. As it is not certain that all leader peptides of Ana o 3 are degraded in the edible nut, we choose to take along the leader peptide in our epitope analysis.
2. Secondly, Ana o 1, 2 and 3 all three function as seed storage proteins. In general, allergens are not enzymatic or structural proteins that are metabolically required in the plant [50].

3. Next, most allergens have a molecular mass between 10 and 60kDa [51, 52], a molecular weight range that includes Ana o 1 (50kDa), 2 (53kDa) and 3 (12.6kDa).
4. Also, high abundance of a protein (>1% of the total protein content [53] is an often occurring characteristic of allergenic proteins that is also observed for Ana o 1 (±4%), Ana o 2 (±52%), and Ana o 3 (±25%) (**Chapter 4**). It is likely that this high abundance increases the likelihood of (part of) the protein surviving the digestive tract, and coming into contact with the immune system, possibly setting off a sensitization reaction.
5. Besides the high abundance of allergenic proteins within the total protein content, also the total protein content is often relatively high (>20%) in allergenic foods [54]. Many allergenic foods with a protein content <30% often have high levels of fat, while non-allergenic foods with a protein content <30% often have high levels of carbohydrates. The food matrix might interfere with the digestion process by competition (protein) [55], and by protecting the proteins inside lipid structures [56]. Cashew nuts have both a high protein content (±16.8%) (**Chapter 4**), and a high lipid content (±50%) [57-60].
6. This before-mentioned characteristic (high protein content) is also related with higher levels of glycation (Maillard reaction) [61]. Roasting of cashew nuts increases the presence of advanced glycation end products (AGEs) from 6730kU/100g to 9807kU/100g cashew nuts. The formation of AGEs is stronger during high heat low moisture processing steps (dry roasting, frying), compared to heating steps in the presence of water (e.g. steaming, boiling) [61]. In **Chapter 3** we described that, of the cashew nut allergens, only Ana o 1 is glycosylated. Upon heating, no glycation of the cashew nut proteins was observed in our study (**Chapter 4**), however, the heating was applied to cashew nuts within the shell, shielding the nuts from the heat. Indeed, upon removal of the shell, there was no difference in colour between raw and heated (20min 100˚C steamed, 2min 180˚C fried, 8min 150˚C drum roasted) cashew nuts [27]. In the raw, as well in as the heated cashew nuts, a strongly glycated 3kDa cashew nut protein was detected [26, 62]. This 3kDa protein has not been recognised as an allergen by western blotting, possibly this 3kDa protein is too small to contain multiple epitopes [56]. Glycation may also influence a proteins' allergenicity. Huby *et al*. stated that many allergens are glycosylated [63]. The effect of glycation on the allergenic reaction is not yet fully determined. For example, in peanut, glycation of raw Ara h 1 (7S vicilin) did not alter its' immunoreactivity (reverse EAST inhibition assay) [64], while glycation of Ara h 2 and Ara h 6 (2S albumins) reduced IgE binding (ELISA) [65], glycated rAra h 2 increased IgE binding (ELISA, EAST assay) [66], and glycation of a total peanut protein extract increased IgE binding (western blot) [67].

7. A second to last characteristic that is often observed in protein allergens, is stability of their immunological properties to (*in vitro*) digestion [56, 68, 69]. For example peanut allergens have been shown to be still immunologically active after gastric digestion [70]. However, resistance to digestibility does not always provide a clear distinction between allergenic and non-allergenic proteins [71], and results can differ between studies [72]. Overall, 2S albumins (Brazil nut, mustard, soy, sunflower seed) seem to be consistently stable against pepsin digestion [72]. In **Chapter 4**, we reported that, of the three allergens, the 2S albumin Ana o 3 is most resistant to pepsin digestion. After 60min of pepsin digestion, using physiologically relevant levels of pepsin [56], Ana o 3 was still detectable on SDS-PAGE, while Ana o 1 and 2 were not. No difference in pepsin digestibility was observed between proteins obtained from cashew nuts that had been exposed to different heat treatments in-shell. When the cashew nut proteins were reduced by urea and DTT before pepsin digestion, disrupting the disulphide bridges and thereby the protein conformation, Ana o 3 was more easily digested. With this experiment the critical role of the disulphide bridges for digestion stability is demonstrated.
   Mattison *et al*. [33], also observed a higher pepsin digestion stability of Ana o 3 compared to Ana o 1 and 2. In their study mass spectrometric analysis revealed the presence of 5-15kDa polypeptides from both Ana o 2 and Ana o 3 after pepsin digestion [33]. As we have not performed such an analysis on our digested proteins, we cannot confirm if protein fragments of Ana o 2 were obscured by the protein bands of Ana o 3. However, it should be noted that Mattison *et al.* used about ten times less pepsin in this analysis. Also, we did not observe protein bands above 12kDa (**Chapter 4**), and in our data the protein pattern on SDS-PAGE in the 6- 10kDa range of the digested proteins, is identical to the non-digested protein pattern of Ana o 3 in this same 6-10kDa range. Mattison *et al.* also studied cashew nut protein digestion, using both pepsin and trypsin, each for 30 minutes [33]. Only 10kDa proteins were visible after this digestion step, containing mostly Ana o 3 and minor amounts of Ana o 2 polypeptides [33].
   It is important to realise that a protein that is more easily digested, is expected to be less allergenic as the chance of sensitization is less than for digestion-stable proteins. Once a person is allergic, however, the allergic reaction often occurs quickly after consumption of the allergen [73], before the proteins reach the digestive enzymes. For peanut it was shown that uptake of immune-reactive peanut polypeptides, occurred already 10 minutes after ingestion of the peanuts [74]. The absorption of immune-reactive peanut polypeptides seemed to already start in the mouth, no swallowing of the peanuts seems necessary [74].

8. Lastly, stability of their immunological properties upon (heat-) processing, is a frequently observed characteristic of allergenic proteins [70]. As summarized by Rahaman *et al.* [75], processing may enhance, reduce, or have no effect on protein allergenicity [75]. Processing-induced modifications, such as unfolding, aggregation, glycosylation, and enzymatic hydrolysis, can lead to the disruption and the exposure of linear and conformational epitopes [75]. For example, roasting of peanut can increase IgE binding (ELISA) to the total protein [76]. Roasting of purified Ara h 1 (7s globulin) did not change its IgE binding capacity (EAST assay) [64], while boiling purified Ara h 1 induced the formation of large aggregates and decreased its IgE binding capacity (EAST assay, histamine release by RBL cells) [64]. For cashew nut proteins, dark roasting strongly increased the solubility of Ana o 3, and decreased the solubility of Ana o 1 and 2 as shown on SDS-PAGE and confirmed by ELISA (human IgE, polyclonal rabbit anti-cashew) [28]. Blanching, γ -irradiation, microwaving, and dry roasting did not affect Ana o 1 and 2 binding by a mAb [29]. IgG binding to Ana o 1 and 2 was however affected after 20-30minutes autoclaving, and after incubation at pH 1 and 13. IgG binding of the mAb to Ana o 3 was reduced after blanching, roasting, and autoclaving [29]. However, the specific mAb used in that study targets a conformational epitope of Ana o 3, so no conclusion can be drawn on the effect of these heat treatments on the linear epitopes of Ana o 3. When studying the effect of blanching on the protein composition of food items, it should be kept in mind that proteins can leak into the boiling water. In Figure 6.2, both a raw and a blanched (10min 100˚C) cashew nut protein extract, as well as the concentrated blanching water are presented. Multiple cashew nut proteins elute into the cooking water (lane blanching water), but as Ana o 2 is the most prevalent protein, this is also the most prevalent protein in the cooking water. About 0.2% of the proteins eluted from the cooking water in this experiment. Venkatachalam *et al.* indicated that Ana o 1, 2 and 3 all elute into the cooking water (western blotting, mAb) [29]. Elution of allergens into the cooking water has also been described for peanut [77].

   Cashew nuts always undergo multiple heating steps before they are ready for consumption. As described in **Chapter 1**, the nut is first heated within its shell, and a second heat treatment is later applied to the de-shelled cashew nuts [78]. In **Chapter 4** we have described the effect of this first (in-shell) heat treatment on the cashew nut proteins. Only small differences in protein composition (1D and 2D electrophoresis) were detected between cashews subjected, in-shell, to steaming, frying, and drum roasting. For example the quantity of extracted Ana o 1 was lower in steamed cashew nuts than in fried cashew nuts, the quantity of extracted Ana o 2 was higher in fried cashew nuts than in raw cashew nuts, and the quantity of Ana o 3 was lower in fried cashew nuts than in raw or drum roasted cashew nuts [27]. No

difference in IgE binding was noted between the differently heat-treated cashew nuts. The higher level of soluble Ana o 3 in the drum roasted cashews was seen before [28], however, the increase observed by Mattison *et al.* was much higher (40% [28]) compared to our results (3%). This difference can be explained by the shorter roasting time applied in our study, and by the moderating effect of the thick cashew nut shell in our study, as we applied 8min 150˚C to in-shell cashew nuts [27], and Mattison *et al.* applied 24min 149˚C to de-shelled cashew nuts [28].

Table 6.2 *Allergenic characteristics of Ana o 1, Ana o 2, and Ana o 3. Data obtained from literature, and from* ***Chapter 3*** *[26] and* ***Chapter 4*** *[27] from this thesis. Symbols indicate high (✓),intermediate (±), and no (✗) correlation with allergenicity.*

Characteristic	**Ana o 1**		**Ana o 2**		**Ana o 3**	
Homology to other allergens	80% identical to Pis v 3 [39], 27% identical to Ara h 1 [20]	✓	48% identical to Pis v 2 [40], 43% identical to Ara h 3, 42$ identical to Ara h 4 [19]	✓	64% identical to Pis v 1 [40]	✓
Multiple epitopes	11 linear [20]	✓	22 linear [19] 1 conformational [30, 31]	✓	12-16 linear [18, 26]	✓
No daily used protein	7S globulin, seed storage protein	✓	7S globulin, seed storage protein	✓	2S albumin, seed storage protein	✓
Protein 10-60kDa	50kDa [20, 26]	✓	53kDa (30+21kDa subunits) [19, 26]	✓	12.6kDa [18, 26]	✓
High abundance specific protein (>1%)	4% [27]	✓	52% [27]	✓	25% [27]	✓
High total protein levels, high fat levels	±16.8% protein [27], ±50% lipid [57-60]	✓	±16.8% protein [27], ±50% lipid [57-60]	✓	±16.8% protein [27], ±50% lipid [57-60]	✓
Glycation	Yes [26]	✓	No [26]	✗	No [26]	✗
Digestion stability	Low [27, 33]	✗	Middle [27, 33]	±	High [27, 33]	✓
In-shell processing stability (steam, fry, drum roast)	Only minor effects on protein solubility [27]	✓	Only minor effects on protein solubility [27]	✓	Only minor effects on protein solubility [27]	✓
De-shelled processing stability (Blanching, γ - irradiation, microwaving, dry roasting)	No effect on linear epitopes (antibody binding) [29]	✓	No effect on linear epitopes (antibody binding) [29]	✓	Decreased IgG binding to a conformational epitope [29]	±
De-shelled processing stability (autoclaving, pH 1 or 13)	Decreased mAb binding to linear epitopes [29]	✗	Decreased mAb binding to linear epitopes [29]	✗	Decreased IgG binding to a conformational epitope [29]	±

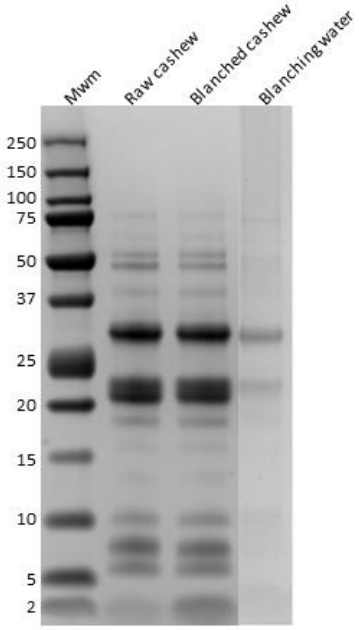


Figure 6.2 *Total protein extract of raw and blanched (10min 100˚C) acetone-defatted cashew nuts (0.1 M ammonium bicarbonate, 0.5M NaCl, 1:5 w/v o/n), 15µg protein loading. Blanching water is 20 times concentrated by evaporation. The molecular weights are indicated on the left.*

Taken together, Ana o 1, 2 and 3 are important and highly stable cashew nut allergens. In **Chapter 4** we have proven that cashew nuts, derived from different origins, all contain Ana o 1, 2 and 3 in similar quantities (as judged from 1D and 2D electrophoresis). Also, no differences were observed in the allergenic isoforms present in the cashew nuts of different origins, as determined by 2D electrophoresis [27]. These results indicate that for the available literature on cashew nuts, each study using different cashew nuts, these parameters are not likely to explain differences in characteristics of cashew nut allergens between different studies.

Both the cDNA sequences for the recombinant production of Ana o 1, 2 and 3, as well as the molecular weight of the three allergens were based on publications by the group of dr Kenneth Roux (Florida State University). However, some conflicting results between our results and their publications were obtained and described here:
First of all, *Ana o 1* is labelled as a major allergen as IgE from 50% (10/20) of a group of cashew allergic patients bound recombinant Ana o 1 [20]. In our studies, we have not performed dot blotting or western blotting experiments with purified or recombinant Ana o 1 with single sera. We have always used serum pools in order to minimize the amount of serum needed for the experiments. The patients in this study were children, and therefore limited amount of blood was drawn at the medical centres. When pooling sera, specific IgE is diluted resulting in a less strong IgE binding pattern, but high levels of specific IgE from one person can skew the IgE binding pattern from the pool towards a single allergen. In our study we have had difficulties with some serum pools to detect IgE binding to Ana o 1, while IgE binding to Ana o 2 and Ana o 3, both also major allergens [18, 19] was almost always detected. A difference in population or age of the patients might explain these differences in

IgE binding patterns [79, 80]. Furthermore, the IgE levels of the cashew allergic children (positive DBPCFC) to Ana o 1 are lower (2.0 kU/l) than the IgE levels to Ana o 2 (6.3 kU/l) and Ana o 3 (13.0 kU/l) [2]. Possibly the pooling of the sera, combined with the low quantity of Ana o 1 in the protein extracts (4% **Chapter 4**), lead to the occasional difficulty in detecting Ana o 1 on western blot.

Secondly, in **Chapter 3**, we have tentatively identified a 53kDa protein from reducing SDS-PAGE as the monomeric form of *Ana o 2*. Under non-reducing conditions, more of this 53kDa protein is detected on SDS-PAGE. However, it was expected that reduced Ana o 2 would only be present in the form of acidic and basic subunits, and not as a 53kDa primary translation product. Wang *et al.* also detected Ana o 2 in this 53kDa protein band, as inhibition western blotting, using *E.coli*-expressed rAna o 2 as inhibitor, decreased IgE binding to this 53kDa protein [19]. Possibly not enough β-mercaptoethanol was added to reduce Ana o 2 fully. On the other hand, since as much as 10% β-mercaptoethanol is added to reduce the proteins, and as purified Ana o 2 does not display this 53kDa protein band, it is not expected that this is the reason. Most likely not all Ana o 2 is cleaved into subunits during the formation of this protein. This also implies that this 53kDa monomeric Ana o 2 is not present in the purified Ana o 2 since it is not detected on SDS-PAGE.

Thirdly, the study of *Ana o 3* resulted in some results that conflicted with data from literature. The first puzzle that needed to be solved was the molecular weight of Ana o 3. Ana o 3 is reported as a 12.6kDa protein [18], while on reducing SDS-PAGE Ana o 3 is detected as three protein bands of 10, 8 and 6kDa (Figure 6.1). A 2S albumin is usually synthesized as a precursor polypeptide, which is subsequently cleaved into a large (8-10kDa) and a small subunit (3-4kDa) [36]. We were curious what subunits of Ana o 3 are present at what molecular weight on SDS-PAGE, since 8 and 4kDa subunits were expected, while 10, 8 and 6kDa subunits were detected. Furthermore, Robotham *et al.* reported that all three protein bands were large subunit polypeptides of Ana o 3, as all three had a similar N-terminal amino acid sequence. The small subunit was not mentioned in this study. As described in **Chapter 3**, we detected polypeptides of the large subunit in the 10 and 8kDa protein bands, while small subunit polypeptides were detected in the 6kDa protein band. Furthermore, N- and C-terminal micro-heterogeneity (terminal clipping) was detected in the small subunit of Ana o 3, only minimal C-terminal micro-heterogeneity was detected in the large subunit. These results are in contrast with the data from Robotham *et al.* as they observed large subunit N-terminal amino acids in the 6kDa protein band. As the three protein bands are very close together on SDS-PAGE, it might be that proteins from the 8kDa protein band, but not the 6kDa protein band was analysed by Robotham *et al.* No N-terminal clipping was reported by Robotham *et al.*, matching our results that no N-terminal clipping occurs in the large subunit.

For major fish allergen parvalbumin it was suggested, based on protein sequence comparisons and an allergen database, that protein sequence micro-heterogeneity is related to increased protein isoform allergenicity [81]. Of course, micro-heterogeneity might also reduce IgE binding [82], for example when occurring in the epitopes of the allergen. As reported in **Chapter 4**, the N- and C-terminal micro-heterogeneity of Ana o 3 in cashew nuts from 8 different origins was highly similar. Therefore no difference in allergenicity is expected between Ana o 3 proteins from different origins. Also it should be noted that most micro-heterogeneity for Ana o 3 was detected in the N-terminus of the small subunit. We have not detected any epitopes in that region (**Chapter 3**), but Robotham *et al.* [18] did.
The second puzzle concerning Ana o 3, concerns the IgE binding to the small subunit (6kDa protein band) of Ana o 3. Despite the presence of 3 (**Chapter 3**) to 7 [18] epitopes on the small subunit of Ana o 3, no IgE binding was observed in any of the multiple western blots we have performed. In literature, only Teuber *et al.* depicts a western blot showing clear IgE binding to all three protein bands of Ana o 3 [83]. This difference in IgE binding can be explained by a difference in serum used on the western blot, our study population apparently binds stronger to the large subunit of Ana o 3 than to the small subunit.

In **Chapter 4**, we have reported that Ana o 1, Ana o 2, and Ana o 3, all contain multiple isoforms. Based on 2D electrophoresis and western blot of both the purified allergens and a total cashew nut protein extract, the isoforms of Ana o 1, 2 and 3 were tentatively identified. We detected 9 isoforms of Ana o 1, 29 isoforms of Ana o 2 (11 isoforms of the acidic subunit, 18 spots of the basic subunit), and 8 isoforms of Ana o 3 (4 isoforms of the 10kDa large subunit, 4 isoforms of the 8kDa large subunit, the small subunit was not observed on 2D electrophoresis). Occurrence of these different isoforms is expected to be due to post-translational modifications such as phosphorylation, glycosylation (Ana o 1), acetylation, etc. Upon western blotting IgE was bound by all isoforms of Ana o 1, all isoforms of the large subunit of Ana o 2, 11 out of 18 isoforms of the small subunit of Ana o 2, and 4 out of the 8 isoforms of the large subunit of Ana o 3. The absence of IgE binding to some of the isoforms of Ana o 2 and 3 might be patient-specific as only one plasma pool was used for these western blots, but it is also possible that the epitopes of these specific isoforms are different from the other isoforms. A last reason for the lack of IgE binding to some of the allergen isoforms is that these might be different proteins, which were not detected on 1D SDS-PAGE because of the similarity in molecular weight with the purified allergen. Future experiments should include protein sequencing experiments in order to confirm the identity of these tentatively identified allergen isoforms.

The allergens Ana o 1, 2 and 3 were purified in **Chapter 3**. The purity of these samples was 96% for Ana o 1, 93% for Ana o 2, and 99% for Ana o 3, as based on SDS-PAGE analysis, and

88% for Ana o 3 as based on HPLC analysis. Based on 2D electrophoresis and 2D western blot, Ana o 1 seems indeed of high purity; all 50kDa protein spots bound IgE, indicating, but not yet confirming, all protein spots are isoforms of Ana o 1. A minor impurity of 30kDa was detected both on 1D (**Chapter 3**) and 2D electrophoresis (**Chapter 4**). This impurity is most likely the acidic subunit of Ana o 2 as it is bound by IgE upon western blotting. On inhibition western blot (total cashew nut protein western blot with plasma pre-incubated with purified Ana o 1), IgE binding to Ana o 1, but also partially to Ana o 2 and 3, was inhibited by purified Ana o 1. The impurity of Ana o 2 in the purified Ana o 1 extract explains the inhibition of IgE binding to Ana o 2. The impurity might be only a small percentage, but for inhibition a large quantity of protein was used (1mg/ml). Next, the inhibition of Ana o 3 can be caused by possible cross-reactivity between 7S globulins and 2S albumins [84].
The purity of Ana o 2 is high based on 1D SDS-PAGE (93%), however, it should be noted that the 53kDa monomeric Ana o 2 is not detected in the purified protein sample (**Chapter 3, 4**). Furthermore, western blotting of a 2D electrophoresis gel of purified Ana o 2 indicates that not all protein spots bind IgE (**Chapter 4**). These non-IgE reactive proteins could be Ana o 2 with modifications in the epitope region, otherwise these proteins might be impurities with the same molecular weight as Ana o 2. The purity of this sample should be confirmed by HPLC analysis, and protein identification analysis should be performed on the 2D electrophoresis protein spots.
The highest purity was observed for Ana o 3 (99%). On 1D SDS-PAGE no impurities were observed, while on 2D electrophoresis not every protein spot was bound by IgE. Again, as for Ana o 2, either these protein spots are isoforms of Ana o 3 with a difference in the epitopes, or these protein spots are impurities of the same molecular weight as Ana o 3.

Besides the known allergens Ana o 1, 2 and 3, it is highly likely that also other allergens are present in cashew nuts. In peanut for example, 11 allergens have been detected: 7S vicilin Ara h 1, 11S globulin Ara h 3 and 4, 2S albumin Ara h 2, 6, and 7, profilin Ara h 5, pathogenesis related (PR) Ara h 8, nonspecific lipid transfer protein (nsLTP) Ara h 9, and oleosin Ara h 10 and 11 [85]. In tree nuts, such as almond, chestnut, hazelnut, and walnut also similar allergens (nsLTP, oleosin, profilin, PR) have been identified [86, 87]. On our western blots (**Chapter 3, 4**) several IgE-binding protein bands are observed that are not Ana o 1, 2 or 3: E.g. 80, 60, 40, 17, and 13kDa protein bands. Based on the protein size comparison with the peanut allergens, the 13kDa cashew nut protein might be a profilin, and the 17kDa protein could be a PR protein. The nsLTP proteins have a similar molecular weight as Ana o 3 (7-9kDa [35]), which makes the identification of this protein difficult. The 60kDa IgE-binding cashew nut protein is of comparable molecular weight as Ara h 1 (7S vicilin) and Ara h 3 and 4 (11S globulins). Oleosins are fat-soluble proteins, and as cashew nuts are often defatted prior to protein extraction, such proteins are likely not to be detected on western

blot. Also in the experiments in this thesis the cashew nuts were defatted prior to protein extraction, therefore it is expected that no oleosins were present on SDS-PAGE or western blot. It is highly recommended that further protein identification studies should be performed concerning these unknown allergens from cashew nut.

Co-sensitization

Of the children that participated in the IDEAL study, 50% reported having experienced an allergic reaction towards pistachio, 7% to mango, 22% to hazelnut, and 34% to peanut, after consumption of these foods. Based on serum IgE (sIgE) measurements in all cashew nut sIgE-sensitized children, 98% children were co-sensitized to pistachio, 21% to mango, 69% to hazelnut, and 62% to peanut. Co-sensitization of children when measured by skin prick test (SPT) in cashew nut sensitized children (positive cashew nut SPT), resulted in 92% co-sensitization to pistachio, 19% to mango, 44% to hazelnut, and 59% to peanut [1]. As can be observed from Figure 6.3, mango (*Mangifera indica*, fruit) and pistachio (*Pistacia vera*, nut/seed) belong, together with cashew nut (*Anacardium occidentale*), sumac (*Rhus coriaria*, spice), and pink pepper (*Schinus molle* and *Schinis terebinthifolius*, berry used as spice), to the Anacardiaceae family. Besides the Anacardiaceae family, also other tree nuts beside cashew nut and pistachio are indicated in this figure. Peanut is not included in this figure as it is not a tree nut.

	7S globulin	**11S globulin**	**2S albumin**	**Profillin**	**Chitinase**	**Other**		
Cashew	Ana o 1 (N+R 50kDa) [20, 26]	Ana o 2 (N 53kDa, R 21,30kDa) [19, 26, 33]	Ana o 3 (N 13kDa, R 10, 8, 6kDa) [26, 96]					
Pistachio	Pis v 3 (R 45kDa) [39]	Pis v 2 (R 32kDa) [40] Pis v 5 (R 36kDa) [48]	Pis v 1 (R 7kDa) [40]			Pis v 4 (R 26kDa, manganese superoxide dismutase,[48, 97])		
Mango				Man i 3.01, Man i 3.02 [48]	Unnamed chitinase (R 46kDa) [48, 98]	Unnamed Bet v 1-like protein (R 14kDa) [48]	Man i 1 (R 40kDa) [48, 99]	Man i 2 (R 30kDa) [48, 99]
Pink pepper	Unknown							

Table 6.3 *Allergens known from cashew nut, pistachio, mango and pink pepper. Allergen with molecular weight in native (N) or reduced (R) state. Besides these allergens the protein sequence of several other proteins, e.g. rubisco and energy metabolism related proteins, are known for all four mentioned foods (NCBI, protein index).*

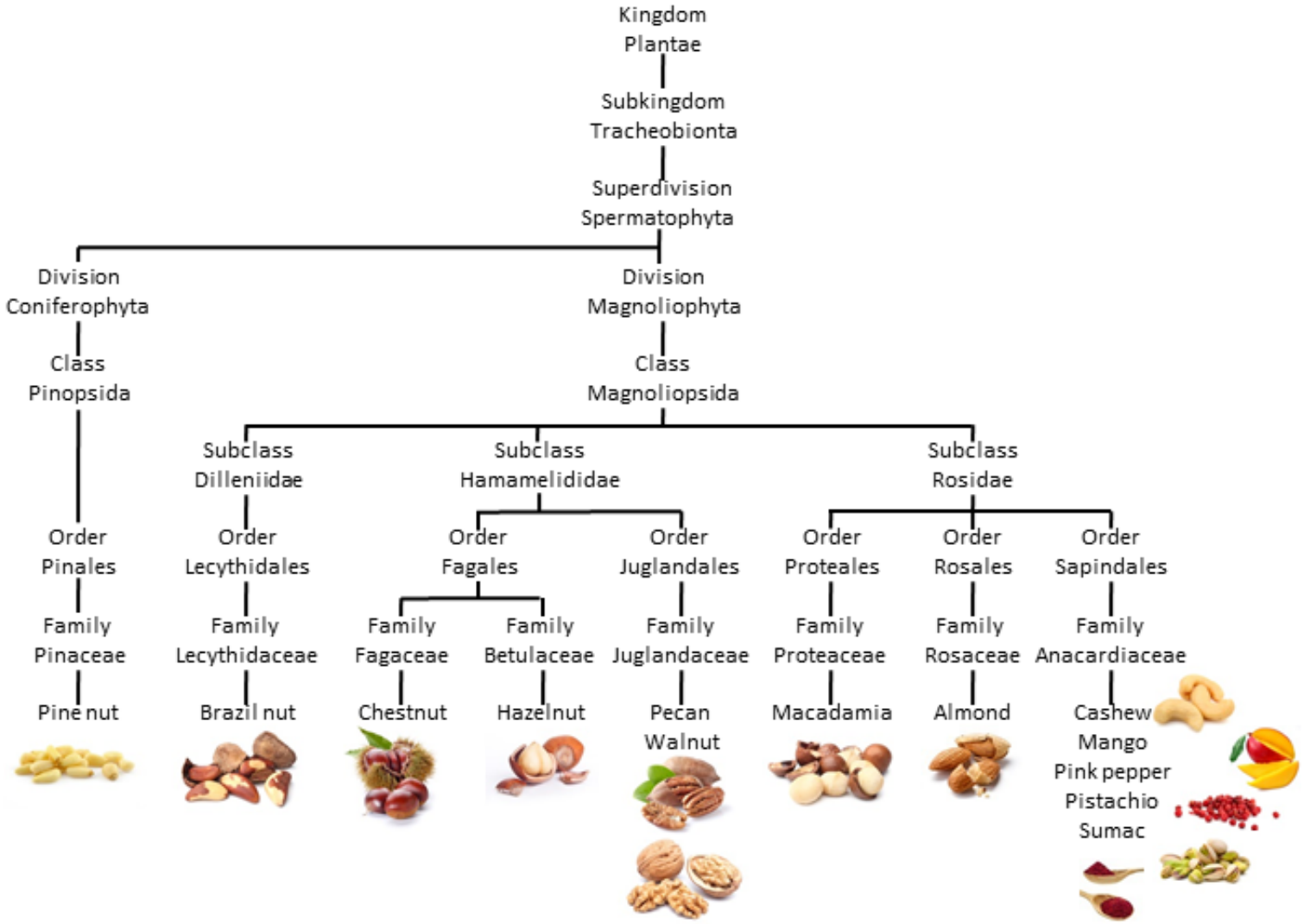


***Figure 6.3**. Taxonomic classification of Anacardiaceae family members and tree nuts. Taxonomy data adapted from de Leon et al. [47], mod3fied using data from plants.usda.gov.*

Because of the high co-sensitization rate of cashew nut allergic patients to pistachio and hazelnut that was observed in the children in our study, we initiated a sub-study into co-sensitization towards multiple tree nuts and Anacardiaceae family members: cashew nut, mango, pink pepper, pistachio, sumac, pine nut, Brazil nut, chestnut, hazelnut, pecan, walnut, macadamia, and almond.

The chance of having an allergy to more than one tree nut is considerable: 35-37% based on clinical history [88, 89], or 14% (2/14) based on food challenges [90]. Allergic cross-reactivity between cashew nut and pistachio has been well-described [41-43, 91], and avoidance of both nuts is advised when diagnosed with cashew nut allergy [92]. More specifically, the cashew nut allergen Ana o 1 and the pistachio allergen Pis v 3, both vicilin proteins, display substantial cross-reactivity [39]. Considering cross-reactivity within the Anacardiaceae family, less research has been done compared to tree nut cross-reactivity. A few cases have been reported where mango allergic individuals were also allergic to pistachio [93], cashew [94], or cashew apple [95]. Other studies on cross-reactivity between cashew and other members of the Anacardiaceae family, such as pink pepper or sumac, could not be found. An overview of all allergens described from cashew, pistachio and mango is presented in Table 6.3.

For this cross-allergenicity study, protein was extracted from raw nuts and fruits (cashew nut, mango, pink pepper, pistachio, sumac, pine nut, Brazil nut, chestnut, hazelnut, pecan, walnut, macadamia, almond) into an urea buffer (8M Urea, 20mM Sodium Phosphate pH7, 1mM NaCl), and into a tris buffer (20mM Tris pH7.6, 150mM NaCl, 1mM EDTA). The tris and urea protein extracts were mixed 1:1 based on equal protein concentrations. Dot blots were spotted in duplicate with 250ng protein from the mixed tris/urea protein extracts. The sera of 57 IDEAL-study patients were studied on dot blot. Of these 57 patients, 14 tested negative for cashew nut allergy in the DBPCFC [1], 41 tested positive, and 2 were undecided. Based on the dot blots, the 57 patients could be categorized into four groups:

- Group 1, Anacardiaceae sensitized: positive for cashew nut, positive for at least one other member of the Anacardiaceae family, positive for up to one tree nut outside the Anacardiaceae family; 7 patients (7/7 positive DBPCFC): serum pool 1.
- Group 2, cashew nut sensitized: positive for cashew nut, not positive for other members of the Anacardiaceae family, positive for up to two tree nuts outside the Anacardiaceae family; 18 patients (14/18 positive DBPCFC).
- Group 3, Anacardiaceae and tree nut sensitized: positive for cashew nut, positive for at least one other members of the Anacardiaceae family, positive for more than two tree nuts outside the Anacardiaceae family; 11 patients (8/11 positive DBPCFC): serum pool 3.

- Group 4, not cashew nut sensitized: not positive for cashew nut; 21 patients (12/21 DBPCFC positive, 2/21 undecided DBPCFC).

In group 4, false negative results were obtained in 12 out of 20 patients that did not show IgE binding to cashew nut proteins on dot blot but are allergic to cashew nuts as diagnosed by DBPCFC. These results might be caused by the use of raw cashew nuts for the dot blots, while in the DBPCFC roasted cashew nuts were incorporated into muffins. Furthermore these muffins contained sugar and were heated, quite likely causing Maillard reactions, possibly forming glycated cashew nut proteins, which may increase allergic reactivity [66, 67]. Also it is possible that during the protein extraction certain allergens, such as fat soluble oleosins, are missed, and are therefore not present on dot blot while having been present in the muffin. Lastly, the possibility was considered that these patients experienced mainly oral or abdominal complaints during the DBPCFC. Experience of mostly oral complaints could be an indication that these patients responded to unstable proteins, which were denatured by the low pH of the stomach, and which were then possibly also denatured by the high levels of urea in one of the extraction buffers. Experience of mostly abdominal complaints could be an indication that these patients responded to epitopes which are hidden inside the proteins, exposed only after the denaturing effect of the stomach, which might not occur in the protein extraction buffer for the dot blot. However, no significant difference was detected in the occurrence of oral or abdominal complaints (nausea, vomiting, and diarrhoea) between the patients with false negative dot blot results and the patients responding positive on both dot bot and DBPCFC. The false positive results obtained in this experiment (7/36 patients) are expected to be due to the presence of IgE in sensitized, but not allergic, individuals.

Western blotting (Figure 6.4) was performed on cashew nut, pistachio, mango, pink pepper, and sumac proteins, using two serum pools; serum pool 1 (patient group 1, Anacardiaceae sensitized) and serum pool 2 (patient group 3, Anacardiaceae and tree nut sensitized).

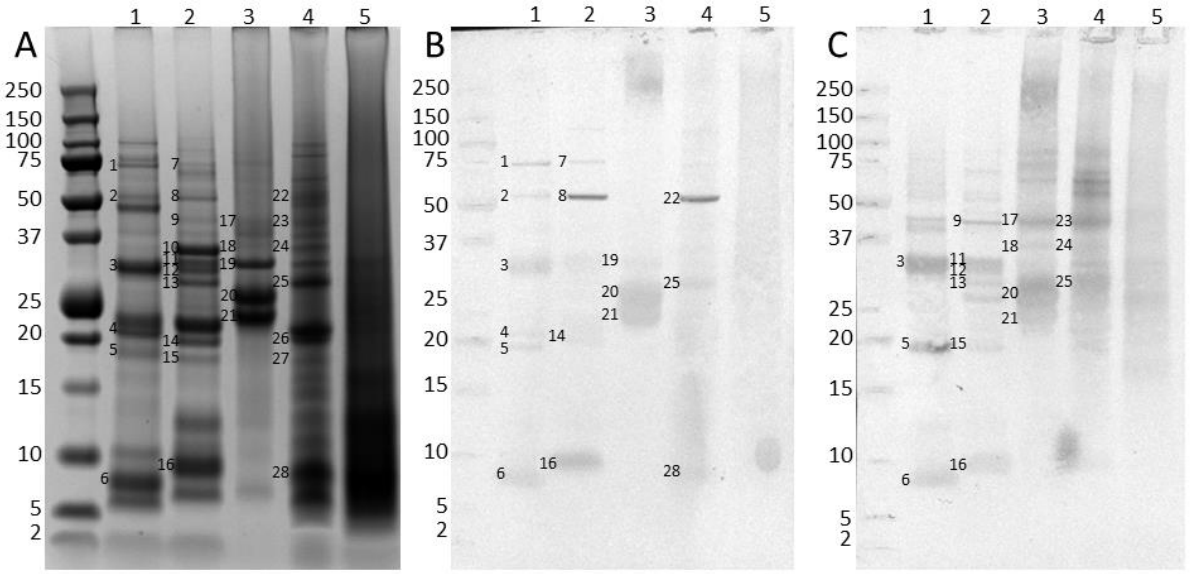


Figure 6.4 *SDS-PAGE (A) and western blot (B, C) of cashew (1), pistachio (2), mango (3), pink pepper (4), and sumac (5) proteins. For western blot B serum pool 1 was used, for western blot C serum pool 3 was used. The protein bands are indicated on the left in the five lanes of SDS-PAGE and western blot.*

The IgE of the two groups of cashew-sensitized patients bound the cashew nut, pistachio, mango, pink pepper and sumac proteins in a distinct pattern. IgE from serum pool 1 binds larger molecular weight proteins (>50kDa), while the IgE from serum pool 3 binds more to proteins in the 15-50kDa range. Both serum pools bind equally to the small molecular weight proteins of 8-10kDa (presumed 2S albumins; Ana o 3 and Pis v 1). The proteins from sumac did not separate well on SDS-PAGE. It is unsure if the lack of IgE binding to sumac (Figure 6.4, B and C, lane 5) on western blot is caused by the lack of sumac protein-specific IgE or by the lack of proper separation of the sumac proteins.

Based on Table 6.3, some of the protein bands, detected on SDS-PAGE and western blot, were tentatively named. Lane 1, cashew nut, shows IgE binding to Ana o 2 (protein band 2 the 53kDa monomeric protein, protein band 4-5 the basic and acidic subunits) and Ana o 3 (8kDa, band 6 the large subunit). Lane 2, pistachio, shows IgE binding to 7S globulin Pis v 3 (41kDa, band 9), 11S globulin Pis v 5 (34kDa, band 10), 11S globulin Pis v 2 (32kDa, band 11), Pis v 4 (26kDa, band 13), and 2S albumin Pis v 1 (9kDa, band 16). IgE binding to mango proteins, lane 3, occurred to Man i 1 (42kDa, band 17), and Man I 2 (33kDa, band 19). As no allergenic proteins are known from pink pepper, no bands could be identified from western blot, however, the molecular weight of band 22 and band 28 could possibly indicate these to be a 11S globulin (52kDa, band 22) and a 2S albumin (8kDa, band 28). The occurence of seed storage proteins such as 11S and 2S albumins in the pink pepper berry is possible as these seed storage proteins are recognised as allergens in various plants such as tomato, kiwi, soy,

lupine, rapeseed, and chickpea [48]. It is likely that the high levels of co-sensitization between cashew nut and pistachio, observed in the children participating in the IDEAL study, is due to the presence of 7S, 11S and 2S albumins in both nuts. Whether or not pink pepper should also be avoided by these children is not yet fully determined. First SPT and sIgE measurements, and possibly DBPCFC tests should be performed to confirm our *results in vivo*.

The presence of IgE against mango proteins in the serum of cashew nut allergic children is interesting, however the question does remain how clinically relevant these allergens are, as serum IgE levels and skin prick test results for mango were significantly lower than for cashew in both patient groups ($p < 0.05$, ANOVA, LSD, SPSS-22). Besides, in a sub-group of the IDEAL patient population (11 children), no children reacted positive during an open food challenge with mango, 17 other children indicated to be able to eat mango at home without any allergic reaction. Of these 28 children, 12 were sensitized to mango as determined by sIgE levels and SPT.

The observed difference in IgE-protein binding profile might indicate a possible difference in primary food allergen (e.g. cashew or pistachio), or a possible difference in primal allergen family (e.g. 2S albumin or 11S globulin). In our study no difference in severity of the allergic reaction towards cashew nut was noted between the two groups. These data should be extended with mass spectrometric protein identification methods to better understand the difference in IgE binding patters between patient groups, and the western blots should be repeated with either more patient serum pools or, preferably, unpooled sera.

Subdivision of patients into groups based on sensitization profile, might, in the future, lead to a better distinction between patients with different co-sensitization and co-allergenicity profiles. The need for more invasive and costly diagnostic methods such as the DBPCFC might decrease when co-allergenicity profiles can be predicted based on serum IgE profile (by western blot or sIgE measurement).

Future directions

In the IDEAL study both the clinical as well as the molecular characteristics of cashew nut allergy was studied. Based on the DBPCFC, SPT results, and the sIgE measurements (using our purified cashew nut allergens), a predictive model was established by van der Valk *et al.* [25]. This model provides an indication whether or not a person should be diagnosed by DBPCFC, or that sIgE and SPT measurements alone can be used for the diagnosis. Perhaps in a later stage other determinants, such as a mediator release assay, can be incorporated into this model. The addition of a mediator release assay might discriminate better between the levels of severity of the allergy. However, first experiments should be performed using various patient sera to see if the level of cell degranulation (using e.g. RBL or LAD2 cells)

correlates with the severity of the patients' allergy. It might also be useful to measure sIgE of the separate pistachio allergens of all cashew nut allergic patients. This might give more information regarding cross-allergenicity between the separate allergens of cashew nut and pistachio. As described in this chapter, different co-sensitization profiles towards cashew, pistachio, pink pepper and mango, were observed for different patient groups. Possibly these sensitization profiles can be linked to true co-allergenicity between cashew nut and other nuts. If such a diagnostic profile can be developed, better advice might be provided concerning the avoidance of certain foods in the diet. However the DBPCFC will still need to be used as the golden standard for allergy diagnostics.
The current study used cashew nut allergy as a showcase allergy. The results from this study might also be used as a basis for other food allergens. A model, similar to the model produced by van de Valk *et al.* was described in 2012 by Klemans *et al.*for peanut allergy [100]. It can be envisioned that such a model might also be prepared for other tree nuts such as pistachio, hazelnut, etc. For such a model again DBPCFC, SPT and sIgE measurements to specific allergens should be performed. sIgE measurements might be performed with native purified allergens as was done in this study, or with recombinantly produced allergens.

The patients in this study were children of 2 to 17 years old. Tree nut allergy is less prevalent among adults than among children [101]. It is unclear if the same sIgE levels to Ana o 1, 2 and 3 can be expected in adults as all studies have been performed in children [1, 37, 38] Furthermore it is unknown if the same co-sensitization and co-allergenicity can be expected in this different age group. Additionally it is unknown if the same patient characteristics such as sIgE levels are similar in children from a different geographic location.
Regarding the cashew nut allergens Ana o 1, 2 and 3 several opportunities for more research remain. First of all two isoforms of Ana o 1 and one isoform of both Ana o 2 and Ana o 3 have been identified, while for the same types of allergens in peanut multiple isoforms have been detected [48], therefore it can be expected that also multiple isoforms can still be detected in cashew nut, especially when keeping in mind the 2D electrophoresis results described in **Chapter 4**. Furthermore it highly likely that beside Ana o 1, 2 and 3, also other allergens (e.g. Ana o 4, 5, 6), such as profilins, pathogenesis related proteins, and oleosins are present in the cashew nut. Further identification of the allergens in cashew nuts would improve the possibilities of diagnostic measurements in cashew nut sensitized and allergic patients.

Several attempts have been made to decrease the allergenicity of cashew nuts [102, 103]. This is a difficult tasks for multiple reasons: first of all, as described in **Chapter 4**, no difference in soluble allergen content was detected between cashew nuts of 8 different

origins. Besides, since Ana o 1, 2 and 3 are present at high levels in the cashew nut, it would be difficult to remove them while maintaining viable and tasty cashew nuts.

The in-shell heat treatments described in **Chapter 4** indicated that no large effect on allergen solubility occurs in the pre-treatment of cashew nuts prior to the removal of the shell. Heat treatments performed after removal of the shell, showed that dark roasting decreased the solubility of Ana o 1 and 2 [28], and that autoclaving can decrease monoclonal antibody binding to linear epitopes of Ana o 1 and 2 [29]. Treatment of cashew nut protein extracts with sodium oleate and sodium sulphite could also reduce IgE binding to the cashew nut allergens [102, 103]. However it is questionable how tasty such autoclaved or dark roasted cashew nuts are, or what the effect is of sodium oleate or sodium sulphite on whole cashew nuts. It might, however, be relevant to study the combined effect of pre-shelling heat treatments with subsequent roasting steps under conditions generally applied to cashew nuts meant for consumption.

Conclusions

In this thesis we have characterised the cashew nut allergens Ana o 1, 2 and 3, both as native purified proteins (**Chapter 3**), but also as the *P. pastoris* recombinantly produced proteins (**Chapter 5**). Furthermore we have compared the cashew nut allergens from cashew nuts of different origins **(Chapter 4**), and subjected to different heat treatments in-shell (**Chapter 4**). Furthermore, we have provided a background on the cashew nut proteins (**Chapter** 1), cashew nut allergy (**Chapter 1**), and the pathway of (allergenic) protein transport in sensitized and non-sensitized persons (**Chapter 2**).

The results from this thesis have already been applied to expand the clinical patient characterisation with measurements of IgE levels to purified cashew nut allergens. These results have subsequently been used in a predictive model that can be used to evaluate the need for DBPCFC in children with possible cashew nut allergy. This model, when confirmed to work in other (geographic) patient populations, might be used to avoid unnecessary food provocations, and perhaps similar models should be composed for other food allergens or patient populations. In this thesis the current knowledge regarding the cashew nut allergens has been expanded with purification protocols and protocols for the recombinant purification of the allergens from *P. pastoris*, identification of N- and C-terminal micro-heterogeneity of native Ana o 3, 2D electrophoretic profiles for Ana o 1, 2 and 3, and information regarding the presence of the three allergens in multiple cashew nut origins.

References

[1] van der Valk, J.P., et al., Multicentre Double-Blind Placebo-Controlled Food Challenge Study in Children Sensitised to Cashew Nut. PLoS One, 2016. 11(3): p. e0151055.

[2] van der Valk, J.P., et al., sIgE Ana o 1, 2 and 3 accurately distinguish tolerant from allergic children sensitised to cashew nuts. Clin Exp Allergy, 2016.

[3] van der Valk, J.P., et al., Failure of introduction of cashew nut after a negative oral food challenge test in children. Pediatr Allergy Immunol, 2016. 27(6): p. 654-8.

[4] van der Valk, J.P., et al., No difference in health-related quality of life, after a food challenge with [cashew nut in children participating in a clinical trial. Pediatr Allergy Immunol, 2016.

[5] Blom, W.M., et al., Threshold dose distributions for 5 major allergenic foods in children. J Allergy Clin Immunol, 2013. 131(1): p. 172-9.

[6] Davoren, M. and J. Peake, Cashew nut allergy is associated with a high risk of anaphylaxis. Archives of Disease in Childhood, 2005. 90(10): p. 1084-1085.

[7] Hourihane, J.O., et al., An evaluation of the sensitivity of subjects with peanut allergy to very low doses of peanut protein: A randomized, double-blind, placebo-controlled food challenge study. Journal of Allergy and Clinical Immunology, 1997. 100(5): p. 596-600.

[8] Sampson, H.A., Food allergy. Part 1: immunopathogenesis and clinical disorders. J Allergy Clin Immunol, 1999. 103(5 Pt 1): p. 717-28.

[9] Agunod, M., et al., Correlative study of hydrochloric acid, pepsin, and intrinsic factor secretion in newborns and infants. Am J Dig Dis, 1969. 14(6): p. 400-14.

[10] Mason, S., Some Aspects of Gastrif Function in Newborn. Archives of Disease in Childhood, 1962. 37(194): p. 387-&.

[11] Xu, R.J., Development of the newborn GI tract and its relation to colostrum milk intake: A review. Reproduction Fertility and Development, 1996. 8(1): p. 35-48.

[12] Vance, G.H.S., et al., Exposure of the fetus and infant to hens' egg ovalbumin via the placenta and breast milk in relation to maternal intake of dietary egg. Clinical and Experimental Allergy, 2005. 35(10): p. 1318-1326.

[13] Vadas, P., et al., Detection of peanut allergens in breast milk of lactating women. JAMA, 2001. 285(13): p. 1746-8.

[14] Flokstra-de Blok, B.M. and A.E. Dubois, Quality of life measures for food allergy. Clin Exp Allergy, 2012. 42(7): p. 1014-20.

[15] Turnbull, J.L., H.N. Adams, and D.A. Gorard, Review article: the diagnosis and management of food allergy and food intolerances. Aliment Pharmacol Ther, 2015. 41(1): p. 3-25.

[16] Luengo, O. and V. Cardona, Component resolved diagnosis: when should it be used? Clin Transl Allergy, 2014. 4: p. 28.

[17] Borres, M.P., et al., Recent advances in component resolved diagnosis in food allergy. Allergol Int, 2016. 65(4): p. 378-387.

[18] Robotham, J.M., et al., Ana o 3, an important cashew nut (Anacardium occidentale L.) allergen of the 2S albumin family. Journal of Allergy and Clinical Immunology, 2005. 115(6): p. 1284-1290.

[19] Wang, F., et al., Ana o 2, a major cashew (Anacardium occidentale L.) nut allergen of the legumin family. Int Arch Allergy Immunol, 2003. 132(1): p. 27-39.

[20] Wang, F., et al., Ana o 1, a cashew (Anacardium occidental) allergen of the vicilin seed storage protein family. J Allergy Clin Immunol, 2002. 110(1): p. 160-6.

[21] Alcocer, M.J., et al., The disulphide mapping, folding and characterisation of recombinant Ber e 1, an allergenic protein, and SFA8, two sulphur-rich 2S plant albumins. J Mol Biol, 2002. 324(1): p. 165-75.

[22] Wallowitz, M., et al., Jug r 4, a legumin group food allergen from walnut (Juglans regia Cv. Chandler). J Agric Food Chem, 2006. 54(21): p. 8369-75.

[23] Kleber-Janke, T. and W.M. Becker, Use of modified BL21(DE3) Escherichia coli cells for high-level expression of recombinant peanut allergens affected by poor codon usage. Protein Expr Purif, 2000. 19(3): p. 419-24.

[24] Burks, A.W., et al., Recombinant peanut allergen Ara h I expression and IgE binding in patients with peanut hypersensitivity. J Clin Invest, 1995. 96(4): p. 1715-21.

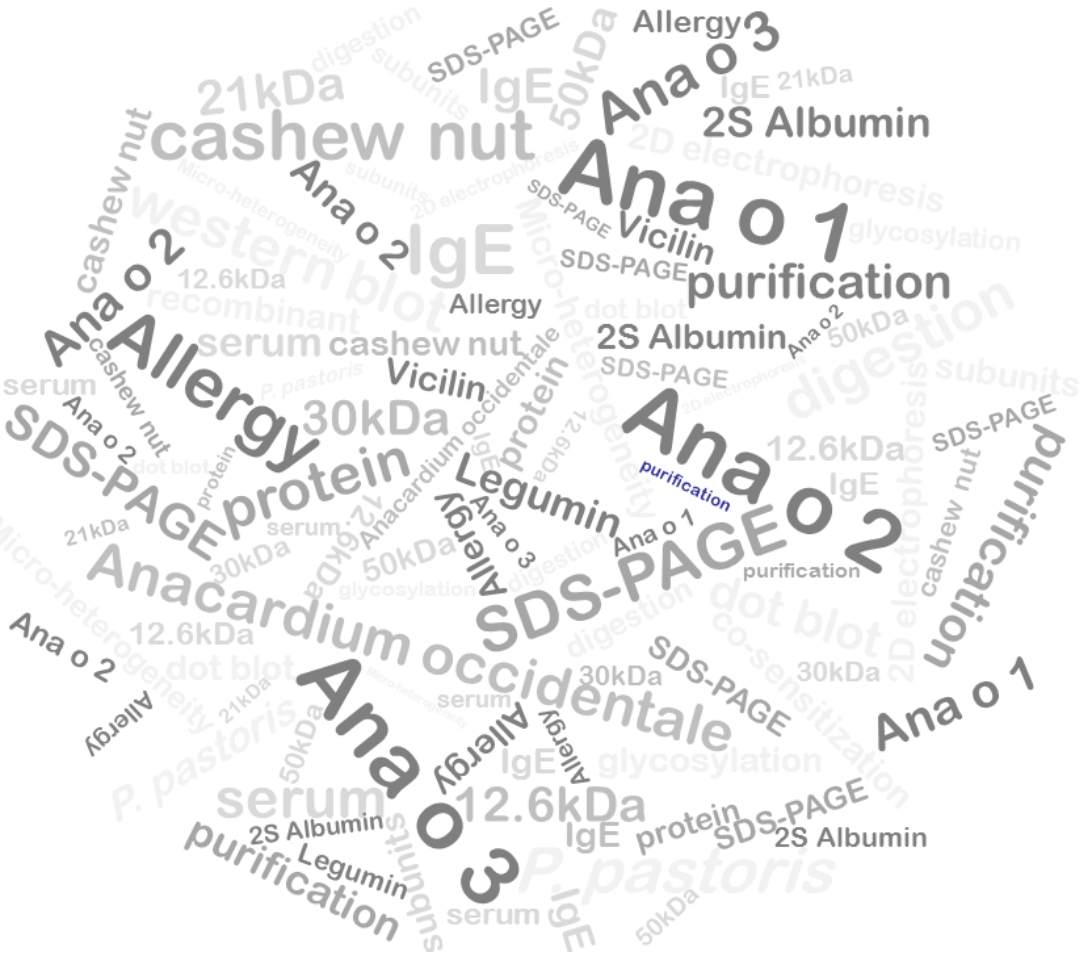
[25] van der Valk, J.P.G.v.W., R.; Vergouwe, Y.; Steyerberg, E. W.; Reitsma, M.; Wichers, H. J.; Savelkoul, H. F.; Vlieg-Boerstra, B.; de Groot, H.; Dubois, A. E.; de Jong, N. W., Prediction of cashew nut allergy in sensitised children. PhD thesis, J. P. van der Valk, Erasmus Universiteit Rotterdam, , 2016.

[26] Reitsma, M., et al., Purification and Characterization of Anacardium occidentale (Cashew) Allergens Ana o 1, Ana o 2, and Ana o 3. J Agric Food Chem, 2016. 64(5): p. 1191-201.
[27] M. Reitsma, S.B.-N., L. Sijbrandij, E. de Weert, S. Sforza, H. F. J. Savelkoul, N. W. de Jong, H. J. Wichers, Protein characterization of cashew nuts (Anacardium occidentale) from diverse origins and subjected to various processing methods. Unpublished.
[28] Mattison, Cp. P. M., B. Vant-Hullc, A. M. Vargasa, R. L. Wassermand, C. C. Grimm, Heat-induced alterations in cashew allergen solubility and IgE binding. Toxicology Reports, 2016(3): p. 244-251.
[29] Venkatachalam, M., et al., Effects of processing on immunoreactivity of cashew nut (Anacardium occidentale L.) seed flour proteins. J Agric Food Chem, 2008. 56(19): p. 8998-9005.
[30] Xia, L., et al., Mapping of a conformational epitope on the cashew allergen Ana o 2: a discontinuous large subunit epitope dependent upon homologous or heterologous small subunit association. Mol Immunol, 2010. 47(9): p. 1808-16.
[31] Robotham, J.M., et al., Characterization of a cashew, 11S globulin (Ana o 2), conformational epitope. Mol Immunol, 2010. 47(9): p. 1830-8.
[32] Reitsma, M., Bastiaan-Net, S., J. P. M. van der Valk, H. F. J. Savelkoul, N. W. de Jong, H. J. Wichers, Pichia pastoris recombinantly expressed Anacardium occidentale (cashew nut) allergens Ana o 1, Ana o 2 and Ana o 3. Unpublished.
[33] Mattison, C.P., C.C. Grimm, and R.L. Wasserman, In vitro digestion of soluble cashew proteins and characterization of surviving IgE-reactive peptides. Mol Nutr Food Res, 2014. 58(4): p. 884-93.
[34] Hummel, M., et al., Simple, Rapid, and Selective Isolation of 2S Albumins from Allergenic Seeds and Nuts. J Agric Food Chem, 2015.
[35] Mills, E.N.C., Shewry, P. R., Plant food allergens. 2004: Blackwell science Ltd. 219.
[36] Moreno, F.J. and A. Clemente, 2S Albumin Storage Proteins: What Makes them Food Allergens? Open Biochem J, 2008. 2: p. 16-28.
[37] Lange, L., et al., Ana o 3-specific IgE is a good predictor for clinically relevant cashew allergy in children. Allergy, 2016.
[38] Savvatianos, S., et al., Sensitization to cashew nut 2S albumin, Ana o 3, is highly predictive of cashew and pistachio allergy in Greek children. J Allergy Clin Immunol, 2015. 136(1): p. 192-4.
[39] Willison, L.N., et al., Pistachio vicilin, Pis v 3, is immunoglobulin E-reactive and cross-reacts with the homologous cashew allergen, Ana o 1. Clin Exp Allergy, 2008. 38(7): p. 1229-38.
[40] Ahn, K., et al., Identification of two pistachio allergens, Pis v 1 and Pis v 2, belonging to the 2S albumin and 11S globulin family. Clin Exp Allergy, 2009. 39(6): p. 926-34.
[41] Noorbakhsh, R., et al., Pistachio allergy-prevalence and in vitro cross-reactivity with other nuts. Allergol Int, 2011. 60(4): p. 425-32.
[42] Parra, F.M., et al., Pistachio nut hypersensitivity: identification of pistachio nut allergens. Clin Exp Allergy, 1993. 23(12): p. 996-1001.
[43] Goetz, D.W., B.A. Whisman, and A.D. Goetz, Cross-reactivity among edible nuts: double immunodiffusion, crossed immunoelectrophoresis, and human specific igE serologic surveys. Ann Allergy Asthma Immunol, 2005. 95(1): p. 45-52.
[44] Fernandez, C., et al., Allergy to pistachio: crossreactivity between pistachio nut and other Anacardiaceae. Clin Exp Allergy, 1995. 25(12): p. 1254-9.
[45] Funes, E., Milán, J. M., Pagán, J. A., López, J. D., Garcia, F. J., Negro, J. M., Hernández, J., Polo, F., Rico, P., Allergy to Anacardiaceae. Identification of allergens. Alergol. Immunol. Clin, 1999. 14(2): p. 82-89.
[46] Besler, M., Paschke, A., Rodriguez, J. Allergen data collection: Mango (Mangifera indica). in Internet Symposium on Food Allergens. 2001. www.food-allergens.de.
[47] de Leon, M.P., J.M. Rolland, and R.E. O'Hehir, The peanut allergy epidemic: allergen molecular characterisation and prospects for specific therapy. Expert Rev Mol Med, 2007. 9(1): p. 1-18.
[48] www.allergome.org.
[49] Martoglio, B., Intramembrane proteolysis and post-targeting functions of signal peptides. Biochem Soc Trans, 2003. 31(Pt 6): p. 1243-7.
[50] Mills, E.N.C., et al., Food allergens of plant origin - their molecular and evolutionary relationships. Trends in Food Science & Technology, 2003. 14(4): p. 145-156.
[51] Samartin, S., A. Marcos, and R.K. Chandra, Food hypersensitivity. Nutrition Research, 2001. 21(3): p. 473-497.
[52] Sicherer, S.H., Food allergy. Lancet, 2002. 360(9334): p. 701-10.
[53] Metcalfe, D.D., et al., Assessment of the allergenic potential of foods derived from genetically engineered crop plants. Crit Rev Food Sci Nutr, 1996. 36 Suppl: p. S165-86.

[54] Lehrer, S.B., R. Ayuso, and G. Reese, Current understanding of food allergens. Ann N Y Acad Sci, 2002. 964: p. 69-85.
[55] Schulten, V., et al., A food matrix reduces digestion and absorption of food allergens in vivo. Mol Nutr Food Res, 2011. 55(10): p. 1484-91.
[56] Wickham, M., R. Faulks, and C. Mills, In vitro digestion methods for assessing the effect of food structure on allergen breakdown. Mol Nutr Food Res, 2009. 53(8): p. 952-8.
[57] Fetuga, B., G. Babatunde, and V. Oyenuga, Composition and nutritive value of cashew nut to the rat. J Agric Food Chem, 1974. 22(4): p. 678-82.
[58] Toschi, T.G., et al., A Study on Cashew Nut Oil Composition. Journal of the American Oil Chemists Society, 1993. 70(10): p. 1017-1020.
[59] Nagaraja, K.V., Lipids of high-yielding varieties of cashew (Anacardium occidentale L.). Plant Foods Hum Nutr, 1987. 37(4): p. 307-11.
[60] Brufau, G., J. Boatella, and M. Rafecas, Nuts: source of energy and macronutrients. Br J Nutr, 2006. 96 Suppl 2: p. S24-8.
[61] Uribarri, J., et al., Advanced glycation end products in foods and a practical guide to their reduction in the diet. J Am Diet Assoc, 2010. 110(6): p. 911-16 e12.
[62] Sathe, S.K., et al., Solubilization and electrophoretic characterization of select edible nut seed proteins. J Agric Food Chem, 2009. 57(17): p. 7846-56.
[63] Huby, R.D.J., R.J. Dearman, and I. Kimber, Why are some proteins allergens? Toxicological Sciences, 2000. 55(2): p. 235-246.
[64] Blanc, F., et al., Boiling peanut Ara h 1 results in the formation of aggregates with reduced allergenicity. Mol Nutr Food Res, 2011. 55(12): p. 1887-94.
[65] Vissers, Y.M., et al., Effect of heating and glycation on the allergenicity of 2S albumins (Ara h 2/6) from peanut. PLoS One, 2011. 6(8): p. e23998.
[66] Gruber, P., W.M. Becker, and T. Hofmann, Influence of the maillard reaction on the allergenicity of rAra h 2, a recombinant major allergen from peanut (Arachis hypogaea), its major epitopes, and peanut agglutinin. J Agric Food Chem, 2005. 53(6): p. 2289-96.
[67] Chung, S.Y. and E.T. Champagne, Allergenicity of Maillard reaction products from peanut proteins. J Agric Food Chem, 1999. 47(12): p. 5227-31.
[68] Astwood, J.D., J.N. Leach, and R.L. Fuchs, Stability of food allergens to digestion in vitro. Nat Biotechnol, 1996. 14(10): p. 1269-73.
[69] Sampson, H.A., Update on food allergy. J Allergy Clin Immunol, 2004. 113(5): p. 805-19; quiz 820.
[70] Verhoeckx, K.C., et al., Food processing and allergenicity. Food Chem Toxicol, 2015. 80: p. 223-40.
[71] Fu, T.J., U.R. Abbott, and C. Hatzos, Digestibility of food allergens and nonallergenic proteins in simulated gastric fluid and simulated intestinal fluid-a comparative study. J Agric Food Chem, 2002. 50(24): p. 7154-60.
[72] Bogh, K.L. and C.B. Madsen, Food Allergens: Is There a Correlation between Stability to Digestion and Allergenicity? Crit Rev Food Sci Nutr, 2016. 56(9): p. 1545-67.
[73] Pumphrey, R.S., Lessons for management of anaphylaxis from a study of fatal reactions. Clin Exp Allergy, 2000. 30(8): p. 1144-50.
[74] Dirks, C.G., et al., Does absorption across the buccal mucosa explain early onset of food-induced allergic systemic reactions? Journal of Allergy and Clinical Immunology, 2005. 115(6): p. 1321-1323.
[75] Rahaman, T., T. Vasiljevic, and L. Ramchandran, Effect of processing on conformational changes of food proteins related to allergenicity. Trends in Food Science & Technology, 2016. 49: p. 24-34.
[76] Maleki, S.J., et al., The effects of roasting on the allergenic properties of peanut proteins. Journal of Allergy and Clinical Immunology, 2000. 106(4): p. 763-768.
[77] Mondoulet, L., et al., Influence of thermal processing on the allergenicity of peanut proteins. J Agric Food Chem, 2005. 53(11): p. 4547-53.
[78] Azam-Ali, S.H., Judge, E. C. Small-scale cashew nut processing. 2001 2001; Available from: http://www.fao.org/3/a-ac306e.pdf.
[79] Vereda, A., et al., Peanut allergy: Clinical and immunologic differences among patients from 3 different geographic regions. Journal of Allergy and Clinical Immunology, 2011. 127(3): p. 603-607.
[80] Ballmer-Weber, B.K., et al., IgE recognition patterns in peanut allergy are age dependent: perspectives of the EuroPrevall study. Allergy, 2015. 70(4): p. 391-407.
[81] Lapteva, Y.S., V.N. Uversky, and S.E. Permyakov, Sequence microheterogeneity of parvalbumin, the major fish allergen. Biochim Biophys Acta, 2013. 1834(8): p. 1607-14.

[82] Breiteneder, H., et al., Four recombinant isoforms of Cor a I, the major allergen of hazel pollen, show different IgE-binding properties. Eur J Biochem, 1993. 212(2): p. 355-62.
[83] Teuber, S.S., et al., Characterization of the soluble allergenic proteins of cashew nut (Anacardium occidentale L.). J Agric Food Chem, 2002. 50(22): p. 6543-9.
[84] Bublin, M., et al., IgE cross-reactivity between the major peanut allergen Ara h 2 and the nonhomologous allergens Ara h 1 and Ara h 3. J Allergy Clin Immunol, 2013. 132(1): p. 118-24.
[85] Saiz, J., et al., Peanut allergens: an overview. Crit Rev Food Sci Nutr, 2013. 53(7): p. 722-37.
[86] Willison, L.N., S.K. Sathe, and K.H. Roux, Production and analysis of recombinant tree nut allergens. Methods, 2014. 66(1): p. 34-43.
[87] Roux, K.H., S.S. Teuber, and S.K. Sathe, Tree nut allergens. Int Arch Allergy Immunol, 2003. 131(4): p. 234-44.
[88] Sicherer, S.H., A.W. Burks, and H.A. Sampson, Clinical features of acute allergic reactions to peanut and tree nuts in children. Pediatrics, 1998. 102(1): p. e6.
[89] Ewan, P.W., Clinical study of peanut and nut allergy in 62 consecutive patients: new features and associations. BMJ, 1996. 312(7038): p. 1074-8.
[90] Bock, S.A. and F.M. Atkins, The natural history of peanut allergy. J Allergy Clin Immunol, 1989. 83(5): p. 900-4.
[91] Hasegawa, M., et al., Clinical features of four cases with cashew nut allergy and cross-reactivity between cashew nut and pistachio. Allergol Int, 2009. 58(2): p. 209-15.
[92] van der Valk, J.P., et al., Systematic review on cashew nut allergy. Allergy, 2014. 69(6): p. 692-8.
[93] Renner, R., et al., Identification of a 27 kDa protein in patients with anaphylactic reactions to mango. J Investig Allergol Clin Immunol, 2008. 18(6): p. 476-81.
[94] Quercia, O., et al., Unexpected anaphylaxis to cashew nut. Allergy, 1999. 54(8): p. 895-7.
[95] Hegde, V.L. and Y.P. Venkatesh, Anaphylaxis following ingestion of mango fruit. J Investig Allergol Clin Immunol, 2007. 17(5): p. 341-4.
[96] Robotham, J.M., et al., Ana o 3, an important cashew nut (Anacardium occidentale L.) allergen of the 2S albumin family. J Allergy Clin Immunol, 2005. 115(6): p. 1284-90.
[97] Noorbakhsh, R., et al., Cloning, expression, characterization, and computational approach for cross-reactivity prediction of manganese superoxide dismutase allergen from pistachio nut. Allergol Int, 2010. 59(3): p. 295-304.
[98] Diaz-Perales, A., et al., Cross-reactions in the latex-fruit syndrome: A relevant role of chitinases but not of complex asparagine-linked glycans. J Allergy Clin Immunol, 1999. 104(3 Pt 1): p. 681-7.
[99] Paschke, A., et al., Characterization of allergens in mango fruit and ripening dependence of the allergenic potency. Food and Agricultural Immunology, 2001. 13(1): p. 51-61.
[100] Klemans, R.J., et al., The diagnostic value of specific IgE to Ara h 2 to predict peanut allergy in children is comparable to a validated and updated diagnostic prediction model. J Allergy Clin Immunol, 2013. 131(1): p. 157-63.
[101] Sicherer, S.H., et al., US prevalence of self-reported peanut, tree nut, and sesame allergy: 11-year follow-up. J Allergy Clin Immunol, 2010. 125(6): p. 1322-6.
[102] Mattison, C.P., et al., Decreased immunoglobulin E (IgE) binding to cashew allergens following sodium sulfite treatment and heating. J Agric Food Chem, 2014. 62(28): p. 6746-55.
[103] Chung, S.Y., et al., Treatment with oleic acid reduces IgE binding to peanut and cashew allergens. Food Chem, 2015. 180: p. 295-300.

Summary


Allergy
Ana o 3
2S Albumin
cashew nut
Ana o 1
Ana o 2
IgE
Vicilin
purification
Allergy
2S Albumin
serum
cashew nut
Vicilin
Ana o 2
Allergy
protein
30kDa
SDS-PAGE
Legumin
Ana o 1
purification
Anacardium occidentale
Ana o 2
Allergy
Ana o 3
12.6kDa
serum
IgE
protein
SDS-PAGE
2S Albumin
Legumin
purification

Cashew nut allergy can be a severe food allergy of which the prevalence appears to be increasing. The aim of this thesis was a comprehensive molecular and serological characterisation of the cashew nut allergens Ana o 1, 2 and 3 for improved diagnosis and characterisation of patient populations.

Chapter 1 in this thesis provides background information on cashew nuts, allergy, the allergens Ana o 1, 2 and 3, the effect of heat treatments on cashew nut proteins, the digestibility of cashew nut proteins, cross-reactivity between cashew nut proteins and other tree nuts, and the detection of cashew nut in food products. Subsequently, in **Chapter 2**, a review is presented on the topic of epithelial protein and allergen transport. This review describes multiple pathways of intestinal protein transport, sums up existing experimental data concerning protein and peptide transport, and presents different methods to study this. Interestingly, the pathway of (allergenic) protein transport can differ between sensitized and non-sensitized persons. In sensitized persons, protein transport occurs transcellularly via enterocytes, and paracellularly with the involvement of mast cells, while in non-sensitized persons microfold cells and enterocytes are considered most important.

In the next three chapters, cashew nut allergens were studied. Cashew nut allergy and cashew nut allergens were chosen because of a high number of undiagnosed cashew nut allergic children reported at the children's hospital "Kinderhaven", in Rotterdam, an outpatient clinic that is involved in this study. **Chapter 3** describes a protocol for the purification of Ana o 1, 2 and 3 from cashew nuts. Ana o 1 and 3 were purified by protein extraction, salt precipitation and filtering over a 30kDa molecular weight membrane. Ana o 2 was purified by protein extraction followed by gel filtration chromatography. These purified proteins were characterised by SDS-PAGE, western blot, glycoprotein stain, and protein identification. In this chapter also more in-depth analysis was performed on the N- and C-termini of the large and small subunits of Ana o 3. These N- and C-termini of Ana o 3, as well as the SDS-PAGE protein profiles were compared between cashew nuts of different origins in **Chapter 4**. In this chapter also the effects of different heat treatments on the electrophoretic behaviour of cashew nut allergens from various origins were studied, using both 1D and 2D electrophoresis. In these data no significant differences were detected between the electrophoresis patterns of Ana o 1, 2 or 3 in the various origins of cashew nuts. Some small but significant differences in Ana o 1, 2 and 3 content, however, were detected between the differently heated cashew nuts. No major differences in N- and C-terminal micro-heterogeneity were detected between cashew nuts of different origins.

Next, in **Chapter 5**, the cashew nut allergens Ana o 1, 2 and 3 were produced as recombinant proteins using a yeast (*P. pastoris*) production system. This procedure was used as recombinant allergens often produce higher yields of higher purity compared to native purified allergens. The recombinant proteins were compared to the native cashew nut

proteins for their glycosylation pattern, IgE binding capacity, and 2D electrophoresis profile. In **Chapter 6**, the major findings of this thesis are discussed. An overview of the protein characteristics (e.g. 1D and 2D electrophoresis profile, glycosylation, IgE binding, pepsin-digestibility) was provided, as well as a discussion on the clinical benefits that can be derived from the results obtained in this thesis. Also some additional results are presented, studying the serologic cross-reactivity between cashew nuts and other tree nuts and Anacardiaceae nuts and fruits.

This thesis provides an in-depth study regarding the protein characteristics of the cashew nut allergens Ana o 1, 2 and 3. Using the allergens that were purified in this thesis project, the serum IgE levels of Ana o 1, 2 and 3 could be measured in cashew nut-allergic children. The allergens were also recombinantly produced to obtain higher quantity of allergens for regular use in diagnostics of cashew nut allergy. The results from this thesis can potentially expand clinical patient characterisation with measurements of IgE levels to purified and recombinantly produced major cashew nut allergens. These results might have applications for other food allergens or patient populations.

Acknowledgements

Toen ik mijn PhD traject meer dan 4 jaar geleden begon, waren mijn twee hoofddoelen het ontwikkelen van mezelf van student naar werknemer en het aanleren van nieuwe technieken in het lab. Ondertussen heb ik niet alleen deze doelen volbracht, maar ook vele andere dingen geleerd: data interpretatie, artikelen schrijven, presenteren, plannen, studenten begeleiding, omgaan met stress en lastige resultaten, en uiteraard nog veel meer. Uiteraard was dit allemaal niet mogelijk zonder de steun van iedereen om mij heen. Daarom zou ik hier graag mijn dank willen uitspreken.

Allereerst natuurlijk een woord van dank richting mijn promotoren en begeleiders: *Harry, Nicolette, Huub en Shanna*. Dank voor de kans die jullie me gaven om aan dit project te beginnen. Nog veel meer zou ik jullie willen bedanken voor jullie hulp, tijd, begeleiding, inzichten, en denkwerk van de afgelopen 4 jaar. *Harry* bedankt dat je me gefocust hield op het duidelijk overbrengen van de boodschap binnen de artikelen. *Nicolette*, bedankt voor je enorme berg positieve energie, dit vond ik altijd erg motiverend. *Huub* bedankt voor je begeleiding tijdens het doorhakken van knopen richting het einde van mijn PhD traject. Als laatste zou ik *Shanna* willen bedanken voor alle tijd en energie die je, vaak naast je "gewone" werk, in mijn onderzoek hebt gestoken. Zonder jou zou ik nooit mijn weg hebben gevonden in het lab met alle klonerings experimenten. Daarnaast heb je veel oog voor detail, wat mij scherp hield tijdens het opschrijven van mijn data in de vorm van artikelen. Hierbij ook meteen een woord van dank aan mijn mede-promovendus *Hanna* voor de vele serum samples waar ik mee aan de slag kon. Bedankt ook aan de vele *bedrijven* die bij dit IDEAL project betrokken waren, bedankt voor jullie interesse en enthousiasme.

I've had a great time at FBR, with *all colleagues* and the many *(PhD) students* that came by over the years. The cheerful and positive atmosphere helped me through the less happy moments in my PhD. I would also like to thank *Book* for her positive energy and support during my master thesis, making me enjoy lab work while I was never a lab-person beforehand! *Lonneke*, jij verdient wat extra waardering. Ik vond het erg gezellig om jou als bureau-buurvrouw te hebben en ik ben blij dat we samen de leuke en minder leuke dingen van het leven/PhD traject konden delen. Bedankt dat je mijn paranimf wil zijn, en bedankt voor je vriendschap! Also a big thanks to *Food Chemistry*, I may not have been very involved in the group, but I really did enjoy the PhD trips, the occasional hallway talks, and the social part of supervising FRAAI.

During my PhD I've supervised a number of students, each of you taught me skills in supervision and guidance, and each of you made my efforts worthwhile with their hard work and good results: *Jack Yang, Jean-Francois Faivre, Yvonne Veltman, Eveline de Jong, Gracian Camps Ramon, Danny (Ying Deng), Mauro Wijnands, Lutske Sijbrandij,* and *Evelien de Weert*.

Bedankt ook aan alle meiden om me heen: vanuit de middelbare/basisschool *Verena* en *Elise*; de chickies *Lirije, Karlijn, Martine, Anneke* en *Marijke*; en mijn mede-voedingsmiepjes *Willeke* en *Annelies*. Daarnaast heb ik ook vele nieuwe vriendschappen gevonden de afgelopen jaren: *Edwin*, *Stefan* en *Wolf* met jullie mannen avonden en jullie prinsessenspel. *Elise*, bedankt dat je mijn paranimph wil zijn!

Wolf en *Marieke*, het was enorm gezellig met jullie als buren en als goede vrienden. Ondertussen zijn we dan geen buren meer, maar gelukkig zijn we, als rasechte stalkers, een stukje jullie richting op verhuisd. En tot slot natuurlijk *Ilse* en *Gerard*, ik vind het altijd erg gezellig met z'n allen, deze vriendschap gaan we na het afronden van alle PhD's zeker voortzetten! Dank ook voor jullie veelvuldige hulp bij het oppassen op Elsa.

Uiteraard ook dank aan mijn familie (*pap, mam, Arnout*) en schoonfamilie (*Jelle, Martje, Judy, Willem*) voor jullie interesse, steun en geduld tijdens deze 4 drukke jaren. Ik vond het altijd erg leuk om in " Jip en Janneke taal" uit te leggen hoe het nu zit met die cashew noten. *Pap, mam*, bedankt voor het goede voorbeeld wat jullie me altijd gegeven hebben.

En dan tot slot natuurlijk Jan. *Jan*, bedankt voor je steun en liefde. Ik denk dat terwijl ik mijn lab- en schrijf-skills ontwikkelde, jij je opvrolijk-skills enorm hebt ontwikkeld. Bedankt voor je luisterend oor, het gedachteloos bankhangen, en alle lol. Dankzij jouw invloed (een overload aan puppy foto's) hebben we eindelijk de stap durven om een hond in huis te nemen: *Elsa*. Wat een prachtig iets is het om samen een hond op te voeden, en wat doet ze het goed! Ons altijd-enthousiaste-tikkie-eigenzinnige-mormel houdt me altijd vrolijk en is een heerlijke afleiding na een lange dag werken.

About the author

Curiculum vitae Marit Reitsma

I was born in Capelle aan den IJssel on December 26 1988. After completing secondary school at the Emmauscollege in Rotterdam (VWO, profiel natuur gezondheid), I started my bachelor Nutrition and Health at the Wageningen university in 2007, which I finished in 2010. I continued my education with the master Nutrition and Health, specialisation molecular nutrition and toxicology at the Wageningen university. During my master I did a thesis on the subject of the immunomodulating effects of β-glucans on THP-1 macrophages at the Cell Biology and Immunology group of the Wageningen university. Next I studied the endocrine disrupting effects of thioxanthone photoinitiators at the research institute RIKILT in Wageningen. I finished my master education with an internship at knowledge institute TNO in Zeist on the subject of protein transport across the small intestine. I graduated my master cum laude in 2012.

Directly after graduating I started as a PhD candidate at the Wageningen university in a combined project of the Food Quality and Health Effects group of Wageningen University Food and Biobased Research, the chairgroup Food Chemistry of the Wageningen University, and the chairgroup Cell Biology and Immunology of the Wageningen University. My research focussed on cashew nut allergy; purifying the three main cashew nut allergens, producing these allergens in a recombinant yeast expression system, and characterising these proteins in a molecular and serological manner. I have also studied mediator release assays, and tested co-sensitization of cashew nut allergic patients to nuts and fruits related to the cashew nut. During my PhD I have joined the educational programme of the VLAG graduate school, supervised the theses of several bachelor and master students, supervised Wageningen University laboratory practicals, and attended multiple (inter)national conferences. Moreover, I obtained a grant to learn 2D electrophoresis at the Manchester University of Biotechnology, and published a technical note on the epitope mapping of Ana o 3 in collaboration with the company PEPperPRINT.
November 2016 I started working as a project assistant at the contract research organisation PPD in Bennekom.

Contact information:

Email: MaritReitsma1@gmail.com
Linkedin: https://www.linkedin.com/in/maritreitsma

List of publications

Peer reviewed publications

van der Valk JP, Gerth van Wijk R, Vergouwe Y, Steyerberg EW, **Reitsma M**, Wichers HJ, Savelkoul HF, Vlieg-Boerstra B, de Groot H, Dubois AE, de Jong NW. sIgE Ana o 1, 2 and 3 accurately distinguish tolerant from allergic children sensitized to cashew nuts. *Clin Exp Allergy.* 2017, 47(1), 113-120.

Reitsma M, Bastiaan-Net S, Sforza S, van der Valk JP, van Gerth van Wijk R, Savelkoul HF, de Jong NW, Wichers HJ. Purification and Characterization of Anacardium occidentale (Cashew) Allergens Ana o 1, Ana o 2, and Ana o 3. *JAFC*. 2016, 64, 1191-1201.

van der Valk JP, Gerth van Wijk R, Dubois AE, de Groot H, **Reitsma M**, Vlieg-Boerstra B, Savelkoul HF, Wichers HJ, de Jong NW. Multicentre Double-Blind Placebo-Controlled Food Challenge Study in Children Sensitised to Cashew Nut. *PLoS One*. 2016, 11 (3).

Reitsma M, Westerhout J, Wichers HJ, Wortelboer HM, Verhoeckx KC. Protein transport across the small intestine in food allergy. *Mol. Nutr. Food Res.* 2014, 58, 194–205.

Reitsma M, Bovee TF, Peijnenburg AA, Hendriksen PJ, Hoogenboom RL, Rijk JC. Endocrine-Disrupting Effects of Thioxanthone Photoinitiators. *J. Toxicol. Sci.* 2013, 132 (1), 64–74.

Chanput W, **Reitsma M**, Kleinjans L, Mes JJ, Savelkoul HF, Wichers HJ. β-Glucans are involved in immune-modulation of THP-1 macrophages. *Mol. Nutr. Food Res*. 2012, 56, 822–833.

Application note

Reitsma M, Bastiaan-Net S. IgE Epitope Mapping of the Cashew Nut Allergen Ana o 3. Application note in collaboration with Pepperprint Germany, 2016.

Overview of completed training activities

Discipline specific courses and conferences	year
Meeting, IDEAL project half-yearly meeting, The Netherlands [b]	2012-2016
Conference, INFOGEST, London, UK [b]	2012
Conference, EMBRN, Berlin, Germany	2012
Conference, EMBRN, Udine, Italy	2013
Conference, ISMA, Vienna, Austria	2013
Conference, INFOGEST [a, c]	2014
Conference, EMBRN, Marseille, France	2015
Symposium immunomodulating effects of food compounds, [a]	2012
Symposium mucosal factors regulating allergy (WIAS) [a]	2013
Course, bioinformatics (EPS) [a]	2013
Course, flow cytometry, BD Bioscience, Erembodegem, Belgium	2015
Course, BD Accuri flow cytometry, BD Bioscience, Leiden, The Netherlands	2015
Experimental work abroad: 2D electrophoresis at the Manchester University of Biotechnology, Manchester, UK [b]	2016

General courses and activities	year
VLAG PhD week, Baarlo, The Netherlands	2012
Course, Information Literacy for PhD Including Endnote (WGS) [a]	2013
Course, competence assessment (WGS) [a]	2013
Course, teaching and supervising thesis students (WGS) [a]	2013
Course, Techniques for writing and presenting a scientific paper (WGS) [a]	2014
Course, career assessment (WGS) [a]	2015
Course, reviewing a scientific paper (WGS) [a]	2015
Course, Adobe InDesign essential training (WGS) [a]	2015
Course, Career orientation (WGS) [a]	2015
Mini symposium, how to write a world-class paper [a]	2013
Mini symposium, mobilising your network (young AFSG) [a]	2014
Workshop, workshop carousel (WGS) [a]	2014

Optional courses and activities	year
Preparation of research proposal	2012
PhD excursion FCH, Singapore and Malaysia [b, c]	2012
PhD excursion FCH, Germany, Denmark, Sweden and Finland [b, c]	2014
Research group meetings, FQHE-FBR group meeting [a, b]	2012-2016
Research group meetings, CBI group meeting [a, b]	2012-2016

Teaching obligation, courses	**year**
Food Related allergies and Intolerances (FCH-21806) [a]	2012-2014
Immunomodulation by food and feed (CBI-50806) [a]	2012-2014

Teaching obligation, students	**year**
Bachelor students, 4 months each, 3 students	2012-2016
Master students, 6 months each, 3 students	2012-2016
Internship students, 3-6 months each, 3 students	2012-2016
High school research project (“profielwerkstuk”), 2 students	2014

VLAG: graduate school for food, nutrition, biobased, and biomolecular sciences
WGS: Wageningen graduate school
WIAS: graduate school, Wageningen institute of animal sciences
EPS: graduate school, experimental plant sciences
FCH: chairgroup food chemistry
CBI: chairgroup cell biology and immunology
FQHE: Food Quality Health Effects group
FBR: Food and Biobased Research

[a] Wageningen, The Netherlands
[b] oral presentation
[c] poster presentation

Colophon

The work in this thesis was performed at Wageningen UR FBR (FQHE group), the laboratory of Food Chemistry, and the laboratory of Cell Biology and Immunology. This research was financially supported by Technology Foundation STW, Stichting Voedselallergie, Siemens Healthcare Diagnostics, HAL Allergy, Intersnack Nederland B.V., ALK-Abello B.V., and the Nederlands Anafylaxis Netwerk. Financial support from Wageningen University and Research for printing this thesis is gratefully acknowledged.

The cashew allergens: a molecular and serological characterisation