Research Article

Narrow Leafed Lupin Beta-Conglutin Proteins Epitopes Identification and Molecular Features Analysis Involved in Cross-Allergenicity to Peanut and Other Legumes

Elena Lima-Cabello[1], Paula Robles-Bolivar[1], Juan D. Alché[1], Jose C. Jimenez-Lopez[1,2,*]

[1] Plant Reproductive Biology Laboratory; Department of Biochemistry, Cell and Molecular Biology of Plants; Estacion Experimental del Zaidin (EEZ); Spanish National Research Council (CSIC); 1 Profesor Albareda Street, Granada 18008 Spain
[2] The UWA Institute of Agriculture, The University of Western Australia, 35 Stirling Highway, Crawley, Perth WA 6009 Australia

[*] Correspondence: Phone: +34958181600, ext. 137; Fax: +34958129600; Email: josecarlos.jimenez@eez.csic.es

Received 2016-02-28; Accepted 2016-09-26

ABSTRACT

The use of narrow leafed lupin - NLL (*Lupinus angustifolius* L.) as a new food is resulting in an increasing number of allergic reactions cases, particularly in atopic patients with other pre-existing legume allergies. In the current study, we have performed an extensive *in silico* analysis of the NLL seed β-conglutin proteins, a new family of major allergen proteins identified in NLL, and a comparison to other relevant food allergens such as peanut Ara h 1. We analysed the variability of surface residues involved in conformational IgE-binding epitopes, lineal B- and T-cell epitopes, and changes in 2-D structural elements and 3D motives, with the aim to investigate cross-allergenicity among lupin, peanut, and other different legumes. Our results revealed that considerable structural differences exist, particularly affecting 2-D elements (loops and coils), and numerous micro-heterogeneities are present in fundamental residues directly involved in epitopes differential variability. Thus, variability of residues involved in IgE-binding epitopes might be a major contributor to the observed differences in cross-allergenicity among legumes.

KEYWORDS

B-/T-Cell Epitopes; Computational Biology; Food Allergen Proteins; Homology Modelling; IgE-binding Epitopes; Immunotherapy; *Lupinus albus*; *Lupinus angustifolius*; *Lupinus luteus*; Legume; PULSE; Vicilins

INTRODUCTION

Nowadays, considerable interest has been focused towards legume seed proteins. Among legume seeds, lupin, a legume crop that belongs to the Fabaceae family, and particularly the edible species called "sweet lupins" [*Lupinus angustifolius* L., or Narrow-leafed lupin (NLL) or blue lupin; *L. albus* or white lupin; and *L. luteus* or yellow lupin], have low content in alkaloids [1]. Narrow-leaf Lupin is currently under widespread cultivation in Europe [2], but most commonly found in Mediterranean countries. NLL is getting more and more recognition as a potential human healthy food with specific nutritional attributes [3] associated with high protein and dietary fibre content [4]. Additionally it has potential healthy pharmaceutical attributes since it plays a role similar to the hypoglycaemic drug metformin which is used for the treatment of type 2 diabetes (T2D), particularly in overweight and obese people [5].

Four main families of seed storage proteins (SSPs), collectively called conglutins, have been identified in NLL through the lupin genome sequencing project [6] the alpha (3 genes), the beta (7 genes), the gamma (2 genes) and the delta conglutin family (4 genes). NLL conglutin genes exhibit different relative transcript abundances and proteins translation in grains [7].

Sweet lupin flour is increasingly used as a novel food ingredient, which is derived from the endosperm tissue of the lupin seed. It contains 40–45% protein, 25–30% fibre and negligible sugar and starch. It is commonly used as an ingredient in baked foods, partially replacing wheat flour in foods such as gluten-free products, bread and pasta, which still remain palatable and are acceptable to consumers [8]. Although, the majority of edible legume grains and seed proteins from lupin species can cause allergy in a small percentage of the population, lupin products are increasingly included in human food, especially in Europe where lupin was approved as a food ingredient in 1997. 'Lupin allergy' occurs either separately or together with peanut allergy or allergy to other legumes [9]. Peanut-lupin cross allergy has been reported in which IgE antibodies recognize peanut allergens and also cross react with NLL conglutins [10]. It has been proposed that all lupin conglutin families are candidate allergens. Currently, the main protein identified and classified as allergen belongs to the vicilin-like proteins family (Cupin superfamily): beta conglutins (~65kDa) or Lup an 1 allergen. Vicilin proteins are a very important class of allergens reported from 43 sources (http://www.meduniwien.ac.at/allfam/). One out of six of these proteins come from two main superfamilies (Cupin and Prolamin) of seed storage proteins [11]. Recently, cross-reactivity has been reported between different vicilin proteins from legume and non-legume species [12], whilst patients who were allergic specifically to lupin and not peanut had serum IgE that bound beta conglutins [13]. However, alpha (~70kDa) and gamma (~47kDa; two subunits of 15 and 35kDa) proteins (the legumin and basic globulin families, respectively) from *Lupinus albus* have been

Publisher: Kernel Press. Copyright © (2016) the Author(s). This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

also considered responsible of allergy to some extent. In this regard, allergenic proteins of 71, 59 and 34 kDa [14] have been identified and isolated using sera from lupin allergen patients, with protein sizes matching alpha, beta and gamma families, respectively. Lupin allergy may arise either by primary sensitisation, or by clinical cross-reactivity in peanut-allergic persons [15, 16]. Various reports suggest the common clinical pattern of lupin allergy is the triggering of an allergic reaction via ingestion of lupin in peanut-allergic individuals. However lupin allergy triggered via ingestion, inhalation and occupational exposure in individuals without peanut allergy has also been reported. Patients allergic to lupin but not to peanut displayed IgE binding predominantly to proteins of approximately 66 kDa, and weak binding to 14 and 24 kDa proteins [9].

The prevalence of allergy to lupin is not clear, but it has been estimated that 30%-40% [17, 18] of peanut-allergic individuals react to lupin. The prevalence of allergy to lupin in the absence of allergy to peanut is currently not known [18]. Considering the increasing number of clinical cases of lupin allergy reported in the literature, lupin was added in 2008 to the list of foods that must be labelled in pre-packaged foods as advised by the European Food Safety Authority (EFSA) (http://www.efsa.europa.eu/). Given the increase in the number of cases of lupin allergy and the assumed frequency of cross-reactivity with other legumes, the involvement of individual major lupin and other protein allergens in these processes is of great interest, i.e. β-conglutins, and their counterparts from other legumes in cross-allergy is of major concern and of great interest to investigate.

Among the four families of NLL seed proteins, β-conglutin proteins are the most abundant protein in NLL seeds, and the most variable (polymorphic) family in terms of gene and protein sequence [6, 7], which may be reflected in differential functionalities and in allergenicity. In the present study, we performed an extensive analysis of the conformational IgE-binding, allergen structure modelling based epitopes (T- and B-cells) identification, providing a comprehensive understanding of the broad cross-reactivity and specific allergy reactions to lupin seed proteins. This is an important step to gain knowledge about the interacting surface of these seed proteins, and for a better understanding of immune responses, helping in the design and development of rational and effective immunotherapy strategies for the treatment of lupin-related food allergy.

METHODS

Allergen sequences

We retrieved allergen sequences necessary for the present study from Gen-Bank/EMBL or Uniprot Database: β1-conglutin or Lup an 1 (F5B8V9), β2- to β7-conglutins (F5B8W0, F5B8W1, F5B8W2, F5B8W3, F5B8W4, and F5BW5), Ara h 1 (P43237, P43238), Len c 1 (Q84UI0, Q84UI1), Gly m 5 or Gly m 5.0101 (O22120), Gly m β-conglycinin or Gly m 5.0302 (P25974), Vig r 2 (Q198W3, B1NPN8).

Phylogenetic analysis of food allergen sequences

Allergen protein sequences from legumes (lupin, peanut, soybean, Mung bean, lentil, chickpea, pea, carob tree) were retrieved and used to perform a phylogenetic analysis. Sequences alignments were performed by using ClustalW multiple sequence alignment tools according to Jimenez-Lopez et al. [19]. Trees were visualized using Treedyn [20].

Template assessment

All allergen sequences were searched for homology in the Protein Data Bank (PDB). Suitable homologous templates were selected by using Swiss-Prot database and BLAST server employing fold recognition [21].

Proteins homology modelling

Sequences were modelled through SWISS-MODEL via the ExPASy web server [22] by using the top PDB closest template structures previously assessed. Models refinement of 3D structural errors, and structural assessment were performed using stereo-chemical and energy minimization parameters [23].

Structural comparison and evolutionary conservational analysis

Allergen proteins structure comparison was performed by superimposition to calculate average distance between their Cα backbones. Protein models were submitted to ConSurf server [24] to generate evolutionary related conservation scores, in order to identify functional regions in the proteins. Functional and structural key residues were con-firmed by ConSeq server [25]. 2-D and 3D were visualized and analyses using PyMol software [26].

Solvent accessible surface area and Poisson–Boltzmann electrostatic potential

Solvent accessible surface area (SASA), defined as the percentage of surface area of a biomolecule that is accessible to a solvent for each residue was calculated by using the GETAREA v1.1. software [27]. The electrostatic Poisson-Boltzmann (PB) potentials for the built structures were obtained [28].

Allergenicity profile assessment

Allergenicity of lupin and other legume allergen sequences was checked by a full FASTA alignment in the Structural Database of Allergenic Proteins (SDAP) [29]. Allergenicity profile was assessed by combination of different parameters: hydrophobicity, antigenicity and SASA [21]. Values of absolute surface area (ASA) of each residue were also calculated by DSSP program [30] and transformed to relative values of ASA and visualized by ASAView [31].

IgE-binding epitopes identification

AlgPred server [32] was used to identify IgE-binding epitopes for lupin and the other legume species 7S proteins. This is a similarity-based approach, whereby a protein is predicted to be an allergen if it has a region/peptide identical to a known IgE epitope.

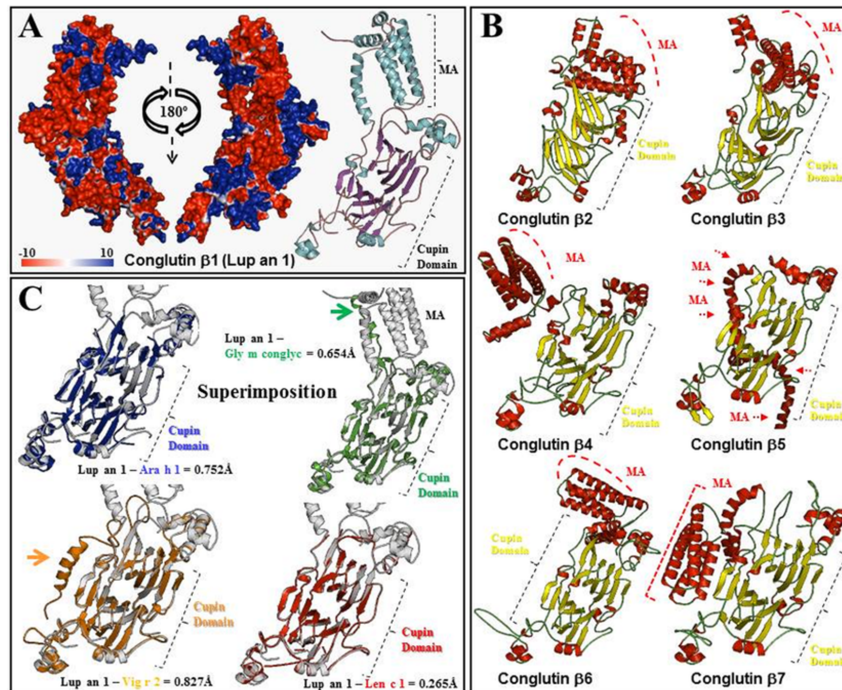

Figure 1: Lupin β-conglutins structural analysis. A) Cartoon and surface representation views of β1-conglutin rotated 180°, showing the surface electrostatic potential clamped at red (-10) or blue (+10). 2-D elements were depicted in cartoon model. B) β2- to β7-conglutins 3D structures depicted as a cartoon diagram. α-helices, β-sheets and coils are depicted in red, yellow and green respectively. C) Super-impositions showed the close structural relationship with allergens from other legumes (Ara h 1, β-conglycinin, Vig r 2, Len c 1). Å= Armstrong; MA = mobile arm.

Sensibility cut-off was established as 80, 60 and 50 for epitopes having <10, between 10 and 15, and >15 residues, respectively.

Linear and conformational B-cell epitopes analysis

For determination of linear (continuous) epitopes, the allergen protein sequences were submitted to ABCpred (uses artificial neural networks) BepiPred 1.0b (based on hydrophobicity scale with a Hidden Markov Model), BCPREDS (uses support vector machine), Bcepred (based on a combination of physico-chemical properties), and COBEpro (uses support vector machine) [33]. Linear and discontinuous antibody epitopes based on a protein antigen's 3D structure were predicted using Ellipro [34] discontinuous epitopes were defined based on PI values and were clustered based on the distance R (Å) between the center of mass of the residues (Immune-epitope and Discotope web-servers).

The epitopes identified frequently by most of the tools were selected [35, 36].

T-cell epitopes identification and analysis

The identification of MHC Class-II (HLA class II) binding regions for all the allergen sequences were performed by using neuronal networks and quantitative matrices derived from published literature. Promiscuous peptides binding to multiple HLA class II molecules were selected. The analysis was made by using the TEPITOPE [37] software, with a threshold of 5% for the most common human HLA-DR alleles [DR1, DR3, DR4, DR7, DR8, DR5 and DR2] among Caucasian population [21], and covering a large proportion of the peptides that bind human HLA.

RESULTS

Searching for allergen proteins templates

We used the Swiss-model server to identify the best possible templates to build allergen structures, finding high scores and very low E-values (ranging 12E^{-34} to 7E^{-42}) for the tem-plates retrieved from Protein Data Bank (PDB) database and used for homology modelling: lupin β-conglutins (1uijA, 2eaaB), Ara h 1 (3s7i, 3s7e), Gly m 5 (1uijA), Len c 1 (1uijA), Gly m β-conglycinin (1uijA), Vig r 2 (2eaaB).

Figure 1 shows that lupin β-conglutins are characterized by a negatively charged surface, a domain from the Cupin superfamily constituted by 2 barrels of 8-10 α-helices each, and a mobile arm, which position may be different depending of the β-conglutin form. One of these barrels is followed the Rossmann fold structure, typically found in oxidoreductase enzymes.

2-D elements comparison to Cupin superfamily domain by superimposition among allergens showed a low value (<1Å) of structural differences, since the mobile arm is absent in these allergens. Overall, β-conglutins were found structurally close to Len c 1 and most distantly related to the allergen Gly m 5.

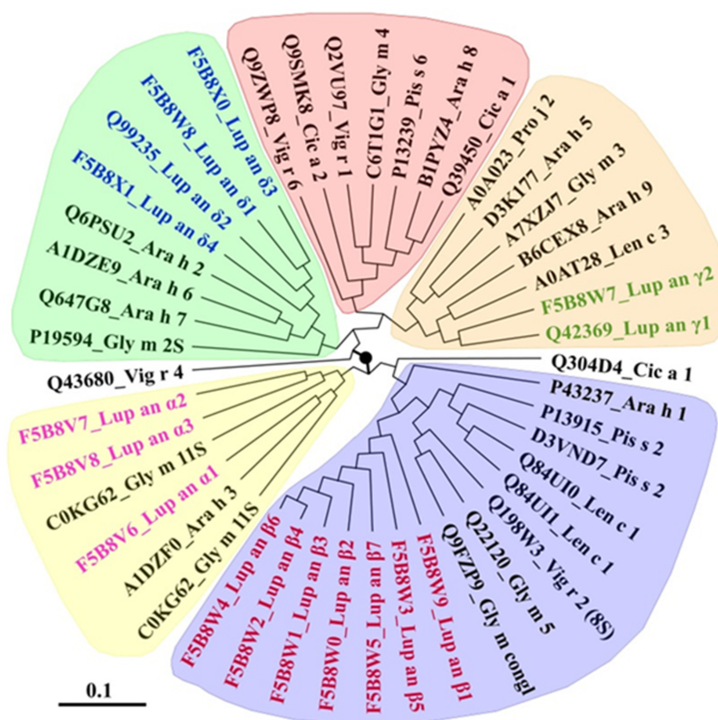

Figure 2: Phylogenetic analysis of food allergens. Neighbour-joining (NJ) method was used to perform a phylogenetic analysis of 45 legume allergens from lupin (conglutins α1-3, β1-7, γ1-2, and δ1-4), peanut (Ara h 1, 2, 3, 5, 6, 7, 8, and 9), soybean (Gly m 5, and Gly m conglycinin), lentil (Len c 1, and 3), pea (Pis s 2, and 6), Mung bean (Vig r 1, 2, 4, and 6), and chickpea (Cic a 1, and 2).

Structural assessment of the β1- to β7-conglutins, Gly m 5, Len c 1, Gly m conglycinin, and Vig r 2 structural models.

Different molecular tools (e.g. stereochemistry and energy minimization) were used to assess the quality of the models built for this study. A general quality assessment parameter as QMEAN displayed adequate values for all models in the comparable range of the templates that were used to build these structures. Most of the residues of the main chain of built models were located in the acceptable regions of the Ramachandran plot shown by Procheck analysis. In addition, Z-scores returned from ProSa indicated that structures showed negative interaction energy and were within the lowest energy range. In addition, the Z-scores were within the range usually found for templates used for allergen structure modelling.

Phylogenetic analysis

We analysed the relationships between lupin β-conglutin proteins and allergens from other species. The data clearly revealed five established groups/clusters. We identified 5 main groups, where β-conglutins were grouped with allergens of 7S-globulin nature (Figure 2).

Within this group, we also found the main allergen of peanut Ara h 1, *Glycine max*. Gly m 5 and Gly m conglycinin. Interestingly, other group contained lupin β-conglutin and peanut Ara h 3; gamma conglutin and Ara h 5 and Ara h 9; and delta conglutin together to Ara h 2, 6 and 7 allergenic proteins from peanut.

Identification of highly antigenic regions in 7S seed proteins

Physicochemical parameters such as hydrophobicity, accessibility, exposed surface, and antigenic propensity of polypeptide chains have been used to identify continuous epitopes (see methods section). In our analysis, antigenicity determinants were assigned by locating the positive peaks in hydrophilicity plots, and identifying the regions of maximum potential of antigenicity.

We identified up to 8 regions in lupin β-conglutins, with high potential of antigenicity, 7 regions in Ara h 1, 7 regions in Gly m 5, 8 regions in β-conglycinin, 7 regions in Vig r 2, 4 in Len c 1, and 5 in Pis s 2. These regions with high antigenicity correlated well with the lineal T- and B-cell and conformational epitopes identified and analysed in the present study. The highest differences in terms of antigenicity region polymorphism correspond to lupin β-conglutins, while the lowest variable allergen was Len c 1.

Analysis of IgE-binding epitopes

IgE-binding epitopes were predicted for Lup an 1 and each one of the legume 7S seed sequences analysed in the current study (Table 1). This analysis using a wide representation of 7S proteins allowed reaching numerous conclusions: i) the number of predicted IgE epitopes ranged from 1 to 8 concerning the species studied, ii) IgE-binding epitopes 1 (12 out of 15 sequences = 80%) and 2 (9 out of 15 sequences = 60%)

Species	Allergen name	Uniprot ID	IgE1	IgE2	IgE3	IgE4	IgE5	IgE6	IgE7	IgE8
Lupinus angustifolius	Conglutin b6	F5B8W4	NFRLLGFGIN	-	RRYSARLSEG	SYFSGFSRNT	-	-	-	-
Lupinus angustifolius	Conglutin b4	F5B8W2	NLRLLGFGIN	KGLTFPGSTE	RRYSARLSEG	SYFSGFSRNT	-	-	-	-
Lupinus angustifolius	Conglutin b3	F5B8W1	NLRLLGFGIN	KGLIFPGSAE	RRYSARLSEG	-	-	-	-	-
Lupinus angustifolius	Conglutin b2	F5B8W0	NLRLLGFGIN	KQLTFPGSVE	RRYNAKLSEG	SYFNGFSRNT	-	-	-	-
Lupinus angustifolius	Conglutin b7	F5B8W5	NLRLLGFGIN	-	RRYSARLSEG	-	-	-	-	-
Lupinus angustifolius	Conglutin b5	F5B8W3	NLRLLGFGIN	KELIFPGSAE	RSYNARLSEG	-	-	-	-	-
Lupinus angustifolius	Conglutin b1 or Lup an 1	F5B8V9	NLRLLGFGIN	KELTFPGSIE	-	-	-	-	-	-
Glycine max	Gly m conglycinin	Q9FZP9	-	-	-	SYLQGFSKNI	QRNFLAGSKD	-	-	-
Glycine max	Gly m 5	O22120	-	-	-	SYLQGFSRNI	-	-	-	-
Lens culinaris	Len c 1	Q84UI0	DLNLIGFGIN	KELAFPGSSR	-	-	-	-	-	-
Lens culinaris	Len c 1	Q84UI1	DLNLIGFGIN	KELAFPGSSR	-	-	-	-	-	-
Pisum sativum	Pis s 2	D3VND7	DLNLIGFGIN	KELAFPGSSH	-	-	-	-	-	-
Pisum sativum	Pis s 2	P13915	NLNLLGFGIN	-	-	-	-	KQSHFASAEP	NKFGKLFEIT	GQRERGRQEG
Arachis hypogaea	Ara h 1	P43237	ELHLLGFGIN	KDLAFPGSGE	RRYTARLKEG	SYLQGFSRNT	HRIFLAGDKD	RESHFVSARP	NNFGRLFEVK	GERTRGRQPG
Vigna radiata	Vig r 2 (8S)	Q198W3	-	-	-	-	QRNFLAGEKD	SESHFVDAQP		

Table 1: IgE epitopes identification and analysis

are present in most of the sequences, iii) only peanut allergen protein Ara h 1 displayed the eight IgE epitopes; iv) only lupin and peanut sequences displayed the first four IgE-binding epitopes, and one sequence (*Glycine max* – Gly m 5) displayed only one IgE-binding epitope.

Four main IgE binding epitopes were predicted for lupin β-conglutins, corresponding to the epitope 1 to 4 (Table 1). These epitopes are 10 amino acids long, sharing higher similarity among β-conglutin proteins (9-10 out of 10 residues = 90-100% for epitope 1, 6-10 out of 10 residues= 60-100% for epitope 2; 7-10 out of 10 residues = 70-100% for epitope 3; 9-10 out of 10 = 90 -100% for epitope 4).

Epitopes 3 and 4 of lupin β-conglutin proteins exhibited the highest similarity to Ara h 1 from peanut, while epitopes 1 and 2 are the most polymorphic compared to peanut allergen proteins.

Analysis of B-cell epitopes

12 antigenic lineal regions prone to B-cell binding were identified in β1-conglutin, 7 for β2 and β7, 5 for β3, and 4 for β4, β5 and β6. In addition, we identified 6 antigenic regions in Ara h 1, 7 in Gly m 5, 11 in β -conglycinin, 4 in Len c 1, 10 in Pis s 2, and 8 in Vig r 2 (Figure 3). Comparative analysis of these regions showed that 5 lineal epitopes in conglutin β1 are located in the mobile arm, 3 of them overlapping with a big conformational epitopic area (black colour, Figure 3) and 2 lineal independent epitopes. Furthermore, β2 and β5 present 3 conformational epitope areas, 1 in β3, β6 and β7, 2 in β4, related to the differential mobile arm structure.

The biggest structural difference between the β-conglutins and the other legume allergens is the presence of the mobile arm in N-terminal region of the lupin β-conglutins and the epitopes which integrate. The number of epitopes and polymorphism analysis of lineal and conformational B-cell epitopes in other legume allergens showed a wide range of variability in both the number and the sequence identity of these epitopes.

Identification of T-cell epitopes

We identified a variable number of anchor motifs to HLA-DR in the sequences of lupin β-conglutins (8 main T-cell epitopes), and their counterparts in five species of legumes.

T1 was the "solo" epitope in the mobile arm of β-conglutins (Figure 4), exhibiting a large surface orientation. This epitope is common for other legume allergens such as peanut (Ara h 1), soybean (β-conglycinin), Mung bean (Vig r 2), and pea (Pis s 2). The rest of epitopes identified in β-conglutins were located in the globular (Cupin-like) domain of these proteins.

Some of these epitopes were differentially shared with other legume allergens, i.e. T2 is the most commonly shared epitope, and T8 only commonly located in allergens of *Glycine max*. In addition, each of the allergen analysed has specific epitopes non-found in other species (Figure 4). Most of these lineal epitopes displayed 50% or more of their residues non-exposed to the surface (T2 to T8).

DISCUSSION

Knowledge about cross-reactivity between lupin seed proteins and other plant food sources at molecular level is scarce. Hypersensitivity to food encompasses IgE-mediated and non-IgE-mediated reactions. However, the first one account for the majority of food-induced immune reactions. Generally, storage proteins contained in legume seeds, grains and nuts are the causative of allergy reactions upon ingestion, mainly due to the high stability under extremes of pH and temperature, and variable similarity in their primary sequence among these allergens [38, 39].

In *Lupinus angustifolius* L., Lup an 1 (β1-conglutin), a vicilin-like (7S-type) protein has been recently identified as a major allergen using proteomic analysis and was recognized by serum IgE from most of 12 lupin-allergic patients' sera [13]. This knowledge provides opportunities to further characterize the linear and conformational epitopes of lupin major allergens. In addition, it will help identifying the triggering features at molecular level for clinical diagnosis (trials) of lupin allergy, as well as for the development and implementation of molecular tools able to identify the presence of lupin allergens as ingredients in plant-derived food. In this regard, rapid targeted proteomic approach based on liquid chromatography-electrospray ionization-tandem mass spectrometry (LC-ESI-MS/MS) analysis could be a reliable approach for determination of allergens, i.e. the major lupin allergens, conglutins, in pasta and biscuits [40]. This identification may prevent lupin allergy reactions and cross-allergenic reactions in sensitized patients. Thus, sequence homology between lupin major allergens and other legume allergens support cross-reactivity among legume seed proteins.

The aim of the present study was to identify and analyse the conformational IgE-binding, and lineal

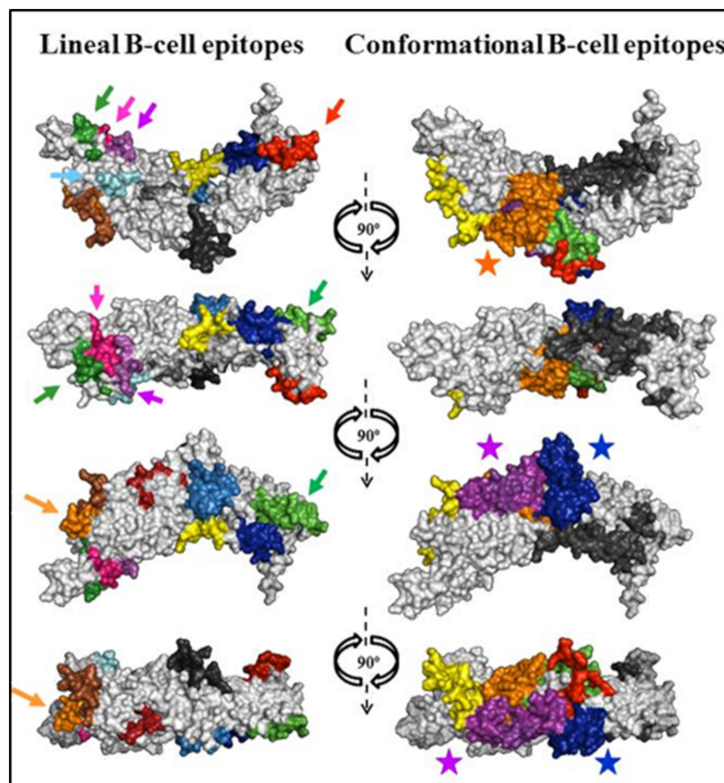

Figure 3: B-cell epitopes analysis in lupin β-conglutins and their legume proteins counterparts. Cartoon representation of Lup an 1 allergen showing in various colours lineal and conformational B-cell epitopes in its surface. Different colours represent the localization in the surface of the Lup an 1 protein for each specific epitope type, linear and conformational. Arrows and stars represent specific lineal and conformational epitopes, respectively, which do not overlap with each other.

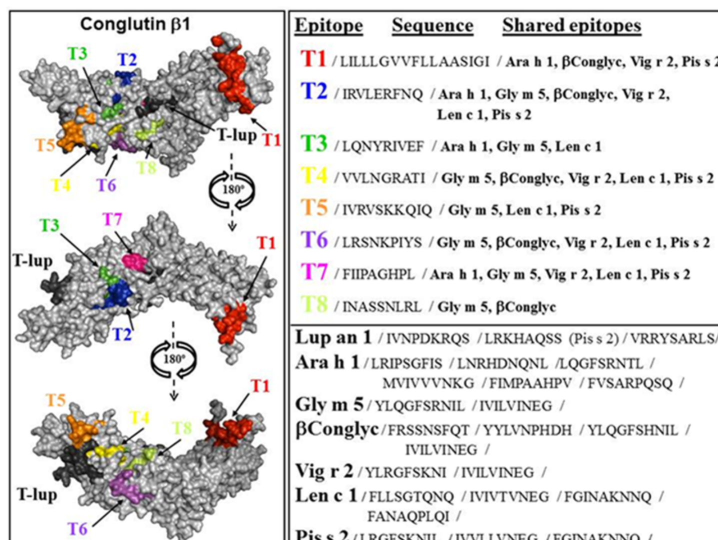

Figure 4: T-cell epitopes comparison between lupin β-conglutins and their legume proteins counterparts. T-cell epitopes depicted on the three views rotated 180° β1-conglutin protein surface, respectively, following a colour code for epitopes identification (T1 to T8, upper right square). Epitopes identified belonging to exclusive legume species have been listed in the figure (bottom Right Square).

B- and T-cell epitopes, providing a comprehensive understanding of the main molecular features responsible of the cross-reactivity between lupin and allergen proteins from other legumes. Many of these features are located in the globular domain characteristic of the 7S and 11S seed storage proteins belonging to the Cupin superfamily [41].

Molecular modelling and sequence polymorphisms characterization helped identifying specific regions, potentially candidates for the development of immunotherapeutic agents for lupin allergy, while conserved regions could be responsible of the cross-reaction between lupin and other legume allergen proteins. Epitope prediction based on knowledge from structural derived surface features such as increased solvent accessibility, backbone flexibility, and hydrophilicity were found to correlate well in the present study [19, 35, 36]. Indeed, contrasting theoretical information about epitopes prediction and available experimental results (IEDB http://www.iedb.org), for allergens analysed in the current study showed that e.g. all predicted specific epitopes for Ara h 1 (Figure 4) have been confirmed experimentally and identified as allergens ID 99142, 98985, 98887, 99364 [42–44], and 148649, 148770 [43, 44]. Searching Lup an 1 specific epitopes (Table 1) in the Ara h 1 epitopes database IEDB allowed us to identify one experimentally confirmed epitope from Ara h 1 (ID 148579) [45, 46] that matched with the Lup an 1 epitope VRRYSARLS in 7 out of 9 positions ($\sim$78%), which may be enough to induce cross-reactivity between both proteins. This finding highlights the relevance that specific epitopes may also have cross-reactivity when epitopes from both allergen proteins share high percentage of identity.

On the other hand, the large number of epitopes that have been identified in β1-conglutin may be one of the reasons why the *L. angustifolius* allergen Lup an 1 is currently one of the main allergen in NLL, particularly the epitopes located in the N-terminal region of protein integrated only by α-helices. We show that this region constitutes a mobile arm in NLL of β-conglutins, which is a unique feature of these proteins in comparison to the vicilin proteins from other legume species [47]. Epitopes located in the globular region of Lup an 1 are the ones commonly shared to other legume proteins. In this regard, a large number of epitopes (thirty-six peptides) have been also identified for the peanut allergen Ara h 1 as T-cell epitopes vaccines candidates by using *in silico* predictions and MHC class II binding assays [44]. Several of these epitopes may be involved in cross-reactivity between Ara h 1 and Lup an 1 [44].

Therefore, a variable degree of polymorphism among allergens has been considered as a major contributor for the considerable differences in allergenicity responses mediated by B- and T-cell, since this polymorphism affect the differential recognition of antigenic epitopes [48]. We have identified surface patterns (conformational epitopes), as well as multiple regions (B- and T-cell epitopes) in legume allergens, including lupin, exhibiting differences in length and variability. Furthermore, we have found shared common B- and T-cell epitopes among these legume allergens, as well as epitopes differentially distributed in specific allergens. Linear B-and/or T-cell epitopes may play the most important roles in cross-reactivity between food allergens, and between *lupinus* sp. and other legume allergens, since food processing or digestion may increase the number or the accessibility of IgE binding epitopes.

Depending on the location of these polymorphic residues, recognition by IgE/IgG may be also affected [49]. Thus, the variability in their surface residues might contribute to generate areas of the protein able of being differentially recognized as Th2- inducing antigens.

We propose that the presence of several of these epitopes (T- and B-cell) is the main reason for cross-reactivity among legume proteins, which however react differentially with lupin β-conglutins forms and between them. The extension of the reactions may be directly linked to the residue variability of these epitopes. It has been reported serological cross-reactivity among legume allergens [50, 51], and Lup an 1 (Ara h1, Len c 1 and Pis s1) [51, 52].

Therefore, detection of allergen-specific IgE antibodies identifies sensitization to a particular allergen and does not allow a decisive differentiation between clinically relevant IgE reactivity (i.e., IgE antibody–allergen complex capable to cross-link FceRI receptors and subsequent effector cell response leading to allergic symptoms) and IgE reactivity not ac-companied by clinical symptoms (i.e. reactivity without an effector cell response). IgE-positivity to several legumes basically indicates immunological cross-reactivity via linear epitopes (sequence homology) or similar conformational epitopes without allowing estimation or prognosis of the clinical significance. This has been observed in lupin, peanut and pea [53, 54]. In this regard, we have found that six T-cell epitopes are shared between Lup an 1 and Len c 1. From these, four epitopes are commonly found in Ara h 1 and Pis s 1 as well. Furthermore, one of these four epitopes is the "T- solo" or T1 located in the mobile arm of β-conglutins. This epitope may play a key role in specific cross-reactivity between legume seeds proteins and lupin β-conglutins as one of the four main families (α, β, γ, δ) of seed storage proteins in lupin [55].

At structural level, antigenic determinants may be integrated by 2-D structure elements, which protrude from the surface of the protein, such as coils and loops [21]. Our results have shown that conformational epitopes are partially or totally integrated by 2-D structure elements, mostly by short α-helices and coils (Figures 1 and 3). Variability in sequence and length of these 2-D elements might additionally justify the differences and the extension of the cross-reactivity between legume allergens [56].

On the other hand, linear B- and/or T- cell epitopes may play the most important roles in cross-allergenicity between novel foods [57]. From the clinical point of view, it is important to know the patterns of cross-reactivity because they often reflect the pattern of clinical sensitivities and reactivity. Food processing or digestion may increase the number or the accessibility of IgE binding epitopes. Thermal processing may influence plant protein allergenicity to a different extent,

either increasing or decreasing IgE immune-reactivity, which has already been shown for legumes other than lupine, e.g. peanut, where thermal processing (roasting) enhanced the allergenic potency [54]. Thus, some food allergens have been described to lead to a loss of some or all the B-cell epitopes (but not the T-cell epitopes) by denaturalization/digestion [58]. In a similar fashion, vicilin-like allergens such as the major peanut and lupin allergen, Ara h 1 and Lup an 1, respectively, also share thermal stability. B- and T-cell responses have a defining and differential recognition of antigenic epitopes, and their localization in the allergen does not necessarily coincide. T-cell receptor recognizes only the linear amino acid sequence [59]. In contrast, B-cell epitopes are either linear or conformational and are located on the surface of the molecule accessible to antibodies. The extension of the epitope may range from 5 to 8 (or even more) amino acids for IgE to be able of binding to the epitope [60–63]. However, we have found linear B-cell epitopes in lupin β-conglutins and the other legume allergens with a wide range of amino acid lengths, and overlapping with conformational epitopes.

ACKNOWLEDGEMENTS

This research was supported by the European Research Program MARIE CURIE (FP7-PEOPLE-2011-IOF), under the grant reference number PIOF-GA-2011-301550 to JCJ-L, and JDA; and by ERDF co-funded projects P2010-AGR-6274 and P2011-CVI-7487 (Junta de Andalucía, Spain). JCJ-L thanks to the Spanish Ministry of Economy and Competitiveness for the grant ref. number RYC-2014-16536 (Ramon y Cajal Research Program).

AUTHOR CONTRIBUTIONS

JCJ-L conceived, designed and performed the study. JCJ-L, EL-C, PR-B and JDA analysed, discussed and assessed the resulting data. JCJ-L and JDA contributed rea-gents/materials/analysis tools. JCJ-L, EL-C, PR-B, and JDA wrote the paper.

CONFLICT OF INTEREST

The authors confirm that this article content has no conflicts of interest.

REFERENCES

1. French RJ, Buirchell BJ. **Lupin: the largest grain legume crop in Western Australia, its adaptation and improvement through plant breeding**. Australian Journal of Agricultural Research. 2005;56(11):1169–1180. doi:10.1071/ar05088.
2. Lucas MM, Stoddard F, Annicchiarico P, Frias J, Martinez-Villaluenga C, Sussmann D, et al. **The future of lupin as a protein crop in Europe**. Frontiers in Plant Science. 2015;6:705. doi:10.3389/fpls.2015.00705.
3. D'Agostina A, Antonioni C, Resta D, Arnoldi A, Bez J, Knauf U, et al. **Optimization of a Pilot-Scale Process for Producing Lupin Protein Isolates with Valuable Technological Properties and Minimum Thermal Damage**. Journal of Agricultural and Food Chemistry. 2006 Jan;54(1):92–98. doi:10.1021/jf0518094.
4. Belski R, Mori TA, Puddey IB, Sipsas S, Woodman RJ, Ackland TR, et al. **Effects of lupin-enriched foods on body composition and cardiovascular disease risk factors: a 12-month randomized controlled weight loss trial**. International Journal of Obesity. 2010 Oct;35(6):810–819. doi:10.1038/ijo.2010.213.
5. Magni C, Sessa F, Accardo E, Varoni M, Morazzoni P, Scarafoni A, et al. **Conglutin gamma, a lupin seed protein, binds insulin in vitro and reduces plasma glucose levels of hyperglycemic rats**. The Journal of Nutritional Biochemistry. 2004 Nov;15(11):646–650. doi:10.1016/j.jnutbio.2004.06.009.
6. Foley RC, Gao LL, Spriggs A, Soo LY, Goggin DE, Smith PM, et al. **Identification and characterisation of seed storage protein transcripts from Lupinus angustifolius**. BMC Plant Biology. 2011;11(1):59. doi:10.1186/1471-2229-11-59.
7. Foley RC, Jimenez-Lopez JC, Kamphuis LG, Hane JK, Melser S, Singh KB. **Analysis of conglutin seed storage proteins across lupin species using transcriptomic, protein and comparative genomic approaches**. BMC Plant Biology. 2015;15(1):106. doi:10.1186/s12870-015-0485-6.
8. Bähr M, Fechner A, Krämer J, Kiehntopf M, Jahreis G. **Lupin protein positively affects plasma LDL cholesterol and LDL:HDL cholesterol ratio in hypercholesterolemic adults after four weeks of supplementation: a randomized, controlled crossover study**. Nutrition Journal. 2013;12(1):107. doi:10.1186/1475-2891-12-107.
9. Peeters KABM, Nordlee JA, Penninks AH, Chen L, Goodman RE, Bruijnzeel-Koomen CAFM, et al. **Lupine allergy: Not simply cross-reactivity with peanut or soy**. Journal of Allergy and Clinical Immunology. 2007 Sep;120(3):647–653. doi:10.1016/j.jaci.2007.05.032.
10. Dooper M, Plassen C, Holden L, Lindvik H, Faeste C. **Immunoglobulin E cross-reactivity between lupine conglutins and peanut allergens in serum of lupine-allergic individuals**. J Investig Allergol Clin Immunol. 2009;19(4):283–291.
11. Breiteneder H, Radauer C. **A classification of plant food allergens**. Journal of Allergy and Clinical Immunology. 2004 May;113(5):821–830. doi:10.1016/j.jaci.2004.01.779.
12. Jimenez-Lopez JC, Zafra A, Palanco L, Florido JF, Alché JdD. **Identification and Assessment of the Potential Allergenicity of 7S Vicilins in Olive (Olea europaeaL.) Seeds**. BioMed Research International. 2016;2016:1–11. doi:10.1155/2016/4946872.
13. Goggin DE, Mir G, Smith WB, Stuckey M, Smith PMC. **Proteomic Analysis of Lupin Seed Proteins To Identify Conglutin Beta as an Allergen, Lup an 1**. Journal of Agricultural and Food Chemistry. 2008 Aug;56(15):6370–6377. doi:10.1021/jf800840u.
14. Parisot L, Aparicio C, Moneret-Vautrin DA, Guerin L. **Allergy to lupine flour**. Allergy. 2001 Sep;56(9):918–919. doi:10.1034/j.1398-9995.2001.00254.x.
15. Holden L, Egaas E. **Quantitative Sandwich ELISA for the Determination of Lupine (Lupinus spp.) in Foods**. Journal of Agricultural and Food Chemistry. 2005 Jul;53(15):5866–5871. doi:10.1021/jf050631i.
16. Quaresma RR, Viseu R, Martins LM, Tomaz E, Inácio F. **Allergic primary sensitization to lupine seed**. Allergy. 2007 Dec;62(12):1473–1474. doi:10.1111/j.1398-9995.2007.01508.x.
17. Vieths S, Scheurer S, Ballmer-Weber B. **Current Understanding of Cross-Reactivity of Food Allergens and Pollen**. Annals of the New York Academy of Sciences. 2006 Jan;964(1):47–68. doi:10.1111/j.1749-6632.2002.tb04132.x.
18. Koplin JJ, Martin PE, Allen KJ. **An update on epidemiology of anaphylaxis in children and adults**. Current Opinion in Allergy and Clinical Immunology. 2011 Oct;11(5):492–496. doi:10.1097/aci.0b013e32834a41a1.
19. Jimenez-Lopez JC, Morales S, Castro AJ, Volkmann D, Rodríguez-García MI, Alché JdD. **Characterization of Profilin Polymorphism in Pollen with a Focus on Multifunctionality**. PLoS ONE. 2012 Feb;7(2):e30878. doi:10.1371/journal.pone.0030878.
20. Chevenet F, Brun C, Bañuls AL, Jacq B, Christen R. **TreeDyn: towards dynamic graphics and annotations for analyses of trees**. BMC Bioinformatics. 2006;7(1):439. doi:10.1186/1471-2105-7-439.
21. Jimenez-Lopez JC, Rodríguez-García M, Alché JD. **Analysis of the Effects of Polymorphism on Pollen Profilin Structural Functionality and the Generation of Conformational, T- and B-Cell Epitopes**. PLoS ONE. 2013 Oct;8(10):e76066. doi:10.1371/journal.pone.0076066.
22. Arnold K, Bordoli L, Kopp J, Schwede T. **The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling**. Bioinformatics. 2005 Nov;22(2):195–201. doi:10.1093/bioinformatics/bti770.
23. Jimenez-Lopez JC, Kotchoni SO, Hernandez-Soriano MC, Gachomo EW, Alché JD. **Structural functionality, catalytic**

mechanism modeling and molecular allergenicity of phenylcoumaran benzylic ether reductase, an olive pollen (Ole e 12) allergen. Journal of Computer-Aided Molecular Design. 2013 Oct;27(10):873–895. doi:10.1007/s10822-013-9686-y.

24. Ashkenazy H, Erez E, Martz E, Pupko T, Ben-Tal N. **ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids**. Nucleic Acids Research. 2010 May;38(Web Server):W529–W533. doi:10.1093/nar/gkq399.
25. Berezin C, Glaser F, Rosenberg J, Paz I, Pupko T, Fariselli P, et al. **ConSeq: the identification of functionally and structurally important residues in protein sequences**. Bioinformatics. 2004 Feb;20(8):1322–1324. doi:10.1093/bioinformatics/bth070.
26. Schrödinger, LLC. **The PyMOL Molecular Graphics System, Version 1.8**; 2015.
27. Fraczkiewicz R, Braun W. **Exact and efficient analytical calculation of the accessible surface areas and their gradients for macromolecules**. Journal of Computational Chemistry. 1998;19(3):319–333. doi:10.1002/(SICI)1096-987X(199802)19:3<319::AID-JCC6>3.0.CO;2-W.
28. Jo S, Vargyas M, Vasko-Szedlar J, Roux B, Im W. **PBEQ-Solver for online visualization of electrostatic potential of biomolecules**. Nucleic Acids Research. 2008 May;36(Web Server):W270–W275. doi:10.1093/nar/gkn314.
29. Ivanciuc O. **SDAP: database and computational tools for allergenic proteins**. Nucleic Acids Research. 2003 Jan;31(1):359–362. doi:10.1093/nar/gkg010.
30. Singh H, Ahmad S. **Context dependent reference states of solvent accessibility derived from native protein structures and assessed by predictability analysis**. BMC Struct Biol. 2009;9(1):25. doi:10.1186/1472-6807-9-25.
31. Ahmad S, Gromiha M, Fawareh H, Sarai A. **ASAView: Database and tool for solvent accessibility representation in proteins**. BMC Bioinformatics. 2004;5(1):51. doi:10.1186/1471-2105-5-51.
32. Saha S, Raghava GPS. **AlgPred: prediction of allergenic proteins and mapping of IgE epitopes**. Nucleic Acids Research. 2006 Jul;34(Web Server):W202–W209. doi:10.1093/nar/gkl343.
33. Sharma P, Gaur SN, Arora N. **In silico Identification of IgE-Binding Epitopes of Osmotin Protein**. PLoS ONE. 2013 Jan;8(1):e54755. doi:10.1371/journal.pone.0054755.
34. Ponomarenko J, Bui HH, Li W, Fusseder N, Bourne PE, Sette A, et al. **ElliPro: a new structure-based tool for the prediction of antibody epitopes**. BMC Bioinformatics. 2008;9(1):514. doi:10.1186/1471-2105-9-514.
35. Gao D, Jimenez-Lopez JC, Iwata A, Gill N, Jackson SA. **Functional and Structural Divergence of an Unusual LTR Retrotransposon Family in Plants**. PLoS ONE. 2012 Oct;7(10):e48595. doi:10.1371/journal.pone.0048595.
36. Jimenez-Lopez JC, Kotchoni SO, Rodríguez-García MI, Alché JD. **Structure and functional features of olive pollen pectin methylesterase using homology modeling and molecular docking methods**. Journal of Molecular Modeling. 2012 Jun;18(12):4965–4984. doi:10.1007/s00894-012-1492-2.
37. Zhang L, Chen Y, Wong HS, Zhou S, Mamitsuka H, Zhu S. **TEPITOPEpan: Extending TEPITOPE for Peptide Binding Prediction Covering over 700 HLA-DR Molecules**. PLoS ONE. 2012 Feb;7(2):e30483. doi:10.1371/journal.pone.0030483.
38. Moreno FJ, Clemente A. **2S Albumin Storage Proteins: What Makes them Food Allergens?** The Open Biochemistry Journal. 2008 Feb;2:16–28. doi:10.2174/1874091x00802010016.
39. Sen M, Kopper R, Pons L, Abraham EC, Burks AW, Bannon GA. **Protein Structure Plays a Critical Role in Peanut Allergen Stability and May Determine Immunodominant IgE-Binding Epitopes**. The Journal of Immunology. 2002 Jul;169(2):882–887. doi:10.4049/jimmunol.169.2.882.
40. Mattarozzi M, Bignardi C, Elviri L, Careri M. **Rapid Shotgun Proteomic Liquid Chromatography–Electrospray Ionization–Tandem Mass Spectrometry-Based Method for the Lupin (Lupinus albus L.) Multi-allergen Determination in Foods**. Journal of Agricultural and Food Chemistry. 2012 Jun;60(23):5841–5846. doi:10.1021/jf302105r.
41. Mills ENC, Jenkins J, Marigheto N, Belton PS, Gunning AP, Morris VJ. **Allergens of the cupin superfamily**. Biochemical Society Transactions. 2002 Nov;30(6):925–929. doi:10.1042/bst0300925.
42. Burks AW, Shin D, Cockrell G, Stanley JS, Helm RM, Bannon GA. **Mapping and Mutational Analysis of the IgE-Binding Epitopes on Ara h 1, a Legume Vicilin Protein and a Major Allergen in Peanut Hypersensitivity**. European Journal of Biochemistry. 1997 Apr;245(2):334–339. doi:10.1111/j.1432-1033.1997.t01-1-00334.x.
43. Flinterman AE, Knol EF, Lencer DA, Bardina L, den Hartog Jager CF, Lin J, et al. **Peanut epitopes for IgE and IgG4 in peanut-sensitized children in relation to severity of peanut allergy**. Journal of Allergy and Clinical Immunology. 2008 Mar;121(3):737–743.e10. doi:10.1016/j.jaci.2007.11.039.
44. Ramesh M, Yuenyongviwat A, Konstantinou GN, Lieberman J, Pascal M, Masilamani M, et al. **Peanut T-cell epitope discovery: Ara h 1**. Journal of Allergy and Clinical Immunology. 2016 Jun;137(6):1764–1771.e4. doi:10.1016/j.jaci.2015.12.1327.
45. DeLong JH, Simpson KH, Wambre E, James EA, Robinson D, Kwok WW. **Ara h 1–reactive T cells in individuals with peanut allergy**. Journal of Allergy and Clinical Immunology. 2011 May;127(5):1211–1218.e3. doi:10.1016/j.jaci.2011.02.028.
46. Prickett SR, Voskamp AL, Phan T, Dacumos-Hill A, Mannering SI, Rolland JM, et al. **Ara h 1 CD4+ T cell epitope-based peptides: candidates for a peanut allergy therapeutic**. Clinical & Experimental Allergy. 2013;43(6):684–697. doi:10.1111/cea.12113.
47. Jimenez-Lopez JC, Lima-Cabello E, Melser S, Foley RC, Singh KB, Juan D A. **Lupin Allergy: Uncovering Structural Features and Epitopes of Beta-conglutin Proteins in Lupinus Angustifolius L. with a Focus on Cross-allergenic Reactivity to Peanut and Other Legumes**. Lecture Notes in Computer Science. 2015;9043:96–107. doi:10.1007/978-3-319-16483-0_10.
48. Wangorsch A, Ballmer-Weber BK, Rösch P, Holzhauser T, Vieths S. **Mutational epitope analysis and cross-reactivity of two isoforms of Api g 1, the major celery allergen**. Molecular Immunology. 2007 Apr;44(10):2518–2527. doi:10.1016/j.molimm.2006.12.023.
49. Ferreira F, Hirtenlehner K, Jilek A, Godnik-Cvar J, Breiteneder H, Grimm R, 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.** Journal of Experimental Medicine. 1996;183(2):599–609. doi:10.1084/jem.183.2.599.
50. Burks AW, Cockrell G, Stanley JS, Helm RM, Bannon GA. **Recombinant peanut allergen Ara h I expression and IgE binding in patients with peanut hypersensitivity.** Journal of Clinical Investigation. 1995 Oct;96(4):1715–1721. doi:10.1172/jci118216.
51. Wensing M, Knulst AC, Piersma S, O'Kane F, Knol EF, Koppelman SJ. **Patients with anaphylaxis to pea can have peanut allergy caused by cross-reactive IgE to vicilin (Ara h 1)**. Journal of Allergy and Clinical Immunology. 2003 Feb;111(2):420–424. doi:10.1067/mai.2003.61.
52. Barre A, Borges JP, Rougé P. **Molecular modelling of the major peanut allergen Ara h 1 and other homotrimeric allergens of the cupin superfamily: a structural basis for their IgE-binding cross-reactivity**. Biochimie. 2005 Jun;87(6):499–506. doi:10.1016/j.biochi.2005.02.011.
53. Guillamón E, Rodríguez J, Burbano C, Muzquiz M, Pedrosa MM, Cabanillas B, et al. **Characterization of lupin major allergens (Lupinus albus L.)**. Molecular Nutrition & Food Research. 2010 Nov;54(11):1668–1676. doi:10.1002/mnfr.200900452.
54. Peeters KABM, Koppelman SJ, Penninks AH, Lebens A, Bruijnzeel-Koomen CAFM, Hefle SL, et al. **Clinical relevance of sensitization to lupine in peanut-sensitized adults**. Allergy. 2009 Apr;64(4):549–555. doi:10.1111/j.1398-9995.2008.01818.x.
55. Wassenberg J, Hofer M. **Lupine-induced anaphylaxis in a child without known food allergy**. Annals of Allergy, Asthma & Immunology. 2007 jun;98(6):589–590. doi:10.1016/s1081-1206(10)60741-7.
56. Valenta R, Duchene M, Ebner C, Valent P, Sillaber C, Deviller P, et al. **Profilins constitute a novel family of functional plant pan-allergens.** Journal of Experimental Medicine. 1992;175(2):377–385. doi:10.1084/jem.175.2.377.
57. Aalberse RC, Akkerdaas J, van Ree R. **Cross-reactivity of IgE antibodies to allergens**. Allergy. 2001 Jun;56(6):478–490. doi:10.1034/j.1398-9995.2001.056006478.x.
58. Schimek EM, Zwölfer B, Briza P, Jahn-Schmid B, Vogel L, Vieths S, et al. **Gastrointestinal digestion of Bet v 1–homologous food allergens destroys their mediator-releasing, but not T cell–activating, capacity**. Journal of Allergy and Clinical Immunology. 2005;116(6):1327–1333. doi:http://dx.doi.org/10.1016/j.jaci.2005.09.007.
59. Pomés A. **Relevant B Cell Epitopes in Allergic Disease**. International Archives of Allergy and Immunology. 2010;152(1):1–11. doi:10.1159/000260078.
60. Meno KH. **Allergen structures and epitopes**. Allergy.

2011;66:19–21. doi:10.1111/j.1398-9995.2011.02625.x.

61. Bannon GA, Ogawa T. **Evaluation of available IgE-binding epitope data and its utility in bioinformatics**. Molecular Nutrition & Food Research. 2006 Jul;50(7):638–644. doi:10.1002/mnfr.200500276.

62. Tanabe S. **IgE-binding Abilities of Pentapeptides, QQPFP and PQQPF in Wheat Gliadin**. Journal of Nutritional Science and Vitaminology. 2004;50(5):367–370. doi:10.3177/jnsv.50.367.

63. Asturias JA. **Molecular and structural analysis of the panallergen profilin B cell epitopes defined by monoclonal antibodies**. International Immunology. 2002 Sep;14(9):993–1001. doi:10.1093/intimm/dxf070.