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

Elena Lima-Cabello¹, Paula Robles-Bolivar¹, Juan D. Alché¹, Jose C. Jimenez-Lopez¹,²,∗

¹Plant Reproductive Biology Laboratory; Department of Biochemistry, Cell and Molecular Biology of Plants; Estacion Experimental del Zaidin (EEZ); Spanish National Research Council (CSIC); 1 Profesor Alberada Street, Granada 18008 Spain
²The 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 allergens 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-leaved 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
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-conglycinin, 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-conglycinin 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 F5B8W5), Ara h 1 (P43237, P43238), Len c 1 (Q84U01, Q84U11), 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), BCPRES (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 $\alpha_1$-$3$, $\beta_1$-$7$, $\gamma_1$-$2$, and $\delta_1$-$4$), 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 $\beta_1$- to $\beta_7$-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 $\beta$-conglutin proteins and allergens from other species. The data clearly revealed five established groups/clusters. We identified 5 main groups, where $\beta$-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 $\beta$-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 $\beta$-conglutins, with high potential of antigenicity, 7 regions in Ara h 1, 7 regions in Gly m 5, 8 regions in $\beta$-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 $\beta$-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%)
Table 1: IgE epitopes identification and analysis

<table>
<thead>
<tr>
<th>Species</th>
<th>Allergen name</th>
<th>Uniprot ID</th>
<th>IgE1</th>
<th>IgE2</th>
<th>IgE3</th>
<th>IgE4</th>
<th>IgE5</th>
<th>IgE6</th>
<th>IgE7</th>
<th>IgE8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lupinus angustifolius</td>
<td>Conglutin b4</td>
<td>P08893</td>
<td>NLRLLGFGIN</td>
<td>KQLTFPGSVE</td>
<td>RYNARLSEG</td>
<td>SYFSGFSRNT</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Lupinus angustifolius</td>
<td>Conglutin b5</td>
<td>P08894</td>
<td>NLRLLGFGIN</td>
<td>KGLTFPGSVE</td>
<td>RYNARLSEG</td>
<td>SYFSGFSRNT</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Lupinus angustifolius</td>
<td>Conglutin b6</td>
<td>P08895</td>
<td>NLRLLGFGIN</td>
<td>KGLTFPGSVE</td>
<td>RYNARLSEG</td>
<td>SYFSGFSRNT</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Lupinus angustifolius</td>
<td>Conglutin b7</td>
<td>P08896</td>
<td>NLRLLGFGIN</td>
<td>KGLTFPGSVE</td>
<td>RYNARLSEG</td>
<td>SYFSGFSRNT</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Lupinus angustifolius</td>
<td>Conglutin b8</td>
<td>P08897</td>
<td>NLRLLGFGIN</td>
<td>KGLTFPGSVE</td>
<td>RYNARLSEG</td>
<td>SYFSGFSRNT</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Glycine max</td>
<td>Gly m 5</td>
<td>Q9FZP9</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Glycine max</td>
<td>Gly m 6</td>
<td>Q9FZP9</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Lens culinaris</td>
<td>Len c 1</td>
<td>Q84UI0</td>
<td>DNLIGFGIN</td>
<td>KELAFPGSSH</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Lens culinaris</td>
<td>Len c 2</td>
<td>Q84UI1</td>
<td>DNLIGFGIN</td>
<td>KELAFPGSSH</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Pisum sativum</td>
<td>Pis s 2</td>
<td>Q5B15</td>
<td>NLRLLGFGIN</td>
<td>RRYSARLSEG</td>
<td>SYFSGFSRNT</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Anachora hystoxa</td>
<td>Ara h 1</td>
<td>P13915</td>
<td>ELHLLGFGIN</td>
<td>KDLAFPGSGE</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Vigna radiata</td>
<td>Vign r 2</td>
<td>Q18973</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

- Conjugation (black colour, Figure 3) and their counterparts in five species of legumes.
- The biggest structural difference between the β-conglutin 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 linear 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-conglutin (Figure 4), exhibiting a large surface orientation. This epitope is common for other legume allergens such as peanut (Ara h 1), soybean (β-conglutin), 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.

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-allergic 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 characterise the conformational IgE-binding, and linear
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°/ß1-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, 98885, 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 (~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 reasons why the L. angustifolius have been identified in allergen proteins share high percentage of identity. 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, 98885, 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 (~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 h 1, Len c 1 and Pis s 1) [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 accompanied 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 colls 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 denaturation/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 [50–53]. 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 Andalucia, 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


44. Meno KH. Allergen structures and epitopes. Allergy.

