Homology modelling of the major peanut allergen Ara h 2 and surface mapping of IgE-binding epitopes

Annick Barre, Jean-Philippe Borges, Raphaël Culerrier, Pierre Rougé∗

Surfaces Cellulaires et Signalisation chez les Végétaux, UMR-CNRS 5546, Pôle de Biotechnologie végétale, 24 Chemin de Borde Rouge, 31326 Castanet-Tolosan, France

Received 17 January 2005; received in revised form 7 March 2005; accepted 18 March 2005
Available online 28 April 2005

Abstract

Three-dimensional models built for the peanut Ara h 2 allergen and other structurally-related 2S albumin allergens of dietary nuts exhibited an overall three-dimensional fold stabilized by disulphide bridges well conserved among all the members of the 2S albumin superfamily. Conformational analysis of the linear IgE-binding epitopes mapped on the molecular surface of Ara h 2 showed no structural homology with the corresponding regions of the walnut Jug r 1, the pecan nut Car i 1 or the Brazil nut Ber e 1 allergens. The absence of epitopic community does not support the allergenic cross-reactivity observed between peanut and walnut or Brazil nut, which presumably depends on other ubiquitous seed storage protein allergens, namely the vicilins. However, the major IgE-binding epitope identified on the molecular surface of the walnut Jug r 1 allergen shared a pronounced structural homology with the corresponding region of the pecan nut Car i 1 allergen. With the exception of peanut, 2S albums could thus account for the IgE-binding cross-reactivity observed between some other dietary nuts, e.g. walnut and pecan nut.

© 2005 Elsevier B.V. All rights reserved.

Keywords: Peanut allergen; 2S albumin; Three-dimensional model; Epitope mapping; IgE-binding cross-reactivity

1. Introduction

Peanut allergy has been recognized as the most worrying food hypersensitivity responsible for extremely severe allergic reactions in children, adolescents and adults [1]. In fact, peanut and peanut-containing food products cause more fatal anaphylactic shocks than any other foodstuffs [2]. Sensitization to peanut generally occurs during the infancy and very often persists throughout the life [3]. In this respect, peanut allergy actually differs from other current food allergies, e.g. to milk or egg proteins, which prevail in children and usually vanish in adults [4]. Until now, the strict avoidance of peanut and peanut by-products into the diet constitutes the best way to prevent the allergic manifestations in sensitized individuals. However, peanut proteins often occur as hidden allergens in many (inadequately labelled) food products and the accidental consumption of offending food allergens is very difficult to avoid. In addition, closely-related allergens occurring in other dietary seeds, e.g. tree nut [5], are susceptible to trigger allergic reactions in individuals previously sensitized to peanut proteins. Obviously, this allergic cross-reactivity drastically complicates the food avoidance rules of peanut-sensitive patients [6] and inevitably impacts upon their quality of life [7]. Peanut allergy depends on three allergens Ara h 1 [8], Ara h 2 [9] and Ara h 3 [10] that occur in seeds as storage proteins. Ara h 1 and Ara h 2 are of paramount importance since serum IgE from >90% of peanut-sensitive individuals recognize these two major allergens [11,12]. Linear IgE-binding B-cell epitopes have been identified along...
the amino acid sequence of Ara h 1 [11] and Ara h 2 [12]. Ara h 2 consists of an extremely ubiquitous 2S albumin seed storage protein also present in many other dietary seeds, including soybean [13], walnut [14], cashew nut [15], Brazil nut [16], sunflower [17] or sesame seeds [18] (see Fig. 3). Although these nut allergens are moderately similar (∼35% of amino acid sequence similarity) to Ara h 2, they apparently share a conserved three-dimensional fold [19]. Some of them could therefore exhibit B-cell epitopes very similar to those recognized on the surface of Ara h 2 and thus account for the IgE-binding cross-reactivity observed between peanut and other dietary nuts. To check this hypothesis, three-dimensional models of Ara h 2 and other nut allergens were built by homology modelling and checked for the presence of epitopes with conformation similar to the Ara h 2 B-cell epitopes. The major IgE-binding epitopes of the walnut Jug r 1 and pecan Car i 1 allergens shared a very similar conformation but exhibited no structural homology with the corresponding region of the peanut allergen. Although 2S albumins offer a molecular basis for the IgE-binding cross-reactivity observed between some dietary nuts, e.g. walnut and pecan nut, they apparently do not account for the cross-reactions reported between some dietary nuts, e.g. walnut and pecan nut, they apparently do not account for the cross-reactions reported between peanut and other dietary nuts.

2. Materials and methods

Multiple amino acid sequence alignments were carried out with CLUSTAL-X [20] and displayed with ESPript [21]. The hydrophobic cluster analysis (HCA) [22] was performed to delineate the conserved secondary structural features (stretches of α-helix) along the amino acid sequence of Ara h 2 and other 2S albumins by comparison with the castor bean (Ricinus communis) 2S albumin allergen Ric c 3 [23] used as a model. HCA plots were generated using the program drawhca from L. Canard (http://www.lmcp.jussieu.fr/~soyer/www-hca/hca-form.html).

Molecular modelling of Ara h 2 and other 2S albumins was carried out on a Silicon Graphics O2 R10000 workstation, using the programs InsightII, Homology and Discover 3 (Accelrys, San Diego CA, USA). The atomic coordinates of the castor bean allergen Ric c 3 [23] (RCSB Protein Data Bank code 1PSY) were used to build the three-dimensional model of the allergens. The percentages of both identity (∼35%) and homology (∼75%) Ric c 3 shares with Ara h 2 (Fig. 1) and other dietary 2S seed albumins allowed us to build rather accurate three-dimensional models using the RMN coordinates of the castor bean 2S albumin as a template. The extended N-terminal region, which is extremely desordered in the 20 RMN models, was necessarily omitted in the model building. Accordingly, the model built for Ara h 2 starts at residue Glu5. Two out of the 10 linear IgE-binding epitopes identified along the amino acid sequence of Ara h 2 fell into the N-terminal region that has been neglected in the model building. However, they do not correspond to the three immunodominant epitopes characterized in Ara h 2. Ara h 2 readily differs from Ric c 3 by an extra-loop corresponding to a Pro-rich repeat occurring in the N-terminal portion of the polypeptide. Another less extended extra-loop occurs at the C-terminal end of the polypeptide chain. These two extra-loops were tentatively modelled using the best geometrical fitting with the core domain as a criterion for choosing among the different loop conformations available in the PDB. Steric conflicts were corrected during the model building procedure using the rotamer library [24] and the search algorithm implemented in the Homology program [25] to maintain proper side-chain orientation. The geometry of loop regions was corrected using the refine option of TurboFrodo [26]. An energy minimization of the final models was carried out by 150 cycles of steepest descent using Discover 3. The program TurboFrodo was run to draw the Ramachandran plot and to perform the superposition of the model with the template protein. PROCHECK [27] was used to assess the geometric quality of the three-dimensional models. As an example, 62% of the residues of Car i 1 (54% for 1PSY used as template) were correctly assigned on the best allowed regions of the Ramachandran plot, the remaining residues being located in the generously allowed regions of the plot except for Glu57 which occurs in the non allowed region (result not shown). Cartoons were drawn with PyMOL W.L. DeLano (http://www.pymol.org/).

Electrostatic potentials were calculated and displayed with GRASP using the parse3 parameters [28]. The solvent probe radius used for molecular surfaces was 1.4 Å and a standard
2.0 Å-Stern layer was used to exclude ions from the molecular surface [29]. The inner and outer dielectric constants applied to the protein and the solvent were, respectively, fixed at 4.0 and 80.0 and the calculations were performed keeping a salt concentration of 0.145 M. No even distribution of the net negative charge of the carboxylic group of negatively charged residues was performed between their two oxygen atoms prior to the calculations.

B-cell epitopes were predicted from the hydrophobic profiles as being the most hydrophilic, flexible and surface exposed regions. Different scales of hydrophilicity [30], flexibility [31], exposition to the solvent [32] and antigenicity [33], were used to build the hydrophobic profiles with the MacVector (Kodak) software.

The surface occupied by the previously identified or predicted sequential B-cell epitopes [14] along the amino acid sequence of Ara h 2 was calculated and displayed on the molecular surface of the proteins with PyMOL. The overall conformation of the sequential B-cell epitopes on the molecular surface of the 2S albums was displayed with PyMOL.

3. Results and discussion

Homology-based molecular modelling of Ara h 2 and other 2S seed albumin allergens showed that seed storage 2S albums share a common three-dimensional structural scaffold made of five α-helices arranged in a right-handed superhelix and connected by more or less extended loops (Fig. 2A and B). This three-dimensional conformation, stabilized by four conserved disulphide bridges, is reminiscent of that found in type 1 LTP which constitute another group of important food allergens [34]. However, 2S albums differ from LTP by the topology of the disulphide bridges [23]. Ara h 2 readily differ from other 2S albums by two additional loops located at the N- and C-terminal ends of the polypeptide chain, respectively. The more extended N-terminal extra loop roughly corresponds to a tandem repeat of a Pro-rich hexapeptide DPYPS which has no equivalent in Ric c 3 or coincides with the less conserved region of all other 2S albums sequenced so far. Ara h 2 contains a putative N-glycosylation site Asn88-Gln-Ser. Although well exposed, its location at the end of the main α4 helix suggests it is not actually glycosylated. Other modelled nut allergens and, especially, Jug r 1 and Car i 1, exhibited a very similar three-dimensional fold but the N-terminal extra loop corresponding to the tandem Pro-rich repeat of Ara h 2 is less extended (results not shown).

Molecular mapping of the 10 linear B-cell epitopes previously characterized along the polypeptide chain of Ara h 2 [12] showed a fairly well exposition on the molecular surface of the three-dimensional model (Fig. 2D). This is especially true for the two major epitopes #6 and 7 which correspond to the protruding N-terminal Pro-rich extra loop of Ara h 2. This extra loop has no counter-part in Ric c 3 (Fig. 2C and D) and possess an overall three-dimensional conformation distinct from that found in the corresponding shorter extra loops of other nut allergens, e.g. Jug r 1 or Car i 1. Even though they are better conserved in other nut allergens (see Fig. 3), other minor epitopes #8, 9 and 10 located at the C-terminus of the Ara h 2 polypeptide also differ by their three-dimensional conformation from the corresponding regions of Ric c 3 (Fig. 2E and F) and other allergens (results not shown). Most of these epitopes seems specific for Ara h 2 and hence unsuitable for promoting IgE-binding cross-reactivity with other nut allergens.

Similarly, no conformational similarity could be detected between the experimentally characterized or predicted B-cell epitopes of other nut allergens and the major or minor B-cell epitopes of Ara h 2. In this respect, although B-epitope 8 103RLQGRQQQQFK114 of Ara h 2 overlaps the core domain (bold letters) of the immunodominant linear B-cell epitope: QGLRGEEMEEMV characterized in Jug r 1 [35], both regions exhibit rather distinct amino acid sequences and overall conformations (Fig. 4). However, this linear B-cell epitope of Jun r 1 is well conserved (identical or homologous residues are underlined) in the pecan nut Car i 1 allergen (EGI3AGSVEVRL), the cashew nut Ana o 3 allergen (EIGRGEEMEEMY), and the Brazil nut Ber e 1 allergen (EMQDFPGEAMRTM) (Fig. 3). Accordingly, its overall conformation is very similar to Car i 1 but rather different in Ana o 3 or in the castor bean Ric c 3 allergen (94GQLHGEESVRVA104) (Fig. 4). This absence of epitopic community among 2S albums from distinct plant families does not support the allergenic cross-reactivity observed between peanut (Fabaceae) and Brazil nut (Juglandaceae), showing that pre-incubation of sera from peanut-sensitive patients with a Brazil nut extract resulted in a decrease in the IgE-binding to peanut extract in ELISA measurements [36]. Similarly, no decrease occurred after pre-incubation with a cashew nut (Anacardiaceae) extract.

Seed storage 2S albums thus appear as ubiquitous allergens distributed in plants belonging to distinct monocot and dicot families. Although their overall three-dimensional structure stabilized by disulphide bridges is well conserved among all the members of the 2S albumin superfamily, their amino acid sequences is highly variable, especially in the N-terminal part of the polypeptide chain. Consequently, they do not share common or consensus surface-exposed B-cell epitopes except for those belonging to the same family, e.g. the walnut Jug r 1 and the pecan nut Car i 1, susceptible to account for the currently observed allergenic cross-reactivity between peanut and other tree nuts. Other seed storage protein allergens largely distributed in dietary seeds, namely vicilins, should be responsible for this allergenic cross-reactivity. However, some IgE-binding cross-reactivity among 2S albums from peanut and other tree nuts could depend on IgE directed against cross-reactive carbohydrate determinants (anti-CDD IgE). A single putative N-glycosylation site 94Asn-Gln-Ser exists along the amino acid sequence of Ara h 2 but it is located into helix α3 that occurs at the C-terminal.

Fig. 3. Molecular mapping of the 10 linear B-cell epitopes characterized in Ara h 2. The location of the tandem Pro-rich repeat at the C-terminus is shown by a dotted line and the linkage of epitope #6 to the N-terminal extra loop by a solid line. The plot of the three-dimensional locations of the epitopes superimposes onto the overall conformation of Ara h 2.

Fig. 4. Sequence alignment of the 10 identified B-cell linear epitopes from peanut (Fabaceae) and Brazil nut (Juglandaceae) along with the homologous epitope from the cashew nut (Anacardiaceae).
Fig. 2. Ribbon diagram of the castor bean 2S albumin Ri c 3 allergen (A) and Ara h 2 (B). The two N- and C-terminal extra loops of Ara h 2 are indicated by stars (3 and 4) N and C indicate the N- and C-termini of the polypeptide chains, respectively. Molecular surfaces of Ri c 3 (C) and Ara h 2 (D) showing the areas occupied by the major B-epitopes 3 (pink), 6 (blue) and 7 (green); N-termini are oriented left. Molecular surfaces of Ri c 3 (E) and Ara h 2 (F) showing the areas occupied by the major B-epitopes 8 (orange), 6 (green blue) and 7 (light pink); N-termini are oriented right.

Fig. 3. Structural alignment of Ara h 2 with Ses i from Sesame seeds and other tree nut allergens Ara a 3, Bet v 1, Car i 1 and Jug r 1. The major B-cell epitopes (numbered 3, 6 and 7) and the minor B-cell epitopes (numbered 1, 2, 4, 5, 8–10) of Ara h 2 and the corresponding regions of other 2S allergens are open boxed. The major B-cell epitope of Jug r 1 and the corresponding regions of other allergens is dashed boxed.
end of the molecule. In addition, it is poorly exposed on the molecular surface in such a way that Ara h 2 may be supposed to be unglycosylated. According to [12] the allergenic character of Ara h 2, which do not contain significant amounts of carbohydrate, appears to be solely due to epitopes. Type O-glycans have also been shown to behave as CCD of the Mugwort pollen allergen Art v 1 [37]. However, sera of patients allergic to grass pollen with significant levels of anti-peanut IgE but without positive skin prick tests and no symptom of peanut allergy were extremely poorly reactive when tested for the presence of anti-CDD IgE [38]. Similarly, one cannot exclude that true conformational B-cell epitopes distinct from continuous epitopes participate in some cross-reactivity between Ara h 2 and other tree nut allergens even though their continuous epitopes do not share sequence or structural similarities. In this respect, the use of mimotopes from phage display libraries, allowed to identify compact patches of sequence motifs on the molecular surface of the Bet v 1 [39] and Phi p 5 [40] allergens that most likely represent the true conformational B-cell epitopes. The coalescence of linear epitopes 6–7 and 9–10 on the molecular surface of Ara h 2 (see Fig. 2D and F) could create more surface-extended conformational epitopes responsible for some cross-reactivity with other tree nut allergens.

Acknowledgement

The financial support of CNRS is gratefully acknowledged.

References


Kelly JD, Helft SL. 2S methionine-rich protein (SSA) from sunflower seed is an IgE-binding protein. Allergy 2000;55:556-60.


