Available online at www.sciencedirect.com
SCIENCE DIRECT
Immunology Letters 100 (2005) 153–158
Immunology Letters
www.elsevier.com/locate/

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 albumins could thus account for the IgE-binding cross-reactivity observed between some other dietary nuts, e.g. walnut and pecan nut.

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 but 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

Abbreviations: Ana o 3, *Anacardium occidentale* (cashew) allergen 3; Ara h 2, *Arachis hypogea* (peanut) allergen 2; Ber e 1, *Bertholletia excelsa* (Brazil nut) allergen 1; Cari 1, *Carya illinoinensis* (pecan) allergen 1; CDD, cross-reactive carbohydrate determinant; HCA, hydrophobic cluster analysis; IgE, immunoglobulin E; Jug r 1, *Juglans regia* (walnut) allergen 1; LTP, lipid transfer protein; Ses i 3, *Sesamum indicum* (sesame) allergen 3

* Corresponding author. Tel.: +33 05 62 19 35 02; fax: +33 05 62 19 35 02.
E-mail address: rouge@scsv.ups-tlse.fr (P. Rougé).

0165-2478/$ – see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.imlet.2005.03.014

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 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 of 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 fall 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 Cari 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

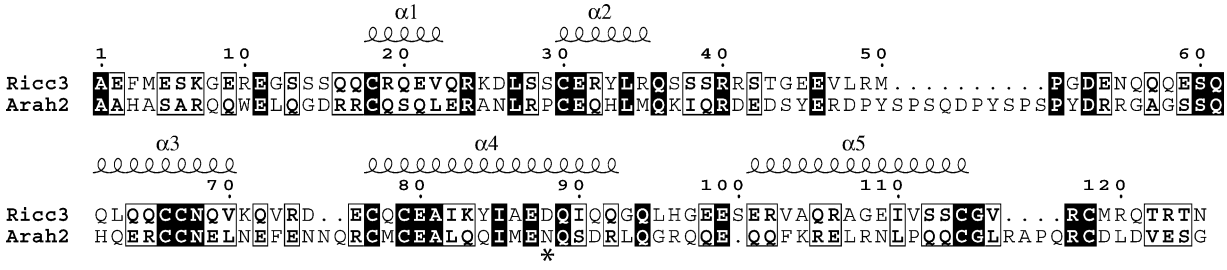

Fig. 1. Structural alignment of Ara h 2 to the castor bean Ric c 3 2S albumin. Stretches of α-helix (αl–α5) occurring along the polypeptide chain of Ric c 3 have been indicated. The putative *N*-glycosylation site of Ara h 2 is indicated by an asterisk (*). The alignment was performed with CLUSTAL-X [20], manually modified according to the HCA plots [22], and representeed with ESPript [21].