

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

Annick Barre, Jean-Philippe Borges, 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 8 September 2004; accepted 23 February 2005

Available online 02 April 2005

Abstract

Three-dimensional models of the major vicilin allergens from peanut (Ara h 1), lentil (Len c 1) and pea (Pis s 1), were built by homology-based modelling from the X-ray coordinates of the structurally closely related soybean β -conglycinin. All the allergen monomers exhibit the typical cupin motif made of two modules related by a pseudo-dyad axis. Each module consists of a β -barrel core domain associated to a loop domain which mainly contains α -helices. The three cupin motifs are assumed to be arranged in a homotrimeric structure similar to that observed in β -conglycinin, phaseolin or canavalin. Most of the sequential B-cell epitopes characterized on the C-terminus of the Ara h 1 allergen are well conserved in both Len c 1 and Pis s 1 allergens. They occupy very comparable areas on the molecular surface of the allergens and exhibit a similar three-dimensional conformation. This antigenic community readily accounts for the IgE-binding cross-reactivity commonly observed between the vicilin allergens from edible legume seeds. The clinical implication of this cross-reactivity is addressed for a definite diagnosis of legume seed allergy.

© 2005 Elsevier SAS. All rights reserved.

Keywords: Peanut allergen; three-dimensional model; B-cell epitopes; T-cell epitopes; allergy

1. Introduction

Soybean and, to a lesser extent, peanut, kidney bean, pea and lentil are among the most commonly used dietary seeds for their good nutritional value and balanced amino acid content. However, IgE-mediated food allergy to legume seed proteins and, especially, to peanut proteins, appears to be on the increase [1] and has currently become a worrying health problem in many countries [2,3]. Peanut allergy is the third most prevalent food allergy [4] that mostly contributes to the fatal

issued anaphylactic reactions in both children and adults [5]. Seed storage proteins [6] have been recognized as being the major peanut allergens Ara h 1 (vicilin) [7], Ara h 2 (2S albumin conglutin) [8], and Ara h 3 (glycinin with a Bowman–Birk trypsin inhibitor activity) [9] responsible for the allergic response of susceptible individuals. Ara h 1 is a homotrimeric protein belonging to the cupin superfamily [10,11] which contains structurally related proteins built from the β -barrel cupin motif but display very different functions [12]. Ara h 1 is particularly important since it is recognized by over 90% of individuals sensitised to peanut [7]. At least 23 IgE-binding sequential B-cell epitopes have been characterized along the polypeptide chain of the Ara h 1 monomer [13,14]. In addition, an IgE-binding cross-reactivity was identified with other legume vicilins from lentil (Len c 1) [15] and pea (Pis s 1) [16]. Other IgE-binding cross-reacting vicilin allergens include β -conglycinin from soybean [17] and Jug r 2 from walnut [18]. This hypersensitivity reactions to multiple dietary seeds has relevant clinical implications, e.g.,

Abbreviations: Ara h1, *Arachis hypogaea* (peanut) allergen 1; HCA, hydrophobic cluster analysis; IgE, immunoglobulin E; Len c 1, *Lens culinaris* (lentil) allergen 1; Pis s 1, *Pisum sativum* (pea) allergen 1; SCR, structurally conserved regions; Vic f 1, *Vicia faba* (faba bean) allergen 1.

* Corresponding author. Tel.: +33 05 62 19 35 58; fax: +33 05 62 19 35 02.

E-mail address: rouge@scsv.ups-tlse.fr (P. Rougé).

to accurately characterize the allergen(s) responsible for the sensitisation or to determine the risk of reaction to related foods in previously sensitised individuals [19]. In this context, the epitopic community of ubiquitous allergens has to be characterized at the molecular level to improve both the accuracy and interpretation of the diagnostic tests (skin prick tests, RASTs, in vivo IgE-binding tests, oral food challenges) by using carefully selected purified allergens or chemically synthesized specific epitopes. Here, we present a homology-based molecular modelling approach used to characterize the structural features responsible for the epitopic community among the major vicilin allergens from peanut (Ara h 1), lentil (Len c 1) and pea (Pis s 1).

2. Material and methods

Multiple amino acid sequence alignments were carried out with CLUSTAL-X [20] and displayed with ESPrnt [21].

The HCA (Hydrophobic Cluster Analysis) [22] was performed to delineate the conserved secondary structural features (strands of β -sheet and stretches of α -helix) along the amino acid sequence of Ara h 1, Len c 1 and Pis s 1 by comparison with the soybean (*Glycine max*) β -conglycinin [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>).

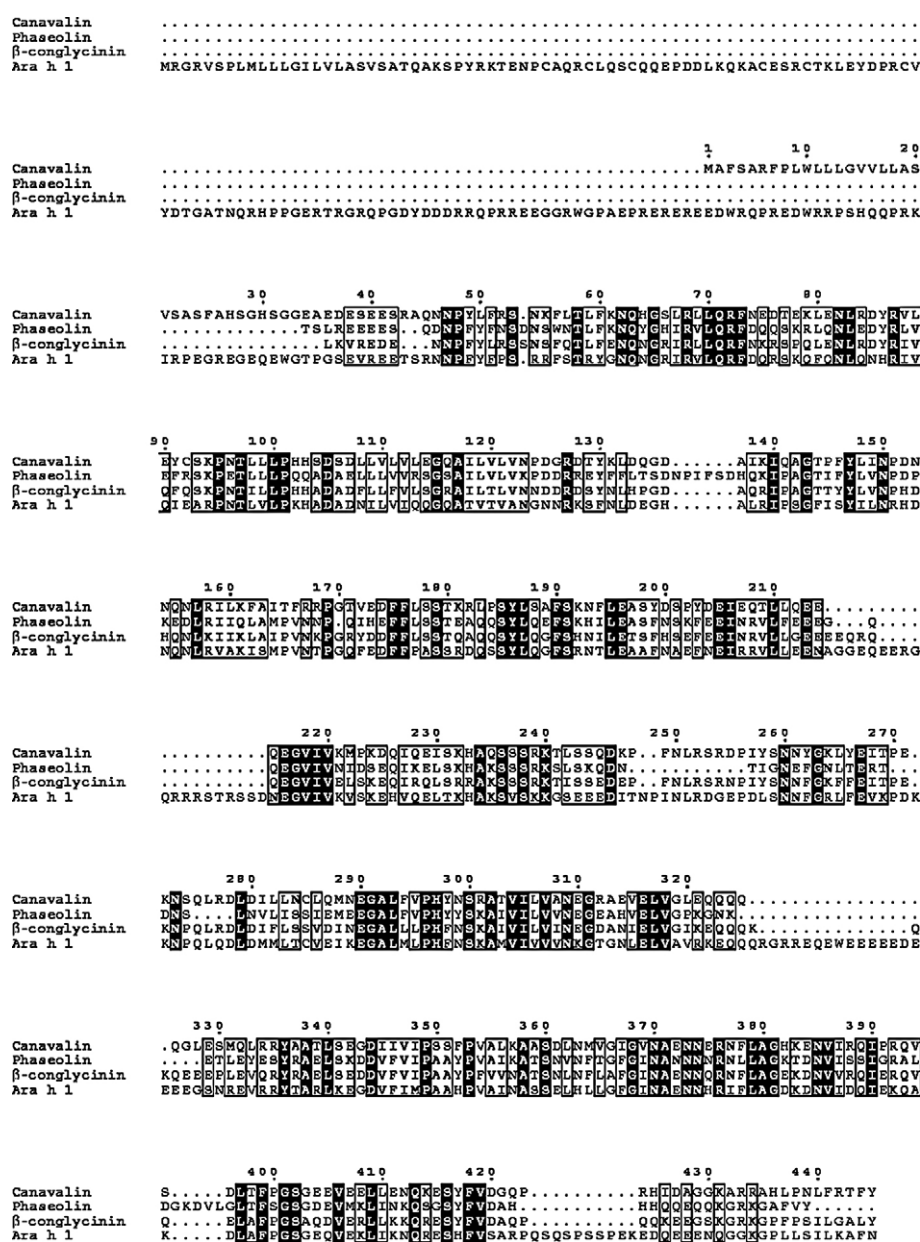


Fig. 1. Amino acid sequence alignment of Ara h 1 and other cupin proteins of known three-dimensional structures: canavalin, β -conglycinin and phaseolin. Identical residues are in black boxes and homologous residues are in open boxes. The alignment was performed with CLUSTAL-X [20] and represented with ESPrnt [21].

Molecular modeling of Ara h 1, Len c 1 and Pis s 1 was carried out on a Silicon Graphics O2 R10000 workstation, using the programs InsightII, Homology and Discover3 (Accelrys, San Diego CA, USA). The three X-ray solved cupin structures available at the PDB including phaseolin from kidney bean [24], canavalin from Jack bean [25] and β -conglycinin from soybean [26], share a high percentages of both identity (~60%) and homology (~90%) with the vicilin monomers of peanut, lentil and pea (Fig. 1) and thus could be used as templates for the molecular building procedure of vicilins. However, since the best sequence identity and structural homologies as shown from both the amino acid sequence alignments and HCA plots occurred with β -conglycinin (Fig. 2), we only used the atomic coordinates of the soybean β -conglycinin [26] (RCSB Protein Data Bank code 1IPJ) to build the three-dimensional model of the allergens. The higher percentages of identity (~60%) and homology (~90%) the vicilin monomers of peanut, lentil and pea share with β -conglycin allowed us to build rather accurate three-dimensional models for the vicilin monomers. The extended loop region corresponding to an X-ray unsolved 288–303 sequence stretch of β -conglycinin was omitted in the models built for other cupins. Steric conflicts were corrected during the model building procedure using the rotamer library [27] and the search algorithm implemented in the Homology program [28] to maintain proper side-chain orientation. The geometry of loop regions was corrected using the refine option of TurboFrodo [29]. An energy minimization of the final models was carried out by 50 cycles of steepest descent using the cvff forcefield of Discover. The program TurboFrodo was run to draw the Ramachandran plot and to perform the superpo-

sition of the model with the template protein. PROCHECK [30] was used to assess the geometric quality of the three-dimensional models. As an example, about 80% of the residues of Ara h 1 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 Asp222 and His410 which occur in the non-allowed region (result not shown). Cartoons were drawn with Molscript [31] and rendered with Bobscrip [32] and Raster3d [33].

Ara h 1 and other vicilin allergens were tentatively assigned as homotrimers displaying an arrangement of the cupin motifs similar to that found in β -conglycinin, phaseolin or canavalin. Most of the amino acid residues occurring at the interfaces of the cupin motifs of the β -conglycinin homotrimer are nicely conserved in the vicilin allergens. Accordingly, no steric clash was observed upon the homotrimeric assembly of the cupin motifs of Ara h 1, Len c 1 or Pis s 1.

Electrostatic potentials were calculated and displayed with GRASP using the parse3 parameters [34]. 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 [35]. The inner and outer dielectric constants applied to the protein and the solvent was, 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.

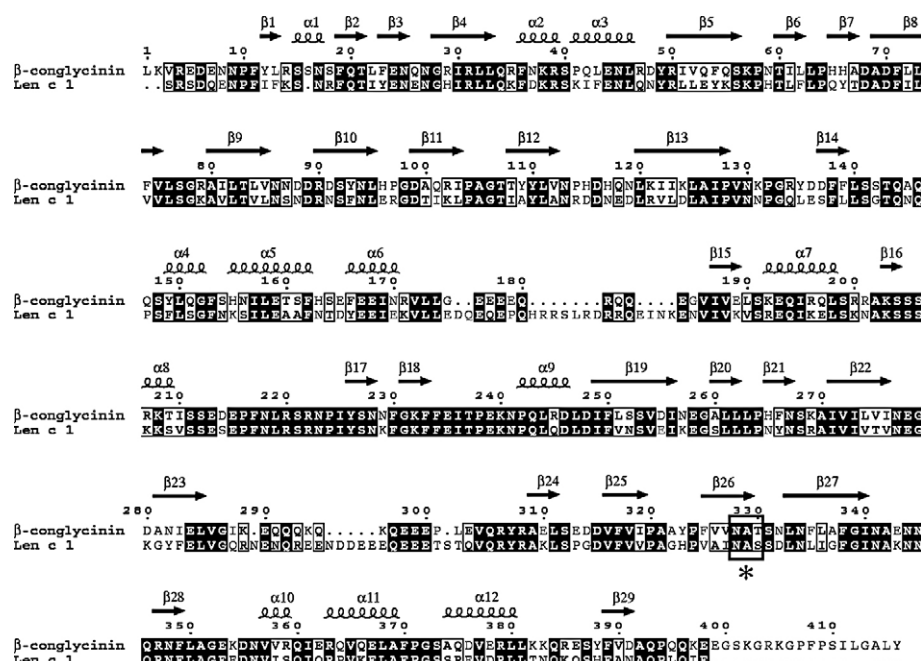


Fig. 2. Structural alignment of Len c 1 to β -conglycinin of soybean seeds. Identical residues are in black boxes and homologous residues are in open boxes. Strands of β -sheet (β 1– β 29) and stretches of α -helix (α 1– α 12) occurring along the polypeptide chain of β -conglycinin have been indicated. The putative *N*-glycosylation sites are open boxed. The alignment was performed with CLUSTAL-X [20], manually modified according to the HCA profiles [22], and represented with ESPript [21].

Epitopes susceptible to interact with T-cells have been predicted along the amino acid sequences of allergens using the MHC class-II binding peptide prediction server (ProPred) (<http://www.imtech.res.in/raghava/propred/index.html>) [36], that predict the MHC-binding oligopeptides derived from an amino acid sequence. Usually, a threshold ≤ 3 was applied for the T-cell epitopes prediction. According to their frequency with respect to the checked HLA alleles, they were considered as major or minor T-epitopes. Sequential B-cell epitopes were predicted from the hydropathic profiles as being the most hydrophilic, flexible and surface exposed regions. Different scales of hydrophilicity [37], flexibility [38], exposition to the solvent [39] and antigenicity [40], were used to build the hydropathic 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 1, Len c 1 and Pis s 1 was calculated and displayed on the molecular surface of the proteins with GRASP [34]. The overall conformation of the sequential B-cell epitopes on the molecular surface of the vicilin proteins was displayed with PyMOL of W.L. DeLano (<http://www.pymol.org>).

3. Results and Discussion

The modelled vicilin monomers Ara h 1, Len c 1 and Pis s 1 all consist of a cupin motif made of two tandemly arrayed modules related by a pseudo-dyad axis (Fig. 3A). Each module consists of a β -barrel core domain built from two walls of antiparallel α -sheet associated to a loop domain which predominantly contains α -helices. An α -helix-containing linker region interconnects the two modules. However, Ara h 1 possesses an extensive N-terminal region of 136 amino acid residues that has no equivalent in the β -conglycinin used as a template protein and could not therefore be modelled. According to the high structural homology they share with other seed proteins from the cupin superfamily, e.g., phaseolin from kidney bean [24], canavalin from Jack bean [25], germin from barley [41] or proglycinin A1aB1b from soybean [42], the three cupin motifs of vicilins were putatively arranged in a homotrimeric structure around a 3-fold symmetry axis in such a way that the cupin motifs occupies the three faces of a flattened prism (Fig. 4A). The additional loop region occurring along the polypeptide chain of the vicilin monomers (which has been omitted during the model building procedure of the vicilin monomers since it partly coincides with a region lacking in the three-dimensional structure of the β -conglycinin template) should extend over the same face (back face) of the homotrimer. The N-terminal parts of the Ara h 1 monomers are similarly supposed to form exposed regions protruding out of the plane formed by the homotrimer.

All these vicilin proteins are rather acidic proteins with a net charge of -7 (Ara h 1) or -8 (β -conglycinin, Len c 1, Pis s 1) and acidic *pI* ranging from 5.6 (Pis s 1) up to 6.6 (Ara h

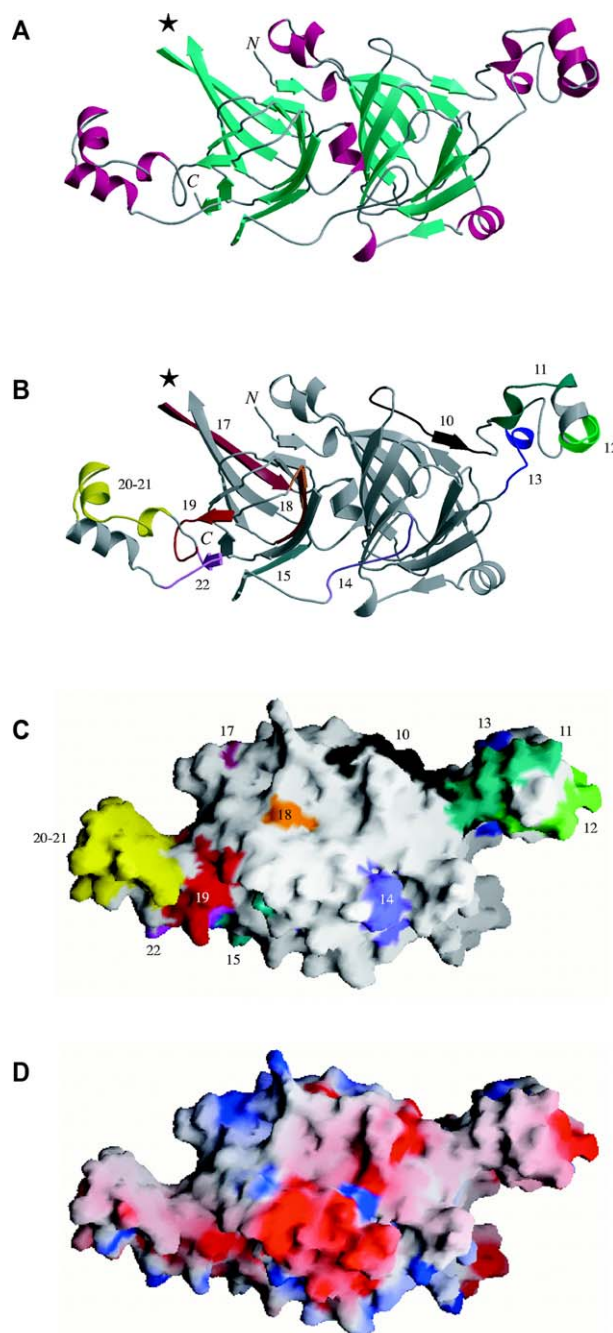


Fig. 3. (A) Ribbon diagram of the three-dimensional model of Ara h 1 monomer. Strands of β -sheet and stretches of α -helix are coloured green and red, respectively. Coil structures and loops are coloured light grey. (B) Location of the characterized sequential B-cell epitopes (numbered 10–22) on the ribbon diagram of Ara h 1. (C) Area occupied by the sequential B-cell epitopes (numbered 10–22) on the molecular surface of Ara h 1. (D) Distribution of the electrostatic potentials on the molecular surface of Ara h 1. The negative potential is coloured red and displayed at -5 *kT* level, the positive potential is coloured blue and displayed at $+5$ *kT* level (1 *kT* = 0.6 kcal). Neutral surfaces are white. The electrostatic potentials were calculated and displayed with GRASP [32].

1). Accordingly, their molecular surface essentially exhibit electronegatively charged areas (Fig. 3D), even though a few

electropositive areas also occur in the best exposed regions of the monomer.

T-cell epitopes were predicted to occur along the amino acid sequence of the Ara h 1 monomer when checked against the 51 HLA alleles of the ProPred database [36]. Three of them, 428MVIVVVNKG TGN439, 168IRVLQR-FDQ176 and 181FQNLQNHRI VQI192 may be considered as major epitopes since using a threshold of 2 they are detected against 46, 38 and 31 over 51 HLA alleles, respectively. Two other T-cell epitopes 239LRIPSGFIS247 and 557FVSAR-PQSQ565 are only detected against 15 and 17 HLA alleles, respectively. All these putative T-cell epitopes contain basic

residues. They correspond to conserved regions located in the β -barrel core domains of the cupin motif and, accordingly, are predicted at homologous positions in the amino acid sequences of Len c 1, Pis s 1 and Vic s 1 allergens.

Mapping of the previously identified sequential B-cell epitopes of Ara h 1 [14] on the three-dimensional models of the monomer (Fig. 3B) and the homotrimer (Fig. 4B) shows that most of the potential IgE-binding sites of Ara h 1 overlap the C-terminal α -helical regions of the cupin module. The sequential B-cell epitopes display more or less extended areas on the molecular surface of both the monomer (Fig. 3C) and the homotrimer (Fig. 4C,D). Some of them contain basic resi-

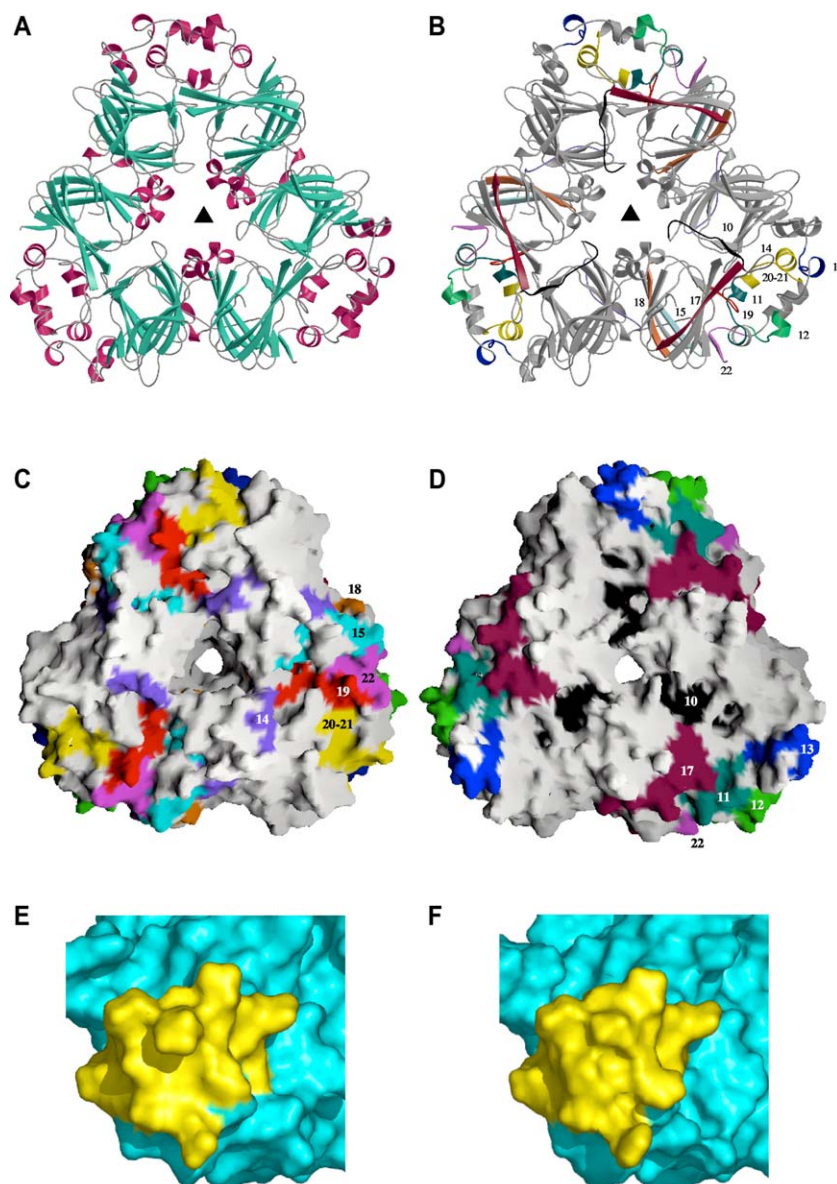


Fig. 4. (A) Ribbon diagram showing the back face of the Ara h 1 homotrimer. The three cupin monomers are similarly coloured as in Fig. 3A. The triangle ● indicates the 3-fold symmetry axis of the homotrimer. (B) Location on the back face of the Ara h 1 homotrimer of the sequential B-cell epitopes (numbered 10–22) previously identified on the C-terminus of the Ara h 1 monomer. The two overlapping 20 and 21 B-cell epitope have been merged and are coloured yellow. (C) Molecular surface of the front face of the Ara h 1 homotrimer showing the surface occupied by the sequential B-cell epitopes (numbered 10–22). (D) Molecular surface of the back face of the Ara h 1 homotrimer showing the exposition of the sequential B-cell epitopes (numbered 10–22). Overall three-dimensional conformation of B-cell epitope 20–21 (yellow) on the molecular surface (marine blue) of Ara h 1 (E) and Len c 1 (F). Both epitopes are similarly oriented and the molecular surfaces were calculated and displayed with PyMOL.

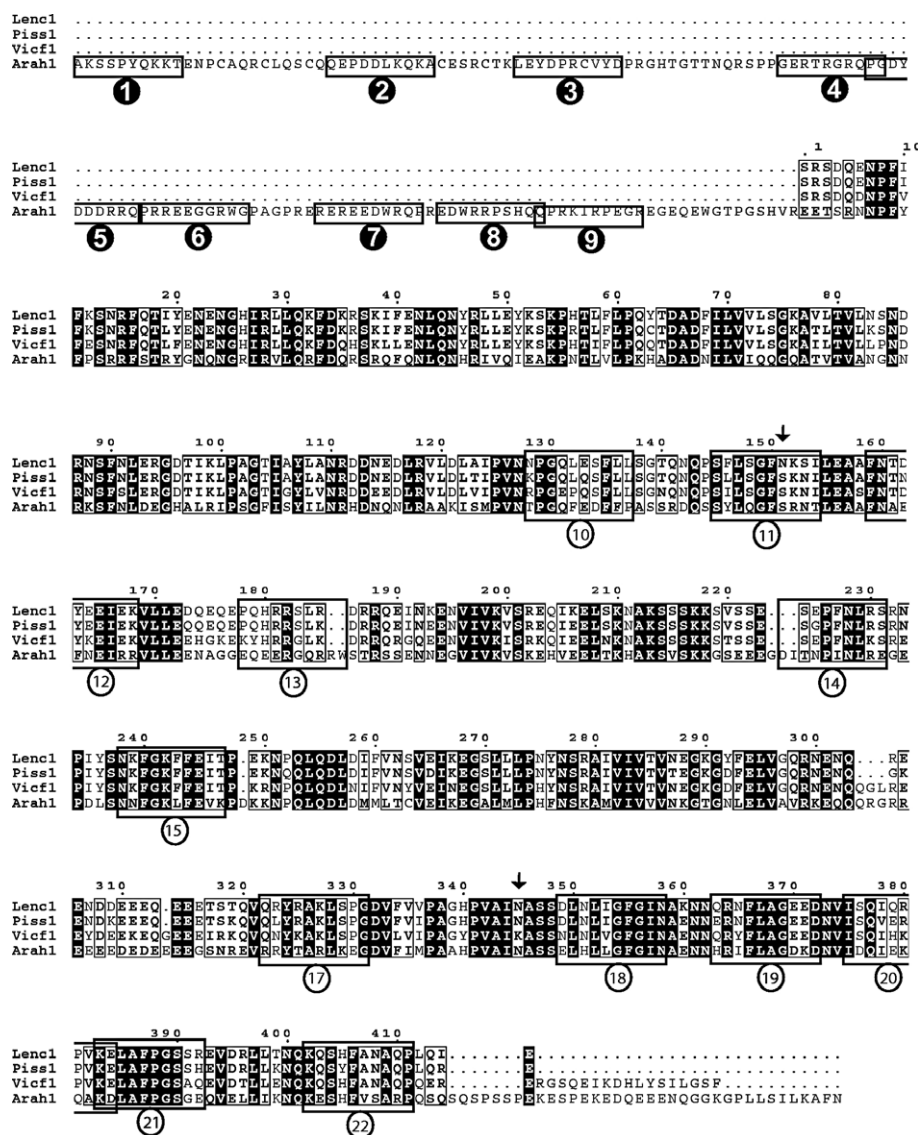


Fig. 5. Multiple structural alignment of Ara h 1, Len c 1, Pis s 1 and Vic f 1. Open boxes numbered 1–9 delineate the sequential B-cell epitopes specific to the N-terminal region of Ara h 1. Open boxes numbered 10–22 delineate the B-cell epitopes of the C-terminal region of Ara h 1 conserved in other vicilins Len c 1, Pis s 1 and Vic f 1. The vertical black arrows indicate the putative *N*-glycosylation sites. The alignment was performed with CLUSTAL-X [20], manually modified according to the HCA profiles [22], and represented with ESPrnt [21].

dues, e.g., epitopes 13, 17 and 20–22, and coincide with the exposed electropositive areas identified on the molecular surface (see Fig. 3D). Even though it is probably well exposed on the surface, the extended loop region that was omitted in the model building procedure since it is X-ray unsolved in the β -conglycinin used as a template, does not contain any B-cell epitope. However, the most exposed IgE-binding regions correspond to epitopes 11–13 and 18–22 of the C-terminal α -helical regions of the cupin module located on the front (epitopes 18–22) and back (epitopes 11–13) faces of the vertices of the prism-shaped homotrimers (Fig. 4C,D). Interestingly, their coalescence on the molecular surface of the cupin domains could create more complex conformational B-cell epitopes susceptible to enhance the specificity of the recognition of the legume vicilins by the corresponding IgE antibodies.

Legume vicilins contain one (Ara h 1, Pis s 1) or two (Len c 1) putative *N*-glycosylation sites (Fig. 4). Except for the *N*-glycosylation site Asn151-Lys-Ser of Len c 1 which belongs to an α -helix region and is not probably glycosylated, the other sites occur in a well-exposed loop and should be actually glycosylated. However, according to its location just before the B-cell epitope 18, the *N*-glycan chain has apparently no masking effect on the other surrounding epitopes susceptible to prevent their recognition by the corresponding IgE antibodies.

Obviously, nine (1–9) out of the 23 sequential B-cell epitopes identified in the N-terminal part of the Ara h 1 monomer, which has no equivalent in other legume vicilins, are specific for the peanut vicilin. The remaining B-cell epitopes characterized in the C-terminal region of Ara h 1 and, especially the best exposed 11–13 and 18–22 epitopes, share a

high degree of homology and three-dimensional conformation with other legume vicilins (Len c 1, Pis s 1) and most probably correspond to B-cell epitopes common to these seed storage protein family (Fig. 5). In this respect, most of these homologous regions were predicted as being exposed B-epitopes along the amino acid sequences of Len c 1 and Pis s 1 using the classical hydrophilicity/exposure/flexibility-based predictive methods. They also appear to be nicely conserved in the soybean β -conglycinin [26] and the vicillin allergen Vic f 1 of the broad bean [43]. However, this epitopic community apparently occurs in other legume seed storage proteins since a conserved IgE-binding epitope has been recently characterized in soybean (G1a and G2a) and peanut (Ara h 3) glycinins [44]. The most pronounced sequence similarity occurs in sequential B-cell epitopes located at the very C-terminal end of the second cupin module (Fig. 5), that are among the best exposed at the vertices of the vicilin homotrimers. Besides this sequence similarity, these B-cell epitopes also share a very similar overall conformation suggesting they could be recognised by the same IgE antibodies (Fig. 3E,F). These closely structurally related B-cell epitopes most probably account for the IgE-binding cross-reactivity towards different seed storage vicilins and readily participate in the allergic response of sensitised individuals to multiple legume seeds. In addition, the remarkable resistance of the homotrimeric arrangement of the cupin monomers to both thermal denaturation and proteolysis [10] preserves the most exposed IgE-binding sites [45] and thus greatly enhance their allergenic propensity.

Our knowledge on the B-cell epitopic communities of seed storage allergens have three main clinical implications: (1) they have to be carefully considered in the diagnostic of food allergy and, especially, in the identification of the sensitising and triggering allergen(s) using skin prick tests or in vivo IgE-binding tests; (2) they can help the clinicians to decide on the opportunity to prevent the consumption of the suspected allergen-containing foods by the susceptible individuals to avoid any anaphylactic shock and (3) the incriminated allergens should be characterized as hidden allergens and properly labelled in the food products to prevent their consumption by sensitised individuals and especially, children or adolescents.

Acknowledgements

The financial support of CNRS is gratefully acknowledged.

References

- [1] W. Burks, Peanut allergy: a growing phenomenon, *J. Clin. Invest.* 111 (2003) 950–952.
- [2] H.A. Sampson, Update on food allergy, *J. Allergy Clin. Immunol.* 113 (2004) 805–819.
- [3] T.T. Perry, M.K. Conover-Walker, A. Pomés, M.D. Chapman, R.A. Wood, Distribution of peanut allergen in the environment, *J. Allergy Clin. Immunol.* 113 (2004) 973–976.
- [4] S.H. Sicherer, A. Muñoz-Furlong, A.W. Burks, H.A. Sampson, Prevalence of peanut and tree nut allergy in the United States of America determined by a random digit dial telephone survey: a 5-year follow-up study, *J. Allergy Clin. Immunol.* 103 (1999) 559–562.
- [5] H.A. Sampson, L.M. Mendelson, J.P. Rosen, Fatal and near-fatal anaphylactic reactions to foods in children and adolescents, *N. Engl. J. Med.* 327 (1992) 380–384.
- [6] A.W. Burks, G. Cockrell, J.S. Stanley, R.M. Helm, G.A. Bannon, Isolation, identification, and characterization of clones encoding antigens responsible for peanut hypersensitivity, *Int. Arch. Allergy Immunol.* 107 (1995) 248–250.
- [7] O.M. Viquez, K.N. Konan, H.W. Dodo, Structure and organization of the genomic clone of a major peanut allergen gene, Ara h 1, *Mol. Immunol.* 40 (2003) 565–571.
- [8] O.M. Viquez, C.G. Summer, H.W. Dodo, Isolation and molecular characterization of the first genomic clone of a major peanut allergen, Ara h 2, *J. Allergy Clin. Immunol.* 107 (2001) 713–717.
- [9] H.W. Dodo, O.M. Viquez, S.J. Maleki, K.N. Konan, cDNA clone of a putative peanut (*Arachis hypogaea* L.) trypsin inhibitor has homology with peanut allergens Ara h 3 and Ara h 4, *J. Agric. Food Chem.* 52 (2004) 1404–1409.
- [10] E.N.C. Mills, J. Jenkins, N. Marigheto, P.S. Belton, A.P. Gunning, V.J. Morris, Allergens of the cupin superfamily, *Biochem. Soc. Trans.* 30 (2002) 925–929.
- [11] H. Breiteneder, C. Radauer, A classification of plant food allergens, *J. Allergy Clin. Immunol.* 113 (2004) 821–830.
- [12] J.M. Dunwell, A. Purvis, S. Khuri, Cupins: the most functionally diverse protein superfamily? *Phytochemistry* 65 (2004) 7–17.
- [13] A.W. Burks, D. Shin, G. Cockrell, J.S. Stanley, R.M. Helm, G.A. Bannon, Mapping and mutational analysis of the IgE-binding epitopes on Ara h 1, a legume vicilin protein and a major allergen in peanut hypersensitivity, *Eur. J. Biochem.* 245 (1997) 334–339.
- [14] D.S. Shin, C.M. Compadre, S.J. Maleki, R.A. Kopper, H. Sampson, S.K. Huang, A.W. Burks, G.A. Bannon, Biochemical and structural analysis of the IgE binding sites on Ara h 1, an abundant and highly allergenic peanut protein, *J. Biol. Chem.* 273 (1998) 13753–13759.
- [15] G. López-Torrejón, G. Salcedo, M. Martí-Esteban, A. Díaz-Perales, C.Y. Pascual, R. Sánchez-Monge, Len c 1, a major allergen and vicilin from lentil seeds: protein isolation and cDNA cloning, *J. Allergy Clin. Immunol.* 112 (2003) 1208–1215.
- [16] M. Wensing, A.C. Knulst, S. Piersma, F. O’Kane, E.F. Knol, S.J. Koppelman, Patients with anaphylaxis to pea can have peanut allergy caused by cross-reactive IgE to cecilin (Ara h 1), *J. Allergy Clin. Immunol.* 111 (2003) 420–424.
- [17] A.W. Burks, J.R. Brooks, H.A. Sampson, Allergenicity of major component proteins of soybean determined by enzyme-linked assay (ELISA) and immunoblotting in children with atopic dermatitis and positive soy challenges, *J. Allergy Clin. Immunol.* 81 (1988) 1135–1142.
- [18] S.S. Teuber, K.C. Jarvis, W.R. Peterson, A.M. Dandekar, A.A. Ansari, Identification and cloning of a complementary DNA encoding a vicilin-like proprotein, Jug r 2, from english walnut kernel (*Juglans regia*), a major food allergen, *J. Allergy Clin. Immunol.* 104 (1999) 1311–1320.
- [19] S.H. Sicherer, Clinical implications of cross-reactive food allergens, *J. Allergy Clin. Immunol.* 108 (2001) 881–890.
- [20] J.D. Thompson, T.J. Gibson, F. Plewniak, F. Jeanmougin, D.G. Higgins, The CLUSTAL-X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tool, *Nucleic Acids Res.* 15 (1997) 4876–4882.

- [21] P. Gouet, E. Courcelle, D.I. Stuart, F. Metoz, ESPript: analysis of multiple sequence alignments in PostScript, *Bioinformatics* 15 (1999) 305–308.
- [22] C. Gaboriaud, V. Bissery, T. Benchetrit, J.P. Mornon, Hydrophobic cluster analysis: an efficient new way to compare and analyse amino acid sequences, *FEBS Lett.* 224 (1987) 149–155.
- [23] A.W. Burks, G. Cockrell, J.S. Stanley, R.M. Helm, G.A. Bannon, Recombinant peanut allergen Ara h I expression and IgE binding in patients with peanut hypersensitivity, *J. Clin. Invest.* 96 (1995) 1715–1721.
- [24] M.C. Lawrence, T. Izard, M. Beuchat, R.J. Blagrove, P.M. Colman, Structure of phaseolin at 2.2 Å resolution. Implications for a common vicilin/legumin structure and the genetic engineering of seed storage proteins, *J. Mol. Biol.* 238 (1994) 748–776.
- [25] T.-P. Ko, J. Day, A. McPherson, The refined structure of canavalin from jack bean in two crystal forms at 2.1 and 2.0 Å resolution, *Acta Crystallogr. D* 56 (2000) 411–420.
- [26] N. Maruyama, M. Adachi, K. Takahashi, K. Yagasaki, M. Kohno, Y. Takenaka, E. Okuda, S. Nakagawa, B. Mikami, S. Utsumi, Crystal structures of recombinant and native soybean beta-conglycin homotrimers, *Eur. J. Biochem.* 268 (2001) 3595–3604.
- [27] J.W. Ponder, F.M. Richards, Tertiary templates for proteins. Use of packing criteria in the enumeration of allowed sequences for different structural classes, *J. Mol. Biol.* 193 (1987) 775–791.
- [28] M.T. Mas, K.C. Smith, D.L. Yarmush, K. Aisaka, R.M. Fine, Modeling the anti-CEA antibody combining site by homology and conformational search, *Proteins Struct. Func. Genet.* 14 (1992) 483–498.
- [29] A. Roussel, C. Cambillau, TURBO-FRODO, in: S.G. Committee (Ed.), *Silicon Graphics Geometry Partners Directory*, Silicon Graphics, Mountain View, California, 1989.
- [30] R.A. Laskowski, M.W. MacArthur, D.S. Moss, J.M. Thornton, PROCHECK: a program to check the stereochemistry of protein structures, *J. Appl. Crystallogr.* 26 (1993) 283–291.
- [31] P.J. Kraulis, Molscript: a program to produce both detailed and schematic plots of protein structures, *J. Appl. Crystallogr.* 24 (1991) 946–950.
- [32] R.M. Esnouf, An extensively modified version of Molscript that includes greatly enhanced coloring capabilities, *J. Mol. Graph.* 15 (1997) 132–134.
- [33] E.A. Merritt, D.J. Bacon, Raster3D photorealistic molecular graphics, *Methods Enzymol.* 277 (1997) 505–524.
- [34] A. Nicholls, K.A. Sharp, B. Honig, Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons, *Proteins Struct. Func. Genet.* 11 (1991) 281–296.
- [35] M.K. Gilson, B.H. Honig, Calculation of electrostatic potential in an enzyme active site, *Nature* 330 (1987) 84–86.
- [36] H. Singh, J.P.S. Raghava, ProPred: prediction of HLA-DR binding sites, *Bioinformatics* 17 (2001) 1236–1237.
- [37] J. Kyte, R.F. Doolittle, A simple method for displaying the hydrophobic character of a protein, *J. Mol. Biol.* 157 (1982) 105–132.
- [38] P.A. Karplus, G.E. Schulz, Prediction of chain flexibility in proteins, *Naturwissenschaften* 72 (1985) 212–213.
- [39] B. Lee, F.M. Richards, The interpretation of protein structures: estimation of static accessibility, *J. Mol. Biol.* 55 (1971) 379–400.
- [40] T.P. Hopp, K.R. Woods, Prediction of protein antigenic determinants from amino acid sequences, *Proc. Natl. Acad. Sci. USA* 78 (1981) 3824–3828.
- [41] E.J. Woo, J.M. Dunwell, P.W. Goodenough, A.C. Marvier, R.W. Pickersgill, Germin is a manganese containing homohexamer with oxalate oxidase and superoxide dismutase activities, *Nat. Struct. Biol.* 7 (2000) 1036–1040.
- [42] M. Adachi, Y. Takenaka, A.B. Gidamis, B. Mikami, S. Utsumi, Crystal structure of soybean proglycinin A1aB1b homotrimer, *J. Mol. Biol.* 305 (2001) 291–305.
- [43] R. Bassuner, N.V. Hai, R. Jung, G. Saalbach, K. Muntz, The primary structure of the predominating vicilin storage protein subunit from field bean seeds (*Vicia faba* L. var. minor cv. Fribo), *Nucleic Acids Res.* 15 (1987) 9609.
- [44] P. Xiang, T.A. Beardslee, M.G. Zeece, J. Markwell, G. Sarath, Identification and analysis of a conserved immunoglobulin E-binding epitope in soybean G1a and G2a and peanut Ara h 3 glycinins, *Arch. Biochem. Biophys.* 408 (2002) 51–57.
- [45] S.J. Maleki, R.A. Kopper, D.S. Shin, C.W. Park, C.M. Compadre, H. Sampson, A.W. Burks, G.A. Bannon, Structure of the major peanut allergen Ara h 1 may protect IgE-binding epitopes from degradation, *J. Immunol.* 164 (2000) 5844–5849.