

Identification of Potentially Cross-Reactive Peanut-Lupine Proteins by Computer-Assisted Search for Amino Acid Sequence Homology

Fabrizio Guarneri^a Claudio Guarneri^a Salvatore Benvenga^{b, c}

^aIstituto di Dermatologia, Policlinico Gaetano Martino, and ^bSezione di Endocrinologia del Dipartimento Clinico Sperimentale di Medicina e Farmacologia, Università di Messina, ^cProgramma Infradipartimentale di Endocrinologia Molecolare Clinica, Azienda Ospedaliera Universitaria Policlinico Gaetano Martino, Messina, Italia

Key Words

Cross-allergy · Food allergy · Peanut · Lupine · Amino acid sequence homology

Abstract

Background: Allergic cross-reactions are an issue of major concern because of implications for public health. The molecular basis of cross-allergy is the similarity of epitopes belonging to proteins of different organisms. Lupine is an emerging cause of food allergy, which has become a 'hot topic' because of recent large-scale introduction into processed foods and frequent cross-reactions with other members of the legume family. However, no lupine allergen has been characterized thus far. Prompted by a recently reported case of peanut-lupine cross-allergy, we wished to identify the possible cross-reactive allergen(s) between the two vegetal species. **Methods:** We used computer-aided amino acid sequence comparison, a well-established technique for the study of protein homology, and followed the FAO/WHO guidelines for the identification of potential allergens. We also performed a three-dimensional modeling of the suspected cross-reactive proteins to compare their molecular surfaces. **Results:** We found a highly significant sequence homology and molecular similarity between allergen Ara h 8 of peanut and pathogenesis-related protein PR-

10 of white lupine. Another protein of lupine, the β -conglutin precursor, was found to be significantly homologous to the Ara h 1 allergen of peanut. The molecular surfaces of Ara h 8 and PR-10 were remarkably similar. **Conclusions:** Our in silico data allow to predict the allergenicity of PR-10 and β -conglutin precursor of white lupine according to FAO/WHO guidelines. Amino acid sequence homology also suggests that these proteins could be responsible, at least in part, for some of the allergic cross-reactions between peanut and lupine reported in the literature.

Copyright © 2005 S. Karger AG, Basel

Introduction

Lupine has recently been widely introduced into foods as a component of beverages, bread, cookies and pasta as an additive to wheat flour and is a 'new' food allergen [1–3]. Concern for public health exists not only because lupine is often a 'hidden' component of processed foods, but also because it has multiple cross-reactivities. The most important cross-reaction of white lupine is with peanut [2, 3], one of the major food allergens. No lupine allergen has been fully characterized until now. Recently, Fæste et al. [1] reported a case of peanut-lupine cross-allergy and tried to elucidate its molecular basis by immu-

Ara h 1	96	QPGDYDDRRQPRREEGGRWGPAGPRE-REEREEDWRQPREDRRPSHQQ-----PRKIRPEGR	152
		+P + + + QPRR+ P RE RE+E++ P R+ +++ R + E R	
LaβCP	40	RPEEREQE EWQPRRQR-----PQSRREEREQE QE QGSPSYPRRQSGYERRQYHERSEQREER	96
Ara h 1	153	EGEQEWGTPGSHVREETSRRNPFYFSPRRFSTRYGNQNGRIRVLQRFQDQSRQFQNLQNHRI	214
		E EQ+ G+P R+ NP++F S+RF T Y N+NG+IRVL+RFDQR+ + +NLQN+RI	
LaβCP	97	EQEQQQGSPSYRRQR-----NPYHFSSQRFQTLYKNRNGKIRVLERFDQRTNRLENLQNYRI	154
Ara h 1	215	VQIEAKPNTLVLPKHADADNILVIQQGQATVTVANGNNRKSFNLDGEGHALRIPSGFISYILN	276
		V+ ++KPNTL+LPKH+DAD +LV+ G+AT+T+ N + R+++NL+ G ALRIP+G SYILN	
LaβCP	155	VEFQSKPNTLILPKHSDADYVLVLNNGRATITIVNPDRRQAYNLEYGDALRIPAGSTSYILN	216
Ara h 1	277	RHDNQNLRVAKISMVNTPGQFEDFFPASSRDQSSYLQGFSRNTLEAAFNAAFNEIRRVLL-	337
		DNQ LRV K+++P+N PG F DF+P+S++DQ SY GFSRNTLEA FN + EI+R++L	
LaβCP	217	PDDNQKL RVV KLAIPINNP GFYDFYPSSTKDQQS YFGFSRNTLEATFNTRYEEIQRIILG	278
Ara h 1	338	-EENAGGEQEERGQRRWSTRSSENNEGVIVKVSKEHVEELTKHAKSVSKKGSEEEGDITNPI	398
		E+ E++ RGQ + S+ +EGVIV VSK+ +++LTKHA+S S K + + P	
LaβCP	279	NEDEQEYEEQRRGQE-----SDQDEGVIVIVSKKQIQKLTAKHAQSSSGDKDKPSD---SGPF	332
Ara h 1	399	NLREGEPDLSNNFGKLFVVKPDKNPQLQDLDMMLTCVEIKEGALMLPHFNSKAMVIVVVNK	460
		NLR EP SN +G +E+ PD +NPQ+QDL++ LT ++I EGAL+LPH+NSKA+ +VVV++	
LaβCP	333	NLRSNEPIYSNKYGNFYEITPD-RNPQVQDLNISLTYIKINEGALLPHYNSKAIYVVVVDE	393
Ara h 1	461	GTGNLELVAVRKEQQQRGRREEEEDDEEEEGSNREVRRYTARLKEGDVFI MPAAHPVAINA	522
		G GN ELV +R +Q+Q+ +EE+E+ EV RY+ARL EGD+F++PA +P++INA	
LaβCP	394	GEGNYELVGIRDQQRQDQE QEEKE-----EVIRYSARLSEGDI FVIPAGYPISINA	445
Ara h 1	523	SSELHLLGFGINAENNHRIFLAGDKDNVIDQIEKQAKDLAFPGSGEQVEKLIKQKESHFVS	584
		SS L LLGFGINA+ N R FLAG KDNVI Q+++ +L FPGS E +E+LIKQ++S+F +	
LaβCP	446	SSNLRLLGFGINADENQRNFLAGSKDNVIRQLDRAVNELTFPGSAEDIERLIKQQQSYFAN	507
Ara h 1	585	ARPQSQSQSPSSPE	598
		+PQ Q Q S E	
LaβCP	508	GQPQQQQQQQSEKE	521

Identity 48%; similarity (identical + chemically similar amino acids) 69%; E = 10⁻¹³⁸

Fig. 1. Comparative amino acid sequence alignments of the homologous segments of allergen Ara h 1 of *A. hypogaea* and β-conglutinin precursor of *L. albus* (LaβCP). The two cupin domains are at positions 183–309 and 398–548 of Ara h 1, and 126–249 and 332–494 of LaβCP. The symbol ‘+’ indicates a conservative amino acid replacement.

nochemical methods. They found that two lupine proteins of 38 and 15 kDa were recognized by anti-peanut IgE antibodies present in the patient’s serum. It is well known that cross-allergy between organisms of different species is due to structurally homologous proteins, which are recognized by the same IgE antibody. The analysis of homologies between proteins by means of computer-assisted amino acid sequence comparison is a long-established technique, which already produced significant results in several immunity-related disorders [4, 5]. We used this technique to detect homologies between known peanut allergens and lupine proteins, so as to at least narrow the spectrum of the possible proteins responsible for the cross-allergy reported by Fæste et al. [1].

Materials and Methods

We probed the database of the Allergen Nomenclature Subcommittee of the International Union of Immunological Societies (<http://www.allergen.org>) for known allergens of peanut (*Arachis hypogaea*). In a search performed on February 2, 2005, we then extracted the amino acid sequences of the 8 peanut allergens (Ara h 1–8) and of the 87 sequenced proteins of white lupine (*Lupinus albus*) from the Entrez Protein database of the National Center for Biotechnology Information (NCBI). Next, we compared each peanut allergen with all lupine proteins, using NCBI BLAST (Basic Local Alignment Search Tool) software [6], which was set to search also in the CDD (Conserved Domain Database) [7]. BLAST performs a residue-by-residue comparison between two proteins, giving each couple of residues a homology score. This score is used to calculate the E parameter, which describes the number of hits that can be ‘expected’ to be found by chance when searching a database of a particular size. Although the E value is significant when <1, we

Ara h 3	27	GGYIETWNPNNQEFECAGVALSRLVLRNALRRPFY-SNAPQEIFIQQGRGYFGLIFPGCPR	87
		G + E N + + ++L+ + + AL P Y S A + + +G G + L+	
LaβCP	346	GNFYEITPDRNPQVQDLNISLTYIKINEGALLLPHYNSKAIYVVVVDEGEGNYELV-----	401
Ara h 3	88	HYEEPHTQGRRSQSQRPPRRLQGEDQSQQQRDSHQKVHRFDEGDLIAVPTGVAFWLYNDHDT	149
		G R Q + Q ++Q +++ + + R EGD+ +P G + +	
LaβCP	402	-----GIRDQQR-----QQDEQEEKEEEVIRYSARLSEGDIFVIPAGYPISINASSNL	449
Ara h 3	150	DVVAVSLTDTNNDN-----QLDQFPRRFNLAGNTE--QEFLRYQQQSRQSRRRSL	198
		++ + N N QLD+ G+ E + ++ QQQS	
LaβCP	450	RLLGFGINADENQRNFLAGSKDNVIRQLDRAVNELTFPGSAEDIERLIKNQQS-----	503
Ara h 3	199	PYSPYSPQSQPRQEEREFSR-GQHSRR	225
		Y QP+Q++++ S + G+ RR	
LaβCP	504	----YFANGQPQQQQQQSEKEGRRGR	527
Identity 20%; similarity (identical + chemically similar amino acids) 38%; E = 2 × 10 ⁻⁶			

Fig. 2. Comparative amino acid sequence alignments of the homologous segments of allergen Ara h 3 of *A. hypogaea* and β-conglutinin precursor of *L. albus* (LaβCP). The two cupin domains are at positions 14–181 and 340–484 of Ara h 3, and 126–249 and 332–494 of LaβCP. The symbol ‘+’ indicates a conservative amino acid replacement.

Ara h 8	1	MGVFTFEDEITSTVPPAKLYNA-MKDADSITPKII-DDVKSVEIVEGNGGPGTIKKLTIVE	59
Bet v 1	1	MGVFNYEDEATSVIAPARLFKSFVLAD N LIPKVAPENVSSAENI E GNGGPGTIKKITFPE	61
LaPR-10	1	MGIFTFEDESTSTVAPARLYKALVKDADTIIPK-AVEAIQSVETVEGNGGPGTIKKLTIE	60
Ara h 8	60	DGETKFILHKVESIDEANYAYNYSVVGVALPPTAEKITFETKLVEGPNGGSIGKLTLYH	120
Bet v 1	62	GSHFKYMKHRVDEIDHANFKYCYSIEGGPLGDTLEKISYEIKIVAAPGGG S ILKITSKYH	122
LaPR-10	61	GGETKYVLHKIEEIDEANLGYNYSIVGGVGLPDTVEKITFETKLVEGVNGGSIGKVITIKIE	121
Ara h 8	121	TKGDAKPDEEELKKGKAKGEGLFRAIEGYVLANPTQY	157
Bet v 1	123	TKGDISLNEEEEIKAGKEKGAGLFAVENYLV AHPNAYN	160
LaPR-10	122	TKGDAKPNEEEGKA AKVRGDAFFKAIENYLSAHPY-YN	158
Ara h 8 versus Bet v 1: Identity 52%; similarity (identical + chemically similar amino acids) 71%; E = 3 × 10 ⁻⁴²			
Ara h 8 versus PR-10 of <i>Lupinus albus</i> : Identity 57%; similarity (identical + chemically similar amino acids) 71%; E = 5 × 10 ⁻⁴⁵			
Bet v 1 versus PR-10 of <i>Lupinus albus</i> : Identity 69%; similarity (identical + chemically similar amino acids) 83%; E = 7 × 10 ⁻⁶²			

Fig. 3. Comparative amino acid sequence alignments of allergen Ara h 8 of *A. hypogaea*, major birch allergen Bet v 1 of *B. pendula* and pathogenesis-related protein 10 of *L. albus* (LaPR10). Amino acids of Bet v 1 critical for IgE binding [see ref. 13–15] are highlighted by bold-face letters.

elected to be more stringent. Accordingly, we considered matches that were highly significant, upon setting the E level at <0.001. Next, we evaluated if the selected proteins could be considered as allergens according to the FAO/WHO guidelines [8], that is an identity greater than 35% on a segment of 80 residues or six consecutive residues conserved in both the known and the potential allergen [8]. Finally, we used the software SWISS-MODEL [9] to obtain a three-dimensional model of the homologous proteins and the Ras-Mol program (version 2.6, which is freely available at <http://www.umass.edu/microbio/rasmol/getras.htm>) to spatially visualize the homologous molecules.

Results

Only Ara h 1 (vicilin), Ara h 3 (glycinin) and Ara h 8 (pathogenesis-related protein 10 or PR-10) showed significant homologies with lupine proteins; these last proteins have not been reported as allergens thus far. Ara h 1 (fig. 1) and Ara h 3 (fig. 2) are homologous to the β-conglutinin precursor of *L. albus* (LaβCP), and the homology concerning Ara h 1 exceeded the aforementioned

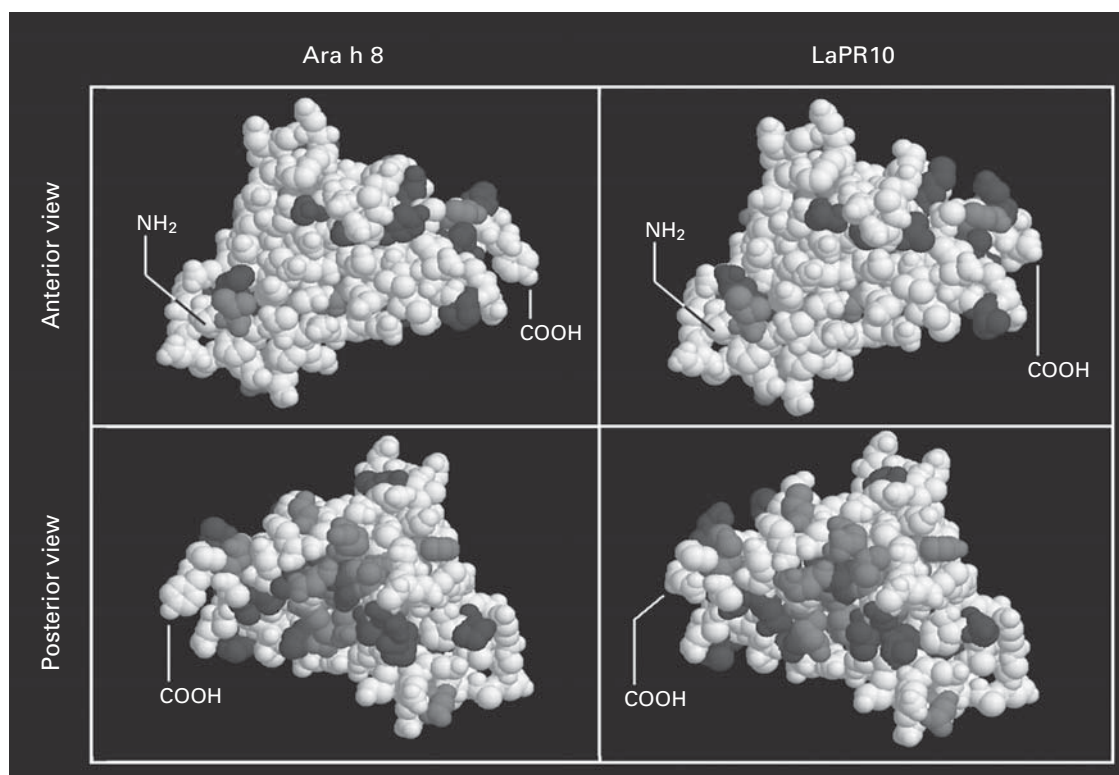


Fig. 4. Molecular surface homology between allergen Ara h 8 of *A. hypogaea* (left) and pathogenesis-related protein 10 of *L. albus* (right) from two opposite views. White, light gray and dark gray indicate identical, chemically similar or chemically dissimilar (nonhomologous) amino acids, respectively.

FAO/WHO criteria [6] (fig. 1). We were unable to obtain complete and reliable three-dimensional models of Ara h 1, Ara h 3 and La β CP using the SWISS-MODEL program [9]. However, all three proteins contain two cupin domains: aa 183–309 and 398–548 of Ara h 1, aa 14–181 and 340–484 of Ara h 3, aa 126–249 and 332–494 of La β CP. The term ‘cupin’ derives from the Latin word *cupa* (small barrel), because of the β -barrel secondary structure. Cupin domains are typical of a family of plant seed storage proteins. Noteworthy, several members of the 7 S vicilin-like or the 11 S legumin-like seed storage globulin families contain cupin domains and are allergens [10]. Ara h 8 is 69% identical and 83% similar to PR-10 of *L. albus* (LaPR10), with an E value of 7×10^{-62} (fig. 3); the homolog segment identified by BLAST spanned almost the entire length (>98%) of both proteins. These results were well beyond FAO/WHO requirements [6]. Also, Ara h 8 and LaPR10 have a long segment (aa 2–153 and 2–154, respectively) which is 100% aligned with Bet v 1, the major pollen allergen of *Betula pendula*. Homol-

ogies between Ara h 8, Bet v 1 and PR10 of *L. albus* are shown in figure 3. Three-dimensional structures of Ara h 8 and LaPR10 were very similar in the prediction made with SWISS-MODEL [9]. The remarkable molecular resemblance of their surfaces is shown in figure 4. Indeed, the spatial similarity between the two proteins was not unexpected, based on their high degree of amino acid sequence homology.

Discussion

Based on the FAO/WHO guidelines [6], our results point to a potential role of the β -conglutinin precursor as an allergen of *L. albus* that cross-reacts with Ara h 1. We also found that LaPR10, which bears homology to Ara h 8, meets the two FAO/WHO criteria for the identification of new allergens [6]. Ara h 8 and LaPR10 belong to the PR-10 family of proteins, which all share the amino acid sequence of Bet v 1. PR10 proteins are synthesized

in response to pathogenic environmental stimuli. As PR-10 proteins are highly conserved across different vegetal species, they frequently cause allergic cross-reactivity [11]. Cross-reactivity between Ara h 8 and Bet v 1 has been demonstrated experimentally [12], which agrees nicely with our in silico data, showing a highly significant sequence homology (fig. 3) between these two proteins. Three amino acids of Bet v 1 are reported as critical for IgE binding [13–15]: two are conserved in LaPR10 and the third is, like in Ara h 8, substituted by a chemically similar residue (fig. 3). Moreover, the molecular weight of LaPR10 is 16,949, close to the 15K of one of the cross-reactive lupine proteins identified by

Fæste et al. [1]. Taken together all data, we hypothesize that PR-10 and β -conglutin precursor of *L. albus* are new allergens that cross-react with Ara h 8 and Ara h 1 of *A. hypogaea*, respectively. Cross-reactivity with other proteins of the same family might also exist. Amino acid sequence comparison and search for conserved domains can facilitate allergological research. Experimental demonstration of IgE binding to candidate allergens remains the definitive proof of allergenicity and cross-reactivity. However, in silico analysis helps predicting potential allergens and/or cross-allergens and narrows down the number of candidate proteins to be tested for IgE binding.

References

- 1 Fæste CK, Løvik M, Wiker HG, Egaas E: A case of peanut cross-allergy to lupine flour in a hot dog bread. *Int Arch Allergy Immunol* 2004; 135:36–39.
- 2 Moneret-Vautrin DA, Guerin L, Kanny G, Flabbee J, Fremont S, Morisset M: Cross-allergenicity of peanut and lupine: The risk of lupine allergy in patients allergic to peanuts. *J Allergy Clin Immunol* 1999;104:883–888.
- 3 Hefle SL, Lemanske RF Jr, Bush RK: Adverse reaction to lupine-fortified pasta. *J Allergy Clin Immunol* 1994;94:167–172.
- 4 Persidis A: Autoimmune disease drug discovery. *Nat Biotechnol* 1999;17:1038–1039.
- 5 Benvenga S, Bartolone L, Trimarchi F: Thyroid iodide transporter: Local sequence homologies with thyroid autoantigens. *J Endocrinol Invest* 1997;20:508–512.
- 6 Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ: Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res* 1997;25:3389–3402.
- 7 Marchler-Bauer A, Anderson JB, DeWeese-Scott C, Fedorova ND, Geer LY, He S, Hurwitz DI, Jackson JD, Jacobs AR, Lanczycki CJ, Liebert CA, Liu C, Madej T, Marchler GH, Mazumder R, Nikolskaya AN, Panchenko AR, Rao BS, Shoemaker BA, Simonyan V, Song JS, Thiessen PA, Vasudevan S, Wang Y, Yamashita RA, Yin JJ, Bryant SH: CDD: A curated Entrez database of conserved domain alignments. *Nucleic Acids Res* 2003;31:383–387.
- 8 FAO/WHO: Allergenicity of genetically modified foods. Rome, FAO/WHO, 2001.
- 9 Schwede T, Kopp J, Guex N, Peitsch MC: SWISS-MODEL: An automated protein homology-modeling server. *Nucleic Acids Res* 2003;31:3381–3385.
- 10 Mills EN, Jenkins J, Marigheto N, Belton PS, Gunning AP, Morris VJ: Allergens of the cupin superfamily. *Biochem Soc Trans* 2002;30: 925–929.
- 11 Breiteneder H, Ebner C: Molecular and biochemical classification of plant-derived food allergens. *J Allergy Clin Immunol* 2000;106: 27–36.
- 12 Mittag D, Akkerdaas J, Ballmer-Weber BK, Vogel L, Wensing M, Becker WM, Koppelman SJ, Knulst AC, Helbling A, Hefle SL, Van Ree R, Vieths S: Ara h 8, a Bet v 1-homologous allergen from peanut, is a major allergen in patients with combined birch pollen and peanut allergy. *J Allergy Clin Immunol* 2004;114: 1410–1417.
- 13 Wiche R, Gubesch M, König H, Fotisch K, Hoffmann A, Wangorsch A, Scheurer S, Vieths S: Molecular basis of pollen-related food allergy: Identification of a second cross-reactive IgE epitope on Pru av 1, the major cherry allergen (*Prunus avium*). *Biochem J* 2005;385:319–327.
- 14 Spangfort MD, Mirza O, Ipsen H, Van Neerven RJ, Gajhede M, Larsen JN: Dominating IgE-binding epitope of Bet v 1, the major allergen of birch pollen, characterized by X-ray crystallography and site-directed mutagenesis. *J Immunol* 2003;171:3084–3090.
- 15 Son DY, Scheurer S, Hoffmann A, Hausteiner D, Vieths S: Pollen-related food allergy: Cloning and immunological analysis of isoforms and mutants of Mal d 1, the major apple allergen, and Bet v 1, the major birch pollen allergen. *Eur J Nutr* 1999;38:201–215.