

ORIGINAL PAPER

In silico identification of potential new latex allergens

F. Guarneri*, C. Guarneri*, B. Guarneri* and S. Benvenga^{†‡}

*Istituto di Dermatologia, Policlinico 'Gaetano Martino', Università di Messina, Messina, Italy, [†]Sezione di Endocrinologia del Dipartimento Clinico Sperimentale di Medicina e Farmacologia, Università di Messina, Messina, Italy and [‡]Programma Infradipartimentale di Endocrinologia Molecolare Clinica, Azienda Ospedaliera Universitaria Policlinico 'Gaetano Martino', Messina, Italy

Clinical and Experimental Allergy

Summary

Background Allergy to latex of *Hevea brasiliensis* is a frequent problem. In spite of the significant progress of research, the identity and cross-reactivity of some latex allergens are unknown.

Objective To identify, among the fully characterized latex proteins, those with a higher probability to be allergenic.

Methods We used *in silico* techniques (amino acid sequence comparison and molecular modelling) to identify potential new allergens among the known proteins of *H. brasiliensis*.

Results Cu/Zn superoxide dismutase, heat shock protein and calmodulin of *H. brasiliensis* show highly significant ($E < 10^{-9}$) amino acid sequence homologies with known allergens.

Conclusions On the basis of our data, Cu/Zn superoxide dismutase, heat shock protein and calmodulin are the most probable allergens among fully characterized proteins of *H.*

brasiliensis, and could potentially explain, at least in part, the multiple cross-reactivities of latex with vegetable foods and other plant-derived products. Consequently, we think that the above proteins should be particularly considered in the future laboratory and clinical research.

Keywords calmodulin, cross-reactivity, Cu/Zn superoxide dismutase, heat shock protein, latex

Submitted 2 January 2006; revised 5 March 2006; accepted 10 April 2006

Correspondence:

Fabrizio Guarneri, Viale Annunziata,
Residence dei Fiori, villa 7, I-98168
Messina, Italy.
E-mail: f.guarneri@tiscali.it

Introduction

Allergic reactions to natural rubber latex (NRL), the milky fluid produced by the *Hevea brasiliensis* tree, are a frequent problem. It is estimated that about 1% of the general population is sensitized to NRL [1], and this percentage is much higher in certain categories, such as occupationally exposed subjects (2–17%) [2] and patients with a history of repeated surgical interventions (~5–55%) [3, 4]. Allergy to latex is often severe and even life-threatening, with the large diffusion of rubber products making prevention difficult [5, 6]. The American and European guidelines have succeeded in reducing the incidence of allergic reactions to NRL in high-risk workers of the medical field, but their efficacy in the general population is unclear [7].

Another issue of latex allergy is the high frequency of cross-reactions with other plant species, which led some clinicians to coin the terms 'latex-fruit syndrome' and 'latex-fruit-pollen syndrome' [7–10]. Despite the great efforts of researchers in the last two decades, our current

knowledge about NRL allergens and their cross-reactivities is still insufficient. In 2000, Petsonk [11] reported that there are nearly 60 allergenic polypeptides in NRL. Of these, only 13 have been characterized as allergens thus far (see the Internet site of the Allergen Nomenclature Sub-committee of the International Union of Immunological Societies, available at <http://www.allergen.org>). Thus, both *in vivo* and *in vitro* diagnosis are often difficult and unreliable [12], and, because symptomatic, therapy is inadequate.

Prompted by these discouraging data, we aimed at identifying potential new allergens among known proteins of *H. brasiliensis*, using *in silico* techniques.

Materials and methods

We extracted the amino acid sequence of all proteins belonging to *H. brasiliensis* (rubber tree), except the 13 known allergens (Hev b 1–13), from public domain databases (PDB, SwissProt, PIR and PRF databases plus

all non-redundant GenBank CDS translations). Duplicate entries were manually deleted. For each of the 102 sequences obtained, amino acid sequence comparison was performed with all known allergens contained in the same databases ($n = 4369$ as of 31 December 2005). To compare proteins, we used the BLAST (Basic Local Alignment Search Tool) program [13]. BLAST performs a residue-by-residue comparison between two proteins, giving each couple of residues a homology score, that is used to calculate the E parameter. This parameter describes the number of hits that can be 'expected' to be found just by chance when searching a database of a particular size. Although the E value is significant when < 1 , we elected to be very stringent and considered matches that were highly significant, namely $E < 10^{-9}$.

All homologies found were then analysed according to the FAO/WHO guidelines for the identification of potential allergens and cross-allergens [14]. According to these criteria, a protein is a potential allergen if it shares more than 35% sequence similarity (on a segment of 80 residues) or an identity of at least six contiguous amino acids with a known allergen [14].

Finally, we used the software SWISS-MODEL [15] to obtain three-dimensional models of the homologous proteins, and the RasMol program (version 2.6, freely available at <http://www.umass.edu/microbio/rasmol/getras.htm>) to visualize the molecules and compare their surface homology.

Results

Only three of the 102 proteins of *H. brasiliensis* resulted in being significantly ($E < 10^{-9}$) homologous to known allergens of other species and can be considered potentially allergenic according to the above FAO/WHO criteria [14]: calmodulin (accession number: AAV88360), Cu/Zn superoxide dismutase (Hb-CuZnSOD, accession number:

AAO14117, heat shock protein (Hb-hsp, accession number: AAQ08597) (Table 1).

Calmodulin matches four allergens, with 39–43% amino acid identity, 61–65% similarity (i.e., identical plus chemically similar amino acids), E values between 8×10^{-10} and 10^{-27} . Hb-CuZnSOD is homologous to only one allergen, Ole e 5 (CuZnSOD) of *Olea europaea*: the shared segment spans virtually the entire length of both proteins (151/152 amino acids), and identity, similarity and E level are particularly high (74%, 85% and 2×10^{-68} , respectively).

Heat shock protein is homologous solely to Asp f 12 (heat shock protein 1) of *Aspergillus fumigatus*; the shared segment spans 440 of the 441 amino acids of Asp f 12, sequence identity is 67%, similarity 84% and E is 10^{-171} .

Three-dimensional molecular modelling of the three aforementioned *H. brasiliensis* proteins produced a complete model only for Hb-CuZnSOD. Thus, only Ole e 5, the protein homologous to Hb-CuZnSOD, was submitted to SWISS-MODEL for molecular modelling. Molecular surface comparison showed a remarkable similarity (Fig. 1).

Discussion

In silico protein analysis is a well-established technique for the identification of candidate antigens and the assessment of immunological cross-reactivity, which has proven useful, in many cases, to predict the involvement of specific proteins in the pathogenesis of some immunomediated diseases [16, 17].

Prediction of allergenicity is an interesting topic in bioinformatics, potentially helpful for clinical research. Unfortunately, in spite of the many efforts made with different methods [14, 18–22], a completely reliable algorithm for allergen identification is unavailable at present. This is mainly due to the partial knowledge about the mechanisms behind the development of allergic reactions,

Table 1. Amino acid sequence homologies between potentially allergenic proteins of *Hevea brasiliensis* and known allergens.

Potentially allergenic <i>H. brasiliensis</i> proteins		Known allergens homologous to <i>H. brasiliensis</i> proteins				
Protein	Segment	Allergen and organism	Segment	Identity	Similarity	E
Calmodulin	7–146	Jun o 4/Jun o 2 (polcalcin) [<i>Juniperus oxycedrus</i>]	21–156	61/140 (43%)	88/140 (62%)	10^{-27}
	11–146	Ole e 8 (calcium-binding allergen) [<i>Olea europaea</i>]	19–157	55/139 (39%)	87/139 (62%)	3×10^{-26}
	1–71	Aln g 4 (polcalcin) [<i>Alnus glutinosa</i>]	1–69	31/71 (43%)	44/71 (61%)	3×10^{-10}
	5–69	Ole e 3 (polcalcin) [<i>O. europaea</i>]	1–66	28/67 (41%)	44/67 (65%)	8×10^{-10}
Cu/Zn superoxide dismutase	1–151	Ole e 5 (Cu/Zn superoxide dismutase) [<i>O. europaea</i>]	1–151	113/151 (74%)	129/151 (85%)	2×10^{-68}
Heat shock protein	260–698	Asp f 12 (heat shock protein hsp1) [<i>Aspergillus fumigatus</i>]	2–441	299/442 (67%)	374/442 (84%)	10^{-171}

'Similarity' refers to identical plus chemically similar amino acids.

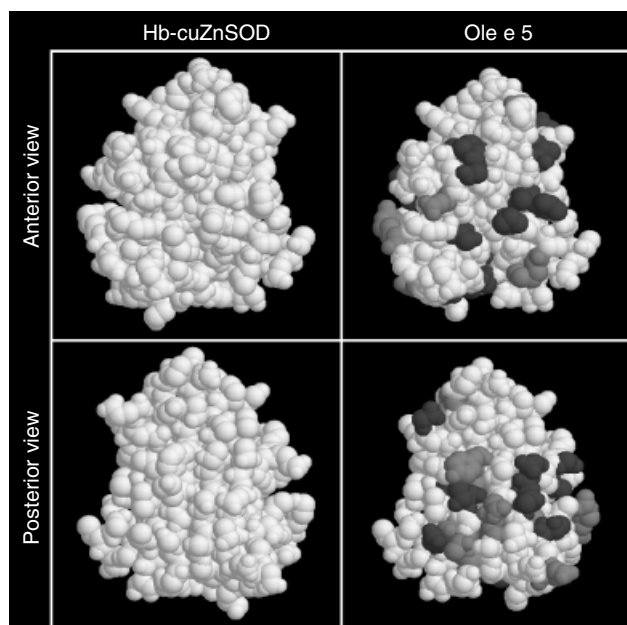


Fig. 1. Molecular surface homology between Cu/Zn superoxide dismutase of *Hevea brasiliensis* (Hb-CuZnSOD) and allergen Ole e 5 (Cu/Zn superoxide dismutase of *Olea europaea*) from two opposite views. White, light grey and dark grey indicate identical, chemically similar or chemically dissimilar (non-homologous) amino acids, respectively.

and, secondarily, to the incompleteness of allergen databases (new allergens are still being discovered). Hence, even approaches that take into account simultaneously many aspects of allergenicity (amino acid sequence, three-dimensional structure, molecular surface analysis) are not, and cannot be, completely accurate. Nevertheless, *in silico* analysis can be an important tool to identify potential new allergens and their cross-reactivity, indicating which proteins should be primarily considered for further investigation by other techniques [22].

Current FAO/WHO guidelines, proposed by a joint expert consultation, are based on amino acid sequence similarity assessment (see Materials and methods). It has been suggested that the FAO/WHO method, although sensible, is not accurate [22], the low accuracy being mainly because of the second criterion (i.e. the presence of a minimum of six consecutive identical amino acids in both the known allergen and the potential allergen). To increase the precision, we pre-selected only proteins that had a very highly significant degree of homology (nine orders of magnitude greater than the minimum *E* value accepted) with known allergens, and considered as potential allergens only those that met both (rather than either) FAO/WHO criteria.

Our results point to a possible role of calmodulin, Cu/Zn superoxide dismutase and heat shock protein as candidate allergens of *H. brasiliensis*. Based on the same data, the hypothesis that these proteins could be responsible for

some of the multiple cross-reactivities reported for NRL appears worthy of further investigation.

The most interesting data concern Hb-CuZnSOD. This protein shows extensive and highly significant homology, in both amino acid sequence and molecular surface (Fig. 1), with the allergen Ole e 5 of *Olea europaea* (olive), and has been identified [23] as an IgE-binding protein. Ole e 5 was discovered in 1998 [24] and is the first Cu/Zn superoxide dismutase identified as an allergen in a pollen source. Ole e 5 has an IgE-binding frequency in *O. europaea*-allergic subjects of about 35% [24, 25], and has been proposed [25] as a possible source of cross-reactivity, due to its widespread presence in plants. This suggestion [25] agrees well with our data.

The homology between heat shock proteins of *A. fumigatus* (allergen Asp f 12) and *H. brasiliensis* is also of interest, not only because it is extensive and highly significant but also because the shared segment of Asp f 12 includes the amino acid sequence DDATDLIPEWLGF (aa 84–96), which is the only known B cell epitope of this allergen [26]. In the corresponding sequence of Hb-hsp (DNCEELIPEYLGF, aa 342–354), eight amino acids are conserved and three are substituted by chemically similar ones.

Calmodulin of *H. brasiliensis* is homologous to many calcium-binding allergenic proteins, most of which are reported as cross-reactive with one another [27–29] and some are considered as true pan-allergens [23]. B cell epitopes have been identified in the N-terminal region of Aln g 4 [29], one of the allergens homologous to calmodulin of *H. brasiliensis*: interestingly, the segment shared by Aln g 4 and calmodulin of *H. brasiliensis* includes aa 1–66 of Aln g 4 (Table 1), further strengthening the hypothesis that the above *H. brasiliensis* protein is an allergen.

The reliability of our research is also supported by a recent study [30], which shows that pollen allergens are restricted to a few protein families (29 out of 7868), including calmodulins and Cu/Zn superoxide dismutases.

Conclusion

The widespread diffusion of latex products, the highly frequent and often severe allergic sensitization, and the numerous cross-reactions with many different plant products (pollens, fruits) call for a complete knowledge about NRL allergens.

The 'gold standard' for the identification of allergens is the demonstration of specific IgE binding. However, in the case of latex (and many vegetal species), performing an extensive search with this technique is neither easy nor economic, due to the high number of polypeptides to be tested. Hence, *in silico* analysis can be a preliminary and complementary valid support to clinical and laboratory research, because it can sort out which proteins deserve

primary attention. Accordingly, our data suggest that future research for new latex allergens should be firstly focused on Cu/Zn superoxide dismutase, heat shock protein and calmodulin of *H. brasiliensis*, which appear to be the most probable allergens among the fully characterized proteins of this plant.

References

- Turjanmaa K, Makinen-Kiljunen S. Latex allergy: prevalence, risk factors, and cross-reactivity. *Methods* 2002; 27:10–4.
- Ahmed SM, Aw TC, Adisesh A. Toxicological and immunological aspects of occupational latex allergy. *Toxicol Rev* 2004; 23: 123–34.
- Capriles-Hulett A, Sánchez-Borges M, Von-Scanzoni C, Medina JR. Very low frequency of latex and fruit allergy in patients with *Spina bifida* from Venezuela: influence of socioeconomic factors. *Ann Allergy Asthma Immunol* 1995; 75:62–4.
- Niggemann B, Buck D, Michael T, Wahn U. Latex provocation tests in patients with *Spina bifida*: who is at risk of becoming symptomatic? *J Allergy Clin Immunol* 1998; 102:665–70.
- Taylor JS, Erkek E. Latex allergy: diagnosis and management. *Dermatol Ther* 2004; 17:289–301.
- Guarneri F. Il lattice: florilegio allergologico. *Ann Ital Dermatol Allergol* 2003; 57:8–12.
- Wagner S, Breiteneder H. Hevea brasiliensis latex allergens: current panel and clinical relevance. *Int Arch Allergy Immunol* 2005; 136:90–7.
- Levy DA, Mounedji N, Noirot C, Leynadier F. Allergic sensitization and clinical reactions to latex, food and pollen in adult patients. *Clin Exp Allergy* 2000; 30:270–5.
- Salcedo G, Diaz-Perales A, Sanchez-Monge R. The role of plant panallergens in sensitization to natural rubber latex. *Curr Opin Allergy Clin Immunol* 2001; 1:177–83.
- Blanco Guerra C. Latex-fruit syndrome. *Allergol Immunopathol (Madrid)* 2002; 30:156–63.
- Petsonk EL. Couriers of asthma: antigenic proteins in natural rubber latex. *Occup Med* 2000; 152:421–30.
- Pridgeon C, Wild G, Ashworth F, Egner W, Ward AM. Assessment of latex allergy in a healthcare population: are the available tests valid? *Clin Exp Allergy* 2000; 30:1444–9.
- 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; 25:3389–402.
- FAO/WHO. *Allergenicity of genetically modified foods*. Rome (Italy): FAO/WHO, 2001.
- Schwede T, Kopp J, Guex N, Peitsch MC. SWISS-MODEL: an automated protein homology-modeling server. *Nucleic Acids Research* 2003; 31:3381–5.
- Persidis A. Autoimmune disease drug discovery. *Nat Biotechnol* 1999; 17:1038–9.
- Benvenga S, Guarneri F, Vaccaro M, Santarpia L, Trimarchi F. Homologies between proteins of *Borrelia burgdorferi* and thyroid autoantigens. *Thyroid* 2004; 14:964–6.
- Zorzet A, Gustafsson M, Hammerling U. Prediction of food protein allergenicity: a bioinformatic learning systems approach. *In Silico Biol* 2002; 2:525–34.
- Ivanciuc O, Schein CH, Braun W. Data mining of sequences and 3D structures of allergenic proteins. *Bioinformatics* 2002; 18: 1358–64.
- Li KB, Issac P, Krishnan A. Predicting allergenic proteins using wavelet transform. *Bioinformatics* 2004; 20:2572–8.
- Riaz T, Hor HL, Krishnan A, Tang F, Li KB. WebAllergen: a web server for predicting allergenic proteins. *Bioinformatics* 2005; 21:2570–1.
- Stadler MB, Stadler BM. Allergenicity prediction by protein sequence. *FASEB J* 2003; 17:1141–3.
- Posch A, Chen Z, Dunn MJ, Wheeler CH, Petersen A, Leubner-Metzger G, Baur X. Latex allergen database. *Electrophoresis* 1997; 18:2803–10.
- Boluda L, Alonso C, Fernandez-Caldas E. Purification, characterization, and partial sequencing of two new allergens of *Olea europaea*. *J Allergy Clin Immunol* 1998; 101:210–6.
- Butteroni C, Afferni C, Barletta B *et al.* Cloning and expression of the *Olea europaea* allergen Ole e 5, the pollen Cu/Zn superoxide dismutase. *Int Arch Allergy Immunol* 2005; 137:9–17.
- Svirshchevskaya EV, Alekseeva IG, Titov VM, Frolova EG, Sapozhnikov AM. Determination of T and B cell epitopes of *Aspergillus fumigatus* ribotoxin and heat shock protein. *Russ J Immunol* 1998; 3:61–8.
- Ledesma A, Villalba M, Vivanco F, Rodriguez R. Olive pollen allergen Ole e 8: identification in mature pollen and presence of Ole e 8-like proteins in different pollens. *Allergy* 2002; 57:40–3.
- Tinghino R, Barletta B, Palumbo S *et al.* Molecular characterization of a cross-reactive *Juniperus oxycedrus* pollen allergen, Jun o 2: a novel calcium-binding allergen. *J Allergy Clin Immunol* 1998; 101:772–7.
- Hayek B, Vangelista L, Pastore A *et al.* Molecular and immunologic characterization of a highly cross-reactive two EF-hand calcium-binding alder pollen allergen, Aln g 4: structural basis for calcium-modulated IgE recognition. *J Immunol* 1998; 161:7031–9.
- Radauer C, Breiteneder H. Pollen allergens are restricted to few protein families and show distinct patterns of species distribution. *J Allergy Clin Immunol* 2006; 117:141–7.