

Primary structure of *Lep d I*, the main *Lepidoglyphus destructor* allergen

Javier VARELA², Pilar VENTAS¹, José CARREIRA¹, Julio A. BARBAS³, Guillermo GIMENEZ-GALLEGU² and Florentino POLO¹

¹ Research Department, Alergia e Inmunología Abelló S. A., Madrid, Spain

² Centro de Investigaciones Biológicas, C. S. I. C., Madrid, Spain

³ Instituto Cajal, C. S. I. C., Madrid, Spain

(Received May 18, 1994) – EJB 94 0750/3

The most relevant allergen of the storage mite *Lepidoglyphus destructor* (*Lep d I*) has been characterized. *Lep d I* is a monomer protein of 13273 Da. The primary structure of *Lep d I* was determined by N-terminal Edman degradation and partially confirmed by cDNA sequencing. Sequence polymorphism was observed at six positions, with non-conservative substitutions in three of them. No potential N-glycosylation site was revealed by peptide sequencing. The 125-residue sequence of *Lep d I* shows approximately 40% identity (including the six cysteines) with the overlapping regions of group II allergens from the genus *Dermatophagoides*, which, however, do not share common allergenic epitopes with *Lep d I*.

Exposure to environmental aeroallergens may result in an allergic immune response in approximately 15% of the population, provoking clinical symptoms that range from extrinsic asthma and allergic rhinitis to atopic dermatitis. The importance of pyroglyphid mites of the order *Acaridae*, and mainly those belonging to the genus *Dermatophagoides*, as a primary cause of respiratory allergy has long been recognized since it was first reported by Voorhorst et al. [1]. However, the role in human allergic pathology of different non-pyroglyphid mites, storage mites, had not been demonstrated until recently [2–5]. Several studies carried out in the past few years have shown that the storage mite *Lepidoglyphus destructor* (*Glycyphagidae* family) is a significant source of proteins that can elicit allergic respiratory symptoms [2–8]. Sensitization to *L. destructor* has been extensively described among rural populations [3–4], but also may occur among urban populations in geographic areas with favorable climate conditions for the growth of this mite [5–7], and it has normally been associated with damp housing conditions [8]. The incidence of allergy to *L. destructor* may even surpass that to *Dermatophagoides* in some regions, such as tropical and subtropical areas. Thus, Fernández-Caldas et al. [6] studied the prevalence of specific IgE to house dust and storage mites in 100 urban dwellers with asthma and/or rhinitis who lived in the subtropical Tampa Bay (Florida, USA) area, and

they found that sensitization to the storage mites *L. destructor* and *Tyrophagus putrescentiae* was more frequent (69%) than sensitization to *Dermatophagoides pteronyssinus* (66%).

Although several polypeptides with IgE-binding ability were identified in *L. destructor* extracts by SDS/PAGE immunoblotting, a protein of approximately 14000 Da was demonstrated to be the principal allergen [9, 10]. In a previous work [11], we purified this protein by mAb-based affinity chromatography, and showed that the protein preparation met the purity requirements of the International Union of Immunological Societies Sub-Committee for Allergen Nomenclature to be termed *Lep d I* [12].

The clinical relevance of this allergen was confirmed by radio allergosorbent test technique (RAST), detecting specific IgE to affinity-purified *Lep d I* in the serum of 87% of *L. destructor*-allergic patients [11]. In addition, *Lep d I* could inhibit up to 75% the binding of IgE to *L. destructor* whole extract in RAST inhibition experiments [11]. In the same study, competition experiments for the binding of human IgE showed that *Lep d I* was not allergenically cross-reactive with allergens from house dust mites belonging to the genus *Dermatophagoides*.

As a number of allergens have been reported to consist of multiple isoforms, and previous observations suggested that this might also happen in the case of *Lep d I* [11], we attempted the sequencing of the allergen by N-terminal Edman degradation in order to detect the possible amino acid substitutions that could give rise to different *Lep d I* isoallergens.

Here we report that native *Lep d I* is a single polypeptide monomer whose amino acid sequence has six polymorphic positions. Primary structure of *Lep d I* appeared 40% identical with the group II allergens from the genus *Dermatophagoides*, though, as remarked above, are allergenically non-cross-reactive with *Lep d I* [11]. The reported data may help

Correspondence to F. Polo, Departamento de Investigación, Alergia e Inmunología Abelló S. A., C/Miguel Fleta, 19, E-28037 Madrid, Spain

Abbreviations. RAST, radio allergosorbent test; 3'-RACE, rapid amplification of cDNA 3' end.

Enzymes. Trypsin (EC 3.4.21.4); chymotrypsin (EC 3.4.21.1); endoproteinase Glu-C (EC 3.4.21.19); endoproteinase Asp-N (EC 3.4.24.33); proline-specific endopeptidase (EC 3.4.21.26).

Note. The novel sequence data published for *Lep d I* have been submitted to the EMBL data bank and are available under accession number X81399.

in the mapping of B-cell and T-cell epitopes involved in the human immune response to this allergen, and thus contribute to the design of immunotherapeutic agents for clinical desensitization.

MATERIALS AND METHODS

Materials

Proteases used for *Lep-d-I*-specific cleavages were all of sequencing grade. Trypsin, chymotrypsin, V8 protease and endoprotease Asp-N were from Boehringer Mannheim. Proline-specific endopeptidase was from ICN Biomedicals. Quick Prep Micro mRNA purification kit was from Pharmacia. *Taq* DNA polymerase was from Perkin-Elmer/Cetus. All other reagents were purchased from commercial sources and were of the highest purity commercially available.

Purification of *Lep d I*

Lep d I was affinity purified via a Le5B5 mAb anti-*(Lep d I)* column, from a 10% (mass/vol) aqueous extract of *L. destructor* mite bodies (>99% purity; Biopol Laboratory Inc.), as described elsewhere [11]. Afterwards, the protein solution was dialyzed extensively against distilled water, lyophilized, and stored at 4°C until use. Concentration of *Lep d I* samples was determined by amino acid analysis.

Size-exclusion HPLC

Lyophilized pure *Lep d I* was dissolved in 100 mM Na₂SO₄, 20 mM NaH₂PO₄, pH 6.8, and chromatographed in a size-exclusion HPLC Ultraspherogel-SEC 2000 (0.75 cm×30 cm) column (Beckman), using the same solvent as the mobile phase at a constant flow rate of 0.5 ml/min. The column had been previously calibrated by injection of a gel-filtration-standard from Bio-Rad. The chromatographic procedure was carried out on HPLC equipment from Waters comprising a Model 680 gradient programmer, a 510 pump, a U6K injector, and a 490 multi-wavelength ultraviolet/visible detector set at 214 nm and 280 nm.

Mass spectrometry

Mass spectrometry was carried out in a Kratos Kompact MALDI 3 V2.0 (Kratos Analytical). Affinity-purified *Lep d I* (100 pmol) was applied to the probe and dried in a stream of air at room temperature.

3,5-Dimethoxy-4-hydroxycinnamic acid was used as the ionizing matrix.

Reductive cleavage and alkylation of *Lep d I*

Affinity-purified *Lep d I* was reduced with 2-mercaptoethanol, and *S*-carboxymethylated with iodoacetic acid following an established procedure [13]. After completion of the reaction, excess reagents were removed by chromatography onto a Sephadex G-25 column equilibrated with 0.1% trifluoroacetic acid. Fractions containing reduced and alkylated allergen were pooled and stored at -20°C until use.

Amino acid analysis and protein sequence analysis

Amino acid analysis was carried out as described [11] using a Dabs-amino acid analysis kit and System Gold equipment supplied by Beckman.

Table 1. Molecular-mass estimations for *Lep d I*. Values of the molecular mass of *Lep d I* were obtained by the following techniques: SDS/PAGE in the presence (+) or absence (-) of the reducing agent 2-mercaptoethanol (values from [11]), gel filtration under non-denaturing conditions (GF HPLC) and MS.

Molecular mass determined by			
SDS PAGE	GF HPLC		MS
	+ 2-mercaptoethanol	- 2-mercaptoethanol	
Da			
16000	14000	13000	13273

N-terminal sequence analysis of the whole protein was performed on a sample of intact reduced and *S*-carboxymethylated *Lep d I*. For internal sequencing *Lep d I* (100–130 µg) was digested with the appropriate specific proteases under the conditions given in the figure legends. Reaction products were loaded onto a 4.5 cm×15 cm Vydac C₁₈ reverse-phase column (5-µm particle size, 30-nm pore size; The Separation Group) equilibrated with 13 mM trifluoroacetic acid, and hydrolysis products separated by elution with a 0 to 100% gradient of acetonitrile/H₂O (7:3) in 11.7 mM trifluoroacetic acid (buffer B), at a flow rate of 0.5 ml/min. Eluted polypeptides were detected by measuring the absorbance of the eluent at 214 nm, and collected manually. A peptide obtained by digestion of *Lep d I* with endoprotease Asp-N, whose sequence had been previously identified as DFYSGTI-PAITPKVKA was used in the case of proline-specific endopeptidase.

N-terminal Edman degradation was carried out on Polybrene-coated filters using an Applied Biosystems Model 477A protein microsequencer. Phenylthiohydantoin derivatives were detected using an on-line analyzer (Applied Biosystems, model 120A).

Poly(A)-rich RNA preparation and cDNA synthesis

Polyadenylated RNA was isolated from live *L. destructor* mites (Alergia e Inmunología Abelló S. A.) using the Quick Prep Micro mRNA purification kit. Briefly, total RNA was extracted from a homogenate of mite bodies in a buffered aqueous solution containing guanidinium thiocyanate and *N*-lauroyl sarcosine. Subsequently, polyadenylated mRNA was isolated by affinity chromatography on an oligo(dT)-cellulose column. First strand cDNA was synthesized from poly(A)-rich RNA using the AP primer and Superscript reverse transcriptase, provided with the 3'-RACE System kit (Gibco-BRL; 3'-RACE, rapid amplification of cDNA 3' end), following the manufacturer's instructions.

Lep-d-I-specific cDNA amplification, cloning and sequencing

After RNase treatment, the first-strand cDNA was amplified to obtain *Lep-d-I*-specific cDNA by two sequential nested amplification reactions using the degenerate oligonucleotides LEP1 and LEP2 as the sense primers and UAP, provided with the 3'-RACE System kit, as the antisense primer in both reactions. LEP1 and LEP2 were synthesized by the phosphoramidite method, based on the nucleotide sequences derived from the N-terminal amino acid sequence of

The complete amino acid sequence of the 125-residue *Lep d I* was initially established by a long N-terminal sequence degradation and determination of the sequences of overlapping peptides generated by several endoproteases. Fig. 1 summarizes the strategy used to determine the full sequence. The abundance of Lys and short repetitive sequences in some stretches of the protein resulted in a very complex overlapping in some regions. Direct N-terminal sequence of the whole protein and several peptides obtained by endoprotease digestion allowed us to establish a block se-

Table 2. Amino acid composition of *Lep d I* allergen. Results of the amino acid composition analysis are from [11]. The analysis was performed on an unmodified sample of the protein and no corrections were made for destruction during hydrolysis. Cysteine was determined as cystine, and the result expressed as half-cystine. In the sequence analysis column, amino acid residues in positions without sequence variations are given; the number of residues of each amino acid involved in sequence polymorphism is indicated in parentheses. n.d., not determined.

Amino acid	No. of residues determined from	
	composition analysis	sequence analysis
Asx	9.2	
Asp		8 (1)
Asn		1 (2)
Glx	9.6	
Glu		7 (1)
Gln		3
Ser	2.0	1 (1)
Thr	11.8	14
Arg	1.7	1
Gly	16.8	15 (1)
Ala	11.6	10 (1)
Pro	4.5	4
Val	12.0	11 (2)
Met	2.1	3
Ile	8.4	7 (3)
Leu	7.0	6
Phe	4.3	4
Cys	5.0	6
Lys	13.7	13
His	4.2	4
Tyr	1.2	1
Trp	n.d.	0

quence comprising 98 amino acids at the N-terminus of *Lep d I*. Another large fragment (25 residues) could be deduced, as shown in Fig. 1, by overlapping the sequence of a set of peptides obtained by endoprotease hydrolysis. Connection of these two blocks was accomplished after sequencing peptide N-7. To improve the overlapping obtained with peptide N-7, this peptide was sequenced from the Ala at position 10 after digestion with proline endoprotease (Fig. 1).

A remarkable feature of the sequence was the identification of different amino acids in six positions (37, 47, 79, 88, 90, 109; Fig. 1). The presence of the internal repetitive sequence (the dipeptide K-V) occurring three times between positions 48 and 57 should be noted as characteristic of *Lep d I*. The molecular mass of *Lep d I* estimated from the sequence ranged between 13041 and 13157 depending on the amino acid substitutions considered. The amino acid composition and the molecular mass of the protein calculated from the sequence agree well with those obtained by direct amino acid analysis (Table 2) and mass spectrometry, respectively.

Since many short and similar peptides were required to establish the sequence region spanning residues 30 – 70 (Fig. 1), the sequence assignment of this stretch was obviously prone to errors. To confirm the proposed sequence, a fragment cDNA spanning this region of *Lep d I* was obtained and sequenced. First-strand cDNA was synthesized from poly(A)-rich RNA isolated from live *L. destructor* mites, and partial cDNA coding for *Lep d I* was specifically amplified using a 3'-RACE/PCR procedure. Two consecutive rounds of PCR amplification were carried out with a nested pair of *Lep d I*-specific 5' primers in conjunction with a 3' primer

(UAP primer) encoding a tag sequence covalently linked 3' to the oligo(dT) in the primer (AP primer) used in first-strand cDNA synthesis. Agarose-gel electrophoresis of the PCR products showed that the amplified fragments had the size expected according to the sequence of Fig. 1. The complete sequence of a 470-bp amplified product was determined (Fig. 2), as can be inferred from the fact that the first five nucleotides in the 5' end (ATTCA) correspond to LEP2 primer; whereas in the 3' end, after the poly-(A) tract, a sequence (GTACTA) complementary to part of the UAP primer sequence could be read. The translated amino acid sequence encoded by clone L2 confirmed that obtained by Edman degradation with E, A, V, I, S and I at the polymorphic positions 37, 47, 79, 88, 90 and 109, respectively. Unexpectedly, DNA sequencing disclosed a Glu at the C-terminus of the polypeptide chain that was not detected during the N-terminal sequencing process.

A sequence similarity search showed a significant degree of similarity between *Lep d I* and group II allergens from *Dermatophagoides farinae* and *D. pteronyssinus*. The best alignment (15.3 SD from random) was detected with the translated sequence from clone DFDER1 encoding *Der f II* [17] (Fig. 3). Six Cys residues are conserved between *Lep d I* and *Der f II* sequences. These residues, detected as approximately 2.5 cystines in the amino acid analysis (Table 2), should be involved in disulfide bonds.

DISCUSSION

In recent years, affinity chromatography with specific mAb has proven a valuable tool for the purification of antigens, mainly when they are only available in complex mixtures, such as the allergenic extracts. The availability of an anti-(*Lep d I*) mAb allowed us to purify this allergen in a single chromatographic step from an *L. destructor* mite bodies' extract [11]. This method yielded *Lep d I* in a sufficient amount to attempt the determination of its primary structure. We report here the complete amino acid sequence of *Lep d I* determined by N-terminal Edman degradation, and confirmed by cDNA sequencing.

Sequence polymorphism of *Lep d I* here reported confirm previous observations about the existence of different isoforms of the protein [11]. The six polymorphic sites detected in *Lep d I* may raise to 64 different *Lep d I* isoforms, if all polymorphisms occur independently. However, some of them could be associated, as has been reported to occur in *Amb a I.4*, where three polymorphisms are inherited as a block [18]. This might be the case of the polymorphisms detected at positions 88 and 90 of *Lep d I*, since an isoleucine residue at position 88 was found in all peptides when serine was identified at position 90, whereas an asparagine residue at position 88 along with a glycine at position 90 were found in other peptides. These observations are in agreement with the absence of N-glycosylated substitutions in affinity-purified *Lep d I* (Ventas, P., Polo, F. and Carreira, J., unpublished results), since one of the possible combination at positions 88 and 90 (N-X-S) that was not detected in any sequenced peptide would have represented a potential N-glycosylation site of the allergen. Nevertheless, the presence of this isoform in minute amounts, or even the possibility that glycosylation made *Lep d I* unrecognizable to the mAb used to purify it, cannot be eliminated. Polymorphism may account for the unusual width of the peak observed by mass spectrometry. Nevertheless, it should be considered that the complete range

since epitopes involved in T-cell recognition, believed to play a crucial role in the immune response to allergens [26], are mostly defined by the primary structure of short peptides derived from the intact protein [27, 28]. Such knowledge may lead to more effective treatments for *L. destructor*-sensitive patients.

This work was supported by *Alergia e Imunología Abelló S. A.* We thank Dr J. Moscoso del Prado for helpful discussions, and Dr H. Ipsen, from ALK Research, for performing the database similarity searches. We also thank I. Courinha for her excellent secretarial assistance.

REFERENCES

- Voorhorst, R., Spieksma-Boezeman, M. I. A. & Spieksma, F. T. M. (1964) Is a mite (*Dermatophagoides* sp.) the producer of the house-dust allergen? *Allerg. Asthma* 10, 329–334.
- Wraith, D. G., Cunningham, A. M. & Seymour, W. M. (1979) The role and allergenic importance of storage mites in house dust and other environments, *Clin. Allergy* 9, 545–561.
- Cuthbert, O. D., Brostoff, J., Wraith, D. G. & Brighton, W. D. (1979) 'Barn allergy': asthma and rhinitis due to storage mites, *Clin. Allergy* 9, 229–236.
- Van Hage-Hamsten, M., Johansson, S. G. O., Hoglund, S., Tull, P., Wiren, A. & Zetterstrom, O. (1985) Storage mite allergy is common in a farming population, *Clin. Allergy* 15, 555–564.
- Korsgaard, J., Dahl, R., Iversen, M. & Hallas, T. (1985) Storage mites as a cause of bronchial asthma in Denmark, *Allergol. Immunopathol.* 13, 143–149.
- Fernández-Caldas, E., Trudeau, W. L., García-Ramos, B. A. E., van Hage-Hamsten, M., Johansson, S. G. O. & Lockey, R. F. (1991) Sensitization to house dust and storage mites in urban dwellers, *J. Allergy Clin. Immunol.* 87, 322.
- Luczynska, C. M., Griffin, P., Davies, R. J. & Topping, M. D. (1990) Prevalence of specific IgE to storage mites (*A.siro*, *L. destructor* and *T.longior*) in an urban population and cross-reactivity with the house dust mite (*D.pteronysinus*), *Clin. Exp. Allergy* 20, 403–406.
- Iversen, M. & Dahl, R. (1990) Allergy to storage mites in asthmatic patients and its relation to damp housing conditions, *Allergy* 45, 81–85.
- Johansson, E., Van Hage-Hamsten, M. & Johansson, S. G. O. (1988) Demonstration of allergen components in the storage mite *Lepidoglyphus destructor* by an immunoblotting technique, *Int. Arch. Allergy Appl. Immunol.* 85, 8–13.
- Ventas, P., Carreira, J. & Polo, F. (1991) Identification of IgE-binding proteins from *Lepidoglyphus destructor* and production of monoclonal antibodies to a major allergen, *Immunol. Lett.* 29, 229–234.
- Ventas, P., Carreira, J. & Polo, F. (1992) Purification and characterization of *Lep d I*, a major allergen from the mite *Lepidoglyphus destructor*, *Clin. Exp. Allergy* 22, 454–460.
- Marsh, D. G., Goodfriend, L., King, T. P., Lowenstein, H. & Platts-Mills, T. (1986) Allergen nomenclature, *Bull. W. H. O.* 64, 767–770.
- Giménez-Gallego, G., Conn, G., Hatcher, V. B. & Thomas, K. A. (1986) The complete amino acid sequence of human brain-derived acidic fibroblast growth factor, *Biochem. Biophys. Res. Commun.* 138, 611–617.
- Marchuk, D., Drumm, M., Saulino, A. & Collins, F. S. (1991) Construction of T-vectors, a rapid and general system for direct cloning of unmodified PCR products, *Nucleic Acids Res.* 19, 1154.
- Sanger, F., Nicklen, S. & Coulson, A. R. (1977) DNA sequencing with chain-terminating inhibitors, *Proc. Natl Acad. Sci. USA* 74, 5463–5467.
- Pearson, W. R. & Lipman, D. J. (1988) Improved tools for biological sequence comparison, *Proc. Natl Acad. Sci. USA* 85, 2444–2448.
- Yuuki, T., Okumura Y., Ando, T., Yamakawa, H., Suko, M., Haida, M. & Okudaira, H. (1991) Cloning and expression of cDNA coding for the major house dust mite allergen *Der f II* in *Escherichia coli*, *Agric. Biol. Chem.* 55, 1233–1238.
- Griffith, I. J., Pollock, J., Klapper, D. G., Rogers, B. L. & Nault, A. K. (1991) Sequence polymorphism of *Amb a I* and *Amb a II*, the major allergens in *Ambrosia artemisiifolia* (short ragweed), *Int. Arch. Allergy Appl. Immunol.* 96, 296–304.
- Muthiah, R., Miller, M. & Kagen, S. (1991) Barn allergy: isolation and characterization of the major allergens of storage mites: *L. destructor*, *J. Allergy Clin. Immunol.* 87, 326.
- Van Hage-Hamsten, M., Bergman, T., Johansson, E., Persson, B., Jörnvall, H., Harfast, B. & Johansson, S. G. O. (1992) N-terminal amino acid sequence of principal allergen of storage mite *Lepidoglyphus destructor*, *Lancet* 340, 614.
- Nishiyama, C., Yuuki, T., Takai, T., Okumura, Y. & Okudaira, H. (1993) Determination of three disulfide bonds in a major house dust mite allergen, *Der f II*, *Int. Arch. Allergy Immunol.* 101, 159–166.
- Stewart, G. A., Bird, C. H. & Thompson, P. J. (1992) Do the group II dust mite allergens correspond to lysozyme? *J. Allergy Clin. Immunol.* 90, 141–142.
- Johansson, E., Borga, A., Johansson, S. G. O. & van Hage-Hamsten, M. (1991) Immunoblot multi-allergen inhibition studies of allergenic cross-reactivity of the dust mites *Lepidoglyphus destructor* and *Dermatophagoides pteronyssinus*, *Clin. Exp. Allergy* 21, 511–518.
- Roebber, M., Klapper, D. G., Goodfriend, L., Bias, W. B., Hsu, S. H. & Marsh, D. G. (1985) Immunochemical and genetic studies of *Amb.t. V* (Ra5G), an Ra5 homologue from giant ragweed pollen, *J. Immunol.* 134, 3062–3069.
- Huang, S. K. & Marsh, D. G. (1991) Human T-cell responses to ragweed allergens: *Amb V* homologues, *Immunology* 73, 363–365.
- O'Hehir, R. E., Garman, R. G., Greestien, J. L. & Lamb, J. R. (1991) The specificity and regulation of T-cell responsiveness to allergens, *Annu. Rev. Immunol.* 9, 67–95.
- Berzofsky, J. A., Brett, S. J., Streicher, H. Z. & Takahashi, H. (1988) Antigen processing for presentation to T lymphocytes: Function, mechanisms, and implications for the T-cell repertoire, *Immunol. Rev.* 106, 5–31.
- Rothbard, J. B. & Gefters, M. L. (1991) Interactions between peptides and MHC proteins, *Annu. Rev. Immunol.* 9, 527–565.

Supplementary Material. **Primary structure of *Lep d I*, the main *Lepidoglyphus destructor* allergen.** Fig. S1. Mass spectrometry of *Lep d I*. Fig. S2. Reverse-phase HPLC analysis of a tryptic digest of *Lep d I*. Fig. S3. Reverse-phase HPLC rechromatography of a chymotryptic digest of *S*-carboxymethylated *Lep d I*. Fig. S4. Reverse-phase HPLC analysis of a *Staphylococcus aureus* V8 digest of *S*-carboxymethylated *Lep d I*. Fig. S5. Reverse-phase HPLC analysis of an endoprotease Asp-N digest of *S*-carboxymethylated *Lep d I*. Fig. S6. Reverse-phase HPLC analysis of proline endopeptidase digest of peptide N-7. Fig. S7. Primers used to amplify by PCR a 3' portion of cDNA of *Lep d I*. This information is available, upon request, from the Editorial Office. A total of 14 pages are available.