cDNA analysis of the mite allergen Lep d 1 identifies two different isoallergens and variants

Margit Schmidt[a,b,*], Susanna Olsson[a], Ingeborg van der Ploeg[a], Marianne van Hage-Hamsten[a]

[a]*Department of Laboratory Medicine, Division of Clinical Immunology, Karolinska Hospital, S-171 76 Stockholm, Sweden*
[b]*Department of Pathology and Laboratory Medicine, East Carolina University, School of Medicine, Greenville, NC 27834, USA*

Received 20 June 1995

Abstract For the first time the complete cDNA encoding two isoallergens of the non-pyroglyphid dust mite *Lepidoglyphus destructor*, Lep d 1, allergen has been sequenced. In addition, one of the isoallergens was found to have two variants. Oligonucleotides were designed from known amino acid sequences. The cDNA sequences were obtained by hybridizing these primers to mRNA and enhancement by the RT-PCR technique. To obtain the different complete encoding cDNA sequences and eliminate heteroduplex artifacts, we performed PCR + 1 reactions. Comparison of the amino acid sequence of the allergen shows leader sequences of 16 amino acids for both isoforms.

Key words: Allergy; Mite; cDNA cloning; PCR; Polymorphism

1. Introduction

Lepidoglyphus destructor (*L. destructor*) is an important non-pyroglyphid mite species in Europe [1–3]. It is the predominant allergen in farming environments [3–5]. Allergy to non-pyroglyphid mites has been mostly associated with occupational exposure [2–6], but people living in damp housing conditions may also be exposed [7]. Until recently, the characterization of mite allergens has focused on the house dust mite *Dermatophagoides* species. The major allergens of these species have been studied in detail with physicochemical and recombinant technology. However, only a few studies have been performed on storage mite *L. destructor* allergens [8–11].

We have identified at least twenty IgE-binding components in the molecular weight range of 13–93 kDa [8,9]. The major allergen of *L. destructor*, Lep d 1, is a protein of approximately 15,000 dalton and is recognized by about 90% of sera RAST positive to this mite species. Lep d 1 has been purified and its amino acid sequence has been reported [10,11]. Polymorphism for this protein has been discussed based on peptide analysis; however, Varela et al. [11] did not distinguish different isoallergens and their variants, since their data were based on protease-digested protein fragments. This is the first report of the complete encoding cDNA sequence of two isoallergens and its variants of the Lep d 1 major allergen.

2. Material and methods

2.1. Isolation of mRNA

L. destructor cultures were obtained from Allergon AB (Ängelholm, Sweden). The isolation of mRNA was carried out using the PolyATtract System 1000 (Promega, Scandinavian Diagnostic Services, Falkenberg, Sweden). Briefly, approximately 0.5 g of mite culture was homogenized in 4 ml extraction buffer containing 2% β-mercaptoethanol. After addition of 8 ml of preheated dilution buffer and 500 pmol of biotinylated oligo(dT), the sample was incubated at 70°C for 5 min. The solution was centrifuged at 12,000 × *g* for 10 min at room temperature. The supernatant was transferred and incubated with 6 ml of Streptavidin MagneSphere Paramagnetic Particles at room temperature for 2 min. After washing three times with 0.5× SSC and capturing the magnetic particles with a magnetic stand, the mRNA was eluted with 3 ml water. The nucleic acid was then alcohol precipitated.

2.2. Construction of oligonucleotide primers for PCR

Two degenerate oligonucleotides (primer 1 and primer 2) were designed according to the known amino acids [10,11]. Primer 1 was in sense orientation and corresponded with the N-terminal end, primer 2 was in anti-sense direction according to amino acids 37 to 42. In addition, for the 3′ RACE protocol match primer 3 and oligo(dT1 and dT2) were used; for the 5′ RACE protocol match primers 4, 5 and a 5′ anchor primer (Gibco/BRL, Life Technologies, Täby, Sweden) were utilized. Sense primers (6, 7) were designed for the PCR + 1 reactions; primer 7 shows the same sequence as primer 6 but has an additional restriction endonuclease (RE) site, *Hin*dIII, at the 5′ end. All primers are listed in Table 1.

2.3. PCR amplification

Reverse transcriptase polymerase chain reaction (RT-PCR) was used to amplify the RNA and to obtain three overlapping cDNA fragments using the following primer combinations: primer 1 and primer 2; primer 3, oligo(dT1) and oligo(dT2); primer 4 and primer 5 with 5′ RACE anchor primer.

The amplification of the middle fragment of Lep d 1 allergen cDNA was based on standard protocols. Briefly, poly(A)-selected RNA was added to 1 μmol of oligonucleotide primer 2 in a 10 μl volume, heated to 70°C for 10 min and then chilled on ice. The following solutions were added to the final concentration in 1× first strand synthesis buffer (BRL): 0.5 mM each dNTP and 10 mM dithiothreitol. After 2 min at 42°C, 200 units of Superscript RNAse H$^-$ (BRL) were added and the reaction incubated for 30 min at 42°C. The mixture was then heated to 55°C for 5 min, 2 units of RNAse H added, and incubated for 10 min at 55°C, then chilled on ice. The cDNA was purified using glassmax columns (BRL). The PCR amplification was carried out using primer 1 and primer 2 (each 0.4 μM) in 1× extender buffer (Stratagene, AH Diagnostics AB, Skärholmen, Sweden) containing 0.25 mM each dNTP and overlaid with 1 volume of mineral oil. After an initial denaturation at 94°C for 8 min, 5 units of Taq polymerase (BRL) and an equal amount of Taq extender (Stratagene) were added at 80°C. The mixture was then subjected to 40 cycles of amplification in an MJ Research Minicycler: 1 min at 94°C, 2 min at 55°C, and 2.5 min at 72°C.

The 3′ fragments were PCR amplified using the same conditions as described above with the primers oligo(dT1) for cDNA synthesis, and the primers oligo(dT2) and 3 for the PCR reactions. The 5′ fragments were obtained using primer 4 for the cDNA synthesis, then the fragments were tailed with 1 mM dCTP and 30 units terminal dideoxytransferase according to manufacturer's protocol (BRL). PCR amplifications were performed as described above using primer 5 and 5′ anchor primer (BRL).

2.4. PCR + 1 reactions

Synthesis of cDNA from mRNA was performed exactly under the same conditions as described above for the 3′ PCR fragments. The first

*Corresponding autor. Department of Laboratory Medicine, Division of Clinical Immunology, Karolinska Hospital, S-171 76 Stockholm, Sweden. Fax: (46) (8) 7295926.

SSDI 0014-5793(95)000779-2

part of PCR + 1 [12] reactions were carried out in 50 μl containing one half (20 μl) of the cDNA sample, 0.25 mM of each NTP, 50 μM primer oligo(dT2), 5 μM primer 6 (a sense primer starting 10 nucleotides upstream from the initial codon and ending with nucleotide 11 of the leader sequence) in 1× extender buffer. Taq polymerase and extender were added under hot start conditions as described above. Additionally, we added 1 μl of Perfect match (Stratagene) to increase the specificity of primer annealing. Each of the 30 amplification cycles consisted of: 94°C for 2 min, 50°C for 2 min, 72°C for 3 min. The final extension was for 10 min at 72°C.

The second part of PCR + 1 reactions contained 25 μl of the first reaction sample from above and 50 μM of primer 7. Additionally, Taq polymerase and extender were added under hot start conditions, and a single cycle was performed as described above. The reactions were then cooled to 4°C.

2.5. Cloning and screening of PCR and PCR + 1 reactions

All amplified products were analyzed by electrophoresis, and the positive samples were eluted from the agarose gels. Isolated DNA's from PCR reactions were directly cloned into TA3 plasmid using the TA cloning kit (Invitrogen). After transformation, plasmid DNA minipreparations from bacterial colonies were digested with the appropriate RE(s). Namely, *Eco*RI for PCR fragments and *Hin*dIII for the PCR + 1 fragments. The potential PCR + 1 clones were analyzed by gel electrophoresis according to the expected size. The identification of complete PCR + 1 clones was confirmed by DNA sequencing of the incorporated primer 7, which contains a *Hin*dIII site at the 5′ end, into the PCR + 1 fragments.

2.6. DNA sequencing

Recombinant plasmids were sequenced by the dideoxy chain termination method using the Sequenase 2.0 kit (US Biochemical, Amersham Sweden AB, Solna, Sweden) and SP6 and T7 promotor sequencing primers.

3. Results and discussion

In order to obtain the complete cDNA sequence of Lep d 1 major allergen, we utilized RT-PCR, PCR + 1 and standard cloning techniques. We first amplified a small fragment using degenerated primers 1 and 2, which were designed according to the known amino acid sequence. The 3′ fragments were obtained with both oligo(dT) primers and match primer 3. The amplification of the 5′ fragments were performed using primer 4 for cDNA synthesis followed by the nested primer 5 in combination with the anchor primer (BRL). Three independent PCR reactions were performed for the 3′ RACE and 5′ RACE.

The cloned PCR fragments were sequenced and analyzed using the DNASIS (National Biosciences) and Genetic Computer Group (Wisconsin) software packages. The data obtained from a minimum of three independent clones of the 3′ and 5′ fragments indicated that we had amplified at least two isoallergens. The nucleotide composition differed at some positions in both the coding and the non-coding regions.

The formation of heteroduplex DNA arising from quasi-homologous sequences using PCR has been described [13]. Borriello and Krauter [12] have reported the complications arising from artifactual DNA sequences via mismatch repair during amplification and subcloning into bacteria which will repair the mismatches randomly for replication. This generates sequences from which it is difficult to determine the real sequences. To overcome these artifacts, they developed the PCR + 1 methodology and this PCR technique has been, meanwhile, successfully applied when polymorphism is present [14,15]. We have designed the primers for the 5′ end of the PCR + 1 method to a conserved region we have found after sequencing 9 clones of two independent 5′ RACE PCR reactions. This area stretches from 10 nucleotides before the initial codon into the third amino acid of the leader sequence. For the 3′ end, we have used the oligo(dT1 and dT2) primers. This enabled us to survey a broader range of possible variants of this protein. In order to increase the specificity of the products, we have added Perfect Match (Stratagene) to the reactions.

Sequencing of potential PCR + 1 clones confirmed that 97% of the screened clones contained the restriction site *Hin*dIII at the 5′ end. We have minutely examined 28 PCR + 1 clones of 6 independent reactions and identified two isoallergens, one with two variants. The identification was based on sequences that were identical from clones of at least two independent reactions. We have named these allergens according to the new allergen nomenclature [16] Lep d 1.0101, Lep d 1.0102 and Lep d 1.02. Only 6 out of 28 (21.4%) clones represented Lep d 1.02, whereas the two variants Lep d 1.0101 and Lep d 1.0102 were equally represented (each 39.3%). The data for all three forms further show a common leader sequence corresponding to 16 amino acids which has hitherto never been reported. In Fig. 1 we show the nucleotide sequences of these forms of Lep d 1.

The two variants Lep d 1.0101 and Lep d 1.0102 show differences of 6 nucleotides in the encoding part of the cDNA which are silent mutations. The non-coding 3′ end indicates a block mutation and lengthens the cDNA of the latter by one nucleotide (Fig. 1). Lep d 1.0102 is identical in the coding area with clone L2 described by Varela et al. [11]; however, the non-coding 3′ end reveals differences between the polyA signal and the polyA tail and additionally 15–5 nucleotides upstream of the polyA signal (Fig. 3). These differences may be due to a slight variability in the site of the oligo(dT) primer annealing and/or to the presence of further variants in the mite population.

The mutations in Lep d 1.02 are more numerous (Fig. 1) and lead to 13 changes (10.4%) of the amino acid composition of the protein (Fig. 2). Earlier, Varela et al. [11] have postulated the possible polymorphism of Lep d 1. Their arguments were based on the amino acid sequence of overlapping peptides and

Table 1
List of primers used in PCR and PCR + 1

primer 1	5′-AA(A/G)ATGACNTT(T/C)AA(A/G)GA-3′
primer 2	5′-GCNGC(A/G)AA(T/C)TTNGC(T/C)T-3′
primer 3	5′-GAACTTGACATCACGGGCTGC-3′
primer 4	5′-CCGCAAGCCATGACACCGTGGTCTCCGATC-3′
primer 5	5′-AATTTGGCCTCCAAAGTCATCTTCT-3′
oligo(dT1)	5′-TCTGAATTCTCGAGTCGACATCTTTTTTTTTTTTTTT-3′
oligo(dT2)	5′-TCTGAATTCTCGAGTCGACATCT-3′
anchor primer (BRL)	5′-CUACUACUACUAGGCCACGCGTCGACTAGTACGGGIIGGGIIGGGIIG-3′
primer 6	5′-ACAATTCAAAATGATGAAATT-3′
primer 7	5′-GAGAAGCTTACAATTCAAAATGATGAAATT-3′

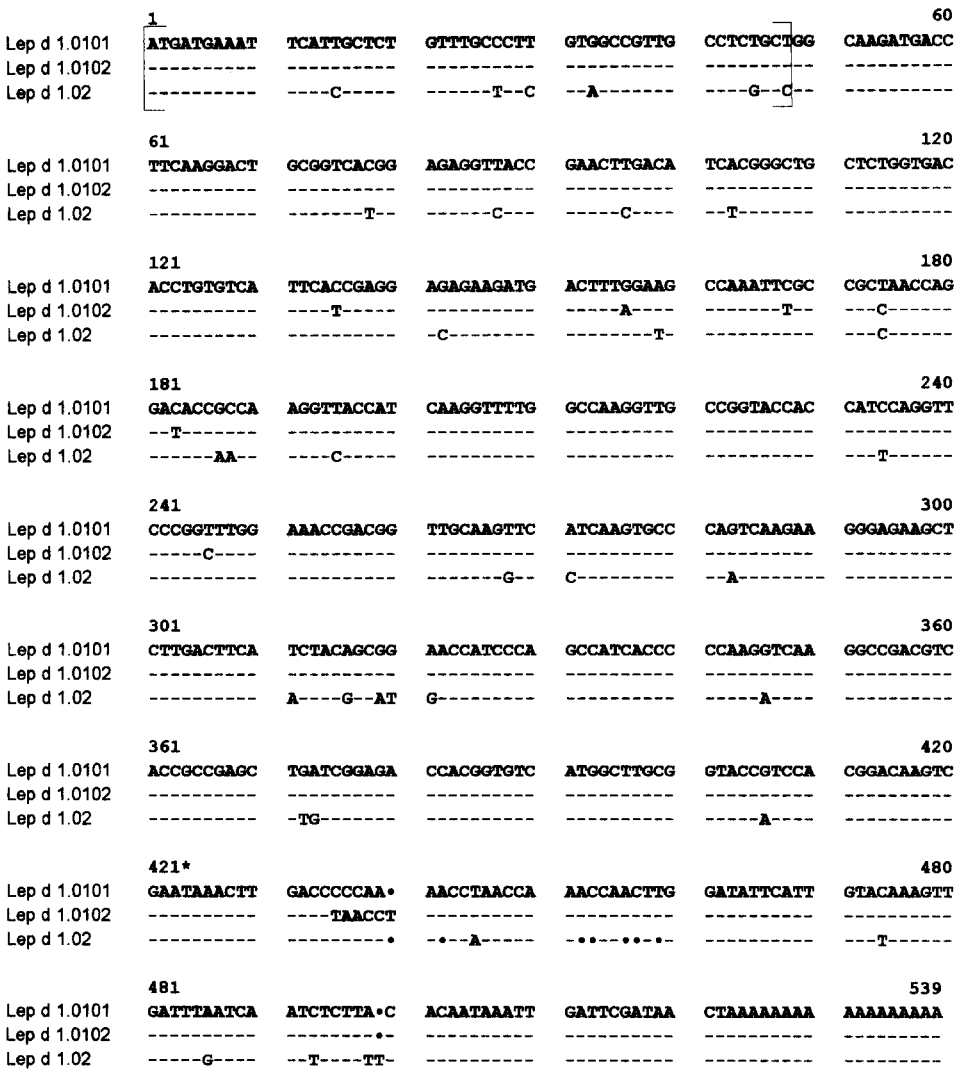

Fig. 1. Comparison of nucleotide sequences of Lep d 1 cDNA clones Lep d 1.0101, Lep d 1.0102 and Lep d 1.02. The sequences begin with the first nucleotide of ATG start codon. Identities with Lep d 1.0101 sequence are indicated with dashes. Gaps are introduces for alignment purposes. Brackets represent the leader sequence and the asterisks indicate the stop codon.

therefore could not distinguish different isoallergens and variants. Their analysis indicated six polymorphic sites, whereas our cDNA data identified 13 such sites; these also include the six previously identified by those authors. Although it is plausible that intermediate forms exist, they probably only exist in minute amounts.

The comparison of the cDNA with other nucleotide sequences using GCG BLASTN [17] showed a significant homology with group II allergens from *Dermatophagoides*. Lep d 1.01 and Lep d 1.02 showed 52% identity with Der p 2 and 57% with Der f 2. In earlier studies, we have shown a limited amount of allergenic cross-reactivity between Lep d 1 and *Dermatophagoi-*

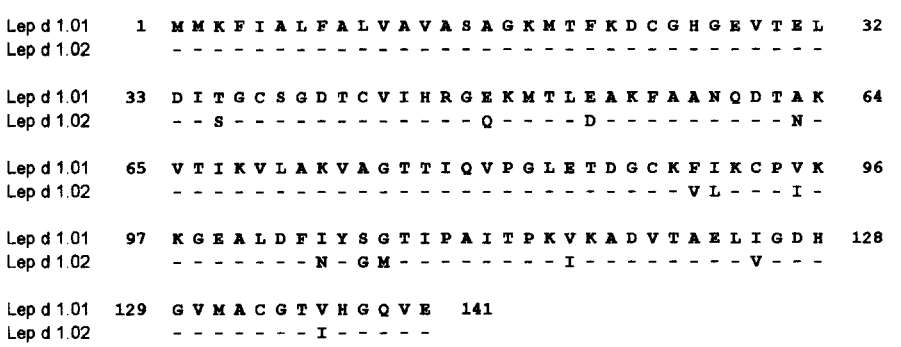

Fig. 2. Comparison of the amino acid composition of Lep d 1.01 and Lep d 1.02. Only those of Lep d 1.02 that differ from Lep d 1.01 are given.

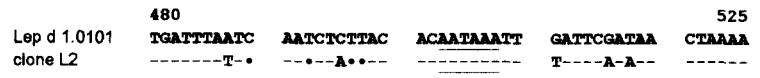

Fig. 3. Partial comparison of the 3′ non-coding region of Lep d 1.0102 and clone L2 [11]. Identities are indicated by dashes, gaps are introduced for alignment purposes, and the polyA signal is underlined.

des pteronyssinus [8]. In regard to the function of group 2 allergens of *Dermatophagoides*, there is accumulating evidence that they are lysozymes [18]. The function of Lep d 1 is so far unclear. However, the localization of Lep d 1 in the digestive tract of the mite and the fecal pellets [19] indicates that the allergen is involved in digestion which makes it reasonable to assume that it is an enzyme.

Lep d 1 is a major allergen of the storage mite *Lepidoglyphus destructor*, and most patients show IgE reactivity to it. This is the first report that substantiates the existence of polymorphic forms for Lep d 1 based on cDNA analysis.

We have successfully cloned and sequenced two isoallergens and variants of that protein. A baculovirus expression system utilizing insect cells should allow us to express the proteins in their native conformation. Knowledge of the sequence and the ability to express the proteins and fragments that include T and B cell epitopes will allow us to study both the different and common effects of the two isoforms. In addition, the expressed proteins will be useful as diagnostic tools and could eventually lead to immunotherapy with a mixture of T cell epitope peptides, rather than with highly allergenic mite proteins.

Acknowledgements: This work was supported by the Swedish Work Environment Fund, the Swedish Medical Research Council (Grant 16X-105), the Swedish Asthma and Allergy Association, the Immunotechnology Program funded by the Swedish National Board for Industrial and Technical Development, the Hedlund Foundation, the Hesselman Foundation, Åke Wibergs Foundation, Lars Hiertas Memorial Foundation, Consul Th. C. Berghs Foundation, the Karolinska Institute and by the National Institute of Allergy and Infectious Diseases, USA (AJ-17162).

References

[1] van Hage-Hamsten, M., Johansson, E., Wirén, A. and Johansson, S.G.O. (1990) Allergy 45, 409–419.
[2] Armentia, A., Tapias, J., Barber, D., Martin, J., de la Fuente, R., Sanchez, P., Salcedo, G. and Carreira, J. (1992) Ann. Allergy 68, 398–403.
[3] Iversen, M., Korsgaard, J., Hallas, T. and Dahl, R. (1990) Clin. Exp. Allergy 20, 211–219.
[4] van Hage-Hamsten, M., Johansson, S.G.O., Höglund, S., Wirén, A. and Zetterström, O. (1985) Clin. Allergy, 555–564.
[5] Cuthbert, O.D., Brostroff, J., Wraith, D.G. and Brighton, W.D. (1979) Clin. Allergy 9, 229–236.
[6] Blainey, A.D., Topping, M.D., Ollier, S. and Davies, R.J. (1989) J. Allergy Clin. Immunol. 84, 296–303.
[7] Iversen, M. and Dahl, R. (1990) Allergy 45, 81–85.
[8] Johansson E., Borgå, Å., Johansson, S.G.O. and van Hage-Hamsten, M. (1991), Clin. Exp. Allergy 21, 511–518.
[9] Olsson, S., Härfast, B., Johansson, S.G.O. and van Hage-Hamsten, M. (1994), Allergy 49, 620–625.
[10] van Hage-Hamsten, M., Bergman, T., Johansson, E., Persson, B., Jörnvall, H., Härfast, B. and Johansson, S.G.O. (1992), Lancet 340, 614.
[11] Varela, J., Ventas, P., Carreira, J., Barbas, J.A., Gimenez-Gallego G. and Polo, F. (1994), Eur. J. Biochem. 225, 93–98.
[12] Borriello, F. and Krauter, K.S. (1990), Nucleic Acid Res. 18, 5481–5487.
[13] Jansen, R. and Ledley, F.D. (1990) Nucleic Acid Res. 18, 5153–5156.
[14] L'Abbé, D., Belmaaza, A., Décary, F. and Chartrand, P. (1992), Imunogenetics 35, 395–397.
[15] Walker, R.A. and McConnell, T.J. (1994), Dev. Comp. Immunol. 18, 325–342.
[16] WHO/IUS Allergy Nomenclature Subcommittee World Health Organization, Geneva, Switzerland (1995), Clin. Exp. Allergy 25, 27–37.
[17] Altschul, S.F., Gish, W., Miller, W., Meyers, E.W. and Lipman, D.J. (1990), J. Mol. Biol. 215, 403–410.
[18] Stewart, G.A., Bird, C.H. and Thompson, P.J. (1992), J. Allergy Clin. Immunol. 87, 141–142.
[19] van Hage-Hamsten, M., Olsson, S., Emilson, A., Härfast, B., Svensson, A. and Scheynius, A. (1995), Clin. Exp. Allergy 25, in press.