

Sequence polymorphisms of the Der p 3 house dust mite allergen

W.-A. SMITH and W. R. THOMAS

TVW Telethon Institute for Child Health Research, Perth, Australia

Summary

Background The trypsin-like protein Der p 3 is a major allergen of *Dermatophagoides pteronyssinus*. Like other vertebrate and invertebrate trypsin-like molecules, isoelectric-focusing studies with the natural Der p 3 protein have indicated that several isoforms exist.

Objective To determine the extent of the sequence variation of the Der p 3 allergen and distinguish at the molecular level, whether the sequence isoforms represent allelic variants or multiple genes of the allergen.

Methods Five cDNA clones of Der p 3 have been isolated from a λ gt10 *D. pteronyssinus* library, using a radiolabelled polymerase chain reaction (PCR) Der p 3 P3WS1 probe and sequenced. Southern blot and inverse PCR analysis of Eco R1 digested genomic DNA was performed.

Results Southern blot analysis of Eco R1 digested genomic DNA showed that the DNA encoding Der p 3 was located on a single 3.5 kb fragment and inverse polymerase chain reaction analysis (PCR) of this DNA showed that there was only a single Der p 3 gene on this 3.5 kb fragment. The nucleotide sequence of one of the clones was identical to the original Der p 3 P3WS1 clone and two clones differed only in their 3' untranslated sequences. The other two contained nucleotide changes which lead to several substitutions at the amino acid level, both conservative and non conservative. Clone 3 had 98.7% identity with Der p 3 P3WS1. One clone for which the full sequence was not available (clone 4) had only 84.4% identity with the original clone and is therefore consistent with an isoallergen.

Conclusions These data along with our previous genomic sequence shows that for the most part, the Der p 3 allergen has only minor sequence variations (variants) although the isoallergen indicated by clone 4 needs further investigation. It is now evident that Der p 3 is encoded by a single gene and that most cDNA clones constructed from commercial mites show only minor sequence variation similar to that observed for the group 1 and group 2 house dust mite allergens.

Keywords: Der p 3, sequence polymorphisms, *Dermatophagoides*, trypsin proteins, house dust mite

Clinical and Experimental Allergy, Vol. 26, pp. 571–579. Submitted 3 July 1995; revised 4 September 1995; accepted 16 October 1995.

Introduction

Sequence polymorphisms have been identified for a number of common allergens including the major group 1 and 2 allergens from the house dust mite. Five cDNA clones of Der p 1 for example, contained one to three polymorphisms

with each clone being unique [1]. The significance of these polymorphisms in terms of antibody and T-cell recognition has not yet been determined but it may have some implications if peptides were to be used to modify T-cell function. In this respect, even single conservative changes inside and outside of the T-cell epitopes can have profound effects on the immune response [2–4]. Similar sequence polymorphisms have been reported for Der f 2. Four Der f 2 polymerase chain reaction (PCR) generated clones were

Correspondence: Dr W. R. Thomas, TVW Telethon Institute for Child Health Research, PO Box 855, West Perth, Western Australia 6872.

compared with two Der f 2 sequences published by Yuuki *et al.* [5]. Five of the six sequences differed from each other. Each sequence differed by only two or three amino acid residues, but two of the polymorphic sites had non-conservative substitutions. Yuuki *et al.* [6] isolated another three unique Der f 2 cDNA clones. Each clone had between one and four differences when compared with the pFL1 sequence [5]. Protein sequencing revealed that the protein variants could be detected in mite extracts [7].

Depending on the location of these sequence polymorphisms recognition of IgE and/or IgG may be affected. Nishiyama *et al.* demonstrated that all three polymorphic Der f 2 variants were equally recognized by IgE from a panel of sera from 14 house dust mite allergic patients. There was, however, a difference observed in the binding ability of some of the monoclonal antibodies between the three variants, possibly because monoclonal antibodies are more discriminatory. In contrast, Breiteneder *et al.* [8] showed that there was a marked difference in the ability of variants of Cor a 1, the major allergen of hazel pollen, to bind IgE from a panel of allergic sera. Each of the variants had identities of 96–99% with each other and yet there were very significant differences observed in their IgE binding patterns.

Isoelectric focusing studies [9] have suggested the existence of as many as eight isoforms of Der p 3 with pIs ranging from 3.7 to 8.6. A genomic DNA sequence encoding a sequence differing by four amino acids from the previously published sequence has also been described [10]. We now report the sequence of five Der p 3 cDNA clones and an analysis of genomic DNA by Southern blot and inverse PCR. The data shows that Der p 3 is encoded by one gene and apart from one interesting sequence requiring further investigation, there are only a few variant amino acids residues.

Materials and methods

Dermatophagoides mite cultures

Whole mite bodies were purchased from the CSL Ltd, Parkville, Australia and the spent mite medium was a gift from the same source.

Oligonucleotide

Oligonucleotides were synthesized using an Applied Biosystem PCR mate.

Extraction of genomic DNA from Dermatophagoides mites

Live mites were homogenized in 6 M guanidine hydrochloride and the soluble extract centrifuged over a 4.8 M

CsCl pad [11]. The DNA at the top of the pad was collected and extracted with phenol.

Southern blot of genomic DNA

Genomic *D. pteronyssinus* DNA was digested with Eco R1 (Toyobo) and then separated by agarose gel electrophoresis prior to transfer onto nylon membrane by a method adapted from Southern [12] using Hybond N+ nylon membrane (Amersham) under alkaline conditions. The hybridizations were performed as outlined below for the isolation of the Der p 3 cDNA isoforms from the λ gt10 cDNA library, with the exception that the air dried nylon membranes did not require baking at 80°C.

PCR conditions and thermal cycling parameters

All the PCR reactions performed included a negative control. The thermal cycling was conducted in a PTC-100 programmable thermal cycler (Bresatec, Australia). The 40 cycle programme was as follows: (5 min at 94°C, 2 min at 60°C, 2 min at 72°C [Elongation step]) 1X; (1 min at 94°C, 2 min at 60°C, 2 min at 72°C) 38X; (1 min at 94°C, 2 min at 60°C, 10 min at 72°C) 1X.

PCR with pfu polymerase

The PCR was performed according to the manufacturer's instructions (Stratagene La Jolla, CA, USA) with the addition of 20 pmoles of primers, 250 μ M dNTPs and up to 250 ng of target DNA.

Preparation of radiolabelled DNA probes by random priming

PCR amplified DNA probes were labelled using a Giga prime random prime kit (Bresatec Ltd, Adelaide, Australia) using α -³²P dATP and the reaction stopped by heating at 65°C for 10 min. The labelled probes were then purified using a Sephadex G50 column.

Isolation of Der p 3 cDNA clones from the λ gt10 cDNA library

Screening of the library [13] was performed with a random prime α -³²P labelled PCR probe. The library was plated at an 25 000 pfu per 150 mm petri dish and the phage lifted onto nitrocellulose filters (Schleicher & Schuell, Dassel, Germany) which were denatured and baked [12] for 1 h. The filters were prehybridized for 2 h in hybridization buffer (50% (v/v) Formamide, 25% (v/v) SSC (20 $\times$), 1% (v/v) Denhardt's solution (100 $\times$), 0.01% (w/v) Poly C, 0.01% (w/v) Poly U and 0.01% (w/v) denatured herring sperm DNA).

Hybridizations were performed in hybridization buffer at 42°C with 10⁶ cpm/mL labelled probe overnight. Filters were rinsed in 2 × SSC for 5–30 min at room temperature (RT) and then washed at RT at high stringency in each of the following buffers progressively at 15 min intervals: 2 × SSC, 0.1% sodium dodecyl sulfate (SDS); 0.5 × SSC, 0.1% SDS; and 0.1 × SSC, 0.1% SDS. A final 10 min wash was performed in 0.1 × SSC, 1% SDS buffer at 50°C. The filters were air dried and exposed to autoradiographic film (Fuji RX-100, Tokyo, Japan) in a Kodak X-Omatic cassette.

Isolation of DNA from λ gt10 Der p 3 clones

Phage DNA was prepared using a polyethylene glycol precipitation procedure [14].

Subcloning and DNA sequencing

λ gt10 phage DNA was digested with Eco R1 (Toyobo) and the released fragment (20–40 ng) ligated to Eco R1 digested M13mp19 or mp18 [12]. Purified *Pfu* PCR products (40–60 ng) were ligated to Sma 1 digested M13mp19 or mp18 [15]. Ligations were performed with T4 ligase (Promega Biotec, Madison, USA). The recombinant DNA was transformed using *E. coli* TG-1 and sequencing performed using the dideoxynucleotide chain termination method [16] and Sequenase (US Biochemicals). The oligonucleotide primers used for sequencing included the universal primers, and those listed in Table 1. The primers were used to sequence the M13 clones in both directions.

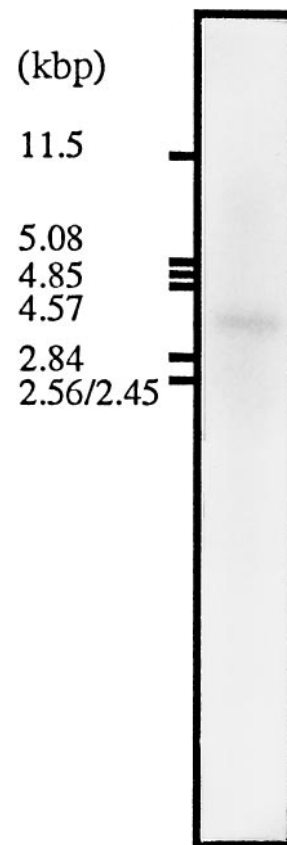


Figure 1. Southern blot analysis of Eco R1 digested *D. pteronyssinus* genomic DNA. Eco R1 digested genomic DNA was separated on a 1% agarose gel and the DNA transferred to a nylon membrane for hybridisation with the Der p 3 P3WS1 cDNA probe. The hybridising band is 3.5 kb in size. The sizes of the λ Pst 1 digested molecular weight markers are indicated in kb.

Table 1. PCR and sequencing oligonucleotides

		Location	Source#
1.	CAGGGAATTCATACACGTGTTGG	787–808	N
2.	TCTGAATTCGACATTCACCAGC	195–174	N
3.	TCAGAAAAAGCTTTGGCTGGTGAATCACCATATCAAAT	162–196	P
4.*	GCGGATCCATTGTTGGTGGTGAAAAAGC	144–169	N
5.	AAAGCTGTTGGATTACCA	456–473	N
6.	GAATCAACACCACCATTAGCAACATCACCACCGCAAATCAT	688–648	P
7.	TACATCCGATCCTTTTGC	491–474	N
8.*	AATCACTGTGAACGTTTGA	850–830	N

* Oligonucleotides used to isolate and sequence the cDNA and genomic DNA for the Der p 3 clones. Oligonucleotides 4* and 8* were used to PCR amplify the genomic DNA. The nucleotide sequence underlined in Oligonucleotide 4* represents a Bam H1 restriction site and the double underlined residues in oligonucleotides 1 and 2 represent an Eco R1 restriction site. Oligonucleotides 1 and 2 were used for inverse PCR. The sequence information, (#) both, nucleotide (N) and protein (P), used to generate these oligonucleotides has been described [13]. All oligonucleotides are types 5'–3' with the location of the residues within the original Der p 3 P3WS1 [13] sequence indicated.

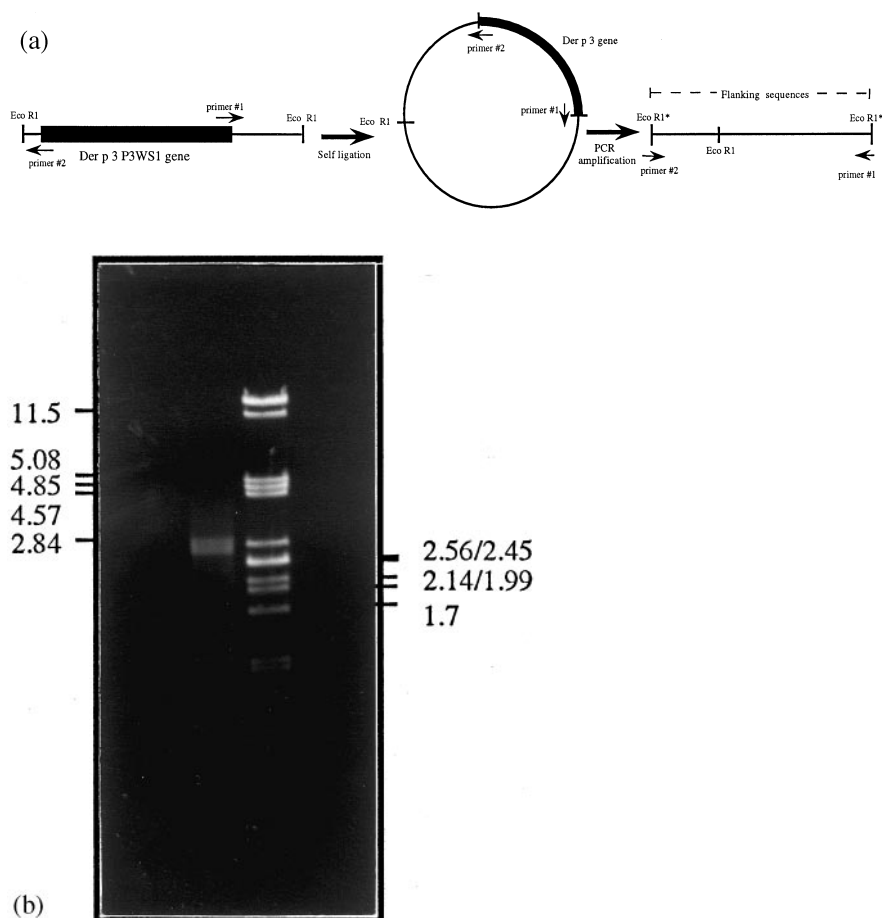


Figure 2. Inverse PCR analysis of the Der p 3 genomic DNA. (a) Diagrammatic representation of the inverse PCR procedure. The oligonucleotide primers both have an Eco R1 site near the 5' termini (*). (b) A 12 μ L aliquot of the PCR product was analysed by electrophoresis on a 1% agarose gel. A λ Pst 1 digested molecular marker was also run on the gel. The size of the fragments in the molecular marker is indicated in kb.

DNA and protein sequence analysis

Sequence analysis was performed using the Mac Vector software, version 3.5 (IBI, New Haven, USA). The pI values for the deduced amino acid sequences were calculated using the Mac Vector software, version 3.5 (IBI, New Haven, USA). Computation for sequence homology with other proteins was performed at the NCBI using the BLAST network service. The version of BLAST used, was BLAST 1.3.1OMP (15 December 1994).

Results

Southern Blot analysis of genomic DNA

The sequence encoding the mature protein Der p 3 P3WS1 clone was amplified by PCR using oligonucleotides 4 and 8 (Table 1). The single 800 bp band was gel purified and labelled with α - P^{32} dATP. Eco R1 digested *D. pteronyssinus* genomic DNA was transferred to a nylon membrane by Southern blot transfer and then hybridised with the radio-labelled Der p 3 P3WS1 probe. The DNA probe bound a

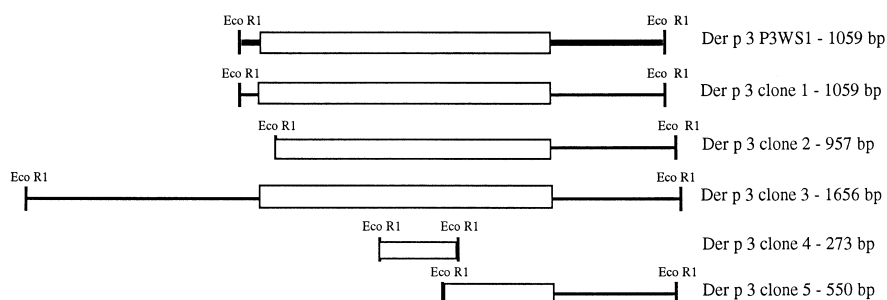


Figure 3. Diagrammatic alignment of the Der p 3 cDNA clones. The Eco R1 digested DNA fragments of the polymorphic clones were aligned to the original Der p 3 P3WS1 clone. The size of each of the fragments is also shown in bp. The boxed areas show the position of the sequence encoding the Der p 3 polypeptide.

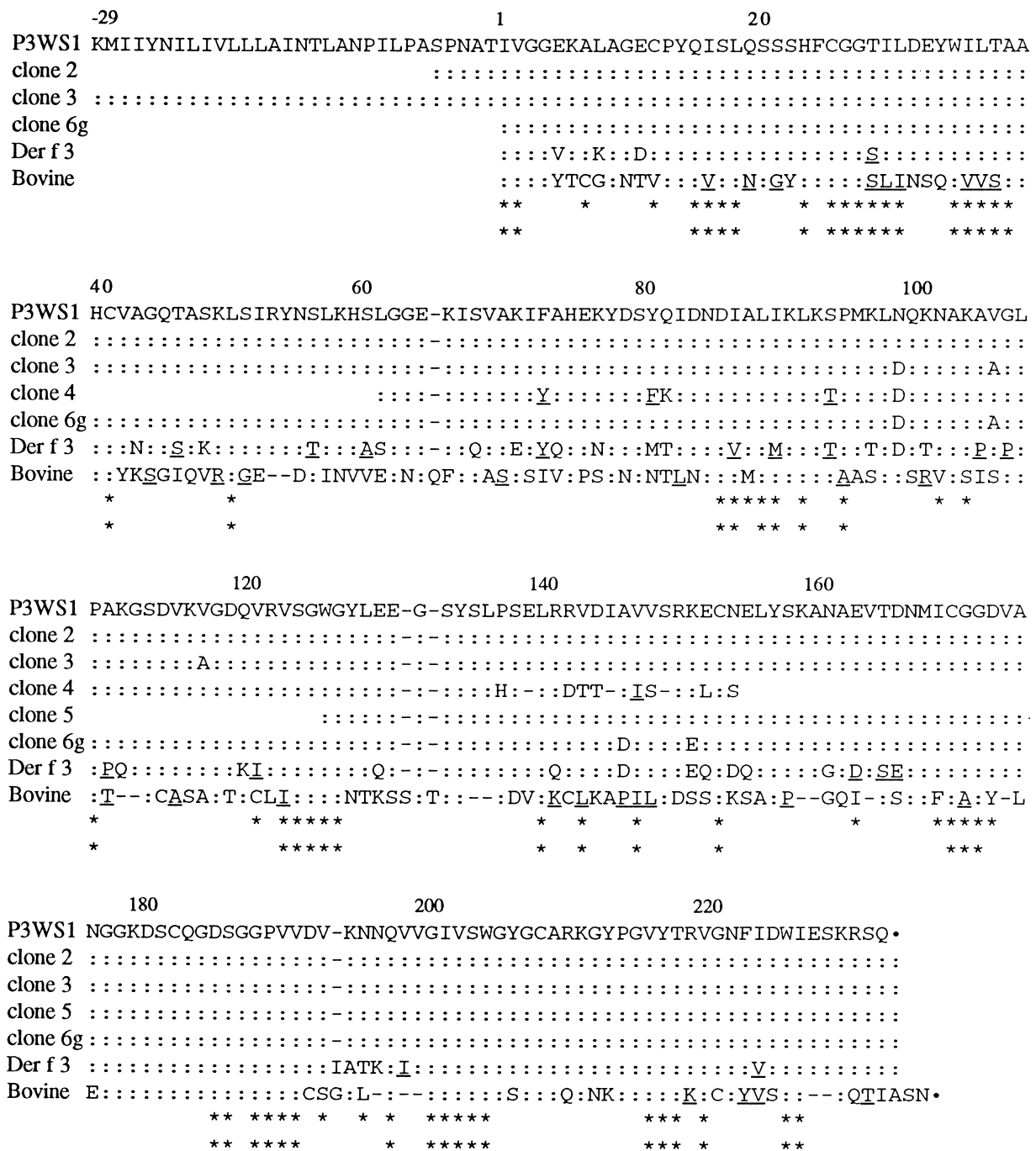


Figure 4. Amino acid sequence alignment of Der p 3 P3WS1 with the Der p 3 cDNA clones, Der p 3-6g [10], Der f 3 [10], and bovine trypsin [24]. The sequences numbering above the sequences is based on the P3WS1 sequence. Identical residues are illustrated by (:), conservative amino acid substitutions by an underlined single letter code for amino acid residues and non-conservative changes by the single letter code. The following groupings were used to determine similarity: A, G, S, T, P, C (unpaired); E, D; N, Q; H, K, R; I, L, M, V; F, Y, W; C_{S-S}. Spaces have been inserted for maximal protein sequence alignment. The (*) residues below the bovine sequence represent the residues in bovine trypsin which are inaccessible to solvents. A double (*) represents those Der p 3 P3WS1 residues predicted to be inaccessible to solvents by comparison with the bovine trypsin sequence (21). A (●) represents a stop codon.

(a)

clone 3 CATTAATTTTTCAGTTCAAAAAAAAAAGAAAAATAAAATTTAAAATTTTAAAAATTTGCAAAATCAAA 68

clone 3 ATCAAAATCAATAAAAAATTAATAAAAAAAATTATTAATGAAAAAAATATTATTATCAGCCAACTTTT 136

clone 3 CTGGTGATGGTGATGATGATAAATTGACCATCCATCCATCCATTCACCGTCCCCCACCCTGATTGTC 204

clone 3 CATCGATTTATTTGGATGGTCTAGCTGTTGTTGTTGTTGAATTTGCTGCAGCATCAACTTTTTTTATGA 272

clone 3 TGTGGTGGTTTAGTTTTACTTTTCTTTTCCCAAACCTGGTTTACTTGTAACACTACGTTTGTTTACAT 340

clone 3 TTTTACCATTGTTATTACTATTGTTGCTTGGTGCTGGTCTTGGAGATCTTGTGTTGTTGTTGTTGCA 408

clone 3 TTAAC TACAACAGCAGTTGAATCATTAAACATTGTTTAAATAAAACAAGTGTACGATAAACAATAAGGT 476

clone 3 TGTAATGAATCGAATTGAAAACATTTTTGAAATCTTCTTTATTCTTCAGTCGAACAATATTTTTTTTCG 544

clone 3 GTTTAAAAGTTGTTGAAAGTTGTTGTTTGATTCTGATTTTGAATAATAATATTTATTTGAAAACATT 612

P3WS1 . T . . . T

clone 1 . T . . . T

clone 3 GAAAAACATCAAAACAAC--TGTGTACAACCACAAAAGATGATCATCTATAATATTTTAATTGTTTTTA 680

P3WS1 TG . G . . . TG -G

clone 1 TG . G . . . TG -G

↓ ↓

clone 3 TTATTGGCCATTAATACATTGGCTAATCCAATTCTACCAGCATCACCAAATGCAACT 737

P3WS1

clone 1

(b)

clone 3 TTGATAAAAAAATTTTCTACATATAAAAAATTGGCAAATGATGCCATTATATGTTATCATTATGTTTC 67

P3WS1 C . A . . . G

clone 1 C . A . . . G

clone 2 G

clone 5 C . A . . . G

clone 3 TGGTGATTCTTTTTCTACTTTTATCATTTTATTTTATATTCGAGTTGAAAACAAACAATCCGGATTG 134

P3WS1 C AT

clone 1 C AT

clone 2 T

clone 5 C AT

clone 3 TATAGGGATTGTGTTTTTGTGTTTTTCAATTTAGAAAAATCCAAAATAAATAAATAAATAAACATTTTC 201

P3WS1 - . G . . - A T

clone 1 - . G . . - A T

clone 2 . . . T - - - -

clone 5 - . G . . - A T . . AAAA

clone 3 AACAAAAAAAAAAAAAAAAAAC 220

P3WS1

clone 1

clone 2 . . G

single DNA fragment of 3.5 kb in Eco R1 digested *D. pteronyssinus* genomic DNA (Fig. 1).

Inverse PCR of the Der p 3 gene

Inverse PCR was used to amplify the 3' and 5' sequences flanking the Der p 3 gene. Two oligonucleotide primers, #1 and #2 (Table 1) were designed from the Der p 3 cDNA sequence to amplify these regions of the Der p 3 Eco R1 genomic fragment as a single molecule. Eco R1 digested *D. pteronyssinus* DNA was self-ligated at low concentrations designed to produce circularised DNA and then amplified by PCR as outlined in Fig. 2. The PCR product was a doublet band of approximately 2.8–2.9 kb in size (Fig. 2).

Sequence analysis of Der p 3 cDNA clones

A further five cDNA clones were obtained by screening the cDNA library with the DNA from the Der p 3 clone P3WS1 (Fig. 3). Clone 1 was identical to P3WS1 and since the library had been amplified, it could be the same clone. Clone 2 had an identical nt sequence for the coding region and clone 3 which was characterised by an extended 5' untranslated region had 6/786 nt differences in the coding region. Two partial clones were obtained one of which, clone 5, was identical to P3WS1 in the coding region and the other, clone 4, had 41/271 nt changes, in the coding region available for comparison. The nt differences in clones 3 and 4 also altered some of the translated amino acid sequence residues (Fig. 4). Clone 3 had three changes, all non-conservative, at positions 98, 105 and 116. Two of these changes were the same as residues in the previously described genomic DNA clone 6 g which had four residues different to P3WS1. Clone 4 was quite different from all the other clones with identity in 76/90 available residues (84.4%). The changes at residues 81, 98, 136, 141, 143, 147, 151 and 153 were non-conservative and the changes at 72, 80, 93 and 146 were conservative.

Nucleotide sequences upstream of the coding region were available for clone 1 (55 nts) and clone 3 (648 nts) (Fig. 5a). The sequence of clone 1 (identical to P3SW1) differed from the corresponding sequence of clone 3 by 9/55 nts. The nucleotide sequences for both the signal peptide and activation peptides for clone 3 was identical to the P3WS1

sequence. Nucleotide sequences downstream were available for clones 1, 2, 3 and 5. In the 210 3' nts available for comparison, clones 1 and P3SW1 differed from clone 5 by 5 nt and from clone 2 and 3 by 17 and 11 nts respectively. Some of the differences of clones 2 and 3 were the same, they differed from each other, by 10 nts.

Discussion

Molecular polymorphisms occur from: posttranslational modifications such as glycosylation, gene families encoding related polypeptides and allelic variation. Der p 3 does not contain potential N-glycosylation sites but this does not exclude other posttranslational modifications. The Southern blot analysis used here shows that all of the DNA hybridizing with a Der p 3 probe was contained within a 3.5 kb Eco R1 fragment. The 1656 bp cDNA for clone 3 leaves little space for more than one Der p 3 gene on this 3.5 kb fragment given that the gene would also contain an intron of approximately 87 bp. The result of the inverse PCR, which analysed the size of the regions flanking the coding sequence of Der p 3 on the Eco R1 fragment, was consistent with these assertions. The major products, a doublet of 2.8/2.9 kb in combination with the 786 bp coding region and 87 bp intron would be approximately equivalent to predicted size of the Eco R1 fragment. The presence of two genes in either relative orientation would give a more complex pattern with smaller bands predominating. The doublet suggests that the more sensitive PCR technique, demonstrates a minor allele which is not detectable by Southern.

The sequence of the coding region of the five newly described cDNA clones, as well as P3SW1 and the genomic DNA clone, 6 g, show that with the exception of clone 4, the Der p 3 polypeptide and nt sequence is largely conserved. All of these Der p 3 cDNA clones have complete identity with the 18 residue N-terminal sequence published by Stewart *et al.* [9] but differ at two out of the 30 residues in the N-terminal sequence published by Smith *et al.* [13]. Clones P3SW1, 1, 2 and 5 had identical amino acid sequences. Clone 3 and the genomic clone 6 g differed by three and four of the 232 amino acid residues, two of the substitutions for each of these clones being identical. This degree of polymorphism is similar to that described for the group 1 and 2 allergens where the different cDNA clone

Figure 5. Nucleotide sequence alignments of the non-coding regions of the Der p 3 molecules. Identical residues are indicated by (.) and spaces have been inserted for maximal protein sequence alignments. (a) Alignment of the 5' untranslated nucleotide sequences. The ATG encodes the +1 residue of the signal peptide. The arrow (↓) highlights the cleavage site of the signal peptide with the residues which encode the activation (or pro) peptide being underlined. The cleavage site between the activation peptide, and the mature protein sequence, is indicated by the arrow (↓). (b) Alignment of the 3' untranslated nucleotide sequences.

differ by between one and three residues. The substitutions are typically between one of two residues at a defined position, for example positions 98 can be D or N and 105 can be A or V, and the substitution is usually to a residue found in Der f 3. The sequence variations between the clones is less than 3% so they would therefore be classified as variants by the IUIS allergen nomenclature guidelines [17]. It is proposed that P3SW1 (and clones 1, 2 and 5 with the same amino acid sequence) is Der p 3 0101, clone 3 with 98.7% identity is Der p 3 0102 and the genomic clone 6 g with 98.3% identity is Der p 3 0103.

Data from isoelectricfocusing analysis have shown that Der p 3 has at least eight isomers with widely varying pI values. The calculated pIs of Der p 3 0101, 0102, and 0103 here are, 8.5, 8.1 and 6.1 respectively. It is therefore possible to have large changes in pI with only minor sequence variation.

With the exception of both clones 1 and P3SW1 which are identical, the Eco R1 fragment of the other cDNA clones differed in length due to different 5' and 3' truncations. From examining the sequence it appears that this is mostly due to incomplete cDNA synthesis. The 3' truncation of clone 4, however, is possibly due to a sequence polymorphism where a single base change from the P3SW1 sequence could have created an Eco R1 site.

The 3' and 5' untranslated regions of the cDNA clones show more sequence variation than the coding region. Heterogeneity in nucleotide sequence and length is expected from the lack of selection pressure [18]. It has also been documented that substitutions in the non-coding regions of eukaryotic DNA can have be as high as 1% [19]. The polymorphisms in these regions are shown here because they may in the future be useful markers for genetic variation in mites and they also give a comparison of the conservation of the coding region. It is noteworthy that clone 2 which had the most 3' polymorphism with respect to P3SW1, had an identical coding sequence at both the nt and amino acid level. In contrast, Chua *et al.* found that for the five cDNA clone of Der p 1 examined, all of them had identical 3' untranslated regions despite having some amino acid sequence difference [1].

Clone 4 which contained 288 nts encoding 90 amino acids is clearly different to the other cDNA and genomic clones. This degree of identity (84%) with P3WS1 if extended over the whole molecule would classify clone 4 as an isoallergen (Der p 3, 0201), although a full clone would be required for this designation. Given that the Southern data obtained by hybridization with P3SW1 DNA shows only a single Der p 3 gene it must be considered as an unusual allele. Also only 3/16 of the substitutions were to the residues found in Der f 3 which itself shows a similar degree of identity (81%) to P3WS1. Given the large sequence variation observed, further study of this allergen would be important since it

could be expected to markedly affect T-cell epitopes and hence immune responsiveness of people with different MHC haplotypes.

Considerable variation in the primary amino acid sequence of trypsin can be tolerated without severely affecting the tertiary structure [20]. The bovine trypsin [21] residues determined to be inaccessible to solvent, i.e. those residues buried within the molecule, have a high degree of identity with the corresponding residues found in the Der p 3 P3WS1 sequence (Fig. 3). This would suggest that the two molecules, which have 46% identity, have a very similar structural core. Perhaps significantly, all of the Der p 3 sequence polymorphisms lie outside of these core residues. The tertiary structures of the Der p 3 molecules would be expected to be very similar to bovine trypsin given that very few insertions or deletions are required for maximal protein sequence alignment. The residues involved in both the catalytic activity and substrate binding, and the residues surrounding these areas are very conserved. These regions form the hydrophobic core of the β -barrels or the loops that comprise the active site and binding pockets of the enzyme [22]. Secondary structure designations by Matthews *et al.* [23] for bovine trypsin identified a trypsin domain connection between residues 98 and 115. Despite only a very limited degree of identity between the Der p 3 and bovine trypsin molecules within this region, the domain connection is of interest because two of the Der p 3 polymorphic sites fall within this region. The first is residue 98 at the N-termini of the connection, with a non-conservative substitution of an asparagine residue (Der p 3, P3WS1, bovine trypsin) to an aspartic acid residue (Der p 3, 3.3, 3.4, 3.6 g and Der f 3). The second, a non-conservative substitution, is at residue 105 in the centre of the domain connection with a change from a valine residue (Der p 3, P3WS1, Der f 3) to an alanine residue (Der p 3, 3.3, 3.6 g).

The data presented here, show that Der p 3 is encoded by a single gene and that most cDNA clones constructed from commercial mites show only minor sequence variation similar to the group 1 and group 2 mite allergens. The variants, Der p 3.0101, 0102 and 0103, show about 98% sequence identity with P3WS1 but give a wide range of calculated pIs which are consistent with the published values obtained by chromatofocussing studies [9]. It should be noted that these three variants may account for three of the eight isoforms detected with affinity purified natural Der p 3 [9], with only a few amino acid differences between the three sequences. On this basis, it is quite feasible therefore that other allelic variants exist with minor sequence variations that could also account for the remaining five isoforms. One cDNA clone out of the 6 cDNA and one genomic clones did, however, show considerably less sequence identity (84%), but as this clone was only a partial sequence it should be investigated further. The

biological effects of the variants can now begin to be examined and in this context it should be noted that this can be variable, with Breiteneder [8] showing marked differences in IgE binding from hazel Cor a 1 variants with 96–99% identity and Nishiyama [7] finding no effect on IgE binding of variant Der f 2 allergens.

Acknowledgments

This work was supported by the National Health and Medical Research Council of Australia.

References

- Chua KY, Kehal PK, Thomas WR. Sequence Polymorphisms of cDNA Clones Encoding the Mite Allergen Der-p-I. *Int Arch Allergy Immunol* 1993; 101:364–8.
- Evavold BD, Allen PM. Separation of IL-4 production from the Th cell proliferation by an altered T cell receptor ligand. *Science* 1991; 252:1308–10.
- Liu Z, Williams KP, Chang YH, Smith JA. Single amino acid substitution alters T cell determinant selection during antigen processing of *Staphylococcus aureus* nuclease. *J Immunol* 1991; 146:438–43.
- Liu Z, Williams KP, Chang YH, Smith JA. Immunodominance: A single amino acid substitution within an antigenic site alters intramolecule selection of T cell determinants. *J Immunol* 1993; 151:1852–8.
- Yuuki T, Okumura Y, Ando T, Yamakawa H. Cloning and sequencing of cDNA's corresponding to mite major allergen Der f II. *Jpn J Allergy* 1990; 39:557–61.
- Yuuki T, Okumura Y, Ando T et al. Cloning and expression of cDNA coding for the major house dust mite allergen Der f II in *Escherichia coli*. *Agric Biol Chem* 1991; 55: 1233–8.
- Nishiyama, C, Yuuki T, Usui Y et al. Effects of amino acid variations in recombinant Der f II on its human IgE and mouse IgG recognition. *Int Arch Allergy Immunol* 1994; 105:62–9.
- Breiteneder H, Ferreira F, Hoffmann-sommergruber K et al. Four Recombinant Isoforms of Cor a I, the Major Allergen of Hazel Pollen, Show Different IgE-Binding Properties. *Eur J Biochem* 1993; 212:355–62.
- Stewart GA, Ward LD, Simpson RJ, Thompson PJ. The group III allergen from the house dust mite *Dermatophagoides pteronyssinus* is a trypsin-like enzyme. *Immunology* 1992; 75:29–35.
- Smith W, Thomas WR. Comparative analysis of the genes encoding the group 3 allergens from *Dermatophagoides pteronyssinus* and *Dermatophagoides farinae*. *Int Arch Allergy Immunol* (in press).
- Stewart GA, Thomas WR. In vitro translation of messenger RNA from the house dust mite *Dermatophagoides pteronyssinus*. *Int. Archs Allergy Appl Immunol* 1987; 83: 384–9.
- Maniatis T, Fritsch EF, Sambrook J. *Molecular Cloning. A laboratory manual*. Cold Spring Harbor: Cold Spring Harbor Laboratory Press, 1982.
- Smith WA, Chua KY, Kuo MC et al. Cloning and sequencing of the *Dermatophagoides pteronyssinus* group III allergen, Der p III. *Clin Exp Allergy*. 1994; 24:220–8.
- Chua KY, Stewart GA, Thomas WR et al. Sequence analysis of cDNA coding for a major house dust mite allergen, Der p1. Homology with Cysteine proteases. *J Exp Med* 1988; 167:175–82.
- Messing J. New M13 vectors for cloning. *Method Enzymol* 1983; 101:20–78.
- Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain terminating inhibitors. *Proc Natl Acad Sci USA* 1977; 74:5463–7.
- King TP, Hoffman D, Lowenstein H et al. Allergen nomenclature. *Int Arch Allergy Appl Immunol* 1994; 105:224–33.
- Bowcock A, Cavalli-Sforza L. The study of variation in the Human genome. *Genomics* 1991; 11:491–8.
- Levitt R. Polymorphisms in the transcribed 3' untranslated region of eukaryotic genes. *Genomics* 1991; 11:484–9.
- Huang R, Hellman L. Genes for mast-cell serine protease and their molecular evolution. *Immunogenetics* 1994; 40: 397–414.
- Jurasek L, Olafsen R, Johnson P, Smillie L. In *Proteolysis and Physiological Regulation*. (D. Ribbons, K. Brew, Eds.) Academic Press Inc, New York, 1976, vol. 11, pp. 93–123.
- Bazan JF, Fletterick RJ. Viral cysteine proteases are homologous to the trypsin-like family of serine proteases: Structural and functional implications. *Proc Natl Acad Sci USA* 1988; 85:7872–6.
- Matthews DA, Smith WW, Ferre RA, Condon B et al. Structure of human rhinovirus 3C protease reveals a trypsin-like polypeptide fold, RNA-binding site, and means for cleaving precursor polypeptide. *Cell* 1994; 77:761–71.
- Chambers JL, Stroud RM. The accuracy of refined protein structures: comparison of two independently refined models of bovine trypsin. *Acta Crystallogr* 1979; 358.