

Cloning and sequencing of the group 6 allergen of *Dermatophagoides pteronyssinus*

B. J. BENNETT and W. R. THOMAS

TVW Telethon Institute for Child Health Research, Perth, Australia

Summary

Background The allergen Der p 6 from *Dermatophagoides pteronyssinus* has been described by substrate affinity as mite chymotrypsin.

Objective The aim of this paper was to describe a cDNA clone encoding the allergen.

Methods cDNA was cloned from a λ library using oligonucleotides published for Der p 6.

Results The clone P6.1.1 had the N-terminal amino acid residues as reported for Der p 6, and at position 189 the Ser was conserved in chymotrypsin for substrate specificity as well as the catalytic triad of His57, Asp102 and Ser195 for serine proteases. The estimated Mr was 24.9 K and it was 37% identical to the trypsin allergen Der p 3.

Conclusion cDNA encoding Der p 6 has been described.

Keywords: Der p 6, house dust mite, *Dermatophagoides pteronyssinus*, cDNA, cDNA sequence, chymotrypsin

Clinical and Experimental Allergy, Vol. 26, pp. 1150–1154. Submitted 14 November 1995; revised 12 March 1996; accepted 20 April 1996.

Introduction

The most significant indoor allergens are those of the house dust mite, which appear to play a major role in the induction of asthma, allergic dermatitis and allergic rhinitis. Many of the allergens are known to be enzymes and, as shown for Der p 1, are probably secreted by the mite during digestion and released in the faeces [1]. Four allergens of *Dermatophagoides pteronyssinus* have proteolytic activity. Der p 1 is a cysteine protease similar to papain [2]. Der p 3 is a trypsin-like serine protease [3,4]. Der p 6 is a chymotrypsin-like serine protease [5] and there is a recently characterized collagenolytic serine protease, Der p 9 [6,7].

Yasueda *et al.* first characterized the presence of a 25–27 kDa protein, purified from spent mite extracts, that contained chymotryptic-like activity [5]. They reported that 41% of 88 mite-allergic patients were IgE-positive for Der p 6, compared with 90% IgE-positive for Der p 1 [3]. King *et al.* used a RAST assay to compare the IgE binding reactivity of 37 patients to Der p 3 and Der p 6. They found that almost 60% of the mite-allergic patients had IgE

reacting to Der p 6, but the binding was, as described by Yasueda, considerably less than to the major allergens Der p 1 and Der p 3 [7].

The N-terminal amino acid sequence of Der p 6 was determined by Yasueda *et al.* [5], and this enabled us to isolate a cDNA clone encoding Der p 6 from a cDNA library. The clone, P6.1.1, was found to contain a sequence consistent with chymotrypsin and with the expected homology to trypsin and other serine proteases.

Materials and methods

Construction of the *D. pteronyssinus* λ gt10 cDNA library

Polyadenylated mRNA was isolated from *D. pteronyssinus* supplied by CSL Ltd (Melbourne, Australia) and was used to synthesize cDNA using a commercially available kit (Amersham Int., Aylesbury, UK). *EcoRI* linkers (New England BioLabs, Beverly, MA, USA) were added to the cDNA, which was subsequently ligated into λ gt10 arms and grown in *Escherichia coli* Y1073 to select for recombinants. *Escherichia coli* JP777 was used as the host strain to generate λ gt10 phage stock for storage and screening purposes [4].

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

Oligonucleotide screening and sequence

Oligonucleotide probes were synthesized using an Applied Biosystem polymerase chain reaction (PCR) mate (Applied Biosystems, Foster City, CA, USA), and the Der p 6 sequencing oligonucleotides were synthesized commercially (Bresatec, Adelaide, Australia) using the preferred codon usage table published by Smith *et al.* [4]. Two Der p 6-specific degenerate oligonucleotides were synthesized: 23 mer (P6-1) 5' GAAGC(A/C/T) GAAGC(A/C/T)CC(A/C)-TT(C/T)CAAAT 3' and (P6-2) 5' CC(A/C)TT(C/T)C-AAAT (C/T)TC(A/G/T)TT(A/G)ATG-AAA 3'. Plaque hybridization was performed as described previously [4]. The filters were washed once at room temperature for 20 min with 6× SSC and then three times at 42°C for 20 min each in 2× SSC, 0.1% sodium dodecyl sulphate (SDS). The filters were air-dried and autoradiographed.

Positive plaques were rescreened until pure. The cDNA from a clone hybridizing with both probes was amplified twice by PCR with high fidelity Pfu polymerase (Stratagene, La Jolla, CA, USA) and gt10 primers, and the DNA subsequently digested with *EcoRI* (Promega, Madison, WI, USA) and ligated into the *EcoRI* site of pUC19. The ligation mixes were transformed into *E. coli* (XL1-Blue).

Sequencing and analysis

The DNA from XLI-Blue clones was purified by alkaline lysis as recommended by Perkin-Elmer (Foster City, CA, USA), and automatically sequenced using dye-labelled dideoxynucleotides and Biosystems 373. Two Der p 6-specific sequencing primers were synthesized: (P6.3) 5' GGGATCCGTTATTGGTGGTCAAGAT 3' and (P6.4) 5' GGGATCCGATCCGGAACCATACAA 3', corresponding to amino acid residues 1–6 and 84–89, respectively. Sequence analysis was carried out using SeqEd (Applied Biosystems) and MacVector (IBI, New Haven, USA) software. Sequence homology was performed at the NCBI using the BLAST network service, and multiple sequence alignments performed using the ClustalW v5.1 public domain software.

Results and discussion

Clone P6.1.1 contained a 995 bp insert that encoded an open reading frame of 281 amino acids (Fig. 1). The published N-terminal sequence of mature Der p 6 is (A)IGXQ(P)AAE(A)(E)APFQI(S)LMK, where residues in parentheses have been tentatively assigned [3]. The corresponding sequence in clone 6.1.1 was VIGGQD-AAEAEAPFQISLMK, which was identical to the published sequence except that the N-terminal, tentatively assigned alanine is a valine in clone P6.1.1 and the tentatively assigned proline is aspartate. The mature form of clone P6.1.1 had a

calculated Mr of 24.9 K and a pI of 5.2. The calculated Mr agrees well with the 25 K determined by Yasueda *et al.* [5], and it should be noted the 231 residues and calculated Mr of 25 K is the same as Der p 3 which migrates as a 30 K molecule on SDS–polyacrylamide gel electrophoresis (PAGE).

Homology searches of the GenBank database showed that the open reading frame of clone P6.1.1 had the expected homology to chymotrypsin, trypsin and other serine proteases. P6.1.1 aligned with bovine (chymotrypsin A) [8], cod [9], shrimp [10], blowfly [11] and human chymotrypsins [12], together with bovine trypsin [8] (Fig. 2). Bovine chymotrypsin A and bovine trypsin were included in the alignment because their 3-dimensional structure has been determined and comparison with these sequences enabled us to comment on the structure of P6.1.1. Most of the cysteine residues are remarkably conserved in the chymotrypsins of the different species. Based on the known structure of the bovine chymotrypsin A and trypsin [8], the cysteine pairs that form disulphide bonds could be assigned in P6.1.1, namely cysteines 30/46, 155/171 and 180/206. Der p 6 does not have the expected cysteine pair at positions 123/190. This cysteine pair appears in bovine, cod and human chymotrypsins but is absent in the shrimp and blowfly chymotrypsin sequences (Fig. 2).

P6.1.1 contained a serine residue at position 178, which corresponds to serine 189 of bovine chymotrypsin A. This serine residue is conserved in all chymotrypsins and is critical for substrate specificity. The small, hydroxyl side chain of serine protrudes into the binding pocket and permits the hydrophobic bulk of the phenylalanine, tyrosine and tryptophan side chain of the peptide substrate to enter the active site. Instead of a serine residue at position 189, trypsins have an aspartate residue. This aspartate residue is highly conserved in all trypsins sequenced to date. The carboxylic acid side chain of aspartate binds to the arginine or lysine side chain of the peptide substrate, thus conferring a different substrate specificity to the trypsin active site. The three residues that form the catalytic charge relay system (His57, Asp102 and Ser195 of bovine chymotrypsin A) [9,11] were also present in P6.1.1 (His45, Asp91 and Ser184).

P6.1.1 was compared with Der p 3 (mite trypsin), and the mature forms of the allergens had a 37.2% (86/231) identity (Fig. 3). The cleavage site of the pro sequence of Der p 6 was bordered by a glycine residue instead of a trypsin substrate residue, such as arginine or lysine which is found in most other chymotrypsins. Der p 3 also lacked the trypsin substrate residue at the proenzyme cleavage site and instead had a threonine.

Several authors have reported that Der p 3 or Der f 3 binds IgE in 80–90% of allergic sera [1, 7, 12, 13], which is greater than the 40% reported for Der p 6. Moreover, King *et al.* [7] found a large difference in the degree of binding by

```

-50                                -40
caattcttataacaatATG ATT AAA ATT ATA ACT ACA ATT ATC CTG ATT ATA ACC GTT GTT GTT 64
Met Ile Lys Ile Ile Thr Thr Ile Ile Leu Ile Ile Thr Val Val Val
-30                                -20
GAT TGT CGT TTT CCA CGT ATT CTT CAA CCA AAA TGG TCA TAT CTA GAT TCA TTA CCG GCT 124
Asp Cys Arg Phe Pro Arg Ile Leu Gln Pro Lys Trp Ser Tyr Leu Asp Ser Leu Pro Ala
-10                                -1 1
AGT TCA TCA ATG ATG AAT GAT AAT AGT AGT CCG ATT GCT GGT GTT ATT GGT GGT CAA GAT 184
Ser Ser Ser Met Met Asn Asp Asn Ser Ser Pro Ile Ala Gly Val Ile Gly Gly Gln Asp
10                                20
GCT GCC GAA GCT GAA GCA CCA TTT CAA ATT TCC TTG ATG AAA GAT TAT CTT ATT ATG AAA 244
Ala Ala Glu Ala Glu Ala Pro Phe Gln Ile Ser Leu Met Lys Asp Tyr Leu Ile Met Lys
30                                40
AGT CAT ATG TGT GGT GGT TCA TTG ATT TCA GCA TCG ACT GTC ATT ACA GCT GCA CAT TGT 304
Ser His Met Cys Gly Gly Ser Leu Ile Ser Ala Ser Thr Val Ile Thr Ala Ala His Cys
50                                60
ACT GAT GGA CAA AAA GCA TCA TCA CTT TCA ATA CGT TAT GGT ACA AAT AAA CGA ACA TCA 364
Thr Asp Gly Gln Lys Ala Ser Ser Leu Ser Ile Arg Tyr Gly Thr Asn Lys Arg Thr Ser
70                                80
TCC AAA TAT GCT GAT CTT TCA ATC AAA CGT ATT GTT CAA CAT GAA TCA TAT GAT CCG GAA 424
Ser Lys Tyr Ala Asp Leu Ser Ile Lys Arg Ile Val Gln His Glu Ser Tyr Asp Pro Glu
90                                100
ACC ATA CAA AAT GAT ATA TCA TTA TTG ATT TTG TCA CAA CCG GTT AAA CCA AGT TCA AAT 484
Thr Ile Gln Asn Asp Ile Ser Leu Leu Ile Leu Ser Gln Pro Val Lys Pro Ser Ser Asn
110                                120
GTT GAT TTT ATT GAT ATT GAA ACC AAA GAT ATT GGT GAT GGT GAA AAA GTA ACT ATT TAT 544
Val Asp Phe Ile Asp Ile Glu Thr Lys Asp Ile Gly Asp Gly Glu Lys Val Thr Ile Tyr
130                                140
GGT TGG GGT TTA ACC GAT GGT AAT ACA AAT AAT TTA CCC GAT AAT TTA CAA AAA GGT TCA 604
Gly Trp Gly Leu Thr Asp Gly Asn Thr Asn Asn Leu Pro Asp Asn Leu Gln Lys Gly Ser
150                                160
ATG ACA ATT GTC GGT AAT GAT AAA TGT AAT GAT AAA TGG GGT TCG GTT AAT ACT ATT CAT 664
Met Thr Ile Val Gly Asn Asp Lys Cys Asn Asp Lys Trp Gly Ser Val Asn Thr Ile His
170                                180
CCG GGT ATG ATT TGT GCA TTG GAT GCA ACA CAA TCT GGT TGT AAT GGT GAT TCT GGT GGA 724
Pro Gly Met Ile Cys Ala Leu Asp Ala Thr Gln Ser Gly Cys Asn Gly Asp Ser Gly Gly
190                                200
CCA TTA GTT TCG GCT GAT CGT AAA TTA ACC GGT ATT GTA TCA TGG GGA CCA AGT AAA TGT 784
Pro Leu Val Ser Ala Asp Arg Lys Leu Thr Gly Ile Val Ser Trp Gly Pro Ser Lys Cys
210                                220
CCA CCA GGT GAA TAT ATG AGT GTA TTT ACA CGA CCA CAA TAT TAT AGT GAT TGG ATT GCT 844
Pro Pro Gly Glu Tyr Met Ser Val Phe Thr Arg Pro Gln Tyr Tyr Ser Asp Trp Ile Ala
230
AAA AAT ATT GTT CAA TAAacgatttttggttaaatccaaccaaccaaatccaacaatatcttttgtaattgg 918
Lys Asn Ile Val Gln *

aattacttgaattttaaacaagccaaatataacatttttcggaattttttaattggaaaaaaaaaaaaaaaaaaaaa 995

```

Fig. 1. Sequence of Der p 6 cDNA clone. Full nucleotide sequence with the aligned translated protein sequence. The first 16 amino acids encode a putative signal sequence, the next 34 amino acids encode a proenzyme sequence including an early cysteine residue (also present in chymotrypsins of other species), and 231 amino acids that encode the mature form of Der p 6. Residues involved in substrate binding are in bold, the stop codon is TAA and marked with *, and the AATAAA polyadenylation signal is underlined.

RAST assay. However, some differences exist in the data reported for the allergenicity of Der p 3. Yasueda *et al.* found a similarly low frequency and degree of IgE binding for Der p 3 and 6 [5]. Genetic differences or differences in the recruitment of patients may account for these discrepancies, and the sequences and recombinant material now available may help to resolve the relative responses at the T- and B-cell level. Although both these allergens are serine proteases they do not cross-react with hyperimmune serum [5] and the 37% sequence identity is not consistent with T- or B-cell cross-reactivity. It would nevertheless be interesting to compare the relative responses to these two related allergens from the same source.

Acknowledgement

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

References

- 1 Stewart GA, Lake FR, Thompson PJ. Faecally derived hydrolytic enzymes from *Dermatophagoides pteronyssinus*: physicochemical characterisation of potential allergens. *Int Arch Allergy Appl Immunol* 1991; 95:248–56.
- 2 Chua, KY *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.

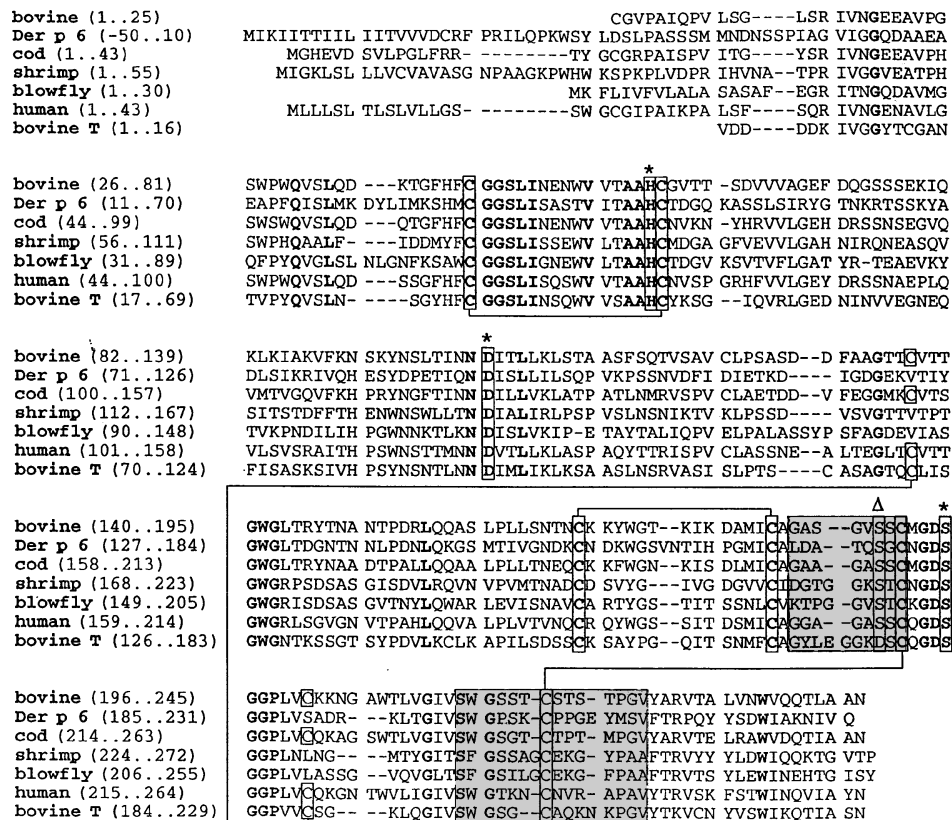


Fig. 2. Multiple sequence alignment. Multiple sequence alignment of six chymotrypsin sequences including Der p 6, bovine (chymotrypsin A), cod, shrimp, blowfly and human, and one bovine trypsin sequence. The conserve residues are in bold, dashes represent alignment gaps, residues binding to substrate are boxed and labelled (*), the residue conferring substrate specificity is boxed and labelled (Δ), residues forming the substrate binding pocket are shaded, and cysteine residues forming disulphide bonds are boxed and joined with a line.

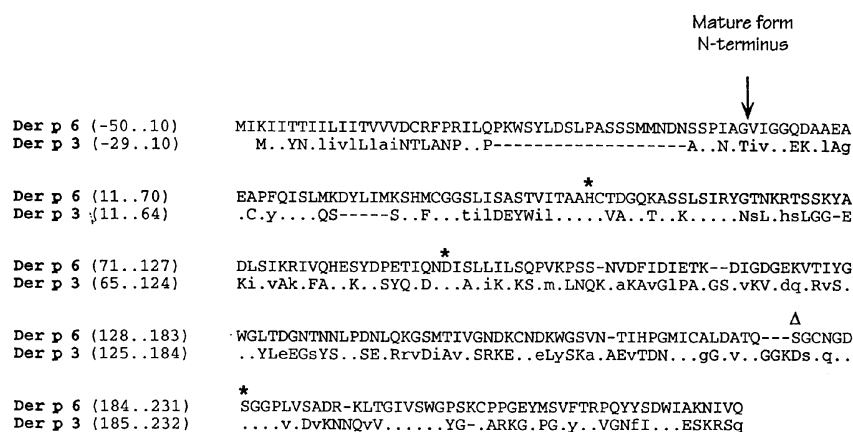


Fig. 3. Sequence comparison of Der p 3 and Der p 6. Sequence alignment of Der p 3 (mite trypsin) to Der p 6 (mite chymotrypsin). Identical residues in Der p 3 are represented by a dot, dashes are alignment gaps, conservative residues are in lower case, and non-conservative residues are in upper case. Catalytic residues are marked * and the residue conferring substrate specificity is marked Δ .

- 3 Stewart GA, Ward LD, Simpson RJ, Thompson PJ. The group III allergens from the house dust mite *Dermatophagoides pteronyssinus* is a trypsin-like enzyme. *Immunology* 1992; 75:29–35.
- 4 Smith W, Chua KY, Kuo MC, Rogers BL, Thomas WR. Cloning and sequencing of *Dermatophagoides pteronyssinus* group III allergen Der p 3. *Clin Exp Allergy* 1994; 24:220–8.
- 5 Yasueda H, Mita H, Akiyama K, Shida T, Sugiyama S. Allergens from *Dermatophagoides* mites with chymotryptic activity. *Clin Exp Allergy* 1993; 23:384–90.
- 6 Stewart GA, Kollinger MR, King CM, Thompson PJ. A comparative study of three serine proteases from *Dermatophagoides pteronyssinus* and *D. farinae*. *Allergy* 1994; 49:553–60.
- 7 King CM, Simpson RJ, Moritz RL, Thompson PJ, Stewart GA. The isolation and characterisation of a novel serine protease allergen from the house dust mite *Dermatophagoides pteronyssinus*. *ANZ J Med* 1995; 25:272.
- 8 Hartley BS. Homologies in serine proteases. *Phil Trans Roy Soc Lond B*. 1970; 257:77–87.
- 9 Guthmundsdottir A, Oskarsson S, Eakin AE, Craik CS, Bjarnason JB. Atlantic cod cDNA encoding a psychrophilic chymotrypsinogen. *Biochim Biophys Acta* 1994; 1219: 221–14.
- 10 Sellos D, van Wormhoudt A. Molecular cloning of a cDNA that encodes a serine protease with chymotryptic and collagenolytic activities in the hepatopancreas of the shrimp *Penaeus vanamei* (Crustacea, Decapoda). *FEBS Lett* 1992; 309:219–24.
- 11 Casu RE, Pearson RD, Jarmey JM, Cadogan LC, Riding GA, Tellam RL. Excretory/secretory chymotrypsin from *Lucilia cuprina*: purification, enzymatic specificity and amino acid sequence deduced from cDNA. *Insect Mol Biol* 1994; 3:201–11.
- 12 Thomas WR, Chua KY, Smith W. Molecular polymorphisms of house dust mite allergens. *Exp Appl Acarology* 1992; 16:153–64.
- 13 Ando T, Homma R, Ino Y *et al.* Trypsin-like protease of mites: purification and characterisation of trypsin-like protease from mite faecal extract *Dermatophagoides farinae*. Relationship between trypsin-like protease and Der f III. *Clin Exp Allergy* 1993; 23:777–84.