Published in final edited form as:

J Mol Biol. 2009 February 20; 386(2): 520–530. doi:10.1016/j.jmb.2008.12.049.

Crystal structures of mite allergens Der f 1 and Der p 1 reveal differences in surface exposed residues that may influence antibody binding

Maksymilian Chruszcz¹, Martin D. Chapman², Lisa D. Vailes², Enrico A. Stura³, Jean-Marie Saint-Remy⁴, Wladek Minor^{1,*}, and Anna Pomés^{2,*}

¹ Department of Molecular Physiology and Biological Physics, University of Virginia, 1340 Jefferson Park Avenue, Charlottesville, VA 22908, USA

² INDOOR Biotechnologies, Inc., 1216 Harris Street, Charlottesville, VA 22903, USA

³ Département d'Ingénierie et d'Etudes des Protéines, Commissariat à l'Energie Atomique Saclay, Gif-sur-Yvette, France

⁴ Center for Molecular and Vascular Biology, University of Leuven, Leuven, Belgium

Summary

The Group 1 mite allergens, Der f 1 and Der p 1, are potent allergens excreted by *Dermatophagoides farinae* and *Dermatophagoides pteronyssinus*, respectively. The human IgE antibody responses to the Group 1 allergens show more cross-reactivity than the murine IgG antibody responses which are largely species-specific. Here, we report the crystal structure of the mature form of Der f 1, which was isolated from its natural source, and a new, high-resolution structure of mature recombinant Der p 1. Unlike Der p 1, Der f 1 is monomeric both in the crystalline state and in solution. Moreover, no metal binding is observed in the structure of Der f 1, despite the fact that all amino acids involved in Ca²⁺ binding in Der p 1 are completely conserved in Der f 1. Although Der p 1 and Der f 1 share extensive sequence identity, comparison of the crystal structures of both allergens revealed structural features which could explain the differences in murine and human IgE antibody responses to these allergens. There are structural differences between Der f 1 and Der p 1 which are unevenly distributed on the allergens' surfaces. This uneven spatial arrangement of conserved *versus* altered residues could explain both the specificity and cross-reactivity of antibodies against Der f 1 and Der p 1.

Keywords

mite allergy; Der f 1; Der p 1; antibody; asthma

Introduction

Inhalation of house dust mite allergens is one of the most important risk factors associated with the development of allergic disease, including rhinitis, atopic dermatitis and asthma. The two principal mite species *Dermatophagoides pteronyssinus* and *D. farinae* belong to the

*Corresponding authors: Wladek Minor (E-mail: wladek@iwonka.med.virginia.edu), Anna Pomés (E-mail: apomes@inbio.com).

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Pyroglyphidae family and excrete allergens including Der p 1 and Der f 1 (Group 1), respectively, which are carried to the lung in mite fecal particles. Humidity is the most important limiting factor for mite growth and deserts or areas of high altitude have low levels of mites.¹ *D. pteronyssinus* is broadly distributed in houses in Western Europe, Japan, North America, New Zealand and Australia. In continental regions of Europe and most of the U.S., both species frequently co-exist, but *D. farinae* can tolerate drier environmental conditions.² Most mite allergic patients (>80%) have IgE antibodies against Group 1 mite allergens^{3,4}. Der p 1 and Der f 1 are cysteine proteases of the clan CA⁵ and family C1. They share both 81% sequence identity and antigenic cross-reactivity.

Despite the high amino acid sequence identity between Group 1 allergens (81%), most of the monoclonal antibodies (mAb) raised against either Der p 1 or Der f 1 were species-specific (3–6%).^{6–8} In contrast, the degree of cross-reactivity of human IgE antibody responses to Group 1 allergens, although variable, is higher (34–90%).⁸ One murine cross-reacting epitope was identified using anti-Der f 1 mAb 4C1.⁷ This mAb inhibited IgE antibody binding to Der p 1 by ~40%, suggesting that the epitopes for 4C1 mAb and a human IgE ab in Der p 1 overlap. These monoclonal antibodies provide tools to study the antigenic determinants in Group 1 allergens by X-ray crystallography.

The crystal structures of the proenzyme and mature forms of recombinant Der p 1 were recently determined.^{9,10} Here we report the crystal structure of mature natural Der f 1 obtained from mite culture and a new, high-resolution structure of recombinant Der p 1. Both allergens are secreted with an N-terminal pro-region that is auto-catalytically cleaved under acidic conditions upon enzyme maturation. The pro-region blocks not only the catalytic activity but also conformational IgE antibody binding epitopes.¹¹ Reports have indicated that proteolytic activity contributes to allergenicity, mostly in the case of Der p 1. Disruption of tight junctions in lung epithelium, and cleavage of receptors (CD23, CD25) favor a Th2 response and induction of release of pro-inflammatory cytokines from bronchial epithelial cells, mast cells and basophils.¹² These effects may promote IgE antibody synthesis and inflammation on lung epithelium, which could explain why mite allergens are strongly associated with asthma. Although a reduction of skin barrier function by proteolytic activity of Der f 1 has been reported, much less is known about its pro-inflammatory effects.¹³

Results and Discussion

Overall structure of Der f 1

Der f 1 was crystallized in space group P4₁ with three protein molecules (chains A, B and C) in the asymmetric unit. The protein is monomeric. The overall fold of Der f 1 is characteristic for papain-like cysteine proteases, and similar to that observed for Der p 1, as expected from their high sequence identity (Fig. 1). The Der f 1 molecule consists of two globular domains connected by a flexible linker. Residues 1–223 could be traced in the electron density of all protein chains, with exception of Ala3 from chain C. Superposition (using secondary-structure matching¹⁴ as implemented in COOT¹⁵) of Der f 1 (chain A) on mature Der p 1 (PDB code: 2AS8, chain A) gave Ca RMSD values of 0.6 Å (over 222 residues) (Fig. 1C), while superposition of Der f 1 and proDer p 1 (PDB code: 1XKG) gave Ca RMSD value 0.5 Å (over 221 residues).

The pattern of disulfide bonds observed in Der f 1 (Cys4–Cys118, Cys32–Cys72 and Cys66–Cys104; numbering of amino acids in whole text is based on Der f 1 sequence) is the same as in Der p 1. Der f 1 shows only 5 polymorphisms (Fig. 1A) compared to 23 in Der p 1.¹⁶ Analysis of the Der f 1 structure reveals that the amino acid differences between the crystallized Der f 1.0101 and other isoforms are located on the surface of the allergen. Almost all mutations, except His162Arg, modify residues identical in both Der f 1 and Der p 1. However, the mutation

His162Arg (variant 1.0102) increases the surface/sequence similarity between the two allergens. Analysis of the electron density shows that the form of Der f 1 crystallized is that of the most abundant variant, that is the variant 1.0101.

Der f 1 has one N-glycosylation motif in its mature form. Electron density observed near residues Asn53 of A and B was interpreted as an N-acetylglucosamine; for chain C the density was too weak to build a model for the carbohydrate moiety. According to mass spectrometry results (data not shown), the molecular weight of the Der f 1 used for crystallization (25,462 Da) is about 270 Da greater than the calculated molecular weight of mature Der f 1, which corresponds to the approximate mass for one N-acetylglucosamine residue. In two of the three Der f 1 molecules in the asymmetric unit (chains A and B), the Asn53 residues are exposed to solvent such that carbohydrate group(s) fit without disrupting the crystal packing. In chain C, the space for a carbohydrate is limited by crystal packing. Thus spatial constraints may limit the population of molecules that form the crystal to those with shorter carbohydrates. It may also explain why the mean temperature factor of chain C is higher, and its corresponding electron density is of lower quality compared to chains A and B. The results presented here, along with those previously published, suggest that different types of short carbohydrates are probably present at the N-glycosylation site in Der f 1. The N-glycosylation site in the pro-domain of Der p 1 is absent in Der f 1. The role of carbohydrate content is not well understood for either allergen. Different effects on enzyme maturation have been reported,^{17–19} but all the studies agree that non-glycosylated mutants appear to have lower solubility, and that N-glycosylation does not influence antibody binding.^{18,20,21}

New crystal form of Der p 1

In addition to the previously reported orthorhombic crystal form¹⁰ of Der p 1, a new crystal form was obtained, which also contains two molecules in the asymmetric unit (space group C2). The relative arrangement of the allergen molecules in the asymmetric units of both crystal forms is similar and the structures superpose with a Ca RMSD value of 0.2 Å. The overall structure is the same except for the conformation of few surface amino acids. The higher resolution data has allowed us to better identify the metallic divalent cation as Ca^{2+} ²² rather than Mg^{2+} as previously reported (Fig. 2).⁷ Moreover, in the new crystal form the catalytically active Cys35 residue is oxidized. This observation is consistent with the need for a reducing agent to activate the enzyme before measuring its catalytic activity. Oxygen atoms from sulfinic group interact with Nε2 from Gln28 side chain and nitrogen atom from His170 main chain. Additionally, a significant conformational change of His171 is observed (Fig. 2A). The change of His171 conformation could contribute to the catalytic mechanism of the protease as explained below or is caused by oxidation of Cys35.

Structural comparison between Der f 1 and Der p 1

Active site—The active site of Der f 1 lays between the two globular domains, as in Der p 1, and other enzymes of clan CA⁵ and family C1. The substrate binding and catalytic residues (Gln29, Cys35, His171 and Asn191) lay in a cleft formed by juxtaposition of both domains (Fig. 1B·2A). Although Group 1 mite allergens were purified³ and cloned²¹ over twenty years ago, the detailed mechanism of their action is still not fully understood. Residues Cys35 and His171, which form a thiolate-imidazolium ion pair (Table 1), are probably most critical for the enzymatic activity, as this ion pair is conserved in other cysteine proteases.^{23,24} The roles of the other amino acids located proximally to Cys35 and His171 have not yet been determined.

The active sites of Der f 1 and Der p 1 are very similar, as are their patterns of substrate specificity.^{25,26} Both enzymes prefer to bind small aliphatic residues in position P2, charged residues in position P1, and small hydrophobic or hydrophilic residues in position P1'.^{27,28} The most significant amino acid difference near the catalytic site is Gln152 in Der f 1 versus

Arg151 in the equivalent position of Der p 1. This difference may explain why a plant cystatin, a proteinaceous cysteine protease inhibitor, isolated from chestnut, inhibits Der f 1, but not Der p 1.²⁸ Cystatins block access to the catalytic cleft, but do not bind covalently to the cysteine residue from the active center, as do small molecular inhibitors, such as E-64. Proteolytic activity of Der f 1 and Der p 1 has been reported to significantly influence immunological signaling and contribute to allergic response.¹³ Their action is especially visible in the lungs,²⁶ and Der f 1 has also been reported to affect skin.¹³

The active site in the proenzyme form of Der f 1 and Der p 1 is blocked by the allergen N-terminal prodomain. The prodomains were shown to be competitive inhibitors of the mature forms of the enzymes,²⁹ but in contrast to other cysteine proteases, Der p 1 was able to completely degrade its propeptide. The non-standard features of the Der f 1 and Der p 1 prodomains (as compared to other papain-like proteins), namely their length and lack of well-conserved ERFNIN motifs,⁹ suggest that Group 1 mite allergens may form a new C1 sub-family of cysteine proteases.

Metal binding site—Unlike in Der p 1 structures, no metal binding was observed in the structure of Der f 1 (Fig. 2B). If the Der f 1 initially bound a metal ion, it may have been removed from the protein during crystallization, as the crystallization solution contained both L-arginine and ammonium sulfate ions at high concentrations. Moreover, during optimization of crystallization conditions, EDTA was added to the solution from which the best crystals were obtained. The amino acids involved in metal binding in Der p 1 (Asp57, Leu58, Glu60 and Glu92; Der f 1 numbering) are all conserved in Der f 1. The differences in conformations of the metal binding residues in Der p 1 compared with the equivalent residues in Der f 1, which lacks a metal, are small (Fig. 2B). These results suggest that the presence or absence of metal ions does not influence the overall protein architecture.

The metal ion is located approximately 12 Å from Cys35 and 10 Å from Gln29, and the role of the metal ion in Der p 1 is unknown. Both Mg²⁺ and Ca²⁺ ions were reported³⁰ to increase catalytic activity of papain from *Carica papaya* and presence of the metal ions, according to CD measurements, also did not influence the secondary structure of the enzyme. Surprisingly, the superposition of the Der f 1, proDer p 1 and mature Der p 1 models reveals a large cluster of structurally conserved water molecules (Fig. 2C) being located opposite (relatively to the catalytic residues) to the substrate binding site. The water molecules fill a funnel-shaped cavity, which has its broader end localized near the metal binding site. The narrow end of the funnel is located near Ser192 and forms a hydrogen bond with the hydroxyl group of the serine. Ser192 also interacts through a H-bond with Gln29 (Table 1). In the case of Der f 1, the water molecule bound near Ser192 is 4.3–4.7 Å apart from Cys35 and 3.8–4.5 Å from His171. It is possible that this water molecule may be used during catalytic reactions, especially during the deacetylation step, in which the cysteine-substrate covalent intermediate is hydrolyzed.³¹ Another interesting feature, observed in the crystal structure of human cathepsin K, is that the position of the water molecule bound to the serine equivalent to Ser192 is positioned almost identically to the corresponding water in Der f 1 and Der p 1. It suggests that the water molecule either assists the serine residue in its enzymatic activity or that it is important for the structure of the cysteine proteases.

Oligomeric state—We determined, by gel filtration at pH=7.5, Der f 1 to be monomeric in solution. Gel filtration experiments¹¹ performed in slightly more acidic conditions (pH=6.5), also indicated both Der f 1 and Der p 1 were monomeric in solution. In contrast, Der p 1 was reported to be dimeric in solution (pH=7.5 and pH=8.0) and in the crystal¹⁰ (see Fig. 3), however our analysis of the crystal structures of Der p 1 (PDB codes: 2AS8, 3F5V) showed that the decrease in accessible surface area of Der p 1 upon dimerization (as defined by Ponstingl *et al.*^{32,33}) is about 600–700 Å², which is slightly below the cut-off value (856

Å²) proposed for discrimination between homodimeric and monomeric proteins.³³ The oligomeric analysis of the Der p 1 and Der f 1 structures with the PISA³⁴ and PITA³³ servers also consistently predicts that both proteins are monomeric.

The dimeric form previously proposed for Der p 1¹⁰ most probably corresponds to the inactive form of the protein, because the loop containing residues Tyr166 and Gln167 from one chain sterically blocks the catalytic cleft in the second protein molecule forming the putative oligomeric assembly (Fig. 3B). Analysis of the possible oligomerization interface in terms of amino acid conservation shows that almost all residues forming the interface in Der p 1 are also preserved in Der f 1. It is not clear if the catalytically inactive dimer previously proposed for mature Der p 1 is physiologically relevant or influence stability of the protein. Oligomerization among cysteine proteases is extremely rare and currently there is only one known cysteine protease that forms a physiological oligomer: the tetramer of cathepsin C.³⁵ All other enzymes from this family are monomeric.

Surface properties and antibody binding—As mentioned above, over 80% of the amino acid sequences of Der f 1 and Der p 1 are identical. Comparison of the crystal structures of the enzymes reveals that the sequence differences between Der f 1 and Der p 1 are not distributed evenly in relation to their molecular surfaces (Fig. 4). The differences in the surfaces of Der f 1 and Der p 1 are more dramatic due to lower sequence identity when only surface amino acids are taken into account. The residues that are different are mostly located in loop and α -helical regions and residues identical in Der f 1 and Der p 1 occupy around 70% of the molecules surface area. The surfaces of the Der f 1 and Der p 1 molecules near the catalytic Cys35 residue appear almost identical, whereas other surface regions of the molecules differ significantly in amino acid composition of the molecular surface. Significant structural differences are apparent as several amino acid patches distal from the catalytic site on the surfaces of both allergens. The four largest patches (Fig. 4) were composed of the following residues (numbering according to Der f 1 sequence): patch I (184, 206, 208-210, 216), patch II (11, 13, 124, 126, 127, 179-181), patch III (8, 9, 15, 19, 20, 133) and patch IV (91, 93, 94, 96, 109, 111 and 112). Such an uneven distribution of conserved *versus* variable residues may in part explain the differences in murine and human IgE antibody responses to these allergens.

Structures of Der f 1 and Der p 1 are of similar to structures of human cysteine proteases, with cathepsin K (EC. 3.4.22.38) being the most similar human protein to Der f 1 in terms of overall architecture and sequence. Although the sequence identity between human cathepsin K and Der f 1 is only 36% (similarity 50%), structures of Der f 1 and the human enzyme are similar with low RMSD values over around 200 superposed residues (e.g. 1.3 Å for PDB structure 1U9V, and 1.4 Å for PDB structure 1Q6K). The degree of surface conservation between Der f 1 and Der p 1, is ~ 70% of identical residues, whereas for Der f 1 and cathepsin K the surfaces are only 30% identical. Only a few residues in the surface area, identical between Der p 1 and Der f 1, are also identical with cathepsin K (Fig. 4), occupying 10% of Der f 1 surface. No IgE binding epitopes would in general be expected in common areas between Der f 1 and cathepsin K. However, few residues predicted to be part of the IgE binding epitopes,^{36,37} were also identical to cathepsin K. Epitope mapping, based either on peptide fragments or mimotopes, is not as precise as mapping based on solving the X-ray crystal structure of allergen-antibody complexes.

Analysis of the amino acid sequences of four monoclonal antibodies against Group 1 mite allergens—The variable regions of the heavy and light chains of four monoclonal antibodies (10B9, 5H8, 6A8, 4C1) against Group 1 mite allergens were sequenced (Fig. 5). Two mAb recognize species-specific epitopes in Der p 1 (10B9 and 5H8) and one binds to a species-specific epitope in Der f 1 (6A8).⁶ The 4C1 mAb raised against Der f 1 binds to a common epitope on both allergens. Additionally, 10B9 inhibits the binding of 4C1 to Der

p 1 and vice versa, indicating that the epitopes recognized by both mAb on Der p 1 either overlap, since the binding of one mAb completely prevents the binding of the other, or lie sufficiently close to one another that binding of either mAb induces conformational changes that make the other site inaccessible.⁶ Surprisingly, 10B9 is species-specific and does not inhibit the binding of 4C1 to Der f 1, so structural differences between both epitopes are expected.

Amino acid sequences corresponding to the complementarity determining regions (CDR) from the light chain -loops L1(26-32), L2(50-52) and L3(91-96)-, and from the heavy chain -CDR1 (31-35b), CDR3(50-65) and CDR3(95-102)- were compared among pairs of antibodies. The variability between the light chains of the mAb pairs 4C1-6A8, 5H8-10B9, 4C1-5H8 and 4C1-10B9 is much higher (~11 times) than between each mAb and its highest homolog in the databases, which is zero for at least one loop. The highest variability in amino acid sequence and length is in the CDR3 from the heavy chain, which is thought to mostly determine antibody diversity and contribute to antigen affinity and specificity. Curiously, the CDR2 of 4C1-10B9 has much less variability (4/16 different residues versus 10-11/16-17) than the other three mAb pairs, which could be related to the cross-reactivity between these two mAb (Figure 5, underlined). The fact that the antibodies 4C1 and 10B9 were raised from different mice by immunization to either Der f 1 or Der p 1, respectively, makes these similarities more remarkable. Determination of the primary structure of mAb against Group 1 mite allergens provides the basis for understanding antibody-allergen interaction.

Conclusions

The crystal structure of Der f 1 is the first reported structure of a natural mite allergen. Der f 1 is a monomeric protein, with a metal binding site that is structurally equivalent to the one described for Der p 1, but lacks a bound ion. The structure reveals four surface patches that differ between Der f 1 and Der p 1, and areas in common that are the putative basis for cross-reactivity between the two allergens. The Der f 1 and Der p 1 structures also show the presence of conserved water molecules in a funnel-shaped cavity, close to the metal binding site at one end and the catalytic site at the other. These water molecules could be biologically significant and may serve to catalyze the cysteine protease activity of the allergen.

The analysis of the variable regions of the heavy and light chains of a panel of four monoclonal antibodies against Der f 1 and Der p 1 revealed parts of the variable regions that may be involved in cross-reactivity. The Der f 1 and Der p 1 structures, together with antibody binding studies, will identify the amino acids essential for IgE antibody binding and will provide a basis for further understanding of antibody-allergen interactions, and subsequently, the development of new recombinant vaccines for treatment of mite allergic disease.

Materials and Methods

Purification and crystallization

Der f 1 was purified from spent *D. farinae* mite culture extract (~100 g per 1L PBS) using affinity chromatography through a 4C1 mAb column. Purified Der f 1 protein (Lot 29084) in PBS buffer was stored at -80°C. Der f 1 (1.2 mL of 1.2 mg/mL in PBS) was incubated on ice for 1h with E-64 (40 µL of 10 mg/mL in water; SIGMA). The solution was dialyzed overnight into 0.5 M L-arginine and 0.05 M NaCl at pH 7.5, concentrated using an Amicon Ultra concentrator (Milipore) with a 10,000 Da molecular weight cut-off and applied on Superdex 200 column. After gel filtration the protein was concentrated to 11 mg/mL. The sample was passed through a 0.22 µm filter and used for crystallization. Crystallization was performed at 296 K using the hanging drop vapor diffusion method in NEXTAL plates. The protein solution was mixed with well solution in a 1:1 ratio. An initial hit was obtained with the Wizard I screen

(Emerald Biosystems) in condition #33 containing 2 M $(\text{NH}_4)_2\text{SO}_4$, 0.2 M Li_2SO_4 , and 0.1 M CAPS at pH 10.5. The conditions were subsequently optimized, and the best diffracting crystals were grown from a solution containing 2 M $(\text{NH}_4)_2\text{SO}_4$, 0.12 M Li_2SO_4 , 0.004 M EDTA, and 0.1 M CAPS at pH 10.5. Crystals obtained from the initial screen and optimized crystals were grown to a similar size. Tracking and analysis of the crystallization experiments were performed with the *Xtaldb* crystallization expert system.

Der p 1 was purified and crystallized as reported previously⁷. Prior to data collection; crystals were soaked for 20 minutes in solution composed of 5 μL DMSO, 1 μL 1M β -mercaptoethanol, 5 μL 1M Tris-HCl pH 8.0 and 150 μL 50% w/v monomethyl polyethylene glycol 2,000.

Data collection, structure solution and refinement

For data collection crystals of Der f 1 (of approximate size $20 \times 15 \times 10 \mu\text{m}^3$) were cooled in a cold nitrogen stream (100K) without cryoprotectant. Data collection was performed at beamline 19-ID³⁸ of the Structural Biology Center at the Advanced Photon Source (APS). Data were collected and processed with HKL-2000³⁹ (Table 2). Der f 1 crystallized in the $P4_1$ space group with three molecules in the asymmetric unit. Analysis of the diffraction intensities revealed that all Der f 1 crystals were partially merohedrally twinned. The Merohedral Crystal Twinning Server (<http://nihserver.mbi.ucla.edu/Twinning/twinning>) was used to determine the twinning fraction, and for all crystals the twinning fraction was close to 0.4 (twinning law: 0 1 0 1 0 0 0 -1). Despite the small size of the crystals, they initially diffracted to around 2.5Å, but the data resolution was quickly affected by radiation damage. To obtain a complete data set, data from two crystals grown in the same drop were merged. A partial molecular replacement (MR) solution was obtained with MOLREP⁴⁰ as incorporated into HKL-3000.⁴¹ A search model for MR was prepared using SWISS-MODEL⁴² and the structure of Der p 1 (PDB code 2AS8) was used as a template. The partial model contained two molecules of Der f 1, and after rigid body refinement, it was discovered that additional unidentified electron density corresponded to a third Der f 1 molecule. The third molecule was localized using PHASER⁴³ as implemented in CCP4 package.⁴⁴ Though E-64 was present in the crystallization conditions, no continuous density corresponding to the molecule was observed, and the inhibitor could not be modeled. The initial refinement was done using CNS⁴⁵ with the protocol for twinned data. The CNS simulating annealing protocol was used to remove model bias. Further refinement was performed using SHELXL⁴⁶ and higher resolution data. Program COOT¹⁵ was used for manual adjustment of the model. The quality of the structure was monitored using MOLPROBITY⁴⁷ and ADIT.⁴⁸ The twin fraction was refined to 0.43. Summary of the data collection and refinement statistics for higher resolution data used for refinement are reported in Table 2.

Data from Der p 1 crystal were collected at ID-23-1 (ESRF, Grenoble). The structure of Der p 1 was solved using model of the orthorhombic Der p 1 form (PDB code: 2AS8) and HKL-3000 in combination with MOLREP. Refinement was done using REFMAC and COOT. In the last stages, refinement was performed using TLS groups defined with the TLMSD server.⁴⁹ Structure validation was performed using the same tools used for the structure of Der f 1. Details of data processing and refinement are summarized in Table 2.

Calculations of surface area were performed using MSMS program at StrucTools server (<http://molbio.info.nih.gov/structbio/basic.html>).

Sequencing of the Group 1 allergen-specific monoclonal antibodies from cell lines

The cell lines producing the anti-Group 1 allergen monoclonal antibodies 4C1 B8 3F8, 6A8 B10 D12, 5H8 C12 D8 and 10B9 F6 A12 were grown at the Lymphocyte Culture Center (University of Virginia). The monoclonal antibodies had been raised in BALB/c mice, which

were reported to be the best responder strain to Group 1 mite allergen^{7,50,51} Total RNA (36–100 µg/3×10⁶ cells) was isolated from the cell lines using an RNeasy Mini kit (Qiagen, Valencia, CA). cDNAs encoding for the light and heavy chains of the mAb were obtained by reverse transcription from RNA (SuperScript[™] III, Invitrogen Corporation, Carlsbad, CA), and the DNA was PCR amplified using specific primers, sequenced and analyzed.

A primer mix containing the degenerate N-terminal primers VHa and VHb [VHa: 5'-gag gtt cag ctg cag cag(ct)c-3' and 5'-gag gtg cag ctg gtg ga(ag)tc-3'], and the C-terminal primer CH2 [CH2: 5'-tt agg agt cag agt aat ggt gag cac atc c-3'] was used to reverse transcribe the RNA and amplify the DNA encoding for the heavy chains at annealing temperatures 40–55°C. Middle primers were used to finalize heavy chain sequencing. To amplify the light chains of the antibodies four pairs of primers were initially tested. The N-terminal primer Vκ4: [5'-caa att gtt ctc acc cag tct cca-3'] was used for DNA amplification of the light chain of 4C1 and 6A8, and Vκ10: [5'-gat atc cag atg aca cag act aca-3'] amplified the light chains of 5H8 and 10B9, combined with the C-terminal primer [KappaConstant: 5'-gat gga tac agt tgg tgc-3']. The mAbs were isotyped G1 for the heavy chain and kappa (κ) for the light chain using the IsoStrip (Roche Diagnostics Corporation, Indianapolis, IN). CLUSTAL⁵² and BLAST were used to align and search for homologous sequences in databases, respectively.

Protein Data Bank Accessions code

The coordinates and structure factors for Der f 1 were deposited in PDB with accession code 3D6S, while the coordinates and structure factors for Der p 1 were deposited with accession code 3F5V.

Acknowledgements

The authors would like also thank Dr. Alex Wlodawer for valuable discussions and Kyle Williams for technical assistance. The work described in the paper was supported by GM53163, Indoor Biotechnologies, Inc. internal grant, the University of Virginia Biomedical Engineering Industrial Internship Program and the Virginia Biotechnology Association. The results shown in this report are derived from work performed at Argonne National Laboratory, at the Structural Biology Center of the Advanced Photon Source. Argonne is operated by University of Chicago Argonne, LLC, for the U.S. Department of Energy, Office of Biological and Environmental Research under contract DE-AC02-06CH11357.

References

1. Platts-Mills TA, Thomas WR, Aalberse RC, Vervloet D, Chapman MD. Dust mite allergens and asthma: report of a second international workshop. *J Allergy Clin Immunol* 1992;89:1046–1060. [PubMed: 1349902]
2. Hales BJ, Shen HD, Thomas WR. Cross-reactivity of T-cell responses to *Dermatophagoides pteronyssinus* and *D. farinae*. Studies with group 1 and 7 allergens. *Clin Exp Allergy* 2000;30:927–933. [PubMed: 10848914]
3. Chapman MD, Platts-Mills TA. Purification and characterization of the major allergen from *Dermatophagoides pteronyssinus* -antigen P1. *J Immunol* 1980;125:587–592. [PubMed: 6771329]
4. Heymann PW, Chapman MD, Aalberse RC, Fox JW, Platts-Mills TA. Antigenic and structural analysis of group II allergens (Der f II and Der p II) from house dust mites (*Dermatophagoides* spp). *J Allergy Clin Immunol* 1989;83:1055–1067. [PubMed: 2732406]
5. Barrett AJ, Rawlings ND. Evolutionary lines of cysteine peptidases. *Biol Chem* 2001;382:727–733. [PubMed: 11517925]
6. Chapman MD, Heymann PW, Platts-Mills TA. Epitope mapping of two major inhalant allergens, Der p I and Der f I, from mites of the genus *Dermatophagoides*. *J Immunol* 1987;139:1479–1484. [PubMed: 2442247]
7. Heymann PW, Chapman MD, Platts-Mills TA. Antigen Der f I from the dust mite *Dermatophagoides farinae*: structural comparison with Der p I from *Dermatophagoides pteronyssinus* and epitope

- specificity of murine IgG and human IgE antibodies. *J Immunol* 1986;137:2841–2847. [PubMed: 2428875]
8. Lind P, Hansen OC, Horn N. The binding of mouse hybridoma and human IgE antibodies to the major fecal allergen, Der p I, of *Dermatophagoides pteronyssinus*. Relative binding site location and species specificity studied by solid-phase inhibition assays with radiolabeled antigen. *J Immunol* 1988;140:4256–4262. [PubMed: 3372999]
 9. Meno K, Thorsted PB, Ipsen H, Kristensen O, Larsen JN, Spangfort MD, Gajhede M, Lund K. The crystal structure of recombinant proDer p 1, a major house dust mite proteolytic allergen. *J Immunol* 2005;175:3835–3845. [PubMed: 16148130]
 10. de Halleux S, Stura E, VanderElst L, Carlier V, Jacquemin M, Saint-Remy JM. Three-dimensional structure and IgE-binding properties of mature fully active Der p 1, a clinically relevant major allergen. *J Allergy Clin Immunol* 2006;117:571–576. [PubMed: 16522455]
 11. Takai T, Kato T, Yasueda H, Okumura K, Ogawa H. Analysis of the structure and allergenicity of recombinant pro- and mature Der p 1 and Der f 1: major conformational IgE epitopes blocked by prodomains. *J Allergy Clin Immunol* 2005;115:555–563. [PubMed: 15753904]
 12. Chapman MD, Wunschmann S, Pomés A. Proteases as Th2 adjuvants. *Curr Allergy Asthma Rep* 2007;7:363–367. [PubMed: 17697645]
 13. Nakamura T, Hirasawa Y, Takai T, Mitsuishi K, Okuda M, Kato T, Okumura K, Ikeda S, Ogawa H. Reduction of skin barrier function by proteolytic activity of a recombinant house dust mite major allergen Der f 1. *J Invest Dermatol* 2006;126:2719–2723. [PubMed: 17008873]
 14. Krissinel E, Henrick K. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr Sect D* 2004;60:2256–2268. [PubMed: 15572779]
 15. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. *Acta Crystallogr Sect D* 2004;60:2126–2132. [PubMed: 15572765]
 16. Piboonpocanun S, Malainual N, Jirapongsananuruk O, Vichyanond P, Thomas WR. Genetic polymorphisms of major house dust mite allergens. *Clin Exp Allergy* 2006;36:510–516. [PubMed: 16630157]
 17. van Oort E, de Heer PG, van Leeuwen WA, Derksen NI, Muller M, Huveneers S, Aalberse RC, van Ree R. Maturation of *Pichia pastoris*-derived recombinant pro-Der p 1 induced by deglycosylation and by the natural cysteine protease Der p 1 from house dust mite. *Eur J Biochem* 2002;269:671–679. [PubMed: 11856327]
 18. Takai T, Mineki R, Nakazawa T, Takaoka M, Yasueda H, Murayama K, Okumura K, Ogawa H. Maturation of the activities of recombinant mite allergens Der p 1 and Der f 1, and its implication in the blockade of proteolytic activity. *FEBS Lett* 2002;531:265–272. [PubMed: 12417324]
 19. Takai T, Mizuuchi E, Kikuchi Y, Nagamune T, Okumura K, Ogawa H. Glycosylation of recombinant proforms of major house dust mite allergens Der p 1 and Der f 1 decelerates the speed of maturation. *Int Arch Allergy Immunol* 2006;139:181–187. [PubMed: 16439856]
 20. Best EA, Stedman KE, Bozic CM, Hunter SW, Vailes L, Chapman MD, McCall CA, McDermott MJ. A recombinant group 1 house dust mite allergen, rDer f 1, with biological activities similar to those of the native allergen. *Protein Expr Purif* 2000;20:462–471. [PubMed: 11087686]
 21. Yasuhara T, Takai T, Yuuki T, Okudaira H, Okumura Y. Biologically active recombinant forms of a major house dust mite group 1 allergen Der f 1 with full activities of both cysteine protease and IgE binding. *Clin Exp Allergy* 2001;31:116–124. [PubMed: 11167959]
 22. Zheng H, Chruszcz M, Lasota P, Lebioda L, Minor W. Data mining of metal ion environments present in protein structures. *J Inorg Biochem* 2008;102:1765–1776. [PubMed: 18614239]
 23. Mellor GW, Thomas EW, Topham CM, Brocklehurst K. Ionization characteristics of the Cys-25/His-159 interactive system and of the modulatory group of papain: resolution of ambiguity by electronic perturbation of the quasi-2-mercaptopyridine leaving group in a new pyrimidyl disulphide reactivity probe. *Biochem J* 1993;290:289–296. [PubMed: 8439297]
 24. Dardenne LE, Werneck AS, de Oliveira Neto M, Bisch PM. Electrostatic properties in the catalytic site of papain: A possible regulatory mechanism for the reactivity of the ion pair. *Proteins* 2003;52:236–253. [PubMed: 12833547]

25. Takai T, Kato T, Sakata Y, Yasueda H, Izuhara K, Okumura K, Ogawa H. Recombinant Der p 1 and Der f 1 exhibit cysteine protease activity but no serine protease activity. *Biochem Biophys Res Commun* 2005;328:944–952. [PubMed: 15707969]
26. Deb R, Shakib F, Reid K, Clark H. Major House Dust Mite Allergens *Dermatophagoides pteronyssinus* 1 and *Dermatophagoides farinae* 1 Degrade and Inactivate Lung Surfactant Proteins A and D. *J Biol Chem* 2007;282:36808–36819. [PubMed: 17848554]
27. Schulz O, Sewell HF, Shakib F. A sensitive fluorescent assay for measuring the cysteine protease activity of Der p 1, a major allergen from the dust mite *Dermatophagoides pteronyssinus*. *Mol Pathol* 1998;51:222–224. [PubMed: 9893750]
28. Pernas M, Sanchez-Ramos I, Sanchez-Monge R, Lombardero M, Arteaga C, Castanera P, Salcedo G. Der p 1 and Der f 1, the highly related and major allergens from house dust mites, are differentially affected by a plant cystatin. *Clin Exp Allergy* 2000;30:972–978. [PubMed: 10848919]
29. Zhang J, Hamilton JM, Garrod DR, Robinson C. Interactions between mature Der p 1 and its free prodomain indicate membership of a new family of C1 peptidases. *Allergy* 2007;62:1302–1309. [PubMed: 17919146]
30. Kaul P, Sathish HA, Prakash V. Effect of metal ions on structure and activity of papain from *Carica papaya*. *Nahrung* 2002;46:2–6. [PubMed: 11890048]
31. Ma S, Devi-Kesavan LS, Gao J. Molecular dynamics simulations of the catalytic pathway of a cysteine protease: a combined QM/MM study of human cathepsin K. *J Am Chem Soc* 2007;129:13633–13645. [PubMed: 17935329]
32. Ponstingl H, Kabir T, Thornton JM. Automatic inference of protein quaternary structure from crystals. *J Appl Crystallogr* 2003;36:1116–1122.
33. Ponstingl H, Henrick K, Thornton JM. Discriminating between homodimeric and monomeric proteins in the crystalline state. *Proteins* 2000;41:47–57. [PubMed: 10944393]
34. Krissinel E, Henrick K. Inference of macromolecular assemblies from crystalline state. *J Mol Biol* 2007;372:774–797. [PubMed: 17681537]
35. Dolenc I, Turk B, Pungercic G, Ritonja A, Turk V. Oligomeric structure and substrate induced inhibition of human cathepsin C. *J Biol Chem* 1995;270:21626–21631. [PubMed: 7665576]
36. Greene WK, Thomas WR. IgE binding structures of the major house dust mite allergen Der p 1. *Mol Immunol* 1992;29:257–262. [PubMed: 1371823]
37. Szalai K, Fuhrmann J, Pavkov T, Scheidl M, Wallmann J, Bramswig KH, Vrtala S, Scheiner O, Keller W, Saint-Remy JM, Neumann D, Pali-Scholl I, Jensen-Jarolim E. Mimotopes identify conformational B-cell epitopes on the two major house dust mite allergens Der p 1 and Der p 2. *Mol Immunol* 2008;45:1308–1317. [PubMed: 17964653]
38. Rosenbaum G, Alkire RW, Evans G, Rotella FJ, Lazarski K, Zhang RG, Ginell SL, Duke N, Naday I, Lazarz J, Molitsky MJ, Keefe L, Gonczy J, Rock L, Sanishvili R, Walsh MA, Westbrook E, Joachimiak A. The Structural Biology Center 19ID undulator beamline: facility specifications and protein crystallographic results. *J Synchr Rad* 2006;13:30–45.
39. Otwinowski Z, Minor W. Processing of X-ray diffraction data collected in oscillation mode. *Macromolecular Crystallography* 1997;276(Pt A):307–326.
40. Vagin A, Teplyakov A. MOLREP: an automated program for molecular replacement. *J Appl Crystallogr* 1997;30:1022–1025.
41. Minor W, Cymborowski M, Otwinowski Z, Chruszcz M. HKL-3000: the integration of data reduction and structure solution-from diffraction images to an initial model in minutes. *Acta Crystallogr Sect D* 2006;62:859–866. [PubMed: 16855301]
42. Kopp J, Schwede T. The SWISS-MODEL Repository of annotated three-dimensional protein structure homology models. *Nucleic Acids Res* 2004;32:D230–D234. [PubMed: 14681401]
43. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. Phaser crystallographic software. *J Appl Crystallogr* 2007;40:658–674.
44. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr Sect D* 1994;50:760–763. [PubMed: 15299374]
45. Brünger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice LM, Simonson T, Warren GL. *Crystallography & NMR*

- system: A new software suite for macromolecular structure determination. *Acta Crystallogr Sect D* 1998;54:905–921. [PubMed: 9757107]
46. Sheldrick GM. A short history of SHELX. *Acta Crystallogr Sect A* 2008;64:112–122. [PubMed: 18156677]
47. Lovell SC, Davis IW, Arendall WB 3rd, de Bakker PI, Word JM, Prisant MG, Richardson JS, Richardson DC. Structure validation by Calpha geometry: phi, psi and Cbeta deviation. *Proteins* 2003;50:437–450. [PubMed: 12557186]
48. Yang H, Guranovic V, Dutta S, Feng Z, Berman HM, Westbrook JD. Automated and accurate deposition of structures solved by X-ray diffraction to the Protein Data Bank. *Acta Crystallogr Sect D* 2004;60:1833–1839. [PubMed: 15388930]
49. Painter J, Merritt EA. TLSMD web server for the generation of multi-group TLS models. *J Appl Crystallogr* 2006;39:109–111.
50. Chapman MD, Sutherland WM, Platts-Mills TA. Recognition of two Dermatophagoides pteronyssinus-specific epitopes on antigen P1 by using monoclonal antibodies: binding to each epitope can be inhibited by serum from dust mite-allergic patients. *J Immunol* 1984;133:2488–2495. [PubMed: 6207232]
51. Ovsyannikova IG, Vailes LD, Li Y, Heymann PW, Chapman MD. Monoclonal antibodies to group II Dermatophagoides spp. allergens: murine immune response, epitope analysis, and development of a two-site ELISA. *J Allergy Clin Immunol* 1994;94:537–546. [PubMed: 7521892]
52. Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 1994;22:4673–4680. [PubMed: 7984417]
53. Gouet P, Robert X, Courcelle E. ESPript/ENDscript: Extracting and rendering sequence and 3D information from atomic structures of proteins. *Nucleic Acids Res* 2003;31:3320–3323. [PubMed: 12824317]
54. Kabat EA, Wu TT. Attempts to locate complementarity-determining residues in the variable positions of light and heavy chains. *Ann N Y Acad Sci* 1971;190:382–393. [PubMed: 5290024]
55. Tomlinson IM, Cox JP, Gherardi E, Lesk AM, Chothia C. The structural repertoire of the human V kappa domain. *EMBO J* 1995;14:4628–4638. [PubMed: 7556106]

Figure 1.

A) Sequence alignment of mature Der f 1 and Der p 1. The alignment was done with CLUSTALW⁵² and the figure was prepared using ESPRIPT.⁵³ Blue stars show residues mutated in different Der f 1 variants. Catalytic residues are marked with orange stars, while disulphide bond forming cysteines are labeled using green numbers. Light-blue star shows N-glycosylation site. B) Model of Der f 1 shown in ribbon representation. Cys35 (orange), Asn53 (blue) and cysteines forming disulphide bonds (red) are shown in stick representation. The disulphide bonds are labeled as on Fig. 1A. C) Superposition of the crystal structures of Der f 1 (PDB code: 3D6S; green) and Der p 1 (PDB code: 2AS8; cyan).

Figure 2.

A) Superposition of the Der f 1 (green) and Der p 1 (yellow) active sites. Interactions between residues are marked with dashed lines and interatomic distances (see Table 1) are in Ångstroms. B) Superposition of the putative Der f 1 metal binding site (green) with the metal binding site observed in the proDer p 1 (blue; PDB code 1XKG) and mature Der p 1 (yellow) structures. Residues are numbered according to the Der f 1 sequence. The water molecules coordinating Y^{3+} (cyan sphere) were omitted for sake of picture clarity. Coordination of Ca^{2+} (black sphere) in Der p 1 structure (PDB code: 3F5V) is shown using dashed lines. C). Water molecules (red spheres) conserved in Der f 1 (PDB code: 3D6S), proDer p 1 (PDB code; 1XKG) and mature

Der p 1 crystal structures (PDB code: 3F5V). Black sphere shows position of the metal ion in structure of Der p 1.

Figure 3.

A) The putative dimer formed by mature Der p 1 (PDB codes: 2AS8, 3F5V). Der p 1 molecules are shown in ribbon representation, while Ca^{2+} ions are shown as red spheres. B) Tyr166 and Gln167 (Der f 1 numbering) are blocking the catalytic cleft in a putative dimer of Der p 1. Chain A is shown in surface representation (green), chain B is shown in blue, while catalytic Cys35A is marked in orange.

Figure 4.

The molecular surface of mature Der f 1 (PDB code: 3D6S) shown in two orientations. Residues that differ in Der f 1 and Der p 1 are shown in yellow. The largest patches of different residues are labeled as I, II, III and IV. Identical residues are marked in green and the conserved Cys35 residue is colored orange. Residues conserved in Der f 1, Der p 1 and human cathepsin K are shown in light blue. The N-terminal pro-peptide is pictured in ribbon representation (red).

	CDR1 (31-35b)	CDR2 (50-65)	
4C1-Heavy	EVQLVESGPGLVKPSQSLSLTCTVTIGYSITSDYAWNWIQRFPGNKLEWMGYISY-SGTTIS		59
10B9-Heavy	EVQLVESGPGSLVKPSQTLSLTCSVTGDSITSGF-WNNWIRKFPNGKLEFMGYITY-SGTSY		58
6A8-Heavy	EVQLQQSGAELVKPGASLCKTSSGYITTRYW-IHWVKQRPGQLGWGGEIFPGCTTIH		59
5H8-Heavy	EVQLVLES GPGLVAPSQSLSLTCTVSGFSLTGYG-VNWWVRPPGKGLEWLGMWVG-DGRID		58
	**** :*. ** * . :.:*.:* ::* :.:* :*: ** : * * *		
		CDR3 (95-102)	
4C1-Heavy	<u>YNPSIKS</u> RISITRDTSKNQFFLQINSVTIEDTATYYCGRT-GVYRYP--ERAPYWGQGTIL		116
10B9-Heavy	<u>YKPSIKS</u> RISITRDTSKNQVFLQINSVTAEDTATYYCARR-GFLTWNYYAMDYWGQGTIS		117
6A8-Heavy	<u>YNEKFKG</u> KATLTIDTSSSTAYMQLSSLTSED SAVYFCARSNGLITYF----DFWQGGTT		114
5H8-Heavy	<u>YNLVRSK</u> RLSLSKDNSQSQIFLKMSLQDDTARYYCARAYQRYDYV---AMDYWGQGTIS		115
	*: *.: :*: *.. :.:*.:*:*:*:*:	:*****	
		CDR1 (24-34): Loop 1 (26-32)	CDR2 (50-56): Loop 2 (50-52)
5H8-Light	DIQMTQTSSLSASLGRVTISCRASQDITNYLNWYQKPDGTVKLLIYYTSRLHSGVPS		60
10B9-Light	DIQMTQTSSLSASLGDGVTISCRASQDISNLYNWYQKPDGTVKLLIYYTSRLHSGVPS		60
6A8-Light	QIVLTQSPSSLSASLGERVSLTICRASQDIGSSINWLQEPDGTIKRLIYATSLDLSGVPK		60
4C1-Light	QIVLTQSPFSMYATLGERVTITICRASQDIYSYLSWLQKPKGKSLKTLIYRANRLITGVPS		60
	:* :*: . * :*: :.:*.:***** . ** **: . :* **: . * :*:*.		
		CDR3 (89-97): Loop 3 (91-96)	
5H8-Light	RFSGSGSGTDYSLTISNLEQEDATYFCQQGKTLTP-TFGGGTKLEIKRADAAPT VSI		116
10B9-Light	RFSGSGSGTDYSLTISNLEQEDATYFCQQGNTLPYTFGGGTKLEIKRADAAPT VSI		117
6A8-Light	RFSGSRSGSDYSLTISSELEDFVDYYCLQYASSPWTTFGGGTKLEIKRADAAPT VSI		117
4C1-Light	RFSGSGSGQDYSLTISSELEYEDMGIIYYCLQYDEFYPTTFGGGTKLEMKRADAAPT VSI		117
	***** ** ***** ** ** :* * *	***** ** ***** ** *****	

Figure 5.

Amino acid sequence alignment of the heavy and light chains of the monoclonal antibodies, indicating the CDR ⁵⁴ (red for the heavy chains and blue and orange for the light chains) and loops (blue for light chain). ⁵⁵ The CDR2 of low variability between the cross-reacting antibodies 4C1 and 10B9 is underlined.

Distances between amino acids near the Der f 1 (PDB code: 3D6S) and Der p 1 (PDB code: 3F5V) active sites. Amino acids are numbered with respect to the Der f 1 sequence.

Table 1

	Der f 1			Der p 1		
	A [Å]	B [Å]	C [Å]	A [Å]	B [Å]	
Cys35 Sγ... Nδ1 His171	3.7	3.6	3.4	3.0/5.6 [*]	3.5/5.4 [*]	
His171 Nε2... Oδ1 Asn191	3.0	2.7	2.7	2.8/7.2 [*]	2.9/7.1 [*]	
Gln29 N ε2...O Cys32	2.8	3.0	2.7	2.8	3.0	
Gln29 Oε1... Oγ Ser192	2.9	2.7	2.6	2.7	2.7	
Ser192 Oγ...H ₂ O	2.5	2.9	2.7	2.8	2.8	

* two values refer to the double conformation of His171

Table 2

Data collection and refinement statistics. Ramachandran plot was calculated using MOLPROBITY.

Protein	Der f 1	Der p 1
PDB code	3D6S	3F5V
Data collection		
Beamline	19ID (APS)	ID-23-1 (ESRF)
Wavelength (Å)	0.979	1.072
Unit cell (Å, °)	a = 91.2, c = 77.7	a = 95.2, b = 84.1 c = 75.4, β= 123.8
Space group	P4 ₁	C2
Solvent content (%)	41	50
Number of protein chains in AU	3	2
Resolution range (Å)	50.0-2.0	50.0-1.36
Highest resolution shell (Å)	2.02-2.00	1.38-1.36
Unique reflections	39590(1078)	103499(4795)
Redundancy	6.4(6.6)	3.6(2.9)
Completeness (%)	92.6(100.0)	98.5(91.5)
R _{merge} (%)	8.0(51.4)	6.9(55.1)
Average I/σ (I)	27.0(3.3)	25.3(2.1)
Refinement		
R	21.1	15.8
R _{free}	24.1	18.3
Mean B value (Å ²)	32.3	8.9
B from Wilson plot (Å ²)	25.3	12.6
RMS deviation bond lengths (Å)	0.018	0.017
Number of amino acid residues	668	444
Number of water molecules	311	679
Number of metal ions	0	2
Ramachandran plot		
Most favored regions (%)	95.8	96.2
Additional allowed regions (%)	4.2	3.8