doi:10.1016/j.jmb.2005.02.062 J. Mol. Biol. (2005) 348, 433–444

JMB

Available online at www.sciencedirect.com

SCIENCE DIRECT®

ELSEVIER

Crystal Structure of Cockroach Allergen Bla g 2, an Unusual Zinc Binding Aspartic Protease with a Novel Mode of Self-inhibition

This paper is dedicated to the memory of Dr Venugopal Dhanaraj, who modeled the structure of Bla g 2

Alla Gustchina[1]*, Mi Li[1,2], Sabina Wünschmann[3], Martin D. Chapman[3] Anna Pomés[3] and Alexander Wlodawer[1]

[1]*Macromolecular Crystallography Laboratory National Cancer Institute Frederick, MD 21702, USA*

[2]*Basic Research Program SAIC-Frederick, Frederick MD 21702, USA*

[3]*INDOOR Biotechnologies Inc., 1216 Harris Street Charlottesville, VA 22903 USA*

**Corresponding author*

The crystal structure of Bla g 2 was solved in order to investigate the structural basis for the allergenic properties of this unusual protein. This is the first structure of an aspartic protease in which conserved glycine residues, in two canonical DTG triads, are substituted by different amino acid residues. Another unprecedented feature revealed by the structure is the single phenylalanine residue insertion on the tip of the flap, with the side-chain occupying the S1 binding pocket. This and other important amino acid substitutions in the active site region of Bla g 2 modify the interactions in the vicinity of the catalytic aspartate residues, increasing the distance between them to ~4 Å and establishing unique direct contacts between the flap and the catalytic residues. We attribute the absence of substantial catalytic activity in Bla g 2 to these unusual features of the active site. Five disulfide bridges and a Zn-binding site confer stability to the protein, which may contribute to sensitization at lower levels of exposure than other allergens.

Keywords: active site; allergy; aspartic proteases; metal binding; self-inhibition

Introduction

Cockroaches excrete a potent allergen, Bla g 2, that is associated with IgE production and asthma.[1] Sensitization and exposure to cockroach allergens is a major risk factor for asthma mortality and morbidity among children living in inner-city areas of the United States, and is strongly associated with asthma in other parts of the world.[2–8] Bla g 2 appears to be an especially potent allergen, which elicits IgE responses at exposure levels that are 10–100-fold lower than other common indoor allergens, such as dust mite and cat.[9] Several mite allergens function as proteolytic enzymes, including Der p 1 (cysteine protease), and Der p 3, Der p 6 and Der p 9 (serine proteases).[10,11] A body of evidence suggests that these enzyme allergens play an important role in asthma by potentiating IgE responses, inducing pro-inflammatory cytokine release, and causing direct damage to airway epithelium.[12–15] However, unlike dust mite, none of the cockroach allergens that have been cloned, except Bla g 2, shows homology to proteolytic enzymes.

Bla g 2 is an important allergen that causes IgE responses in ~60% of cockroach allergic patients and has unusual structural features. Bla g 2 is a 36 kDa glycoprotein that shares homology with the family of aspartic proteases.[1] Aspartic proteases catalyze the hydrolysis of a peptide bond utilizing an acid–base mechanism *via* nucleophilic attack by the water molecule which is polarized by two catalytic aspartate residues. They share similar overall fold of a bilobal molecule with a well-defined substrate-binding cleft. Each domain contains a conserved DTG triad with a catalytic aspartate. However, molecular modeling and functional studies using standard enzyme assays showed that Bla g 2 was an inactive aspartic protease, with critical amino acid substitutions around the catalytic residues.[16,17] In Bla g 2, the

Abbreviations used: PAG, pregnancy-associated glycoproteins; EMP, ethyl mercury phosphate.

E-mail address of the corresponding author: alla@ncifcrf.gov

two glycine residues in DTG triads are substituted by threonine and serine, and the conserved tyrosine 75 in the "flap" region of the enzyme is substituted by phenylalanine. Surprisingly, Bla g 2 is related in its primary structure to a group of inactive mammalian aspartic proteases, known as pregnancy-associated glycoproteins (PAG), which are expressed in the chorion of pregnant females from ungulates, including pig, horse, cow, and sheep.[18,19] A common feature of Bla g 2 and most PAGs is the lack of enzymatic activity in standard aspartic protease assays, for which neither the structural basis nor a molecular mechanism have been established.

Here, we report a high-resolution crystal structure of partially deglycosylated, recombinant Bla g 2 that was expressed in the yeast *Pichia pastoris*. The structure reveals a novel mode of self-inhibition that explains why the catalytic activity of this protein is impaired. Unlike any aspartic protease of known structure, Bla g 2 is a zinc-binding protein.

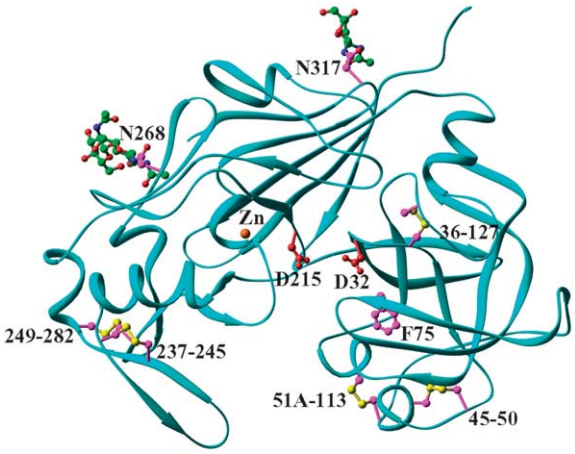

Figure 1. Ribbon representation of the overall fold of Bla g 2. The disulfide bridges and polysaccharides are shown in ball-and-stick representation and their location is marked. The position of Zn is marked as an orange ball.

Results

Crystal structure of Bla g 2

Native Bla g 2 is a glycoprotein that contains three potential N-glycosylation sites involving residues Asn93, Asn268, and Asn317. Since the attempts to crystallize the recombinant native protein were not successful, we tried to reduce the glycosylation state of the molecule by mutating these asparagine residues to glutamine. Several mutants were made, and one of them, the N93Q mutant, was expressed reasonably well and produced the crystals used in this study.

Initial attempts to solve the structure of Bla g 2 by molecular replacement using crystal structures of pepsin (4PEP),[20] chymosin (1CMS),[21] and the theoretical model[16] were unsuccessful, necessitating an *ab initio* approach involving multiple isomorphous replacement (MIR) phasing. The overall fold of the Bla g 2 molecule was, as expected, closely related to the well-known fold of the non-viral aspartic protease family members[22–24] (Figure 1). The molecule is composed of two domains with a large cleft in between, where the two catalytic aspartate residues are located. All secondary structure elements, defined as a structural template for aspartic proteases,[25,26] were also found in the three-dimensional structure of Bla g 2. The two remaining asparagine residues in potential glycosylation sites, Asn268 and Asn317, were found to be *N*-glycosylated, with two polysaccharide units clearly identified as bound to the former, and one to the latter (Figure 1). Five disulfide bridges were found in the structure of Bla g 2: Cys45-Cys50, Cys51A-Cys113, and Cys36-Cys127 are located in the N-terminal domain, while two disulfides, Cys237-Cys245 and Cys249-Cys282, are in the C-terminal domain (Figure 1).

The crystal structure of Bla g 2 revealed an insertion of one residue in the flap region that had not been observed in any other structure of aspartic proteases. The stretch of residues at the tip of the flap in Bla g 2 contains two consecutive phenylalanine residues. One of these, Phe75, is structurally equivalent to Tyr75, a conserved flap residue shown to be important for the catalytic activity of the aspartic protease family,[27] while the other, Phe75A, does not have structural homologs in any other aspartic proteases (Figure 2(a)). As discussed in more detail below, the presence of Phe75A changes only the conformation of the tip of the flap and affects the enzymatic activity of Bla g 2. However, the hydrogen bonding pattern within this secondary structure element is not modified (Figure 2(b)).

Another unusual feature of Bla g 2 that is not found in the other aspartic proteases is a metal-binding site that is occupied by a tightly bound Zn cation (Figure 3(a)). Zinc was not present in any of the solutions used during the Bla g 2 purification or crystallization steps, suggesting that the binding of zinc is not an artifact, but a permanent structural feature related to the function of the molecule. The identity of the bound cation was established by fluorescence scans of a single crystal, performed near the absorption edges of two cations that were compatible with the observed coordination of the metal, namely Ni and Zn. As shown in Figure 3(b), the identity of the bound cation is, without any doubt, Zn. Four residues are involved in the formation of Zn-binding site: two histidine residues (155 and 161) and two aspartate residues (303 and 307). Only Asp303 is conserved in other aspartic proteases, whereas the other three residues are unique to the sequence of Bla g 2 (and some related proteins from other cockroach species). This may explain why zinc binding is a unique property of Bla g 2, and none of the other aspartic proteases with known structure contains bound metals.

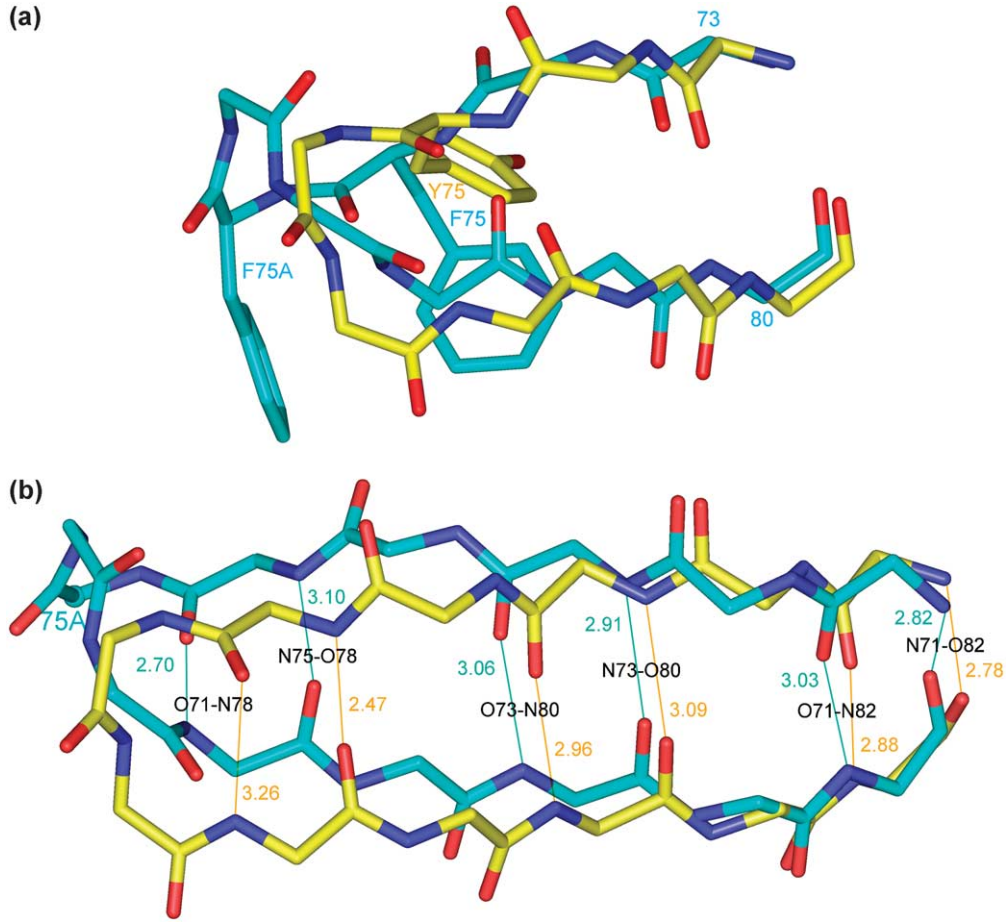

Figure 2. Superposition of the flaps (residues 72–80) in Bla g 2 (blue) and pepsin (yellow). (a) The side-chains of Phe75 and Phe75A (Bla g 2) and Tyr75 (pepsin) are shown in stick representation. (b) Comparison of the pattern of hydrogen bonds within the flaps, also indicating their lengths.

However, all four residues are conserved in the sequences of two other cockroach aspartic proteases, from *Leucophaea maderae* (Lma-p54) and *Periplaneta americana* (Table 1), indicating a possibility of Zn binding in these homologous proteins.

Comparison with other aspartic proteases

Superposition of the C^{α} atoms of Bla g 2 on those of bovine chymosin (1CMS), porcine pepsin (4PEP), yeast proteinase A (1FMU),[28] and human renin (1RNE)[29] leads to rms deviations for the structurally equivalent residues of 1.45 Å (for 290 C^{α} pairs), 1.47 Å (291 pairs), 1.61 Å (294 pairs), and 1.73 Å (298 pairs), respectively. Although the overall fold of Bla g 2 is almost identical with the folds of the other enzymes, several small, but unique deletions and insertions can be seen in the three-dimensional structure. These deletions and insertions can be easily identified in the structurally based sequence alignment, which resulted from this superposition (Figure 4). To make the comparisons easier we used pepsin numbering for Bla g 2, introducing gaps in the case of deletion and adding letter suffixes to the number of the last residue before the insertion for all the residues that follow. An example is Phe75A, inserted between Tyr75 and Asp76. Other single-residue insertions are Asp68A and Gly126A. Two insertions involve pairs of residues (Cys51A and Pro51B, as well as Asp159A and Gly159B), while the longest insertion involves three residues (Thr243A, Arg243B, and Arg243C). Two deletions involve four residues each (9–12 and 293–296), one three residues (62–64), one two residues (22–23), and three additional deletions are single (108, 147 and 209). These insertions and deletions take place predominantly in the loop areas, and they do not affect the overall fold of Bla g 2, but change only the length or the local conformation of the loops (Figure 5).

Structurally based sequence alignment also provides information on the level of sequence similarity between Bla g 2 and other aspartic proteases. All five proteins listed in Figure 4 contain 39 identical residues, while 112 additional residues in Bla g 2 have an identical counterpart in one or more, but not all, of the compared enzymes. Eighty residues in Bla g 2 have homologous substitutions in the sequences of one or more of the other enzymes. The level of sequence similarity between Bla g 2 and each of the enzymes individually is very close; 52% (identity + similarity) for pepsin (25%

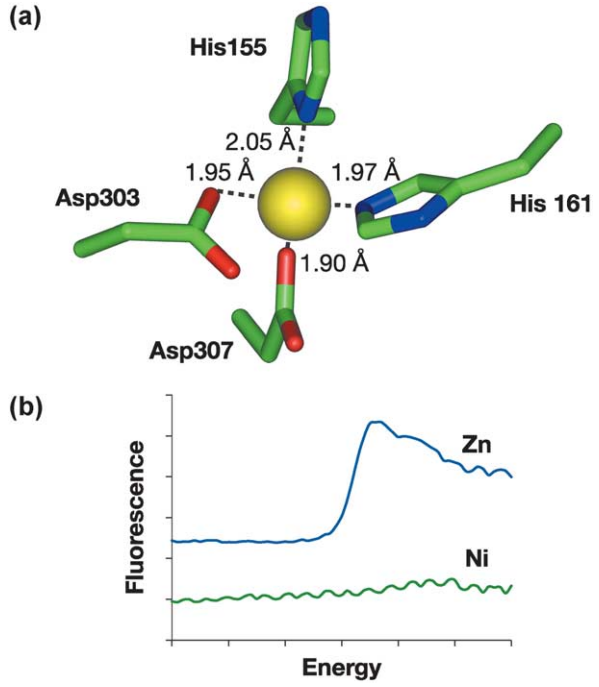

Figure 3. Metal-binding site in Bla g 2. (a) A metal ion and the coordinating residues in Bla g 2. The zinc ion is shown as a ball, and protein residues in stick representation. The marked interatomic distances were not restrained during the refinement process. (b) Fluorescence scans performed on a single crystal around the Ni and Zn absorption edges, showing the presence of the latter metal. The energies were shifted on the abscissa scale to overlap the theoretical absorption edge positions, whereas the ordinate scale is arbitrary.

and 27%), chymosin (27% and 25%) and renin (24% and 28%), and 49% for proteinase A (26% and 23%).

As mentioned above, Bla g 2 contains five disulfide bridges. Only two of them, Cys45-Cys50 and Cys249-Cys282, are conserved within the whole pepsin family. The remaining three disulfides, Cys36-Cys127, Cys51a-Cys113, and Cys237-Cys245, do not have structural analogs in the other aspartic proteases described to date. A disulfide bridge Cys206-Cys210 in pepsin, chymosin, renin and most other aspartic proteases does not have an equivalent in Bla g 2 (or in proteinase A).

The location of glycosylation sites in Bla g 2 also differs from those in proteinase A and renin. Asn67, which is glycosylated in the latter enzymes, is substituted by isoleucine in Bla g 2. This loop also assumes a different conformation due to the deletion (preceding) and insertion (following) Ile67. The second glycosylation site in Bla g 2 is found at Asn268 (Figure 1). This site is close, but not structurally equivalent to the glycosylation site in proteinase A found at Asn268 (that residue number would be 266 if pepsin numbering was applied to proteinase A). Although the residue at position 266 in Bla g 2 is also asparagine, the signature motif NX(S,T) is not present, but shifted two residues forward; thus, Asn268 rather than Asn266 is glycosylated in Bla g 2. Two other glycosylation sites in Bla g 2, Asn93 and Asn317, are unique to this protein. Since Asn93 was mutated to Gln, only Asn317 has an oligosaccharide attached to it, although the carbohydrate bound in this location is not as well ordered as its counterpart attached to Asn268.

Comparison with the previously published theoretical model

The crystal structure of Bla g 2 has been compared with the previously published homology-based model that was created using the X-ray structures of porcine pepsin and bovine chymosin as templates.[16] The overall rms deviation for C^{α} coordinates between the experimental structure and the model is 1.49 Å, which indicates that the secondary structure elements, such as strands and helices, were modeled correctly. However, several loops connecting these structural elements have different lengths and conformations in the crystal structure. The structural features such as the connecting loops between secondary structure elements are well known to vary between the proteins with even the most conserved fold, and their accurate description may be beyond the current limits of the homology modeling techniques. In addition, several structural features that have been revealed by the experimental structure of

Table 1. Alignment of the amino acid sequences in the vicinity of the Zn-binding site in several cockroach allergens

```
                155         161
Blag2     HARFQDGEHFGEIIFGGSDWKYVDGEFTYVPLVGDDSWKFRLDGVKIGDTTVAPA
Lma-p54   HARYPDGEHFGELVLGGVNSKFVVGDFTSAKLQEPDSWKIKLDQVLIGDKNVADD
Per a 2   HGRYPDGQHRGVLVLGGPIPAYYRGDFTYVPLVDQDTWNFKVDSISVGNEVIATD

                                          303   307
Blag2     SQYYIQQNGNLCYSGFQPCGHSDHFFIGDFFVDHYYSEFNWENKTMGFGRSVESV
Lma-p54   AEYYIQRNGDLCYSGFMPS-SLGYYMIGDFFIDHYYTVYNWEKKEMQFAIAREDA
Per a 2   STYQIQQNGDLCYSGFQYS-AGKCFHFGDFFMDNYYGEFDGQNKRMGFAKSVEEL

Lma-p54   Leucophaea maderae
Per a 2   Periplaneta americana
```

Residues that directly coordinate Zn in Bla g 2 are numbered and highlighted.

Figure 4. Structure-based sequence alignment comparing Bla g 2 with porcine pepsin (PEP), bovine chymosin (CMS), human renin (RNE), and yeast proteinase A (pA). White letters on the pink background indicate residues that are conserved in all five enzymes; pink and orange letters are used for residues identical in two or more proteins; green and blue letters denote similar residue types. A unique residue type is shown in black.

Bla g 2 are entirely unprecedented for any members of the family of aspartic proteases and, not surprisingly, are not present in the theoretical model. One of them is the novel conformation of the flap in the vicinity of Phe75A, a single-residue insertion unexpectedly found at the tip of the flap in the crystal structure. The presence of this extra residue also led to the shift in the sequence assignment for the subsequent part of the structure and contributed to wrong assignment of the disulfide bridges, among other problems. As a result, only one out of five disulfide bridges (249-282) was predicted correctly. Another novelty is the fact that Bla g 2 represents the first structure of an aspartic protease with substitution of the Gly residues in the signature triplets DTG in the active site. The side-chains of Thr34 and Ser217, residues which replaced the usual glycine residues, interact with the catalytic aspartate residues, placing them more than 4 Å apart. Thus, although the theoretical model described quite correctly the overall fold of Bla g 2, many of the important details could not be predicted in the absence of an experimental structure.

Active site region

The structure of the active site region in Bla g 2 revealed an unusual hydrogen bonded network, which includes both well-known interactions conserved among the members of the family of aspartic proteases, as well as novel hydrogen bonds uniquely present in Bla g 2 due to amino acid substitutions found in this area (Figure 4). The set of most conserved interactions in the active site of various aspartic proteases, known as "fireman's grip"[30] is also found in the active site of Bla g 2. Both Ser33 and Thr216 participate in highly symmetrical interactions: the side-chain hydroxyl group of Ser33 is hydrogen bonded to the amide nitrogen atom of Thr216 and the carbonyl oxygen atom of Ile214, and Thr216 makes similar interactions with N33 and O31 (Figure 6(a)).

All four carboxylate oxygen atoms of Asp32 and Asp215 are extensively involved in creating a hydrogen bonded network (Figure 6(a)), although the direct hydrogen bond between these two aspartate residues is no longer present. The distance between the two closest ("inner") oxygen atoms of

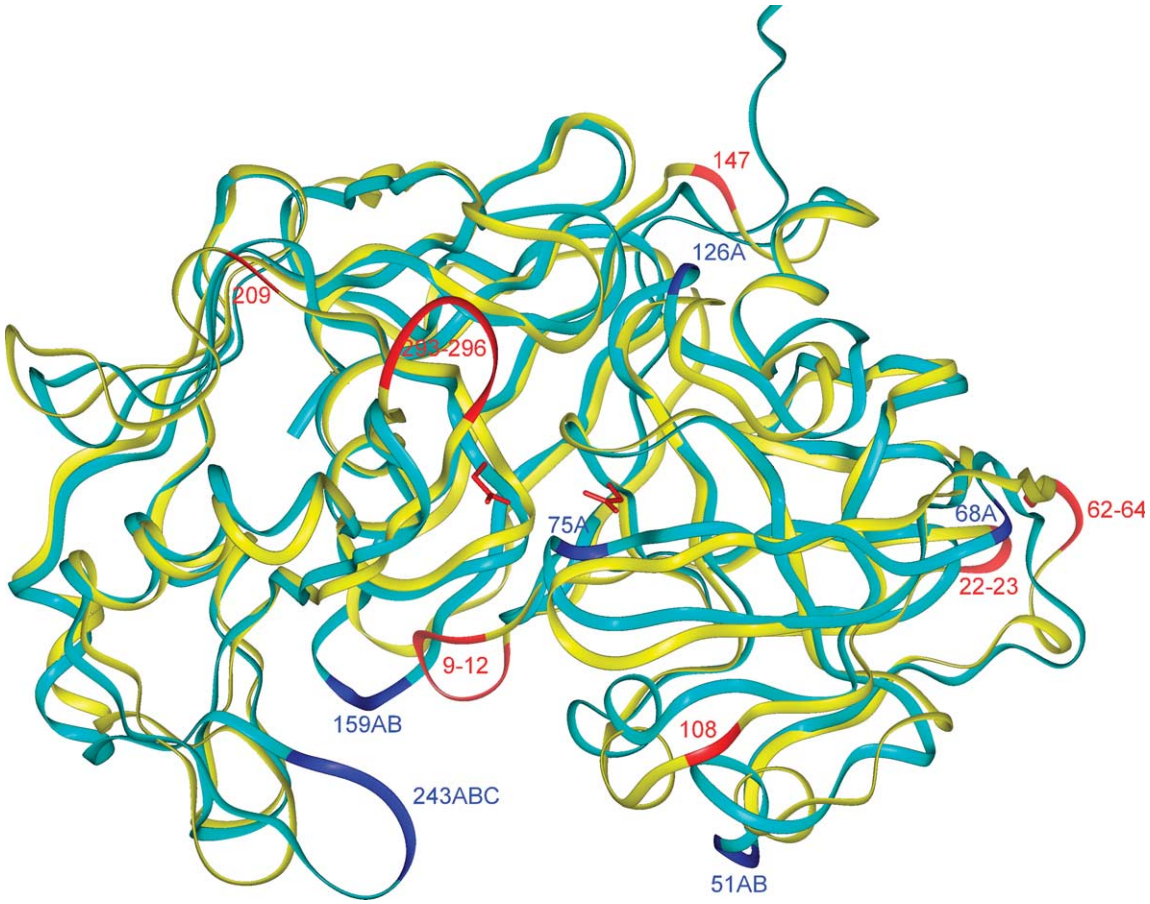

Figure 5. Ribbon diagram of superimposed Bla g 2 (blue) and pepsin (yellow). The deletions and insertions are shown in red for pepsin and in navy blue for Bla g 2, respectively, with the corresponding residue numbers.

these residues is 4.41 Å, significantly too long to permit a hydrogen bond between them, which is crucial for the catalytic mechanism of the aspartic protease family (Figure 6(b)). We attribute the loss of the hydrogen bond between the catalytic aspartate residues to the fact that these two "inner" carboxyl oxygen atoms are hydrogen bonded to the side-chains of Thr34 and Ser217, two "new" residues, which substitute for conserved glycine residues in the canonical triads DTG that contain the catalytic aspartate residues. The presence of these two "new" hydrogen bonds leads to rotation of the side-chains of both aspartate residues out of their usual coplanar orientation and, as a result of such rotation, the distance between the "inner" carboxyl oxygen atoms becomes larger than what is required for hydrogen bond formation. The "outer" carboxyl oxygen atom of Asp32 is hydrogen bonded to Ser35 and the amide nitrogen atom of Phe75A. The former is a conserved interaction that is present in all other aspartic proteases as well, while the latter is a novel hydrogen bond, formed in Bla g 2 due to the unique single-residue insertion of Phe75A in the flap. The carbonyl oxygen atom of Phe75A is involved in the coordination of a water molecule, which is also hydrogen bonded to Ser217 and Lys218, stabilizing their orientation and the conformation of the tip of the flap (Figure 6(a)). The "outer" carboxyl oxygen atom of Asp215 is at the distance of a strong hydrogen bond to Asp76 (2.48 Å). That residue is located at the tip of the flap and is usually Gly in other aspartic proteases, with very few exceptions. The direct contacts of the flap residues with the catalytic aspartate residues are a unique feature of the active site of Bla g 2 and were not observed previously in the other aspartic proteases. Because of these interactions, the flap in Bla g 2 assumes a more closed conformation and is shifted towards the catalytic aspartate residues ~3.5 Å more than in pepsin, making the catalytic aspartate residues less accessible. No water molecules interacting directly with the catalytic aspartate residues were found in the active site of the Bla g 2 crystal structure.

Since neither our biochemical nor structural approaches has as yet yielded crystallizable complexes of Bla g 2 with inhibitors, we had to resort to molecular modeling in order to investigate the mode of ligand binding. An inhibitor bound in the active site of Bla g 2 was modeled by superimposing the coordinates of Bla g 2 and the human renin, the latter with a transition state analog inhibitor present in the active site (pdb code 1RNE[31]). The results of the superposition shown in Figure 7 clearly indicate that the conformation of the active site as observed in our crystals is not conducive to binding of this

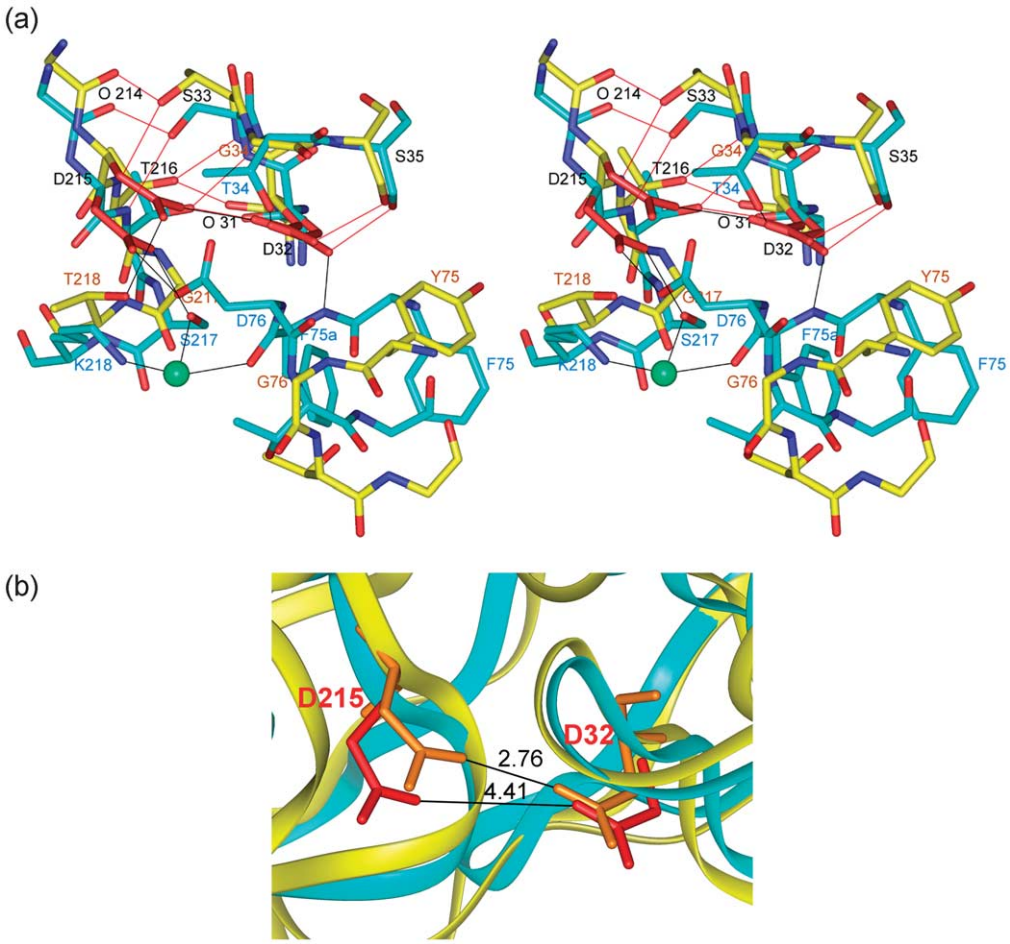

Figure 6. Active sites of Bla g 2 and pepsin. (a) Stereoview of the superposition of the active sites of Bla g 2 (blue) and pepsin (yellow). A water molecule is shown as a ball. Conserved hydrogen bonds are shown in red, while hydrogen bonds unique in Bla g 2 are shown in black. Labels O31 and O214 mark the main chain carbonyl oxygen atoms of the respective residues. (b) A comparison of the conformations of the catalytic aspartate residues in the active sites of Bla g 2 and pepsin (in red and orange, respectively). The distances between the "inner" oxygen atoms of the aspartate residues are indicated.

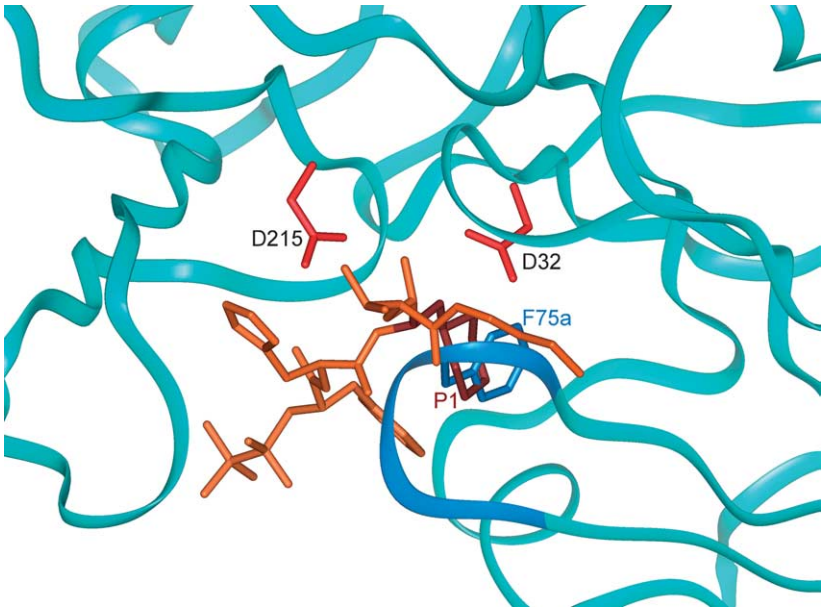

Figure 7. Superposition of Bla g 2 and renin complexed with a transition state inhibitor. The structures were superimposed using C^{α} coordinates. Only Bla g 2 is shown as a light blue ribbon with the tip of the flap in dark blue, whereas the trace of renin is omitted. The inhibitor is shown in orange. The positions of the catalytic aspartate residues and Phe75A in Bla g 2 are also marked.

specific ligand. In particular, the side-chain of Phe75A collides with the side-chain of the residue present at the P1 position in the inhibitor. It should be pointed out that similar observations have been described for other aspartic proteases. The structures of native chymosin[32] and proteinase A[33] revealed self-inhibited states of their active sites that were due to an unusual conformation of Tyr75, with the aromatic ring of this residue occupying the S1 binding pocket. The presence in these structures of a glycine residue at position 76 on the tip of the flap, with glycine-specific torsion angles, is mandatory for adopting such a conformation. Although that glycine residue is substituted by aspartate in Bla g 2, our structure revealed a new self-inhibitory mechanism, when the same S1 pocket is occupied by the aromatic ring of Phe75A, a unique single-residue insertion at the tip of the flap.

Discussion

The structure of Bla g 2 that is reported here was obtained using partially deglycosylated, recombinant enzyme expressed in yeast *Pichia pastoris* and was solved at near-atomic resolution of 1.3 Å, assuring high quality of the resulting model. As expected, the overall structure of Bla g 2 has the bilobal fold of a non-viral aspartic protease with a large cleft located between two structurally homologous domains. However, important interactions in the active site area are significantly modified due to four critical amino acid substitutions in the vicinity of the catalytic aspartate residues and in the flap region. Two such changes involve the highly conserved glycine residues of the canonical DTG triads that contain the catalytic aspartate residues 32 and 215. In Bla g 2, both Gly34 and Gly217 are replaced by polar amino acid residues, threonine and serine, respectively. The presence of these two "new" residues in the otherwise highly conserved active site area had a dramatic effect on the mutual orientation of the catalytic aspartate residues. Two strong hydrogen bonds between the hydroxyl groups of Thr34 and Ser217 on one side, and the two "inner" carboxyl oxygen atoms of Asp32 and Asp215 on the other side, rotate the side-chains of the catalytic aspartate residues out of their usual coplanar orientation, thus increasing the distance between them to 4.41 Å. According to the accepted catalytic mechanism of aspartic proteases,[34] the absence of a hydrogen bond between the aspartate residues must jeopardize the enzymatic properties of the molecule.

Similar substitutions in the active site area are found in other aspartic proteases, which exhibit different level of catalytic activity. Rabbit cathepsin E has a single substitution of the conserved glycine in the first triad that contains the catalytic aspartate to valine, yet it remains an active enzyme.[35] Retroviral protease from feline foamy virus has a glutamine residue instead of glycine and was shown to be proteolytically active.[36] On the other hand, Bla g 2 resembles a large group of inactive (with a single exception[37]) aspartic proteases called pregnancy-associated glycoproteins (PAG) expressed in pregnant females of ungulate mammals. Amino acid substitutions similar to those described above are found in some of these proteins as well. For example, alanine rather than glycine is present in the first triad of boPAG-1 and poPAG-1, and glycine instead of aspartate is found in the second triad of ovPAG-1.[19] No structural data are available for these proteins, and the crystal structure of Bla g 2 provides the first insight on what impact such amino acid substitutions may have on the active site structure.

The structure of Bla g 2 contains a unique single phenylalanine residue insertion, Phe75A, in the highly conserved flap region. Such an insertion has not been seen before in any aspartic protease. The side-chain of Phe75A occupies the S1 substrate-binding site, revealing a novel self-inhibitory mechanism. A motif consisting of two adjacent aromatic residues on the flap, with a phenylalanine residue following the expected tyrosine at position 75, has also been seen in the sequences of two inactive aspartic proteases, Lma-p54 and ovPAG-1.[19,38] However, due to the lack of structural data for these proteins, it is not known if this phenylalanine residue would be at the position structurally equivalent to Phe75A in Bla g 2.

An unprecedented feature of the active site area of Bla g 2 is the direct interaction between flap residues and the catalytic aspartate residues. Two "new" residues located at the tip of the flap, Phe75A (discussed above) and Asp76 are hydrogen bonded to both "outer" oxygen atoms of Asp32 and Asp215, respectively. These interactions place the flap into a completely closed conformation and make the catalytic aspartate residues less accessible. No direct contact between these aspartate residues and the solvent molecules has been observed in Bla g 2 crystals.

Two other novel features of Bla g 2, which are not found in classical aspartic proteases, may contribute to the stability of the allergen structure. One of them is the increased number of disulfide bridges, five in Bla g 2 instead of two or three commonly found in aspartic proteases. The number of disulfide bridges in a protein molecule usually correlates with its stability. The second unusual property of Bla g 2 is the presence of a metal-binding site, in which a Zn cation is tightly bound to two histidine and two aspartate residues, originating from two secondary structure elements, a β-loop and an α-helix, respectively. Only one of the aspartate residues is conserved in other enzymes, while the three remaining residues are unique to Bla g 2, explaining why zinc binding has not previously been seen in the structures of aspartic proteases. However, the same motif, consisting of two histidine residues and two aspartate residues, was found at the appropriate positions in two other cockroach aspartic proteases (from *L. maderae*[38] and *P. Americana* (the sequence was

kindly provided by Dr Chew Fook Tim, The National University of Singapore)), indicating that this property could be a common feature among cockroach aspartic proteases (Table 1).

The stability of the tertiary structure is very important for IgE recognition of the conformational epitopes of respiratory allergens. Disruption of the native structure reduces allergenic activity[39–41] and can be used to design hypoallergenic recombinant allergens with potential use for immunotherapy.[42] The presence of a zinc-binding site in Bla g 2 that provides additional interactions between two neighboring secondary structure elements, as well as the existence of five disulfide bridges, may increase the stability of the allergen structure and influence its persistence in the environment. Chronic exposure to low doses (1–10 μg/year) of this stable cockroach allergen may explain its importance, as well as indicate why sensitization and exposure to Bla g 2 is associated with asthma.[8,9]

It is generally accepted that structural features of allergens are important determinants of their ability to elicit immunological responses. Whether functional properties of allergens, such as ligand binding or enzymatic activity, potentiate their allergenicity is much less well established. The latter hypothesis was primarily derived from studies showing that some allergens are proteases, and their proteolytic activity potentiates allergenicity through additional mechanisms. Examples include proteolytic cleavage by mite allergens of the low-affinity IgE receptor CD23 and the α-subunit of the interleukin-2 receptor CD25, and increase of permeability in the lung epithelium by disruption of the gap junctions.[12–15,43–46] However, none of the cockroach allergens described so far, except Bla g 2, is homologous to any proteolytic enzymes. We have recently shown that despite the pronounced sequence similarity to aspartic proteases, Bla g 2 was inactive in standard aspartic protease assays. Residual aspartic protease activity, at levels 1000-fold lower than pepsin, was only detected in a modified hemoglobin assay that required high concentration of protein and long incubation time (16 hours), and was inhibited by pepstatin, a specific inhibitor of aspartic proteases.[17] These data indicate that, under certain conditions, the self-inhibited state of the active site could be restored to allow Bla g 2 to regain its enzymatic properties. That would require opening of the flap, with simultaneous removal of the side-chain of Phe75A from the S1 pocket. Similar conformational changes were observed in the structures of chymosin[21,32] and proteinase A.[28] In addition, the activation step for Bla g 2 will also require rearrangement in the interactions of the catalytic aspartate residues to overcome the distortion caused by glycine replacement. Much lower enzymatic activity of Bla g 2 may reflect the shift in the equilibrium in solution between the open and self-inhibited state towards the latter. The present study provides a structural basis for the low level of proteolytic activity of Bla g 2, and thus suggests a lack of correlation between enzymatic and allergenic properties of this molecule.

The overall similarity of Bla g 2 to the structures of other well-known aspartic proteases, as well as its homology to PAG, suggest that ligand binding could be involved in the function of this allergen. This places Bla g 2 as a possible member of an important new group of ligand-binding allergens, such as the mammalian lipocalin allergens; Der p 2, a sterol-binding protein; and Fel d 1, a uteroglobin-like molecule with an amphipathic-binding pocket.[47–49] The hydrophobic ligands bound by rat and mouse lipocalin allergens (Rat n 1 and Mus m 1) were identified in crystallographic experiments.[47] Structural data on the complexes of Bla g 2 with ligands and antibodies will help to elucidate the role of this unusual aspartic protease in allergic respiratory disease and facilitate development of vaccines.

Materials and Methods

Expression and purification of a recombinant Bla g 2 mutant

A partially deglycosylated recombinant Bla g 2 mutant (rBla g 2 N93Q) was expressed in *P. pastoris* after substitution of an asparagine residue by glutamine in the putative N-glycosylation motif NLT at position 93. The template used was Bla g 2 cDNA without the signal peptide, inserted into the *P. pastoris* expression vector pGAPZαC (Invitrogen) for constitutive expression of the allergen. Site-directed mutagenesis was performed using a QuikChange™ Kit (Stratagene) and custom-designed primers. The mutated DNA was transformed into XL1-Blue supercompetent *E. coli* cells (Stratagene), and sequenced to confirm the desired amino acid substitution. Transformation of *P. pastoris* (strain GS115) and expression of recombinant proteins was performed as previously described.[50] Recombinant Bla g 2 was purified from culture media by affinity chromatography over a 7C11 monoclonal antibody (mAb) column as described.[16] Bound proteins were eluted with 0.1 M glycine, 0.15 M NaCl, pH 2.5. Eluted fractions were neutralized with 2 M Tris base (pH 8.5) concentrated using Centricon P-80, MWCO 10,000 (Millipore, Bedford, MA) and dialyzed against PBS. The purity of rBla g 2 was >95% judged by SDS-PAGE, and the yield of protein was quantified using Advanced Protein Assay (Cytoskeleton).

Crystallization of Bla g 2 and heavy-atom search

Crystals of Bla g 2 were obtained using the hanging-drop, vapor-diffusion method at room temperature. Each drop contained 2 μl of 12 mg/ml protein in 20 mM Tris buffer, 0.1 M KCl and 1.4 mM β-mercaptoethanol at pH 7.4 and 2 μl of reservoir solution consisting of 20% (w/v) PEG8000, 0.2 M magnesium acetate and 10 mM DTT in 0.1 M sodium citrate buffer at pH 5.8. None of the solutions utilized for crystal growth contained Zn cations. The largest crystal reached the size of 0.3 mm × 0.1 mm × 0.02 mm after the period of growth of two weeks. More than ten different heavy-atom metal reagents at concentration of 0.1–10 mM were used in the heavy-atom derivative search. Two data sets collected from an ethyl mercury phosphate (EMP) derivative with different

Table 2. Statistics of data collection and processing

	Native		EMP		UA	N3
	N1	N2	E1	E2	Uac	Native
Space group			*C*2			
Unit cell parameters						
a (Å)	141.82	141.45	141.84	141.67	141.67	141.6
b (Å)	38.43	38.39	38.54	38.49	39.00	38.6
c (Å)	71.22	71.16	71.13	71.00	71.35	71.5
β (deg.)	100.84	101.20	101.42	101.47	100.96	100.9
Resolution range (Å)	50.0–2.02	50.0–1.9	50.0–2.1	50.0–2.5	50.0–1.99	50.0–1.3
Unique reflections	24,381	30,830	21,296	13,234	25,318	94,625
Redundancy	2.5	8.4	3.7	3.5	5.3	2.9
R_{sym} (%)[a]	6.8 (35.8)	8.3 (35.3)	13.5 (45.2)	13.7 (40.2)	9.3 (47.1)	5.6 (49.4)
Completeness[a] (%)	96.9 (78.1)	99.9 (99.5)	96.0 (79.0)	97.6 (97.0)	94.1(67.2)	98.1 (91.6)

[a] Values for the outermost shell of data are given in parentheses.

soaking times, ten minutes for E1 and ten hours for E2, as well as one from a uranyl acetate derivative (UA), were used in MIR phasing (Table 2).

Data collection

Six data sets were collected and used at different stages of structure determination (Table 2). Data sets N1, N2, E1, E2, and UA were collected at 100 K with a laboratory-based X-ray source (Rigaku H2R generator operated at 50 kV and 100 mA, equipped with a Mar345 image plate and an Osmic monochromator). The data sets N1, E1, E2, and UA were used in solving the structure and phase refinement, whereas data set N2 was used in the model building and refinement (Table 2). The third native data set, N3, used only in the final stage of refinement (Table 3), was collected at beamline SER-CAT 22-ID equipped with MAR225 CCD detector (APS, Argonne National Laboratory, Argonne IL), at the wavelength of 0.971 Å. All data sets were processed with the HKL2000 package.[51] The presence of bound Zn cations in the crystals of Bla g 2 was established by a fluorescence scan of a single crystal, performed at the beamline X-9B of NSLS (Brookhaven National Laboratory, Upton NY) by scanning through Zn and Ni absorption edges.

MIR phasing and structural refinement

The program SOLVE[52] was used in the heavy-atom search. Seven sites were found for the derivative E1, six for E2 and seven for UA, with the combined figure of merit of 0.64 and score of 47.24. All sites were input into SHARP[53] for heavy-atom refinement and phase calculation, which was followed by density modification with SOLOMON in CCP4.[54] An initial model consisting of 65 complete residues and 155 Ala or Gly residues was built by RESOLVE.[55] After several cycles of manual model rebuilding and model refinement performed using O[56] and CNS,[57], respectively, residues from 3 to 330, two carbohydrate residues and 320 water molecules were built into the density. At the final stage, the structure was refined at 1.3 Å with SHELXL97[58] using the dataset N3. The final refinement statistics, listed in Table 3, are acceptable for structures refined at this resolution. A Ramachandran plot showed 93% of main-chain torsion angles in the most-favored region and 6.3% in the additionally allowed region. Ser92 and Asp76 are found in the generously allowed and disallowed regions, respectively, but excellent electron density supports their correct fit.

Table 3. Final refinement statistics

Resolution (Å)	10–1.3
Number of reflections used in refinement	87,692
Number of reflections reserved for R_{free}	4622
Number of molecules in a.u.	1
Number of protein atoms	2567
Number of solvent molecules	478
Number of heteroatoms	43
R_{cryst} (%)	18.0
R_{free} (%)	21.2
r.m.s. deviations from ideality	
Bond lengths (Å)	0.012
Angle distances (Å)	0.029

Protein Data Bank accession code

The coordinates and structure factors have been deposited in the Protein Data Bank (accession code 1YG9).

Acknowledgements

We thank Dr Zbyszek Dauter for the measurements of fluorescence data from a crystal of Bla g 2 and Jerry Alexandratos for assistance in preparation of the Figures. This work was supported in part by Philip Morris USA Inc, and in part with Federal funds from the National Cancer Institute, National Institutes of Health, under contract no. NO1-CO-24000. Diffraction data used in the final refinement were collected at the Southeast Regional Collaborative Access Team (SER-CAT) beamline 22-ID, located at the Advanced Photon Source, Argonne National Laboratory. Use of the Advanced Photon Source was supported by the U. S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract no. W-31-109-Eng-38. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does the mention of trade names, commercial products, or organizations imply endorsement by the U. S. Government.

References

1. Arruda, L. K., Vailes, L. D., Mann, B. J., Shannon, J., Fox, J. W., Vedvick, T. S. *et al*. (1995). Molecular cloning of a major cockroach (*Blattella germanica*) allergen, Bla g 2. Sequence homology to the aspartic proteases. *J. Biol. Chem.* **270**, 19563–19568.
2. Rosenstreich, D. L., Eggleston, P., Kattan, M., Baker, D., Slavin, R. G., Gergen, P. *et al*. (1997). The role of cockroach allergy and exposure to cockroach allergen in causing morbidity among inner-city children with asthma. *N. Engl. J. Med.* **336**, 1356–1363.
3. Pollart, S. M., Chapman, M. D., Fiocco, G. P., Rose, G. & Platts-Mills, T. A. (1989). Epidemiology of acute asthma: IgE antibodies to common inhalant allergens as a risk factor for emergency room visits. *J. Allergy Clin. Immunol.* **83**, 875–882.
4. Call, R. S., Smith, T. F., Morris, E., Chapman, M. D. & Platts-Mills, T. A. (1992). Risk factors for asthma in inner city children. *J. Pediatr.* **121**, 862–866.
5. Gelber, L. E., Seltzer, L. H., Bouzoukis, J. K., Pollart, S. M., Chapman, M. D. & Platts-Mills, T. A. (1993). Sensitization and exposure to indoor allergens as risk factors for asthma among patients presenting to hospital. *Am. Rev. Respir. Dis.* **147**, 573–578.
6. Christiansen, S. C., Martin, S. B., Schleicher, N. C., Koziol, J. A., Hamilton, R. G. & Zuraw, B. L. (1996). Exposure and sensitization to environmental allergen of predominantly Hispanic children with asthma in San Diego's inner city. *J. Allergy Clin. Immunol.* **98**, 288–294.
7. Eggleston, P. A., Rosenstreich, D., Lynn, H., Gergen, P., Baker, D., Kattan, M. *et al*. (1998). Relationship of indoor allergen exposure to skin test sensitivity in inner-city children with asthma. *J. Allergy Clin. Immunol.* **102**, 563–570.
8. Arruda, L. K., Vailes, L. D., Ferriani, V. P., Santos, A. B., Pomés, A. & Chapman, M. D. (2001). Cockroach allergens and asthma. *J. Allergy Clin. Immunol.* **107**, 419–428.
9. Sporik, R., Squillace, S. P., Ingram, J. M., Rakes, G., Honsinger, R. W. & Platts-Mills, T. A. (1999). Mite, cat, and cockroach exposure, allergen sensitisation, and asthma in children: a case-control study of three schools. *Thorax*, **54**, 675–680.
10. Stewart, G. A. & Thompson, P. J. (1996). The biochemistry of common aeroallergens. *Clin. Expt. Allergy*, **26**, 1020–1044.
11. Harris, J., Mason, D. E., Li, J., Burdick, K. W., Backes, B. J., Chen, T. *et al*. (2004). Activity profile of dust mite allergen extract using substrate libraries and functional proteomic microarrays. *Chem. Biol.* **11**, 1361–1372.
12. Hewitt, C. R., Brown, A. P., Hart, B. J. & Pritchard, D. I. (1995). A major house dust mite allergen disrupts the immunoglobulin E network by selectively cleaving CD23: innate protection by antiproteases. *J. Expt. Med.* **182**, 1537–1544.
13. Schulz, O., Laing, P., Sewell, H. F. & Shakib, F. (1995). Der p I,a major allergen of the house dust mite, proteolytically cleaves the low-affinity receptor for human IgE (CD23). *Eur. J. Immunol.* **25**, 3191–3194.
14. King, C., Brennan, S., Thompson, P. J. & Stewart, G. A. (1998). Dust mite proteolytic allergens induce cytokine release from cultured airway epithelium. *J. Immunol.* **161**, 3645–3651.
15. Wan, H., Winton, H. L., Soeller, C., Tovey, E. R., Gruenert, D. C., Thompson, P. J. *et al*. (1999). Der p 1 facilitates transepithelial allergen delivery by disruption of tight junctions. *J. Clin. Invest.* **104**, 123–133.
16. Pomés, A., Chapman, M. D., Vailes, L. D., Blundell, T. L. & Dhanaraj, V. (2002). Cockroach allergen Bla g 2: structure, function, and implications for allergic sensitization. *Am. J. Respir. Crit. Care Med.* **165**, 391–397.
17. Wünschmann, S., Gustchina, A., Chapman, M. D. & Pomés, A. (2005). Cockroach allergen Bla g 2: an unusual aspartic protease. *J. Allergy Clin. Immunol.* In the press.
18. Xie, S. C., Low, B. G., Nagel, R. J., Kramer, K. K., Anthony, R. V. & Zoli, A. P. (1991). Identification of the major pregnancy-specific antigens of cattle and sheep as inactive members of the aspartic proteinase family. *Proc. Natl Acad. Sci. USA*, **88**, 10247–10251.
19. Guruprasad, K., Blundell, T. L., Xie, S., Green, J., Szafranska, B., Nagel, R. J. *et al*. (1996). Comparative modelling and analysis of amino acid substitutions suggests that the family of pregnancy-associated glycoproteins includes both active and inactive aspartic proteinases. *Protein Eng.* **9**, 849–856.
20. Sielecki, A. R., Fedorov, A. A., Boodhoo, A., Andreeva, N. S. & James, M. N. (1990). Molecular and crystal structures of monoclinic porcine pepsin refined at 1.8 Å resolution. *J. Mol. Biol.* **214**, 143–170.
21. Gilliland, G. L., Winborne, E. L., Nachman, J. & Wlodawer, A. (1990). The three-dimensional structure of recombinant bovine chymosin at 2.3 Å resolution. *Proteins: Struct. Funct. Genet.* **8**, 82–101.
22. Andreeva, N. S., Gustchina, A. E., Fedorov, A. A., Shutzkever, N. E. & Volnova, T. V. (1977). X-ray crystallographic studies of pepsin. *Advan. Expt. Med. Biol.* **95**, 23–31.
23. Jenkins, J., Tickle, I., Sewell, T., Ungaretti, L., Wollmer, A. & Blundell, T. (1977). X-ray analysis and circular dichroism of the acid protease from Endothia parasitica and chymosin. *Advan. Expt. Med. Biol.* **95**, 43–60.
24. Hsu, I. N., Delbaere, L. T., James, M. N. & Hofmann, T. (1977). Penicillopepsin: 2.8 Å structure, active site conformation and mechanistic implications. *Advan. Expt. Med. Biol.* **95**, 61–81.
25. Andreeva, N. S. & Gustchina, A. E. (1979). On the supersecondary structure of acid proteases. *Biochem. Biophys. Res. Commun.* **87**, 32–42.
26. Andreeva, N. (1991). A consensus template of the aspartic proteinase fold. In *Structure and Function of the Aspartic Proteinases* (Dunn, B., ed.), pp. 559–572, Plenum Press, New York, NY.
27. Andreeva, N. S. & Rumsh, L. D. (2001). Analysis of crystal structures of aspartic proteinases: on the role of amino acid residues adjacent to the catalytic site of pepsin-like enzymes. *Protein Sci.* **10**, 2439–2450.
28. Li, M., Phylip, L. H., Lees, W. E., Winther, J. R., Dunn, B. M., Wlodawer, A. *et al*. (2000). The aspartic proteinase from *S. cerevisiae* folds its own inhibitor into a helix. *Nature Struct. Biol.* **7**, 113–117.
29. Sielecki, A. R., Hayakawa, K., Fujinaga, M., Murphy, M. E., Fraser, M., Muir, A. K. *et al*. (1989). Structure of recombinant human renin, a target for cardiovascular-active drugs, at 2.5 Å resolution. *Science*, **243**, 1346–1351.
30. Pearl, L. & Blundell, T. (1984). The active site of aspartic proteinases. *FEBS Letters*, **174**, 96–101.
31. Rahuel, J., Priestle, J. P. & Grutter, M. G. (1991). The crystal structures of recombinant glycosylated human renin alone and in complex with a transition state analog inhibitor. *J. Struct. Biol.* **107**, 227–236.

32. Andreeva, N., Dill, J. & Gilliland, G. L. (1992). Can enzymes adopt a self-inhibited form? Results of X-ray crystallographic studies of chymosin. *Biochem. Biophys. Res. Commun.* **184**, 1074–1081.
33. Gustchina, A., Li, M., Phylip, L. H., Lees, W. E., Kay, J. & Wlodawer, A. (2002). An unusual orientation for Tyr75 in the active site of the aspartic proteinase from *Saccharomyces cerevisiae*. *Biochem. Biophys. Res. Commun.* **295**, 1020–1026.
34. Trylska, J., Grochowski, P. & McCammon, J. A. (2004). The role of hydrogen bonding in the enzymatic reaction catalyzed by HIV-1 protease. *Protein Sci.* **13**, 513–528.
35. Kageyama, T. (1993). Rabbit procathepsin E and cathepsin E. Nucleotide sequence of cDNA, hydrolytic specificity for biologically active peptides and gene expression during development. *Eur. J. Biochem.* **216**, 717–728.
36. Winkler, I., Bodem, J., Haas, L., Zemba, M., Delius, H., Flower, R. *et al*. (1997). Characterization of the genome of feline foamy virus and its proteins shows distinct features different from those of primate spumaviruses. *J. Virol.* **71**, 6727–6741.
37. Green, J. A., Xie, S., Szafranska, B., Gan, X., Newman, A. G., McDowell, K. & Roberts, R. M. (1999). Identification of a new aspartic proteinase expressed by the outer chorionic cell layer of the equine placenta. *Biol. Reprod.* **60**, 1069–1077.
38. Cornette, R., Farine, J. P., Quennedey, B., Riviere, S. & Brossut, R. (2002). Molecular characterization of Lma-p54, a new epicuticular surface protein in the cockroach *Leucophaea maderae* (Dictyoptera, oxyhaloinae). *Insect Biochem. Mol. Biol.* **32**, 1635–1642.
39. Lombardero, M., Heymann, P. W., Platts-Mills, T. A., Fox, J. W. & Chapman, M. D. (1990). Conformational stability of B cell epitopes on group I and group II *Dermatophagoides* spp. Allergens. Effect of thermal and chemical denaturation on the binding of murine IgG and human IgE antibodies. *J. Immunol.* **144**, 1353–1360.
40. Takai, T., Yokota, T., Yasue, M., Nishiyama, C., Yuuki, T., Mori, A. *et al*. (1997). Engineering of the major house dust mite allergen Der f 2 for allergen-specific immunotherapy. *Nature Biotechnol.* **15**, 754–758.
41. Smith, A. M., Chapman, M. D., Taketomi, E. A., Platts-Mills, T. A. & Sung, S. S. (1998). Recombinant allergens for immunotherapy: a Der p 2 variant with reduced IgE reactivity retains T-cell epitopes. *J. Allergy Clin. Immunol.* **101**, 423–425.
42. Chapman, M. D., Smith, A. M., Vailes, L. D., Arruda, L. K., Dhanaraj, V. & Pomés, A. (2000). Recombinant allergens for diagnosis and therapy of allergic disease. *J. Allergy Clin. Immunol.* **106**, 409–418.
43. Schulz, O., Sewell, H. F. & Shakib, F. (1998). Proteolytic cleavage of CD25, the alpha subunit of the human T cell interleukin 2 receptor, by Der p 1, a major mite allergen with cysteine protease activity. *J. Expt. Med.* **187**, 271–275.
44. Shakib, F., Schulz, O. & Sewell, H. (1998). A mite subversive: cleavage of CD23 and CD25 by Der p 1 enhances allergenicity. *Immunol. Today*, **19**, 313–316.
45. Dudler, T., Machado, D. C., Kolbe, L., Annand, R. R., Rhodes, N., Gelb, M. H. *et al*. (1995). A link between catalytic activity. IgE-independent mast cell activation, and allergenicity of bee venom phospholipase A2. *J. Immunol.* **155**, 2605–2613.
46. Tomee, J. F., van Weissenbruch, R., de Monchy, J. G. & Kauffman, H. F. (1998). Interactions between inhalant allergen extracts and airway epithelial cells: effect on cytokine production and cell detachment. *J. Allergy Clin. Immunol.* **102**, 75–85.
47. Bocskei, Z., Groom, C. R., Flower, D. R., Wright, C. E., Phillips, S. E., Cavaggioni, A. *et al*. (1992). Pheromone binding to two rodent urinary proteins revealed by X-ray crystallography. *Nature*, **360**, 186–188.
48. Derewenda, U., Li, J., Derewenda, Z., Dauter, Z., Mueller, G. A., Rule, G. S. & Benjamin, D. C. (2002). The crystal structure of a major dust mite allergen Der p 2, and its biological implications. *J. Mol. Biol.* **318**, 189–197.
49. Kaiser, L., Gronlund, H., Sandalova, T., Ljunggren, H. G., Hage-Hamsten, M., Achour, A. & Schneider, G. (2003). The crystal structure of the major cat allergen Fel d 1, a member of the secretoglobin family. *J. Biol. Chem.* **278**, 37730–37735.
50. Pomés, A., Vailes, L. D., Helm, R. M. & Chapman, M. D. (2002). IgE reactivity of tandem repeats derived from cockroach allergen, Bla g 1. *Eur. J. Biochem.* **269**, 3086–3092.
51. Otwinowski, Z. & Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326.
52. Terwilliger, T. C. & Berendzen, J. (1997). Bayesian MAD phasing. *Acta Crystallog. sect. D*, **53**, 571–579.
53. de la Fortelle, E. & Bricogne, G. (1997). Maximum-likelihood heavy-atom parameter refinement for multiple isomorphous replacement and multiwavelength anomalous diffraction methods. *Methods Enzymol.* **276**, 472–494.
54. CCP4. Collaborative Computational Project Number 4. (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallog. sect. D*, **50**, 760–763.
55. Terwilliger, T. C. & Berendzen, J. (1999). Automated MAD and MIR structure solution. *Acta Crystallog. sect. D*, **55**, 849–861.
56. Jones, T. A. & Kjeldgaard, M. (1997). Electron-density map interpretation. *Methods Enzymol.* **277**, 173–208.
57. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W. *et al*. (1998). Crystallography and NMR system: a new software suite for macromolecular structure determination. *Acta Crystallog. sect. D*, **54**, 905–921.
58. Sheldrick, G. M. & Schneider, T. R. (1997). SHELXL: high-resolution refinement. *Methods Enzymol.* **277**, 319–343.

Edited by I. Wilson

(Received 24 January 2005; received in revised form 24 February 2005; accepted 25 February 2005)