Cockroach Allergen Bla g 2

Structure, Function, and Implications for Allergic Sensitization

ANNA POMÉS, MARTIN D. CHAPMAN, LISA D. VAILES, TOM L. BLUNDELL, and VENUGOPAL DHANARAJ[†]

Asthma and Allergic Diseases Center, Department of Internal Medicine, University of Virginia Health System, Charlottesville, Virginia; and Department of Biochemistry, University of Cambridge, Cambridge, United Kingdom

Exposure to German cockroach (*Blattella germanica*) allergens is associated with the development of chronic respiratory diseases, especially asthma. The mechanism by which allergic patients develop specific immunoglobulin E (IgE) responses to environmental allergens is unknown. However, recent reports provided evidence that enzyme activity, especially proteolytic activity, was a major contributor to allergenicity. Bla g 2 is one of the most potent cockroach allergens (prevalence of IgE responses of 60 to 80%) and shows homology to the aspartic proteinase family of enzymes. We investigated whether the allergenicity of Bla g 2 was linked to its putative enzymatic function. A molecular model of Bla g 2, based on the high resolution crystal structures of pepsin and chymosin, showed that the overall three-dimensional structure of Bla g 2 was similar to that of aspartic proteinases with a well-defined binding pocket. However, critical amino acid substitutions in the catalytic triads and in the "flap" region of the molecule suggested that Bla g 2 was inactive and homologous to mammalian pregnancy-associated glycoproteins. This was confirmed experimentally by enzyme assay. The results show dissociation between enzymatic activity and allergenicity for Bla g 2 and suggest that other genetic and environmental factors are important determinants of sensitization.

Keywords: asthma; inflammation; hypersensitivity; allergens; immunotherapy

Immediate hypersensitivity to cockroach (CR) allergens is associated with the increase in asthma mortality and morbidity seen among lower socioeconomic groups living in the United States. Immunoglobulin E (IgE)-mediated sensitization to CR is a risk factor for emergency room asthma admissions, and CR allergen exposure is strongly linked to asthma morbidity among children living in inner cities (1–7). German cockroaches produce several important allergens, including Bla g 1, Bla g 2, Bla g 4, and Bla g 5, that are secreted and accumulate in the environment. Previous studies have shown that Bla g 2 is a potent allergen that elicits IgE responses in 60 to 80% of cockroach allergic patients and gives positive immediate skin tests at concentrations as low as 10^{-5}–10^{-6} μg/ml (7–9). Recent epidemiological studies have shown that 10 to 100-fold lower levels of CR allergens elicit IgE responses when compared with other common indoor allergens, such as dust mite or cat. In a large cohort of U.S. school children, positive immediate skin tests were associated with exposure to median Bla g 2 levels of 0.32 μg/g (range $>$ 0.08–15 μg/g), whereas comparable figures for dust mite were 38 μg/g (range 24–150 μg/g) (10). Exposure to low levels of Bla g 1 and Bla g 2 has also been associated with wheezing among infants in the first 3 mo of life and with increased proliferative T cell responses (11). Sensitization and exposure data suggest that Bla g 2 is an especially potent allergen.

Bla g 2 is a 36-kD protein that shows primary sequence homology to aspartic proteinases (12). The mechanism by which susceptible individuals produce IgE responses to allergens is not known, but an increasing body of evidence suggests that having functional enzyme activity may explain why some allergens, notably the cysteine and serine protease allergens from house dust mite (Der p 1, Der p 3, and Der p 9) and phospholipase A_2 from bee venom, are particularly potent (13–15). Der p 1 may directly promote IgE synthesis through cleavage of the low-affinity IgE receptor (CD23) on B cells and, indirectly, through cleavage of the α-subunit of the interleukin-2 (IL-2) receptor (CD25) on T cells (16–19). Mite protease allergens disrupt the bronchial epithelium and cause release of proinflammatory cytokines from bronchial epithelial cells, mast cells, and basophils (20–23). Der p 1 disrupts tight junctions and facilitates transepithelial allergen delivery and processing (24, 25). Proteolytically active Der p 1 is reported to significantly enhance IgE production in mice compared with enzymatically inactive allergen (26). These lines of evidence suggest that having enzyme activity is a major contributor to allergenicity: the "enzyme hypothesis" (19). Moreover, some enzymes, including Der p 1 and a nonallergenic endopeptidase from ragweed pollen, may contribute to inflammation by inactivating the α_1-proteinase inhibitor, which is the major natural inhibitor of neutrophil enzymes (27–29).

We investigated whether Bla g 2 would be enzymatically active and therefore contribute to the capacity of Bla g 2 to elicit IgE responses. Aspartic proteinases are widely distributed proteolytic enzymes that act through general acid base hydrolysis, via a noncovalent intermediate (30). The three-dimensional structure of several aspartic proteinases has been determined, and the proteolytic mechanism involves the participation of two coplanar aspartic acid residues in the catalytic triads interacting with a water molecule (31, 32). A molecular model of Bla g 2 was constructed based on the crystal structures of pepsin and chymosin, and the specificity pocket was analyzed by comparative studies using peptidomimetic inhibitors. Both molecular and functional studies show that Bla g 2 is an inactive aspartic proteinase and, surprisingly, that Bla g 2 is related to a family of mammalian pregnancy-associated glycoproteins that are thought to be binding proteins.

METHODS

Modeling

The software package Modeller was used to model the cockroach allergen Bla g 2 using the X-ray structures of porcine pepsin (Protein Data Bank [PDB] code: 5pep) and bovine chymosin (PDB code: 4cms) as templates (33–35). These proteins have 25.7 and 28.2% sequence identity with Bla g 2, respectively. Psi-Blast and ClustalW multiple sequence alignments were used to select and align the se-

(*Received in original form April 5, 2001; accepted in final form October 10, 2001*)

This work was supported by National Institutes of Health Grants AI 32557 and AI 34601.

[†]Deceased. This paper is dedicated to the memory of Dr. Venugopal Dhanaraj, scientist, mentor, and friend. Sadly missed.

Correspondence and requests for reprints should be addressed to Anna Pomés, Ph.D., INDOOR Biotechnologies, Inc., 1216 Harris Street, Charlottesville, VA 22903. E-mail: apomes@inbio.com

Am J Respir Crit Care Med Vol 165. pp 391–397, 2002
DOI: 10.1164/rccm2104027
Internet address: www.atsjournals.org

quences before modeling (36, 37). The final model was further manually adjusted using the interactive graphics package O (38). The Procheck suite program was used to assess the stereochemical quality of the model (39). A structural alignment of pepsin, chymosin, and Bla g 2 was obtained using Joy (40).

Study of Bla g 2 Specificity Pocket

Bla g 2 specificity pockets were studied by docking peptidomimetic inhibitors from aspartic proteinase complexes of known tertiary structure and specificity in the cleft of Bla g 2 and comparing them with the original structures. The complexes used for the analysis were those of rhizopuspepsin complexed with pepstatin (PDB code: 6apr at 2.5 Å), the human pepsin 3A complexed with pepstatin (PDB code: 1pso at 2 Å), and the mouse renin complexed with the decapeptide inhibitor CH-66 (PDB code: 1smr at 2.0 Å) (41–43). Amino acid residues involved in the interactions of the inhibitors with Bla g 2 pockets were studied using the program Contact (44). The program Lsqman was used to superpose the different molecules in the same orientation, and these were further displayed using Setor (45, 46).

Purification of Bla g 2 by Affinity Chromatography

German cockroach frass (194 g) was extracted in borate-buffered saline, pH 8.0, and 0.05% sodium azide at 4° C by stirring overnight. The cockroach extract was centrifuged at 12,000 × *g* for 25 min, and the supernatant was centrifuged again at 24,000 × *g* for 25 min. The second supernatant was filtered under vacuum, and the filtrate was dialyzed at least two times against phosphate-buffered saline (PBS) overnight at 4° C (12).

Cockroach extract was absorbed over a monoclonal affinity chromatography column using anti-Bla g 2 mAb 7C11 coupled to CNBr-activated Sepharose 4B (Pharmacia, Piscataway, NJ). The column was washed with PBS, pH 7.0, and eluted either with 0.005 M glycine in 50% ethylene glycol, pH 10, or with 4 mM hydrochloric acid, pH 2.5–3. The basic elution buffer was immediately neutralized with 0.2 M Na_2HPO_4, pH 7.0. The Bla g 2 concentration was measured by enzyme-linked immunosorbent assay as previously described (47).

Milk Clotting Microtiter Plate Assay for the Detection of Enzyme Activity

Enzymatic activity was measured using the milk clotting microtiter plate assay described previously (48). Pepsin and chymosin were used as activity standards while pepstatin and phenylmethylsulfonyl fluoride (PMSF) were used as inhibitors of aspartic and serine proteinase activities, respectively (48, 49). Serial dilutions of samples (20 μl) were incubated in microtiter wells with 12% wt/vol dried milk in 0.2 M sodium acetate, pH 5.4, 10 mM $CaCl_2$, (80 μl) for 1 h at 37° C. The plate was then inverted, and the presence of milk clotting activity was indicated by the formation of a curd at the bottom of a well.

RESULTS

Bla g 2 Model: Validity, Main Substitutions, and Selection of SCRs and SVRs

The Bla g 2 model has a bilobal structure typical of aspartic proteinases, with each lobe comprised of two β-sheets and two short α-helices (Figure 1A). The overall stereochemical quality of the model as assessed by Procheck is favorable. The catalytic site of aspartic proteinases is located in the bottom of the binding pocket or cleft and is formed by two loops, each containing the triad aspartate–threonine–glycine (DTG). Several features of the Bla g 2 binding pockets differ from the active site of catalytically active aspartic proteinases. First, Bla g 2 has amino acid substitutions in the catalytic triads that are absolutely conserved in all active aspartic proteinases: residues DTG 32–34 and DTG 215–217 (using pepsin numbering) are substituted by DST and DTS, respectively (Figures 1B and 2). Second, even though D32 and D215 from the Bla g 2 triads are coplanar (Figure 1B), which is necessary for enzymatic activity, they are inaccessible to water. Threonine 34 is too close to the aspartate 215 (as shown in Figure 1B) and makes this aspartate inaccessible to the water that is essential for activity. Finally, there is a critical substitution in the "flap"region that comprises the residues 72–81 and forms a β-hairpin that partially covers the active site cleft (Figure 1A). Residue tyrosine 75, which is a highly conserved residue among active aspartic proteinases, is substituted by phenylalanine in Bla g 2 (Figure 2).

The regions of Bla g 2 that are structurally conserved (SCRs) with aspartic proteinases are shown in Table 1. Compared with pepsin (5pep), Bla g 2 has five amino acid insertions, one in the SCR1 and four in the structurally variable regions (SVRs) 3, 6, 8, and 13, which are mostly constituted by loops. During evolution, insertions usually occur in the loops and these regions play an important role in specificity, proteolytic susceptibility, folding, function, structure, and antibody recognition (50). The loop that corresponds to the SVR8 is the most variable in length among the aspartic proteinase family and contains a two-amino acid insertion in Bla g 2 (48,

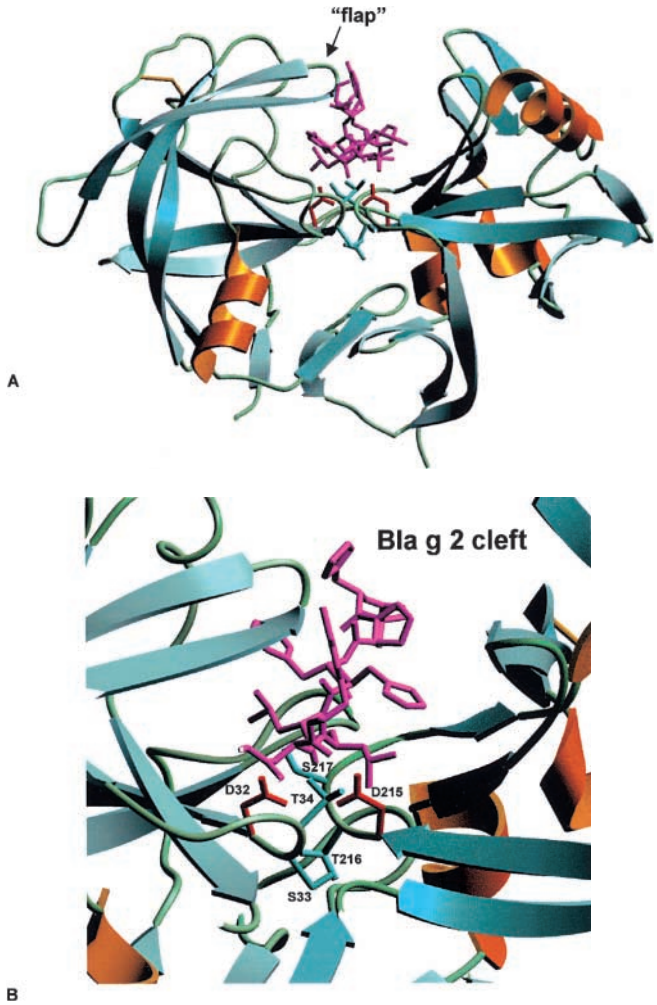

Figure 1. (*A*) Bla g 2 molecular model. Schematic representation of the Bla g 2 model displayed using the program Setor (46). α-Helices are shown in *gold*, β-sheets in *blue* as flat ribbons, and the ropes between these secondary structures represent turn/coil regions in the molecule. The substrate in *fuchsia* (renin inhibitor CH-66) is modeled into the cleft. Two disulfide bonds are shown in *yellow*, and the side chains of the two triads are also shown in the bottom of the cleft: two aspartates in *red* and the serines and threonines in *blue*. (*B*) Bla g 2 area corresponding to the active site of aspartic proteinases. Aspartates 32 and 215 from the two triads are in coplanar disposition. Threonine 34 gets very close to aspartate 215.

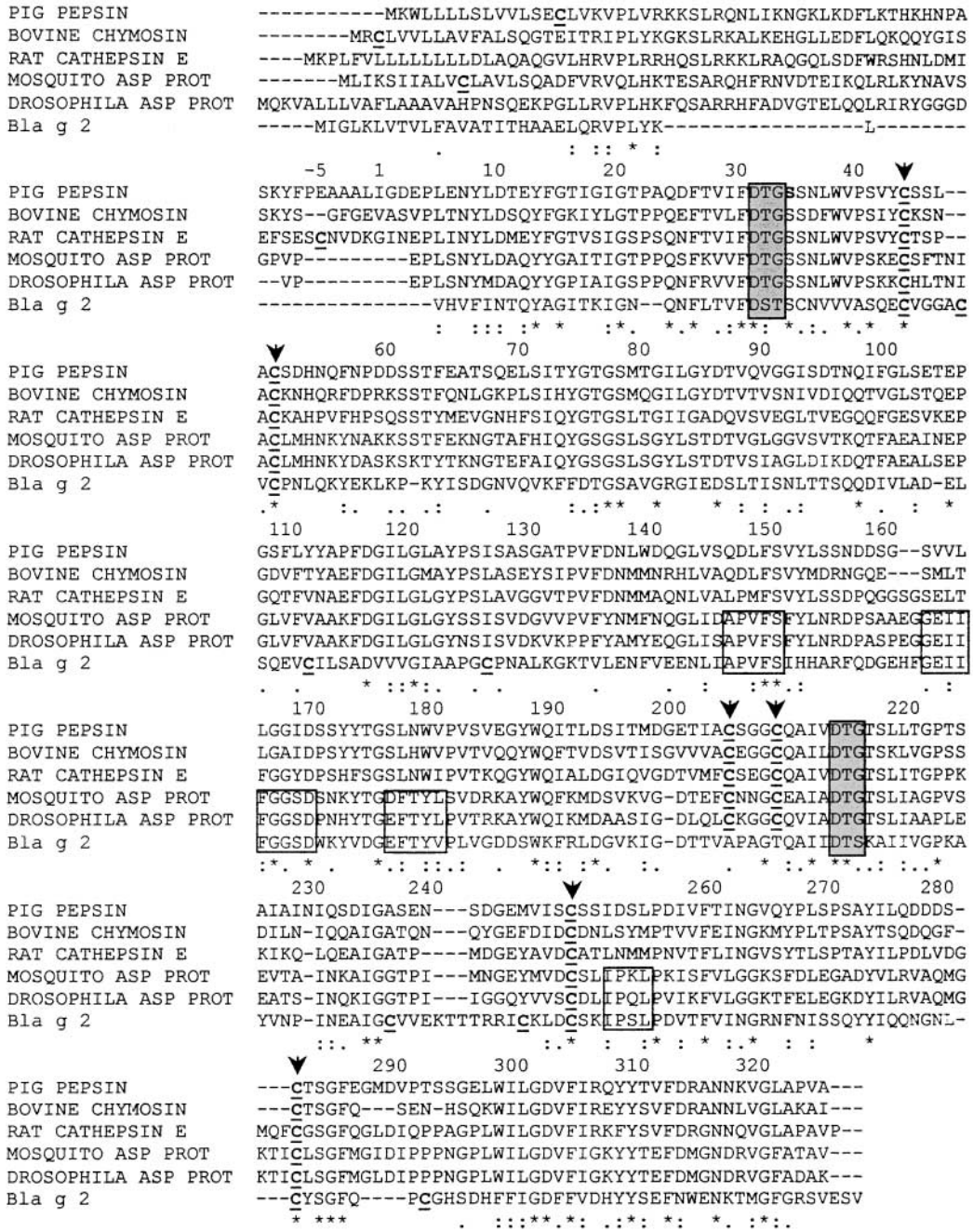

Figure 2. Homologies among Bla g 2-related proteins. Alignment of Bla g 2 with the mosquito lysosomal aspartic proteinase, *Drosophila* aspartic proteinase, pig pepsin, bovine chymosin, and rat cathepsin E (Clustal W alignment). Cysteines are in *bold*, *italics*, and *underlined*. The six cysteines implicated in the three common disulfide bridges in most aspartic proteinases are pointed by *arrows*. *White boxes* indicate regions of homology between the three invertebrate aspartic proteinases (mosquito, *Drosophila*, and Bla g 2). *Gray boxes* indicate catalytic triads, *asterisks* indicate identical or conserved residues in all sequences in the alignment, *colons* indicate conservative substitutions, and *periods* indicate semiconservative substitutions. The amino acid code is the standard single-letter code with a gap denoted by (–).

51). Bla g 2 also has seven amino acid deletions located in regions SCR4, 6, 13, and 16 and the SVR2, 7, and 11 (Table 1).

Cysteine Distribution in Bla g 2

Most aspartic proteinases contain six cysteines that form three disulfide bridges between residues: 45–50, 206–210, and 250–283 (Figures 2 and 3). An unusual feature of Bla g 2 is the presence of 11 cysteines. The second disulfide bridge (i.e., 206–210) is absent in Bla g 2 because of the deletion of cysteine 206 and a substitution of the cysteine 210 by threonine (Figure 3). Each of the other two cysteine pairs is spatially close to two additional cysteines, thus constituting two clusters of four cysteines each. The cysteines 45–50 are close to cysteines 49A and 112, and the cysteines 250–283 are close to cysteines 238 and 248A. From the model, cysteine bonds in each of the two clusters may be interchangeable in Bla g 2. Cysteine 112 corresponds to phenylalanine in chymosin, and leucine in pepsin, and cysteine 238 corresponds to alanine in both aspartic proteinases. However, cysteines 49A and 248A do not have an equivalent amino acid in pepsin and chymosin because they correspond to insertions in Bla g 2. The disulfide bond between cysteines 36 and 128, both in conserved regions of the molecule, is unique to Bla g 2.

Finally, cysteine 294 is not spatially close enough to any of the other cysteines to form an intramolecular disulfide bond (Figure 3). Copurification of a Bla g 2 dimer (in smaller proportion) with a monomer has previously been reported and cysteine 294 would be a good candidate for dimer formation (12, 47). However, addition of 10 nM dithiothreitol (DTT) to natural or recombinant Bla g 2 did not separate it into two

TABLE 1. STRUCTURALLY CONSERVED REGIONS IN THE Bla g 2 MODEL*

SCR Sequence	Characteristics
–43 QRVPLYKLV	SCR1 4 aa insertion
11 NTQYAGITKI	SCR2
	SVR2 2 aa deletion
25 QNFLTVFDSTSCNVVVASQECVG	SCR3 DST triad
	SVR3 2 aa insertion
50 CPNLQKYEKLKPKYISDGNVQVKFF	SCR4 1 aa deletion/Y75→F
79 SAVGRGIEDSLTISNLTTSQQDIVLADE	SCR5
113 CILSADVVVGIAAPGCPNA	SCR6 1 aa deletion
	SVR6 1 aa insertion
134 GKTVLENFVEENLIA	SCR7
	SVR7 1 aa deletion
151 FSIHHAR	SCR8
	SVR8 2 aa insertion
163 GEIIFGGSDWKYVDGEFTYVPLVG	SCR9
190 WKFRLDGVKI	SCR10
203 TTVA	SCR11
	SVR11 1 aa deletion
209 GTQAIIDTSKAIIVGP	SCR12 DTS triad
227 YVNPINEAIGCVVEK	SCR13 1 aa deletion
	SVR13 3 aa insertion
249 DCS	SCR14
253 IPSLPDVTFVINGRNFNISSQYYIQ	SCR15
283 CYSGFQPC	SCR16 4 aa deletion
298 DHFFIGDFFVDHYYSEFN	SCR17
320 TMGFGRSV	SCR18

* Sequences corresponding to the Bla g 2 regions that are structurally conserved (SCRs) with aspartic proteinases. The main charcteristics of the SCRs and some adjacent structurally variable regions (SVRs) compared with the porcine pepsin structure are shown. Numbers indicate the residue position compared with pepsin in the clustal alignment (Figure 2).

monomers; neither did boiling of the sample with addition of up to 100 nM DTT (data not shown). Therefore, cysteine 294 is not responsible for Bla g 2 dimerization.

Specificity Pockets

Comparative analysis of specificity pockets from Bla g 2, pepsin, rhizopuspepsin, and renin for the substrates pepstatin and CH-66 shows that there is a high degree of similarity between amino acid residues found in these pockets. However, the pockets from pepsin, rhizopuspepsin, and renin have more residues close to the inhibitor than Bla g 2. For example, interactions with the aspartic proteinase inhibitor pepstatin are less favorable at the level of the residue 111 for Bla g 2 (valine) and rhizopuspepsin (serine) than for pepsin (phenylalanine) due to the replacement of phenylalanine by a smaller polar residue. Similar substitutions have been described in position 111 in other aspartic proteinases (52, 53). Conversely, other parts of the pocket locally indicate that the Bla g 2 pocket is meant for smaller/different side chains than those in the studied inhibitor. This is the case of isoleucine 10 in Bla g 2 that is too close (< 2 Å) to the renin inhibitor CH-66 (Figure 1). Thus, though the residues that constitute the binding pockets show gross similarity, the shape and size of the pockets show considerable variations. This might indicate that the binding cleft can accommodate a wide range of peptides as in the case of fungal aspartic proteinases.

The polyproline loop in renin (formed by residues in proline 292–proline 297) plays an important role in defining the specificity of the S3′ and S4′ subsites (54). This region does not have an equivalent in pepsin (which has only one proline—proline 293—in this stretch), rhizopuspepsin, and Bla g 2, and, consequently, their S3′ and S4′ specificity pockets are not well defined.

Bla g 2 Is an Inactive Aspartic Proteinase

Bla g 2 was assayed for aspartic proteinase activity using the milk clotting assay. Chymosin and pepsin were used as positive controls (maximum assayed concentrations 10 and 2 μg/well, respectively) and showed milk clotting at concentrations of 0.005 and 0.004 μg/well, respectively. As expected, both pepsin activity and chymosin activity were inhibited by the aspartic proteinase inhibitor pepstatin but not by the serine proteinase inhibitor PMSF (data not shown). Bla g 2 is usually purified over a monoclonal antibody (mAb) affinity column by elution under basic conditions (47). However, because Bla g 2 is secreted into the German cockroach gut where there is a slightly acidic environment, samples eluted under basic or acidic conditions were compared (12, 55). Bla g 2 did not show any milk clotting activity despite the high concentrations of allergen used in the assay (25 μg/well of acid-eluted Bla g 2 and 218 μg/well of base-eluted Bla g 2). German cockroach extracts containing a maximum amount of 1.9 μg Bla g 2 per well showed milk clotting activity that was inhibited by PMSF, but not by pepstatin. Because only serine and aspartic proteinases are known to clot milk, the results prove that cockroach extracts may contain active serine proteases but not active aspartic proteinases.

Bla g 2 Is Similar to the Mammalian Pregnancy-Associated Glycoproteins

The amino acid substitutions in the catalytic triad and the lack of proteolytic activity of Bla g 2 present a scenario similar to that found in mammalian pregnancy-associated glycoproteins (PAGs). The PAGs comprise a group of aspartic proteinases that are secretory products synthesized by the outer epithelial cell layer (chorion) of the placentas of various ungulate species (56, 57). With the exception of PAG-1 from horse (28.0% identity to Bla g 2), most PAGs have amino acid substitutions and deletions that make them inactive (57, 58). In the catalytic triads, bovine PAG-1 and porcine PAG-1 have alanine 34 in-

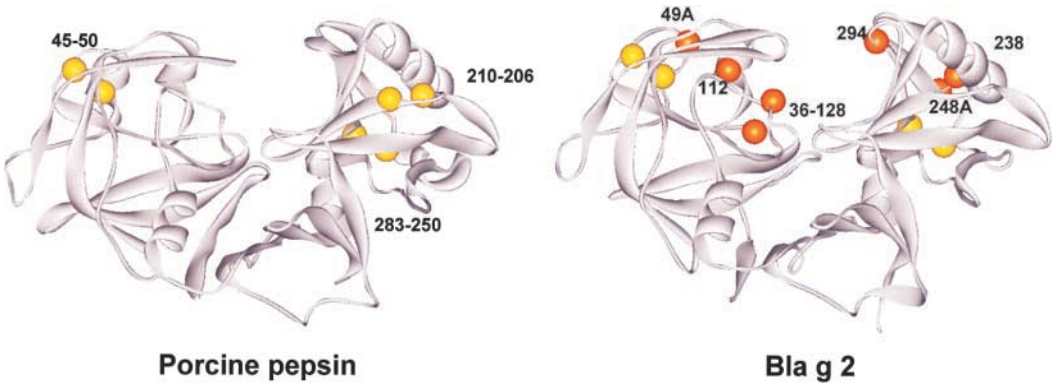

Figure 3. Cysteines distribution in Bla g 2 and pepsin. *Yellow spheres* indicate the cysteines that are conserved in most aspartic proteinases, and *orange spheres* represent the extra cysteines in Bla g 2.

stead of glycine, and ovine PAG-1 has glycine 215 instead of aspartic acid (Figure 4). Substitutions in both triads occur in porcine PAG-1 and the substitution of serine 216 instead of threonine is similar to that of serine 33 in Bla g 2. In the area of the "flap," tyrosine 75 is substituted by a phenylalanine in Bla g 2 and ovine PAG-1 and by proline in porcine PAG-1. Moreover, the substitution of glycine 76, conserved in most aspartic proteinases, by histidine in porcine PAG-1 and by aspartic acid in Bla g 2 may also reduce activity and influence peptide binding (57). Other residue substitutions (115, 117, 189, 290, and 301 by pepsin numbering) or deletions between glutamate 107 and aspartate 118 (as in ovine PAG-2) also contribute to the inactivity of PAGs; Bla g 2 also has a deletion in residue 117 (57).

DISCUSSION

The "enzyme hypothesis" was originally developed to explain why mite allergens are so strongly associated with the development of allergic responses. The finding that most mite allergens were enzymes made this a logical avenue to pursue, and several potentially interesting mechanisms whereby mite protease allergens could enhance IgE responses have been described (14, 16–26). However, our data show that Bla g 2 lacks enzymatic activity, yet is a potent allergen, inducing IgE responses at exposure levels that are often one to two orders of magnitude lower than for Der p 1 (10).

The molecular modeling data provide an explanation for why Bla g 2 is an inactive aspartic proteinase despite sharing the tertiary structure with this group of enzymes. The model predicts that Bla g 2 is similar to pepsin, but has several features that suggest that the allergen is an inactive enzyme: (*1*) the substitutions in the DTG catalytic triads, (*2*) the inaccessibility of aspartates 32 and 215 to water, and (*3*) the substitution of tyrosine 75 by phenylalanine. There are no known examples of aspartic proteinase homologs that have substitutions in the catalytic triads, such as those described for Bla g 2, and still retain proteolytic activity. The catalytic triads must be intact and, specifically, the presence of the small glycine residues is very important for activity. In Bla g 2, the glycine 34 is substituted by the bulkier residue, threonine 34, that gets too close to aspartate 215, displacing the water molecule that is essential for the formation of the intermediate product in the catalytic process of aspartic proteinases (30–32). The only way to reduce the proximity of both residues in the model would be to reorient the side chains of the aspartates; but this would only distort their coplanarity and make them catalytically inactive. The substitution of tyrosine 75 by phenylalanine has been reported to impair the enzymatic activity of recombinant chymosin and *Rhizomucor pusillus* pepsin, since the tyrosyl side chain plays an important role in the stabilization of the tetrahedral intermediate in the catalytic mechanism (59–61). The prediction from the model was experimentally confirmed using the milk clotting enzymatic assay, by showing that neither purified Bla g 2 nor German cockroach extracts had activity that was blocked by the specific aspartic proteinase inhibitor pepstatin A.

Active

		32						215					
Pepsin	(human)	V	F	D	T	G	S	S	I	V	D	T	G	T	S
Cathepsin E	(human)	I	F	D	T	G	S	S	I	V	D	T	G	T	S
Cathepsin D	(human)	V	F	D	T	G	S	S	I	V	D	T	G	T	S
Cathepsin-like	(mosquito)	V	F	D	T	G	S	S	I	A	D	T	G	T	S
Chymosin	(bovine)	L	F	D	T	G	S	S	L	V	D	T	G	T	S
Renin	(human)	V	F	D	T	G	S	S	L	V	D	T	G	A	S
PAG-1	(equine)	I	F	D	T	G	S	A	I	V	D	T	G	T	S

Inactive

PAG-1	(bovine)	V	F	D	T	**A**	S	S	L	V	D	T	G	T	S
PAG-1	(ovine)	V	F	D	T	G	S	S	L	V	**G**	T	G	T	S
PAG-1	(porcine)	I	F	D	T	**A**	S	S	I	L	D	**S**	G	S	A
PAG-2	(ovine)	V	F	D	T	G	S	S	L	V	D	T	G	T	S
Bla g 2	(cockroach)	V	F	D	**S**	**T**	S	C	I	I	D	T	**S**	K	A
				55							239				

Figure 4. Comparison of the active site of aspartic proteinases with the corresponding amino acid triads in Bla g 2 and inactive pregnancy-associated glycoproteins.

Bla g 2 is homologous to the two other insect aspartic proteinases described to date: a lysosomal enzyme from *Aedes aegypti* and a *Drosophila* aspartic proteinase (62, *see* NOTE). Both of these enzymes retain the catalytic triads typical of aspartic proteinases and activity has already been proven for the mosquito protein, whereas Bla g 2 is inactive (63). A surprising feature of the results is that Bla g 2 closely resembles the mammalian PAGs (found in ungulate species), most of which are inactive aspartic proteinases (56–58). Bla g 2 has similar amino acid substitutions in the binding pocket to PAGs and also has a well-defined binding cleft. The function of the PAGs may be related to their ligand-binding capabilities, as has been shown for other aspartic proteinases, for example, from eggs of the hard tick *Boophilus microplus*, which binds heme (64, 65). Comparative analysis of the pockets in pepsin, rhizopuspepsin, renin, and Bla g 2 revealed that changes in pocket composition arise from differences in both sequence and local conformational variability; they are thus determinants of specificity. The fact that the binding site in Bla g 2 does not have stringent specificity might indicate either that the cleft is able to accommodate a wide range of peptides (as with fungal aspartic proteinases) or that the Bla g 2 pocket needs residues bulkier than those in pepstatin and CH-66 for tighter contact with the protein (32).

Several other major allergens provide exceptions to the "enzyme hypothesis." The other cockroach allergens that have been cloned have diverse biological functions and none of them is a proteolytic enzyme (7, 9). Three-dimensional structural studies have shown that Der p 2 has an immunoglobulin fold structure, with primary sequence homology to moth molting protein and human epididymal protein, but no known enzyme function (66, 67). Der p 2 causes sensitization in $>$ 90% of mite allergic patients, at exposure levels that are usually 2 to 10-fold lower than for Der p 1. The functions of other mite allergens, including Der p 5 and Der p 7, are unknown. The major animal allergens (Rat n 1, Mus m 1, Bos d 2, Equ c 1, Can f 1, and Can f 2) are lipocalins, which function as pheromone-binding proteins or transporter proteins (9). This evidence confirms that enzymatic activity of the allergen is not a prerequisite for allergenicity. Recent structural studies have shown that allergens have diverse biological functions—they may be enzymes, structural proteins, binding proteins, and also enzyme inhibitors (9). The fact that several mite allergens are enzymes appears to represent a special case.

Although enzyme function per se is not required for proteins to elicit IgE responses, the results do not rule out the possibility that enzymes inhaled into the respiratory tract could contribute to inflammation. This could occur if particles that are inhaled into the lung carry both allergens and enzymes simultaneously, and clearly both characteristics may coexist in certain molecules, such as Der p 1. The present results show that cockroach extracts contain serine proteases, but not functional aspartic proteinases. Some enzymes may contribute to inflammation by inactivating the α_1-proteinase inhibitor (27–29). The key elements that affect production of IgE Ab appear to be route of exposure, allergen dose, and host im-

mune response genes, all of which preferentially stimulate Th2 responses. The critical effect of allergen dose is emphasized by recent studies that show that high-dose exposure to cat allergen (> 20 μg/g) results in reduced prevalence of IgE Ab responses to Fel d 1 (10, 68). Low-dose exposure to Fel d 1 (0.5–5 μg/g) posed the strongest risk for sensitization. In the same cohort, median exposure to Bla g 2 was 0.32 μg/g, suggesting that Bla g 2 exposure routinely falls in a window that represents the highest risk for sensitization (10). Recognition of low doses of Bla g 2 may be associated with certain HLA DR genes, as has been observed for other allergens, such as Ole e 1, Fel d 1, and the mountain cedar allergen (69–71). Association of the HLA-*DRB1*0101* allele in the Hutterite population and the HLA-*DRB1*0102* allele in African Americans with sensitization to cockroach allergens has been reported (72).

In summary, we carried out comparative modeling and experimental studies to investigate the function of the CR allergen Bla g 2. Knowledge-based molecular modeling of Bla g 2 shows that the overall three-dimensional structure is similar to that of aspartic proteinases, with an intact substrate-binding cleft that could confer a peptide-binding function to this protein. Knowledge of cysteine distribution may help to design recombinant variants for immunotherapy that are able to induce T cell responses and with low IgE binding capacity to avoid undesired secondary effects. The results show that the strong capacity of Bla g 2 to stimulate IgE production following low-dose environmental exposure is unrelated to aspartic proteinase activity, which is also absent in cockroach extracts. Bla g 2 is one of several indoor allergens that show dissociation between enzyme activity and allergenicity. Knowledge of the molecular structure of Bla g 2 will provide a better understanding of the immune response to CR allergens and should allow rational immunotherapeutic strategies for treatment of cockroach allergy to be developed.

Note: The nucleotide sequence for the *Drosophila* gene related to cathepsin D was deposited in the GenBank database by A.W. Page-McCaw, G. Tsang, and G.M. Rubin (December 1999) under GenBank Accession Number AF220040. The amino acid sequence of this protein can be accessed through the NCBI Protein Database, Accession Number AAF23824.

Acknowledgment: We are grateful to Peter Ngo for his technical assistance, and to Wanda Harvey for secretarial support.

References

1. Pollart SM, Chapman MD, Fiocco GP, Rose G, Platts-Mills TAE. Epidemiology of acute asthma: IgE antibodies to common inhalant allergens as a risk factor for emergency room visits. *J Allergy Clin Immunol* 1989;83:875–882.
2. Call RS, Smith TF, Morris E, Chapman MD, Platts-Mills TAE. Risk factors for asthma in inner city children. *J Pediatr* 1992;121:862–866.
3. Gelber LE, Seltzer LH, Bouzoukis JK, Pollart SM, Chapman MD, Platts-Mills TAE. Sensitization and exposure to indoor allergens as risk factors for asthma among patients presenting to hospital. *Am Rev Respir Dis* 1993;147:573–578.
4. Christiansen SC, Martin SB, Schleicher NC, Koziol JA, Hamilton RG, Zuraw BL. Exposure and sensitization to enviromental allergen of predominantly Hispanic children in San Diego's inner city. *J Allergy Clin Immunol* 1996;98:288–294.
5. Rosenstreich DL, Eggleston P, Kattan M, Baker D, Slavin RG, Gergen P, Mitchell H, McNiff-Mortimer K, Lynn H, Ownby D, Malveaux F. The role of cockroach allergy and exposure to cockroach allergen in causing morbidity among inner-city children with asthma. *N Engl J Med* 1997;336:1356–1363.
6. Eggleston PA, Rosenstreich D, Lynn H, Gergen P, Baker D, Kattan M, Mortimer KM, Mitchell H, Ownby D, Slavin R, Malveaux F. Relationship of indoor allergen exposure to skin test sensitivity in inner-city children with asthma. *J Allergy Clin Immunol* 1998;102:563–570.
7. Arruda LK, Vailes LD, Ferriani VPL, Santos ABR, Pomés A, Chapman MD. Cockroach allergens and asthma. *J Allergy Clin Immunol* 2001; 107:419–428.
8. Chapman MD, Vailes LD, Hayden HL, Platts-Mills TAE, Arruda LK. Cockroach allergens and their role in asthma. In: Kay AB, editor. Allergy and allergic diseases. Oxford, UK: Blackwell Science Ltd.; 1997. p. 942–951.
9. Chapman MD, Smith AM, Vailes LD, Arruda LK, Dhanaraj V, Pomés A. Recombinant allergens for diagnosis and therapy of allergic disease. *J Allergy Clin Immunol* 2000;106:409–418.
10. Sporik R, Squillace SP, Ingram JM, Rakes G, Honsinger RW, Platts-Mills TAE. Mite, cat, and cockroach exposure, allergen sensitization, and asthma in children: a case-control study of three schools. *Thorax* 1999;54:675–680.
11. Finn PW, Boudreau JO, He H, Wang Y, Chapman MD, Vincent C, Burge HA, Weiss ST, Perkins DL, Gold DR. Children at risk for asthma: home allergen levels, lymphocyte proliferation, and wheeze. *J Allergy Clin Immunol* 2000;105:933–942.
12. Arruda LK, Vailes LD, Mann BJ, Shannon J, Fox JW, Vedvick TS, Hayden ML, Chapman MD. Molecular cloning of a major cockroach (*Blattella germanica*) allergen, Bla g 2. *J Biol Chem* 1995;270:19563–19568.
13. King C, Brennan S, Thompson PJ, Stewart GA. Dust mite proteolytic allergens induce cytokine release from cultured airway epithelium. *J Immunol* 1998;161:3645–3651.
14. Hewitt CRA, Foster S, Phillips C, Horton H, Jones RM, Brown AP, Hart BJ, Pritchard DI. Mite allergens: significance of enzymatic activity. *Allergy* 1998;53(Suppl 48):60–63.
15. Dudler T, Machado DC, Kolbe L, Annand RR, Rhodes N, Gelb MH, Koelsch E, Suter M, Helm BA. A link between catalytic activity, IgE-independent mast cell activation, and allergenicity of bee venm phospholipase A2. *J Immunol* 1995;155:2605–2615.
16. Hewitt CRA, Brown AP, Hart BJ, Pritchard DI. A major house dust mite allergen disrupts the immunoglobulin E network by selectively cleaving CD23: innate protection by antiproteases. *J Exp Med* 1995; 182:1537–1544.
17. Schulz O, Laing P, Sewell HF, Shakib F. Der p I, a major allergen of the house dust mite, proteolytically cleaves the low-affinity receptor for human IgE (CD23). *Eur J Immunol* 1995;25:3191–3194.
18. Schulz O, Sewell HF, Shakib F. Proteolytic cleavage of CD25, the α subunit of the human T cell interleukin 2 receptor, by Der p 1, a major mite allergen with cysteine protease activity. *J Exp Med* 1998;187:271–275.
19. Shakib F, Schulz O, Sewell H. A mite subversive: cleavage of CD23 and CD25 by Der p 1 enhances allergenicity. *Immunol Today* 1998;19: 313–316.
20. Herbert CA, King CM, Ring PC, Holgate ST, Stewart GA, Thompson PJ, Robinson C. Augmentation of permeability in the bronchial epithelium by the house dust mite allergen Der p 1. *Am J Respir Cell Mol Biol* 1995;12:369–378.
21. Robinson C, Kalsheker NA, Srinivasan N, King CM, Garrod DR, Thompson PJ, Stewart GA. On the potential significance of the enzymatic activity of mite allergens to immunogenicity. Clues to structure and function revealed by molecular characterization. *Clin Exp Allergy* 1997;27:10–21.
22. Tomee JFC, van Weissenbruch R, Monchy JGR, Kauffman HF. Interactions between inhalant allergen extracts and airway epithelial cells: effect on cytokine production and cell detachment. *J Allergy Clin Immunol* 1998;102:75–85.
23. Machado DC, Horton D, Harrop R, Peachell PT, Helm BA. Potential allergens stimulate the release of mediators of the allergic response from cells of mast cell lineage in the absence of sensitization with antigen-specific IgE. *Eur J Immunol* 1996;26:2972–2980.
24. Wan H, Winton HL, Soeller C, Tovey ER, Gruenert DC, Thompson PJ, Stewart GA, Taylor GW, Garrod DR, Cannell MB, Robinson C. Der p 1 facilitates transepithelial allergen delivery by disruption of tight junctions. *J Clin Invest* 1999;104:123–133.
25. Wan H, Winton HL, Soeller C, Gruenert DC, Thompson PJ, Cannell MB, Stewart GA, Garrod DR, Robinson C. Quantitative structural and biochemical analyses of tight junction dynamics following exposure of epithelial cells to house dust mite allergen Der p 1. *Clin Exp Allergy* 2000;30:685–698.
26. Gough L, Shulz O, Sewell HF, Shakib F. The cysteine protease activity of the major dust mite allergen Der p 1 selectively enhances the immunoglobulin E antibody response. *J Exp Med* 1999;190:1897–1902.
27. Kalsheker NA, Deam S, Chambers L, Sreedharan S, Brocklehurst K, Lomas DA. The house dust mite allergen Der p 1 catalytically inactivates α_1-antitrypsin by specific reactive centre loop cleavage: a mechanism that promotes airway inflammation and asthma. *Biochem Biophys Res Commun* 1996;221:59–61.

28. Bagarozzi DA, Pike R, Potempa J, Travis J. Purification and characterization of a novel endopeptidase in ragweed (*Ambrosia artemisiifolia*) pollen. *J Biol Chem* 1996;271:26227–26232.
29. Lindmark B, Svenonius E, Eriksson S. Heterozygous alpha 1-antichymotrypsin and PiZ alpha 1-antitrypsin defficiency. Prevalence and clinical spectrum in asthmatic children. *Allergy* 1990;45:197–203.
30. Blundell, TL, Guruprasad K, Albert A, Williams M, Sibanda BL, Dhanaraj V. The aspartic proteinases. An historical overview. In: James MNG, editor. Aspartic proteinases. New York: Plenum Press; 1998. p. 1–13.
31. Pearl L, Blundell T. The active site of aspartic proteinases. *FEBS Lett* 1984;174:96–101.
32. Dhanaraj,V, Cooper JB. Rational design of renin inhibitors. In: Veerapandian P, editor. Structure-based drug design. New York: Marcel Dekker, Inc.; 1997. p. 321–342.
33. Šali A, Blundell TL. Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol* 1993;234:779–815.
34. Cooper JB, Khan G, Taylor G, Tickle IJ, Blundell TL. X-ray analyses of aspartic proteinases. II. Three-dimensional structure of the hexagonal crystal form of porcine pepsin at 2.3 Å resolution. *J Mol Biol* 1990;214: 199–222.
35. Newman M, Safro M, Frazao C, Kahn G, Zdanov A, Tickle IJ, Blundell TL, Andreeva N. X-ray analyses of aspartic proteinases. IV. Structure and refinement at 2.2 Å resolution of bovine chymosin. *J Mol Biol* 1991;221:1295–1309.
36. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 1997;25:3389–3402.
37. Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 1994;22:4673–4680.
38. Jones TA, Zou JY, Cowan SW, Kjeldgaard M. Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr* 1991;A47:110–119.
39. Laskowski RA, MacArthur MW, Moss DS, Thornton JM. PROCHECK: a program to check the stereochemical quality of protein structures. *J Appl Crystallogr* 1993;26:283–291.
40. Mizuguchi K, Deane CM, Blundell TL, Johnson MS, Overington JP. JOY: protein sequence-structure representation and analysis. *Bioinformatics* 1998;14:617–623.
41. Suguna K, Padlan EA, Bott R, Boger J, Parris KD, Davies DR. Structures of complexes of rhizopuspepsin with pepstatin and other statine-containing inhibitors. *Proteins* 1992;13:195–205.
42. Fujinaga M, Chernaia MM, Tarasova NI, Mosimann SC, James MNG. Crystal structure of human pepsin and its complex with pepstatin. *Prot Sci* 1995;4:960–972.
43. Dealwis C, Frazao C, Badasso M, Cooper J, Tickle I, Driessen H, Blundell T, Murakami K, Miyazaki M, Sueiras-Diaz J, Jones D, Szelke M. X-ray analysis at 2 Å resolution of mouse submaxillary renin complexed with a decapeptide inhibitor Ch-66 based on the 4–16 fragment of rat angiotensinogen. *J Mol Biol* 1994;236:342–360.
44. Collaborative computational project, number 4. "The CCP4 Suite: Programs for Protein Crystallography." *Acta Crystallogr* 1994;D50:760–763.
45. Kleywegt GJ, Jones TA. Where freedom is given, liberties are taken. *Structure* 1995;3:535–540.
46. Evans SV. SETOR: hardware-lighted three-dimensional solid model representations of macromolecules. *J Mol Graphics* 1993;(11):134–138.
47. Pollart SM, Mullins DE, Vailes LD, Hayden ML, Platts-Mills TAE, Sutherland WM, Chapman MD. Identification, quantification, and purification of cockroach allergens using monoclonal antibodies. *J Allergy Clin Immunol* 1991;87:511–521.
48. Nugent PG, Albert A, Orprayoon P, Wilsher J, Pitts JE, Blundell TL, Dhanaraj V. Protein engineering loops in aspartic proteinases: site-directed mutagenesis, biochemical characterization and X-rays analysis of chymosin with a replaced loop from rhizopuspepsin. *Protein Eng* 1996;9:885–893.
49. Emtage JS, Angal S, Doel MT, Harris TJR, Jenkins B, Lilley G, Lowe PA. Synthesis of calf prochymosin (prorennin) in *Escherichia coli. Proc Natl Acad Sci USA* 1983;80:3671–3675.
50. Pickersgill R, Varvill K, Jones S, Perry B, Fischer B, Henderson I, Garrard S, Sumner I, Goodenough P. Making a small enzyme smaller; removing the conserved loop structure of hen lsysozyme. *FEBS Lett* 1994;347:199–202.
51. Bailey D, Cooper JB. A structural comparison of 21 inhibitor complexes of the aspartic protease from *Endothia parasitica. Protein Sci* 1994;3: 2129–2143.
52. Guruprasad K, Törmäkangas K, Kervinen J, Blundell TL. Comparative modelling of barley-grain aspartic proteinase: a structural rationale for observed hydrolytic specificity. *FEBS Lett* 1994;352:131–136.
53. Strop P, Sedlacek J, Stys J, Kaderabkova K, Blaha I, Pavlickova L, Pohl J, Fabry M, Kostka V, Newman M, Frazao C, Shearer A, Tickle IJ, Blundell TL. Engineering enzyme subsite specificity: preparation, kinetic characterization, and X-ray analysis at 2.0-Å resolution of Val111Phe site-mutated calf chymosin. *Biochemistry* 1990;29:9863–9871.
54. Dhanaraj V, Dealwis CG, Frazao C, Badasso M, Sibanda BL, Tickle IJ, Cooper JB, Driessen HPC, Newman M, Aguilar C, Wood SP, Blundell TL, Hobart PM, Geoghegan KF, Ammirati MJ, Danley DE, O'Connor BA, Hoover DJ. X-ray analyses of peptide-inhibitor complexes define the structural basis of specificity for human and mouse renins. *Nature* 1992;357:466–472.
55. Cruden DL, Markovetz AJ. Microbial ecology of the cockroach gut. *Ann Rev Microbiol* 1987;41:617–643.
56. Xie S, Low BG, Nagel RJ, Kramer KK, Anthony RV, Zoli AP, Beckers J-F, Roberts RM. Identification of the major pregnancy-specific antigens of cattle and sheep as inactive members of the aspartic proteinase family. *Proc Natl Acad Sci USA* 1991;88:10247–10251.
57. Guruprasad K, Blundell TL, Xie S, Green J, Szafranska B, Nagel RJ, McDowell K, Baker CB, Roberts RM. 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* 1996;9:849–856.
58. Green J, Xie S, Gan X, Roberts RM. An aspartic proteinase expressed in the equine placenta. *Adv Exp Med Biol* 1998;436:163–167.
59. Suzuki J, Sasaki K, Sasao Y, Hamu A, Kawasaki H, Nishiyama M, Horinouchi S, Beppu T. Alteration of catalytic properties of chymosin by site-directed mutagenesis. *Prot Eng* 1989;2:563–569.
60. Park YN, Aikawa J, Nishiyama M, Horinouchi S, Beppu T. Involvement of a residue at position 75 in the catalytic mechanism of a fungal aspartic proteinase, *Rhizomucor pusillus* pepsin. Replacement of tyrosine 75 on the flap by asparagine enhances catalytic efficiency. *Protein Eng* 1996;9:869–875.
61. Blundell TL, Cooper J, Foundling SI, Jones DM, Atrash B, Szelke M. On the rational design of renin inhibitors: X-ray studies of aspartic proteinases complexed with transition-state analogues. *Biochemistry* 1987;26:5585–5590.
62. Cho WL, Raikhel AS. Cloning of cDNA for mosquito lysosomal aspartic protease. Sequence analysis of an insect lysosomal enzyme similar to cathepsins D and E. *J Biol Chem* 1992;267:21823–21829.
63. Cho WL, Dhadialla TS, Raikhel AS. Purification and characterization of a lysosomal aspartic protease with cathepsin D activity from the mosquito. *Insect Biochem* 1991;21:165–176.
64. Xie S, Green J, Bixby JB, Szafranska B, DeMartini JC, Hecht S, Roberts RM. The diversity and evolutionary relationships of the pregnancy-associated glycoproteins, an aspartic proteinase subfamily consisting of many trophoblast-expressed genes. *Proc Natl Acad Sci USA* 1997; 94:12809–12816.
65. Sorgine MHF, Logullo C, Zingali RB, Paiva-Silva GO, Juliano L, Oliveira PL. A heme-binding aspartic proteinase from the eggs of the hard tick *Boophilus microplus. J Biol Chem* 2000;275:28659–28665.
66. Thomas WR, Chua KY. The major mite allergen Der p 2—a secretion of the male mite reproductive tract? *Clin Exp Allergy* 1995;25:667–669.
67. Mueller GA, Benjamin DC, Rule GS. Tertiary structure of the major house dust mite allergen Der p 2: sequential and structural homologies. *Biochemistry* 1998;37:12707–12714.
68. Platts-Mills T, Vaughan J, Squillace S, Woodfolk J, Sporik R. Sensitization, asthma, and a modified Th2 response in children exposed to cat allergen: a population-based cross-sectional study. *Lancet* 2001;357: 752–756.
69. Cárdaba B, De Pablo R, Vilches C, Martín E, Geller-Bernstein C, De Andrés B, Zaharan Y, Del Pozo V, Gallardo S, De Arruda Chaves E, Waisel Y, Palomino P, Kreisler M, Lahoz C. Allergy to olive pollen: T-cell response from olive allergic patients is restricted by DR7-DQ2 antigens. *Clin Exp Allergy* 1996;26:316–322.
70. Haselden BM, Kay AB, Larché M. Immunoglobulin E-independent major histocompatibility complex-restricted T cell peptide epitope-induced late asthmatic reactions. *J Exp Med* 1999;189:1885–1894.
71. Reid MJ, Nish WA, Whisman BA, Goetz DW, Hylander RD, Parker WA, Freeman TM. HLA-DR4-associated nonresponsiveness to mountain-cedar allergen. *J Allergy Clin Immunol* 1992;89:593–598.
72. Donfack J, Tsalenko A, Danessa MH, Parry R, Solway J, Lester LA, Ober C. HLA-DRB1*01 alleles are associated with sensitization to cockroach allergens. *J Allergy Clin Immunol* 2000;105:960–966.