# Three-Dimensional Modelling of Honeybee Venom Allergenic Proteases: Relation to Allergenicity

Dessislava Georgieva[a], Kerstin Greunke[b], Raghuvir K. Arni[c], and Christian Betzel[a,*]

[a] Institute of Biochemistry and Molecular Biology, University of Hamburg, Laboratory of Structural Biology of Infection and Inflammation, c/o DESY, Notkestrasse 85, Build. 22a, D-22603 Hamburg, Germany. Fax: +49 40 8998 4747. E-mail: Christian.Betzel@uni-hamburg.de
[b] Protein Ligand Structural (PLS)-Design GmbH, Eichenstrasse 42, Hamburg, Germany
[c] Department of Physics, IBILCE/UNESP, Cristóvão Colombo 2265, CEP 15054-000, São José do Rio Preto, SP, Brazil

* Author for correspondence and reprint requests

Z. Naturforsch. **66 c**, 305 – 312 (2011); received July 23, 2010/March 14, 2011

Api SI and Api SII are serine proteases of the honeybee venom containing allergenic determinants. Each protease consists of two structural modules: an N-terminal CUB (Api SI) or a clip domain (Api SII) and a C-terminal serine protease-like (SPL) domain. Both domains are connected with a linker peptide. The knowledge about the structure and function of Api SI and Api SII is limited mainly to their amino acid sequences. We constructed 3-D models of the two proteases using their amino acid sequences and crystallographic coordinates of related proteins. The models of the SPL domains were built using the structure of the prophenoloxidase-activating factor (PPAF)-II as a template. For modelling of the Api SI CUB domain the coordinates of porcine spermadhesin PSP-I were used. The models revealed the catalytic and substrate-binding sites and the negatively charged residue responsible for the trypsin-like activity. IgE-binding and antigenic sites in the two allergens were predicted using the models and programs based on the structure of known epitopes. Api SI and Api SII show structural and functional similarity to the members of the PPAF-II family. Most probably, they are part of the defence system of *Apis mellifera*.

*Key words:* Honeybee Venom, Allergenic Proteases, Protein Modelling

## Introduction

Honeybee stings result very often in IgE-mediated allergic reactions (Müller, 2002). The honeybee venom (HBV) contains a wide diversity of pharmacologically active peptides, proteins, and other organic molecules responsible for these undesirable effects (Hoffman, 2006). Among the venom components, serine proteases form an important group of proteins involved in different physiological processes. Analysis of the *Apis mellifera* genome revealed serine protease and serine protease homolog genes (Zou *et al.*, 2006). A possible involvement of these enzymes in embryonic development and innate immunity was supposed (Zou *et al.*, 2006). However, serine proteases are involved in various physiological processes and other effects can be also expected. At present little is known about the structure-function relationships of the HBV proteolytic enzymes.

The list of HBV allergens, published by the Allergen Nomenclature Sub-Committee of the International Union of Immunological Societies (http://www.allergen.org/Allergen.aspx), includes ten venom proteins: Api m 1 (phospholipase $A_2$), Api m 2 (hyaluronidase), Api m 3 (acid phosphatase), Api m 4 (melittin), Api m 5 (dipeptidylpeptidase allergen IV), Api m 6 (allergen of 8 kDa), Api m 7 (CUB serine protease), Api m 8 (carboxylesterase), Api m 9 (serine carboxypeptidase), and Api m 10 (icarapin variant 2, carbohydrate-rich protein). Several of these allergens were obtained in recombinant forms: Api m 1 (Dudler *et al.*, 1992), Api m 2 (Soldatova *et al.*, 1998), and Api m 3 (Soldatova *et al.*, 2000; Grunwald *et al.*, 2006). Recombinant venom allergens have a potential to improve the diagnosis and treatment of Hymenoptera venom allergy (Hoffman, 2006). However, the two serine proteases, which differ from Api m 7 in their amino acid sequences, are not included in the list mentioned above. For this reason they were named Api SI and Api SII allergens.

Here we describe three-dimensional (3-D) modelling and structure-function relationships of the two HBV allergenic serine proteases Api SI and Api SII. The structural information of the models and programs based on the structure of known epitopes allowed the prediction of allergenic and antigenic sites, as well as the physiological function of both proteases.

## Material and Methods

### *Construction of homology models of Api SI and Api SII*

The sequence alignment technique BLAST (Altschul *et al.*, 1990) was applied to search for primary structure similarities between each serine protease and other proteins. Homology models, based on the sequences of the two HBV enzyme serine protease-like (SPL) domains, were constructed using the crystallographic structure of the prophenoloxidase-activating factor (PPAF)-II (Piao *et al.*, 2005) as a template and applying the SWISS-MODEL server (Schwede *et al.*, 2003). The SPL modules of Api SI and Api SII have 30% and 31% identical residues in comparison to the respective domain of PPAF-II. The homology increased to 40% and 41%, respectively, when conservative substitutions were taken into account. For building of the Api SI CUB domain model the crystallographic structure of the porcine spermadhesin PSP-I (Romero *et al.*, 1997; PSP-I/PSP-II: PDB code 1SPP) was used. Alignment of the PSP-I/CUB domain of Api SI sequences using BLAST search showed 34% identity and 58% homology. The stereochemistry of the models was checked by the program PROCHECK (Laskowski *et al.*, 1993). The surface accessibility was calculated by the program AREAIMOL (Lee and Richards, 1971; Collaborative Computational Project Number 4, 1994).

### *Prediction of allergenicity, IgE binding epitopes, and antigenic determinants*

Prediction of allergenic proteins and mapping of IgE epitopes was performed by the method described in Saha and Raghava (2006) and using the respective programs. The data set applied for testing consisted of 578 allergens and 700 non-allergens. The authors used protein amino acid compositions, a database of known IgE epitopes, and BLAST search against allergen representative peptides. A sequence is predicted to be allergenic by these programs when the score is less than the threshold.

Antigenic determinants of Api SI and Api SII were predicted by the method given in Kolaskar and Tongaonkar (1990).

## Results

Homology modelling is a useful method for predicting structure-function relationships when the protein crystal structure is unknown. This method is based on the observation that the protein three-dimensional structure is more conserved than the amino acid sequence (Marti-Renom *et al.*, 2000). The first and critical step in homology modelling is to find a suitable template for the "target" protein. The quality of the constructed atomic resolution model depends on the degree of sequence identity between the "query sequence" and that of the template, as well as of the quality of the known structure used for comparative modelling. The amino acid sequences of Api SI and Api SII were subjected to a BLAST search (Altschul *et al.*, 1990). The NCBI reference sequence of Api SI is XP_392669 and that of Api SII is XP_623069. PPAF-II (Piao *et al.*, 2005; PDB code 2B9L, chain A) was the protein with a known 3-D structure showing the highest sequence similarity towards the target proteins. The crystal structure of the template protein was solved at 2.0 Å resolution (Piao *et al.*, 2005). Comparison of the sequences of Api SI, Api SII, and PPAF-II is shown in Fig. 1. Both allergens consist of two structural modules each. The first protease is a 39-kDa protein and has an N-terminal CUB domain (residues 44–131) and a C-terminal SPL domain (residues 161–405) connected with a "linker" peptide. The second protease, Api SII, has a molecular mass of 34 kDa, and its polypeptide chain is 52 residues shorter than that of the other enzyme. Api SII consists of peptide-linked clip (residues 29–96) and SPL (residues 99–353) domains. The clip domain of Api SII could not be modeled due to the very low degree of sequence similarity with the respective domain of PPAF-II. A comparison of the PPAF-II and Api SII clip domain sequences showed only 9% identity (Fig. 1). The low identity and the additional 19-residue N-terminal extension in the template makes the structure of the PPAF-II clip domain not suitable for modeling the respective domain in Api SII.

```
Api SI      -MALVNNITTFR--DIVRKVTAALLIFYGSIMLSAKLIEG-QCNYNQNLIPGTTYYIYNP 56
Api SII     -------------------MLIVCLTLIGLLQPLIHVVYAQDQCTTPNQEEGVCINLRSC 41
PPAF-II     GHMAVVNIFGNASEYIPPGYEIVTKAPLGALTALPRCGTGADQGKKVCIVYHRCDGVTNT 60

Api SI      NYPYSYRGSESCVWTVSSDYRVNLTCTDFEIPWSYNCFQDSLTVQINSTTSHRYCGDGGF 116
Api SII     QFLITLLEKE----GLKVKNYLKQSLCRYENNDPFVCCPKNSGRESKIERENSYG----- 92
PPAF-II     VTPEEVINTT-----GEGIFDIRENANECESYLDVCCGLPEGGVLPTPSPTPPVVP---- 111

Api SI      NVVSSSNSMVVTLSSPIWSQGGRFLCEIRAVKRPQDSTNCQCGWNNP----SRIVGG-MD 171
Api SII     -----------PLLPP------------------------QCGFNNIS--HTRVVGG-IP 114
PPAF-II     -----------VLKPS------------------------FCGIRNERGLDFKITGQTNE 136

Api SI      TGVNEFPMMAGI------VDADERAVFCGSTIISVRYVLTAAHCM--TNRNYTRLGVLVG 223
Api SII     AKLGAWPWLTVLGFR-SSLNPSQPRWLCGGSLISARHVLTAAHCA--VRK--DLYVVRIG 169
PPAF-II     AEYGEFPWMVAVLKANVIPGSGEEQLVCGGSLIAPSVVLTGAHCVNSYQSNLDAIKIRAG 196

Api SI      DHDISSGTDTNATMLYRVKKVIVHPNYAHDN-FNDVALLKTRTKMEFGNEVGPACLPFQH 282
Api SII     DLDLSRDDDGAHPIQVEIEDKLIHPDYSTTTFVNDIAVLRLAQDVQFTEYVYPICLPVED 229
PPAF-II     EWDTLTEKERLPYQERKIRQVIIHSNFNPKTVVNDVALLLLDRPLVQADNIGTICLPQQS 256

Api SI      SP--DTFAGSFVQLLGWGTTSFGGPP--SDILQKVTVSVLTNLQCTKFYP--------DL 332
Api SII     NLRNNNFVRNYPFVAGWGSTETRGPA--SDILLEIQLPVINNEQCKQAYSKFKA---AEI 285
PPAF-II     QI----FDSTECFASGWGKKEFGSRHRYSNILKKIQLPTVDRDKCQADLRNTRLGLKFVL 312

Api SI      TPQQMCTYAK--DKDACQMDSGGPVLWQNPTTKR-FVLVGIISMGIGCGD--TAGVNTRV 385
Api SII     DNRVLCAAYRQGGKDACQGDSGGPLML--PQHWY-YYQIGVVSYGYKCAEPGFPGVYTRV 341
PPAF-II     DQTFVCAGGEQ-GKDTCTGDGGSPLFCPDPRNPSRYMQMGIVAWGIGCGDENVPGVYANV 371

Api SI      GAYIDWIVSETADSTYCIIE--- 405
Api SII     TAFLDFIISALK----------- 353
PPAF-II     AHFRNWIDQEMQAKGLSTTPYVE 394
```

Fig. 1. Amino acid sequences of the honeybee venom allergenic proteases Api SI (NCBI reference sequence XP_392669), Api SII (NCBI reference sequence XP_623069), and prophenoloxidase-activating factor (PPAF)-II (Piao *et al.*, 2005). The CUB (residues 44–131) and clip (residues 29–96) domains of Api SI and Api SII, as well as their serine protease-like (SPL) domains (residues 161–405 and residues 99–353, respectively) are underlined.

No suitable template for the Api SII clip domain was found. In the same time, a significant identity/homology between the SPL domains of the target and template proteins was observed (Fig. 1), as it was mentioned before. Also, the two domains share common active sites and identical segments of the polypeptide chains. Due to the absence of a suitable template for modelling the 3-D structure of the whole allergens and the clip domain of Api SII, the SPL domains of the two proteases and the CUB domain of Api SI were built separately using crystallographic data about the respective structural modules of related proteins.

Fig. 1 demonstrates that a more significant identity/homology between the target and template proteins is observed in the SPL domains including the serine protease active site. There is no identity/homology between the CUB/clip domains of the compared proteins. For this reason a CUB domain of spermadhesin PSP-I (Romero *et al.*, 1997; PDB code 1SPP) with a known 3-D structure was used for the further studies. CUB domains are variable in amino acid sequence but share a common fold (Blanc *et al.*, 2007). The SPL domains of the two proteases and the CUB domain of Api SI were modelled separately using

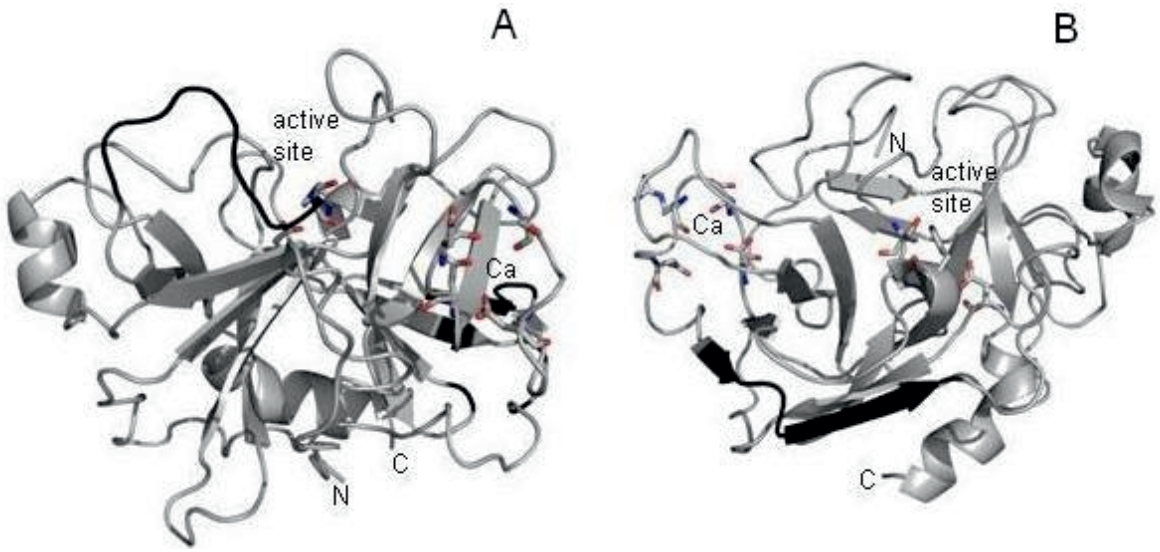


Fig. 2. 3-D models of the honeybee venom allergenic proteases Api SI (A) and Api SII (B) serine protease-like (SPL) modules based on the amino acid sequences of the enzymes and on the $C_\alpha$ coordinates of the prophenoloxidase-activating factor (PPAF)-II SPL domain (Piao *et al*., 2005). Predicted allergenic sites are labelled in black.

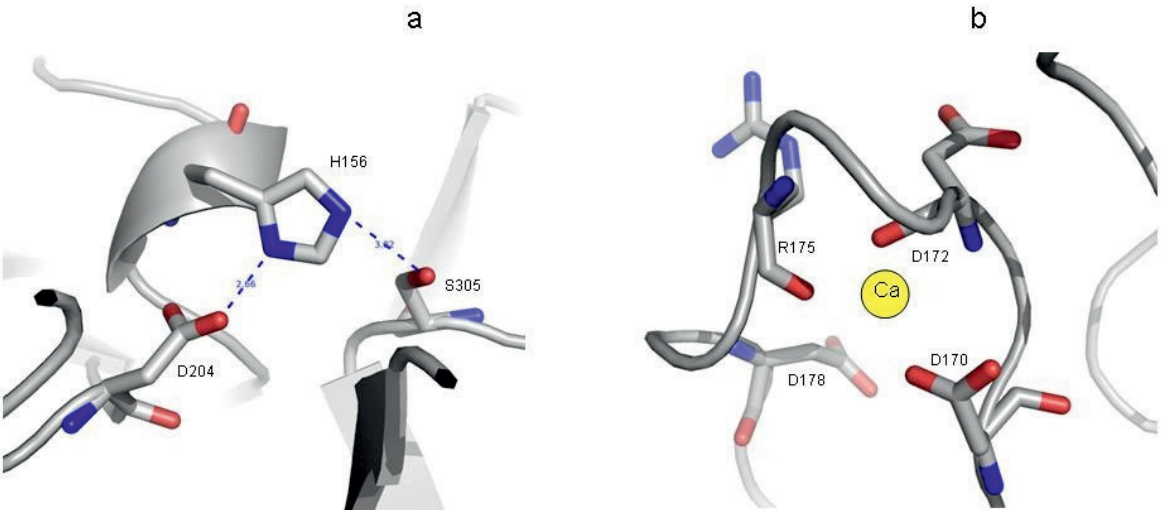


Fig. 3. 3-D models of the (a) catalytic and (b) calcium-binding sites of the honeybee venom protease Api SII based on the amino acid sequence of the enzyme serine protease-like (SPL)domains and the crystallographic coordinates of the prophenoloxidase-activating factor (PPAF)-II SPL domain (Piao *et al.*, 2005).

the coordinates of the PPAF-II SPL module (Piao *et al.*, 2005) and those of the PSP-I CUB domain (Romero *et al.*, 1997).

The 3-D models of the Api SI and Api SII SPL domains, generated from the amino acid sequences and the available structural information about the PPAF-II module, are shown in Fig. 2. The SPL models of the two allergens share common active sites consisting of the catalytic triad His208, Asp257, and Ser349 in Api SI, and His156, Asp204, and Ser305 in Api SII. The Api SII catalytic site is shown in Fig. 3a. There is a cleft in the region of the active sites which most probably serves for substrate binding. Asp343 was identified as the residue responsible for the trypsin-like activity of Api SI and Asp299 for the specificity

of Api SII. The 3-D models predict the presence of $Ca^{2+}$-binding sites in the two SPL domains in which the metal ion is coordinated by the carboxylic groups of Asp224 in Api SI (Asp170 in Api SII), Asp232 in Api SI (Asp178 in Api SII), and the carbonyl oxygen atoms of Asp226 and Ser229 of Api SI (Asp172 and Arg175 in Api SII). The modelled catalytic and calcium-binding sites of the Api SII SPL domain are shown in Figs. 3a and b, respectively. The respective sites of the other protease are similar.

As can be seen in Fig. 2, the secondary structure is dominated by irregular $\beta$-sheets and loops. Elements of the secondary structure are four short and two longer $\alpha$-helices and $\beta$-strands with variable length grouped in two irregular $\beta$-sheets.

CUB domains are specific structures involved in wide diversity of biological functions, found mainly in extracellular and membrane-associated proteins (Blanc *et al.*, 2007). Such structural motifs were observed also in proteases (Blanc *et al.*, 2007 and references therein). The modelled Api SI CUB domain is shown in Fig. 4 and includes residues 44–131. The structure is a typical $\beta$-barrel, has $\beta$-sandwich fold and consists of six $\beta$-strands and loops. The strands form two $\beta$-sheets which are parallel to each other. The first one is a four-stranded $\beta$-sheet composed of the strands $\beta$1, $\beta$6, $\beta$3, and $\beta$4 antiparallel to each other. The second structural element is a two-stranded $\beta$-sheet built by the antiparallel strands $\beta$2 and $\beta$5. The internal part of the CUB module is hydrophobic.

Allergenic determinants were predicted using the HBV protease 3-D models and the programs of Saha and Raghava (2006). The ARPs (allergen representative 24mers peptides) program uses a dataset of 2890 ARPs which have a high similarity in allergenic proteins but not in non-allergenic proteins. The protein is assigned allergen if it has similarity with any ARP (Saha and Raghava, 2006). After BLAST search of the Api SI modules the following hits were found: GEIYYIYNPRYPLPYSGSK-CTW (in the CUB domain) and FCGGTILDEYWILTAAHCVAGO (in the SPL domain). The program based on the amino acid composition characterized the SPL and CUB domains, as well as the N-terminal peptide including residues 1–43 of Api SI as allergens. The result about the linker peptide connecting the two modules was not so definite. The program based on the amino acid composition determined this peptide (residues 132–160) as allergen but the score

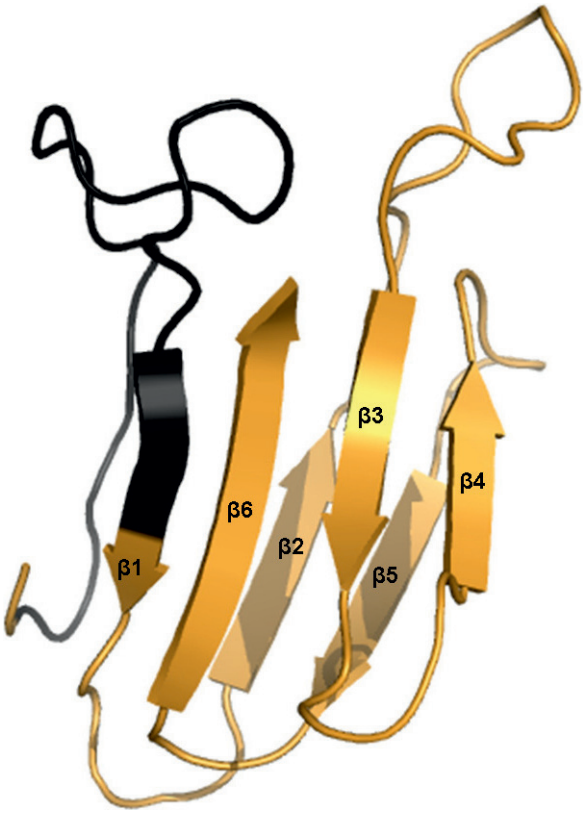


Fig. 4. 3-D model of the honeybee venom protease Api SI CUB domain based on the amino acid sequence and the crystallographic coordinates of the spermadhesin PSP-I CUB domain (Romero *et al.*, 1997). The predicted allergenic site is labelled in black.

of –0.3 was near to the threshold (–0.4). In other words, the presence of IgE epitope in this segment of Api SI is not very probable.

The application of the ARPs program to the SPL domain of Api SII resulted in the following hits: YGCARKGYPGVYTRVGNFVDWIES and GGKDSCQGDSGGPLV. The program determined the module as allergen. The program based on the amino acid composition characterized the SPL and clip domains, as well as the N-terminal peptide (residues 1–28) as allergens. In the last case the score (–0.37) was very near to the threshold of –0.4.

Charged and hydrophobic residues, located on the protein surface, usually participate in protein-protein interactions and often are present in IgE binding epitopes. In this connection lysil residues play an important role (Gehlhar *et*

*al.*, 2006; Rajashankar *et al.*, 2002). Inspection of the 3-D models is important for the prediction of IgE epitopes because these sites are located on the protein surface. The model of the Api SI SPL domain revealed a 3-D cluster of exposed lysines, arginine, and hydrophobic residues: R201, R240, V241, K242, K243, V244, K262, R264, K266. This cluster is a suitable participant in protein-protein interactions. There is a surface loop (residues 340–348) which can also participate in such interactions. The 3-D model of the Api SII SPL domain showed that the surface $\beta$-strand formed by residues 184–192 and containing three acidic residues, a lysine, and four hydrophobic residues can serve as a protein-binding site.

The predicted allergenic sites are labelled in black in the 3-D models of the Api SI and Api SII SPL domains (Fig. 2) and in the modelled structure of the Api SI CUB domain (Fig. 4).

Linear antigenic sites on the structure of the three HBV proteases were predicted using the method of Kolaskar and Tongaonkar (1990) based on experimental antigenic determinants. The 3-D models were examined to assess the accessibility of the predicted structural motifs to the solvent and eventual antigens. The selected antigenic sites responded to the rules of MIF (Molecular Immunology Foundation) Bioinformatics (http://immunax.dfci.harvard.edu/Tools/antigenic.html), namely: the antigenic peptides to contain both hydrophobic and hydrophilic residues, to be accessible to the solvent, and preferably to be in loops, avoiding helical regions. There are three "exposed" segments with a high antigenic propensity in the Api SI CUB domain: residues 64–84, residues 89–102, and residues 115–122. These sequences include $\beta$-strands and a loop. In the SPL module of this allergen only the surface-exposed sequence 237–249, with a very high antigenic propensity, is a suitable candidate for an antigenic epitope, according to the 3-D model. In the Api SII SPL domain three exposed on the protein globule regions were selected as possible antigenic sites: residues 180–195 (a long $\beta$-strand), 211–228, and 311–318 (a loop and part of $\beta$-strand). All regions have a high antigenic propensity.

## Discussion

The information obtained by the 3-D modelling of the two honeybee venom allergenic serine proteases, Api Si and Api SII, is important from several points of view. First of all, knowledge of the HBV components and their properties is necessary due to the strong immunogenic effects of the venom. A lot of people are allergic to the honeybee venom, and the stings can create a dangerous, sometimes life-threatening, anaphylactic reaction. Prediction of allergenic and antigenic sites is useful for the experimental workers to identify epitopes in proteins. This is important for diagnostic purposes and for the development of vaccines. A high IgE antibody reactivity was found by immunoblot studies for components of the honeybee venom with molecular masses between 30 and 39 kDa (Jeep *et al.*, 1996; Winningham *et al.*, 2004 and references therein). Api SI and Api SII belong to this group. Here, we demonstrated that both proteases contain allergenic determinants which were found only in allergenic proteins but not in non-allergenic proteins and should be classified as allergens. However, little is known about the structure-function relationships of the honeybee proteases. The natural substrates and functions of the two *Apis mellifera* allergens are not known. They have modular structures which suggest a complex mechanism of biological activity. Both proteins contain serine protease-like domains with similar three-dimensional structure which suggest the involvement of hydrolysis in their function. Api SI and Api SII should participate in processes requiring proteolysis, such as digestion, development or defence responses. The last function is very probable because the bee venom is used mainly for such purpose. The presence of modules, such as the CUB and clip domains, suggests participation in protein-protein interactions important for the fulfilment of the physiological function. CUB proteins are involved in important physiological reactions, as activation of the complement system (Blanc *et al.*, 2007; Gaboriaud *et al.*, 2004 and references therein), in development processes, and in other activities summarized in Romero *et al.* (1997). CUB serine proteases bind to biological membranes and exert hydrolytic activity in the extracellular matrix and regulate cell development (Misra *et al.*, 1998; Johns *et al.*, 2004; Bork and Beckmann, 1994). The function of venom CUB proteases should be different to that of their counterparts from animal organs. The clip domain also participates in protein-protein interactions as was demonstrated for PPAF-II (Piao *et al.*, 2005). Clip domain serine proteases are com-

ponents of the extracellular signalic cascades and participate in the innate immune responses (Piao *et al.*, 2005). In the bee venom the function of these enzymes should be different in accordance with the purpose of the venomous system. Probably, the two allergenic proteases are components of the honeybee defence system. Their CUB/clip domains are responsible for the association with the target molecules, and in the complex the SPL domain realizes its catalytic function. Allergenicity, which does not depend on the catalytic activity, strengthens the toxic effect of the proteases.

## Conclusions

The three-dimensional models described here provide information about the structural organization and possible function as components of the honeybee defence system of two *Apis mellifera* venom protease allergens. Their modular structure suggests a complex mechanism of biological action in which the N-terminal CUB or clip domain mediates interactions with the target protein, and the second module exerts trypsin-like catalytic activity. The toxic effect results from combination of proteolysis and allergenic reaction. The data can be used for immunological and structural investigations on the HBV allergenic proteases.

*Acknowledgements*

This work was supported by the Deutsche Forschungsgemeinschaft via the Project BE 1443-18-1. Ch. B. acknowledges the RIS Project for providing financial support. R. K. A. acknowledges the support of FAPESP/CNPq/CAPES.

Altschul S. F., Gish W., Miller W., Myers E. W., and Lipman D. J. (1990), Basic alignment search tool. J. Mol. Biol. **215**, 403–410.

Blanc G., Font B., Eichenberger D., Moreau C., Ricard-Blum S., Hulmes D. J. S., and Moali C. (2007), Insights into how CUB domains can exert specific functions while sharing a common fold. J. Biol. Chem. **282**, 16924–16933.

Bork P. and Beckmann G. (1994), The CUB domain. A wide-spread module in developmentally regulated proteins. J. Mol. Biol. **231**, 539–545.

Collaborative Computational Project Number 4 (1994), The CCP4 Suite: Program for protein crystallography. Acta Crystallogr. Sect. D: Biol. Crystallogr. **50**, 760–763.

Dudler T., Chen W. Q., Wang S., Schneider T., Annand R. R., Dempcy R. O., Crameri R., Gmachl M., Suter M., and Gelb M. H. (1992), High-level expression in *Escherichia coli* and rapid purification of enzymatically active honey bee venom phospholipase $A_2$. Biochim. Biophys. Acta **1165**, 201–210.

Gaboriaud C., Thielens N. M., Gregory L. A., Rossi V., Fontecilla-Camps J. C., and Arlaud G. J. (2004), Structure and activation of the C1 complex of complement: unravelling the puzzle. Trends Immunol. **25**, 368–373.

Gehlhar K., Rajashankar K. R., Hofmann E., Betzel Ch., Weber W., Werner S., and Bufe A. (2006), Lysine as a critical amino acid for IgE binding in Phl p 5b C terminus. Int. Arch. Allergy Immunol. **140**, 285–294.

Grunwald T., Bockisch B., Spillner E., Ring J., Bredehorst R., and Ollert M. W. (2006), Molecular cloning and expression in insect cells of honeybee venom allergen acid phosphatise (Api m 3). J. Allergy Clin. Immunol. **117**, 848–854.

Hoffman D. R. (2006), Hymenoptera venom allergens. Clin. Rev. Allergy Immunol. **30**, 109–128.

Jeep S., Paul M., Müller U., and Kunkel G. (1996), Honeybee venom allergy: immunoblot studies in allergic patients after immunotherapy and before sting challenge. Allergy **51**, 540–546.

Johns M. E., Tai P. C., and Derby C. D. (2004), Serine proteases in the spiny lobster olfactory organ: Their functional expression along a developmental axis, and the contribution of a CUB-serine protease. Wiley InterScience DOI 10.1002/neu.20056, 377–391.

Kolaskar A. S. and Tongaonkar P. C. (1990), A semi-empirical method for prediction of antigenic determinants on protein antigens. FEBS Lett. **276**, 172–174.

Laskowski R. A., MacArthur M. W., Moss D. S., and Thornton J. M. (1993), PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Crystallogr. **26**, 283–291.

Lee B. and Richards F. M. (1971), The interpretation of protein structures: estimation of static accessibility. J. Mol. Biol. **55**, 379–400.

Marti-Renom M. A., Stuart A. C., Fiser A., Sanchez R., Melo F., and Sali A. (2000), Comparative protein structure modelling of genes and genomes. Annu. Rev. Biophys. Biomol. Struct. **29**, 291–325.

Misra S., Hecht P., Maeda R., and Anderson K. V. (1998), Positive and negative regulation of Easter, a member of the serine protease family that controls dorsal-ventralpatterning in the *Drosophila* embryo. Development **125**, 1261–1267.

Müller U. R. (2002), Recombinant Hymenoptera venom allergens. Allergy **57**, 570–576.

Piao S., Song Y.-L., Kim J. H., Park S. Y., Lee B. L., Oh B.-H., and Ha N.-C. (2005), Crystal structure of a clip-domain serine protease and functional roles of the clip domains. EMBO J. **24**, 4404–4414.

Rajashankar K., Bufe A., Weber W., Eschenburg S., Lindner B., and Betzel C. (2002), Structure of the

functional domain of the major grass-pollen allergen Phlp 5b. Acta Crystallogr. Sect. D: Biol. Crystallogr. **58**, 1175–1181.

Romero A., Romao M. J., Varela P. F., Kölln I., Dias J. M., Carvalho A. L., Sanz L., Töpfer-Petersen E., and Calvete J. J. (1997), The crystal structures of two spermadhesins reveal the CUB domain fold. Nature Struct. Biol. **4**, 783–788.

Saha S. and Raghava G. P. S. (2006), AlgPred: prediction of allergenic proteins and mapping of IgE epitopes. Nucleic Acids Res. **34**, W202–W209.

Schwede T., Kopp J., Guex N., and Peitsch M. C. (2003), SWISS-MODEL: An automated protein homology-modeling server. Nucleic Acids Res. **31**, 3381–3385.

Soldatova L. N., Crameri R., Gmachl M., Kemeny D. M., Schmidt M., Veber M., and Mueller U. R. (1998), Superior biologic activity of the recombinant bee venom allergen hyaluronidase expressed in baculovirus-infected insect cells as compared with *Escherichia coli*. J. Allergy Clin. Immunol. **101**, 691–698.

Soldatova L. N., Bakst J. B., Hofman D. R., and Slater J. E. (2000), Molecular cloning of a new honey bee venom allergen, acid phosphatise. J. Allergy Clin. Immunol. **105**, S378.

Winningham K. M., Fitch C. D., Schmidt M., and Hoffman D. R. (2004), Hymenoptera venom protease allergens. J. Allergy Clin. Immunol. **114**, 928–933.

Zou Z., Lopez D. L., Kanost M. R., Evans J. D., and Jiang H. (2006), Comparative analysis of serine protease-related genes in the honey bee genome: possible involvement in embryonic development and innate immunity. Insect Mol. Biol. **15**, 603–614.