

Three-dimensional model of the honeybee venom allergen Api m 7: structural and functional insights

Dessislava Georgieva,^a Kerstin Greunke^b and Christian Betzel^{*a}

Received 4th November 2009, Accepted 8th February 2010

First published as an Advance Article on the web 17th March 2010

DOI: 10.1039/b923127g

Api m 7 is one of the major protease allergens of the honeybee venom. It consists of a serine protease-like (SPL) and a CUB domain. The knowledge about the structure and function of Api m 7 is limited mainly to its amino acid sequence. Three-dimensional models of the two structural domains were constructed using their amino acid sequences and the crystallographic coordinates of prophenoloxidase-activating factor (PPAF-II) as a template for the SPL domain and the coordinates of porcine spermadhesin PSP-II for the CUB domain. The structural organization of Api m 7 suggests that the CUB domain is involved in interactions with natural substrates while the SPL domain probably activates zymogens. IgE epitopes and antigenic sites were predicted. Api m 7 shows structural and functional similarity to the members of the PPAF-II family. Possible substrates, function and evolution of the enzyme are discussed in the paper.

Introduction

Honeybee (*Apis mellifera*) venom (HBV) is a rich source of pharmacologically active compounds. It has been applied for years as an anti-inflammatory tool for the treatment of arthritis, rheumatism and tendonitis, for diagnostic tests and immunotherapy of allergic reactions.¹ However, some venom components can create life-threatening allergic effects and even anaphylactic shock in patients sensitive to the bee sting (ref. 2 and references therein). The list of HBV allergens, published by the Allergen Nomenclature Sub-Committee of the International Union of Immunological Societies, updated on 28-09-2009 (<http://www.allergen.org/Allergen.aspx>), includes nine venom proteins: Api m 1 (phospholipase A₂), Api m 2 (hyaluronidase), Api m 3 (acid phosphatase), Api m 4 (melittin), Api m 5 (dipeptidylpeptidase IV), Api m 6, Api m 7 (CUB serine protease), Api m 8 (carboxylesterase) and Api m 9 (serine carboxypeptidase). Several of these allergens were obtained in recombinant forms: Api m 1,³ Api m 2⁴ and Api m 3.^{5,6} Recombinant venom allergens have a potential to improve the treatment of Hymenoptera venom allergy.¹

The honeybee venom contains allergens of 30–39 kDa⁷ and IgE antibody reactivity of proteins with a molecular mass in this region was demonstrated.⁸ The complete amino acid sequence of a protease called Api m 7, containing a CUB domain, was determined.⁹ Studies on the venom proteome by 2-D gel electrophoresis revealed the presence of 49 proteins.² Mass spectrometry analysis resulted in the identification of 6 known proteins and 3 novel compounds.² Serine protease-related genes in the honeybee genome have been analysed for possible involvement in embryo development of *A. mellifera*.¹⁰

Here we describe the 3-D modeling and the structure-function relationships of the HBV allergenic serine proteinase Api m 7. The allergen is a 39 kDa protein and has a CUB- and a serine proteinase-like (SPL) domain connected with a “linker” peptide. Due to the absence of a suitable template for modeling the 3-D structure of the whole allergen, the SPL- and the CUB-domains of Api m 7 were built separately using crystallographic data about the respective structural modules of related proteins.

Experimental

Construction of homology model of Api m 7

The sequence alignment technique BLAST¹¹ was applied to search for primary structure similarities between Api m 7 and other proteins. Homology model, based on the sequence of the honeybee venom enzyme SPL-domain, was constructed using the crystallographic C_α coordinates of the prophenoloxidase-activating factor (PPAF)-II¹² as a template and applying the SWISS-MODEL server.¹³ The SPL module of Api m 7 showed 40% homology to the respective domain of PPAF-II, including the conservative substitutions. For the building of the Api m 7 CUB-domain model the crystallographic C_α coordinates of the porcine spermadhesin PSP-II¹⁴ were used. The stereochemistry of the models was checked by the program PROCHECK.¹⁵ The surface accessibility was calculated by the program AREAIMOL.^{16,17}

The root mean square (r.m.s.) distances between the C_α positions of the modelled protein and the template were calculated by the program LSQKAB.¹⁸

Prediction of allergenicity, IgE binding epitopes and antigenic determinants

Prediction of allergenic proteins and mapping of IgE epitopes were performed by the method of Saha and Raghava¹⁹ and

^a Institute of Biochemistry and Molecular Biology, University of Hamburg, Laboratory of Structural Biology of Infection and Inflammation, c/o DESY, Notkestrasse 85, Building 22a, 22603 Hamburg, Germany. E-mail: Christian.Betzel@uni-hamburg.de; Fax: +49 40 8998 4747; Tel: +49 40 8998 4744

^b Protein Ligand Structural (PLS)-Design GmbH, Eichenstrasse 42, Hamburg, Germany

using the respective programs. The data-set applied for testing consists of 578 allergens and 700 non-allergens. The authors used protein amino acid and dipeptide 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 m 7 were predicted by the method of Kolaskar and Tongaonkar.²⁰

Results

Comparison of the amino acid sequences and conserved sequence motifs among the honeybee serine proteinase and the template protein

The amino acid sequence of Api m 7 from the *Apis mellifera* venom was used in a BLAST search.¹¹ Prophenoloxidase-activating factor¹² was the protein with a known 3-D structure showing the highest sequence similarity towards the target protein. As can be seen in Fig. 1, the homology is restricted to the SPL domain of the HBV allergen. Api m 7 consists of two structural modules, a CUB domain located in the N-terminal part of the polypeptide chain and a serine proteinase-like domain. The two domains are connected with a peptide linker. The N-terminal segment is linked to the CUB domain. There was no identity/homology between the CUB domains of the compared proteins (Fig. 1). For this reason a CUB domain of spermadhesin PSP-II¹⁴ 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.²¹

Building of 3-D homology model of the Api m 7 serine proteinase-like domain based on the amino acid sequence and using the C α coordinates of the PPAF-II SPL domain as a template. Analysis of the model

The SPL and CUB domains of the HBV proteinase are structurally similar to the respective modules of different proteins. The attempts to construct a 3-D model of the whole HBV proteinase using the C α coordinates of one protein as a template were not successful due to variations in the length of loops. For this reason the SPL and CUB domains were modeled separately using the coordinates of related protein modules showing the highest structural similarity. The 3-D model of the Api m 7 SPL-domain (residues 161–405), generated from the amino acid sequence and using the C α coordinates of the respective PPAF-II module, is shown in Fig. 2. The SPL domain of the investigated honeybee venom protease has a catalytic site consisting of the triad His203, Asp257 and Ser349 (Fig. 3). Asp343 determines the trypsin-like specificity of the SPL-domain. There is a cleft in the region of the active site which probably serves as a substrate-binding site. The secondary structure is dominated by irregular β -sheets and loops. Characteristic elements of the structure are four short and two longer α -helices, and β -strands with variable length grouped in two irregular β -sheets. Superposition of the modeled Api m 7 SPL-domain structure on the crystallographic structure of the SPL-domain of PPAF-II gave an average distance between the C α positions of 0.6 Å.

The 3-D model predicts the presence of a Ca²⁺-binding site in the SPL-domain in which the metal ion is coordinated by the carboxylic groups of Glu219 and Glu227, and the carbonyl oxygens of Asp221 and Ser224. In this respect the HBV

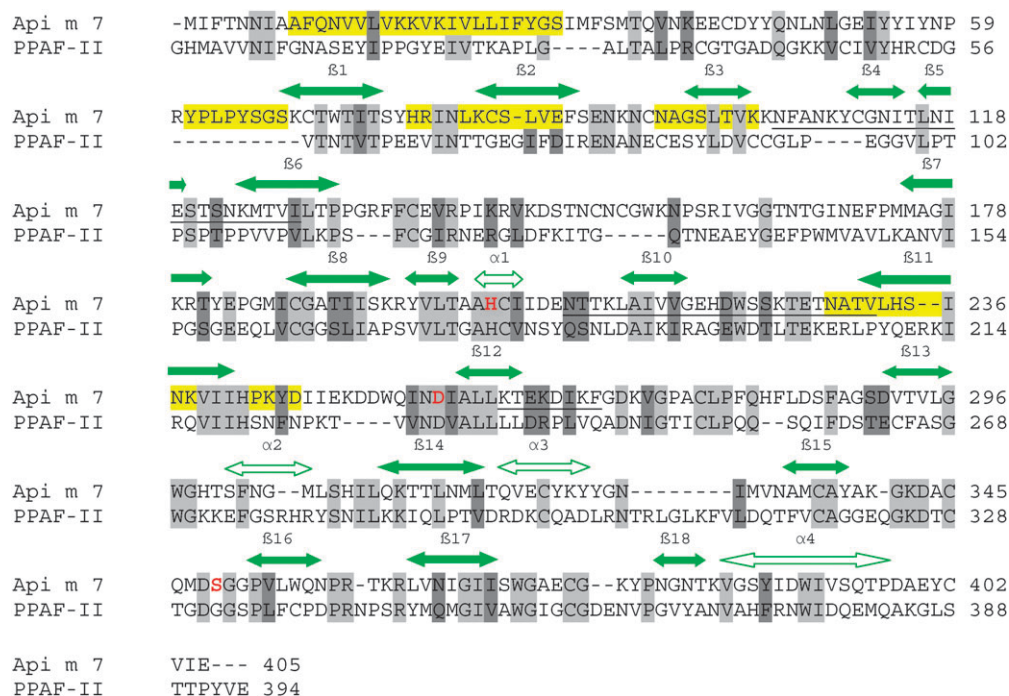


Fig. 1 Amino acid sequences of the honeybee venom allergenic protease Api m 7 and the prophenoloxidase-activating factor (PPAF)-II.¹² Identical residues are highlighted in gray and the conservative substitutions in dark gray. The predicted allergenic determinants are underlined, the antigenic sites are highlighted in yellow and the active-site residues are labelled in red.

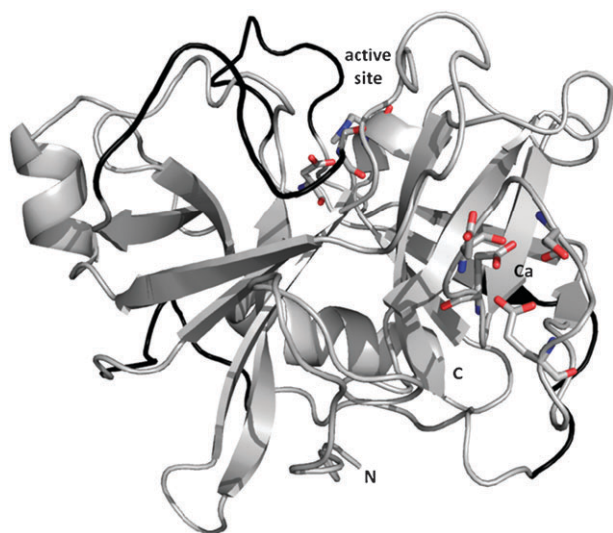


Fig. 2 3-D model of the honeybee venom allergenic protease Api m 7 serine proteinase-like domain based on the amino acid sequence of the enzyme and on the C α coordinates of the prophenoloxidase-activating factor (PPAF)-II serine protease like-domain.¹² Predicted allergenic sites are labelled in black.

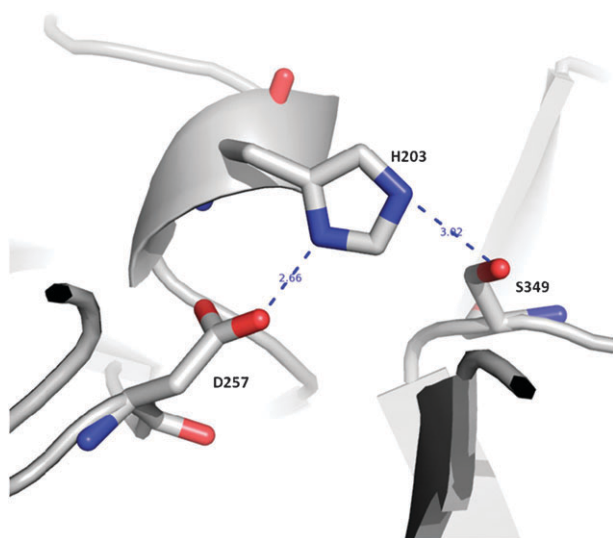


Fig. 3 Catalytic site of the allergenic serine protease Api m 7. The catalytic triad is formed by residues His203, Asp257 and Ser349.

protease SPL-domain shows structural and functional similarity to the members of the PPAF family.¹²

Building of 3-D homology model of the Api m 7 CUB domain based on the amino acid sequence and using the C α coordinates of the porcine spermadhesin (PSP) chain B as a template. Analysis of the model

CUB domains were found in functionally diverse proteins like the components of the complement system, bone morphogenetic proteins, neuronal recognition molecule A5, in extracellular and plasma-membrane-associated proteins including proteinases (ref. 14, 22 and references therein). The modeled Api m 7 CUB domain includes residues Asn48–Thr131.

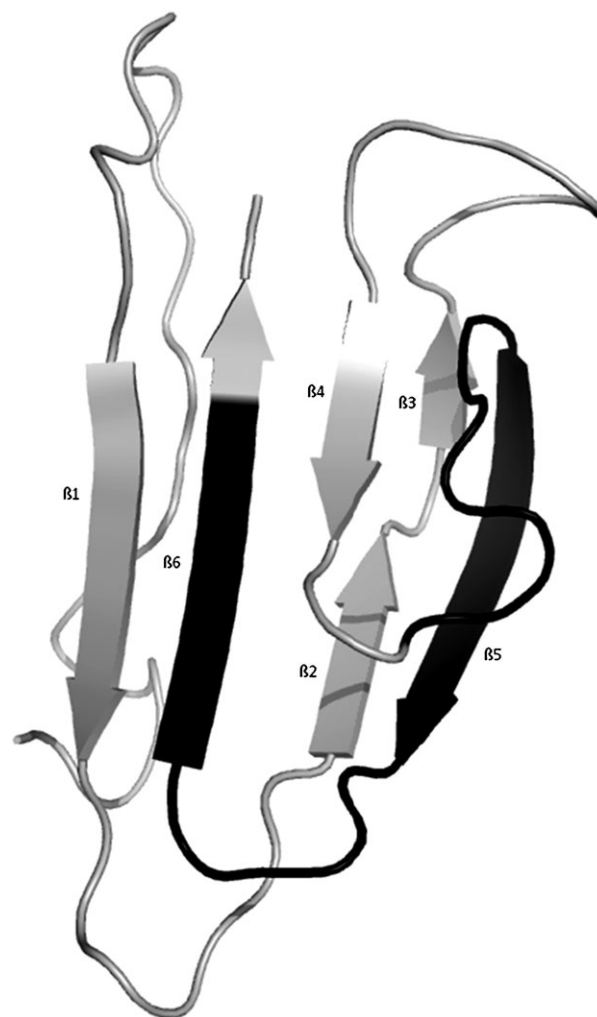


Fig. 4 3-D model of the honeybee venom protease Api m 7 CUB-domain based on the amino acid sequence of the enzyme CUB-domain and the crystallographic coordinates of the spermadhesin PSP-II CUB-domain.¹⁴ Predicted allergenic sites are labelled in black.

It is shown in Fig. 4. The domain lacks α -helices and its secondary structure consists of β -strands and loops; *i.e.* typical β -barrel structure. The Api m 7 module is built of two three-stranded β -sheets which are parallel to each other and are arranged as a sandwich (Fig. 4). Superposition of the modelled Api m 7 CUB-domain on the crystallographic structure of PSP-II showed an r.m.s. of 0.6 Å.

Prediction of allergenic determinants in the molecule of Api m 7

Allergenic determinants were predicted using the HBV protease 3-D models and the programs of Saha and Raghava.¹⁹ The sequence does not contain experimentally proven IgE epitopes. The programs based on the amino acid composition and dipeptide analysis characterized the SPL and CUB domains, as well as the N-terminal peptide including the first 46 amino acid residues of Api m 7 as allergens. The linker peptide (residues 131–160) was also classified as allergen. The program of allergen representative 24mers peptides (ARPs)¹⁹ classified Api m 7 and its domains as allergens and identified the region Asn105–Ile128 from the CUB domain as well as the sequence

Asn209—Val232 from the SPL module as allergenic determinants. The program uses a database of 2890 ARPs which have high similarity in allergenic proteins but not in non-allergenic proteins.¹⁹ All peptide segments mentioned above are exposed on the protein surface and can serve as IgE binding sites.

Linear epitopes of IgE often contain charged or hydrophobic residues (or both types of functional groups which are suitable participants in protein-protein interactions) and many of them at least one lysine (ref. 23 and references therein). The essential role of lysines for the binding of IgE was demonstrated in the case of the major grass pollen allergen PhI p 5b C terminus.²³ Inspection of the 3-D models showed also that the residues Lys262—Phe269 of Api m 7 form an exposed cluster which contains three lysines, two acidic and two hydrophobic residues, and can serve as an allergenic determinant.

Prediction of antigenic determinants on the honeybee venom protease structure

Linear antigenic sites on the polypeptide chain of the HBV protease were predicted using the method of Kolaskar and Tongaonkar²⁰ based on the structures of experimentally determined epitopes. The 3-D models were examined to assess the accessibility of the predicted structural motifs to the solvent and eventual antigens. Also, the rules of MIF Bioinformatics, Molecular Immunology Foundation (<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, were followed. The selected antigenic sites responded to the rules mentioned above. In the Api m 7 N-terminal region and in the CUB-domain the predicted sites are: the residues 9–29 with the highest antigenic propensity; the sequence YLPYSGS (residues 61–68) which is an exposed surface loop; the β 2 strand and several neighbouring residues (His78—Glu88), and the β 3 strand with neighbouring residues (Asn96—Lys103). The best candidate for an antigen-binding site in the Api m 7 SPL-module is the segment Asn229—Asp246 consisting of the β 11 strand and a loop. All regions have a high antigenic propensity.

Discussion

Api m 7 is a protein with a modular structure contributing to the total allergenic effect of the bee sting. The protease represents a combination of structurally and functionally different modules: a CUB domain is connected through a linker peptide with a SPL domain. The enzymatic properties of the allergen are mediated by the SPL domain which is a calcium-dependent trypsin-like protease. Api m 7 has also an N-terminal peptide which can serve as an allergenic determinant. The natural substrates and functions of the *A. mellifera* protease, except allergenicity, are not known. However, the molecular architecture suggests possible biological activities. The modular structure is characteristic for proteins involved in important biological processes. The CUB domain is a structural motif found in the components of the complement system¹⁴

which is the major element of the innate immune system. This module is present in extracellular and plasma-membrane-associated proteins including proteases.¹⁴ CUB-proteins are involved in a number of biological reactions: activation of the complement system (ref. 21, 24 and references therein), in developmental processes (ref. 25 and references therein), tissue repair,²⁶ haemostasis,²⁷ inflammation²⁸ and in a number of other processes (ref. 14 and references therein). However, the role of the CUB domains in proteinases is poorly understood. Some of them are involved in the oligomerization and/or the recognition of substrates and binding partners.¹⁴ It was shown that the N-terminal module is responsible for the dimerization of CUB-proteins.²⁹ It can be supposed that the CUB domain is a key element of Api m 7 mediating its association and interaction with natural substrates. The protease domain probably activates zymogens through cleavage of peptide bonds.

The HBV protease shows the highest sequence similarity with the prophenoloxidase-activating factor and it is reasonable to expect also a functional similarity between the two proteins. PPAF-II activates the prophenoloxidase (ProPO) in the blood plasma to phenoloxidase (PO), which is one of the responses of the immune system in arthropods. PO catalyzes the oxygenation of monophenols to *o*-diphenols and further to *o*-quinones.³⁰ Quinones are involved in the formation of melanin, cytotoxic superoxides and hydroxyl radicals.³⁰ The structural similarity and domain organization suggest similar functional properties of PPAF-II and the HBV protease which makes ProPO a very likely candidate for a substrate of Api m 7. However, excessive melanisation can be fatal for the host.³¹ The combination of allergenicity with possible ProPO activity in the HBV protease can be explained as a strengthening of the effect of bee stings. The predicted antigenic sites can be used to design immunological experiments.

Conclusions

The HBV protease is an important component of the bee defence system but the allergenicity does not depend on its catalytic properties or on the functional potency of the CUB domain. In general, venom toxins most likely evolved from proteins with normal physiological functions and were recruited into the venom proteome early in the evolution.^{32,33} The *Apis mellifera* genome includes 44 serine proteinase, 13 serine proteinase homologs and three prophenoloxidase genes.¹⁰ It can be supposed that Api m 7 evolved from a proteinase with a modular structure (CUB/SPL domains) and during the evolution was recruited from the respective biological system or organ to the venom as a new defence weapon. The 3-D model of Api m 7 provides a basis for investigating structure-function relationships in the HBV allergenic proteases including site-directed mutagenesis.

Abbreviations

Api m 7	Honeybee venom allergen;
SPL	Serine protease-like;
PPAF	Prophenoloxidase-activating factor;
PSP	Porcine spermadhesin;

CUB domain	Structural motif in the components of the complement system;
HBV	Honeybee venom;
ARP	Allergen representative 24mers peptide;
ProPO	Prophenoloxidase;
PO	Phenoloxidase.

Acknowledgements

D. G. thanks the PLS-Design GmbH, Hamburg, Germany, for the financial support. This work was also supported by the Deutsche Forschungsgemeinschaft *via* the Project BE 1443/18-1. Ch. B. acknowledges the RIS Project for providing financial support.

References

- U. R. Müller, *Allergy*, 2002, **57**, 570–576.
- N. Peiren, F. Vanrobaeys, D. C. de Graaf, B. Devreese, J. van Beeumen and F. J. Jacobs, *Biochim. Biophys. Acta, Proteins Proteomics*, 2005, **1752**, 1–5.
- T. Dudler, W. Q. Chen, S. Wang, T. Schneider, R. R. Annand, R. O. Dempcy, R. Cramer, M. Gmachl, M. Suter and M. H. Gelb, *Biochim. Biophys. Acta, Lipids Lipid Metab.*, 1992, **1165**, 201–210.
- L. N. Soldatova, R. Cramer, M. Gmachl, D. M. Kemeny, M. Schmidt, M. Veber and U. R. Mueller, *J. Allergy Clin. Immunol.*, 1998, **101**, 691–698.
- L. N. Soldatova, J. B. Bakst, D. R. Hofman and J. E. Slater, *J. Allergy Clin. Immunol.*, 2000, **105**, S378.
- T. Grunwald, B. Bockisch, E. Spillner, J. Ring, R. Bredehorst and M. W. Ollert, *J. Allergy Clin. Immunol.*, 2006, **117**, 848–854.
- D. R. Hoffman, *Clin. Rev. Allergy Immunol.*, 2006, **30**, 109–128.
- S. Jeep, M. Paul, U. Müller and G. Kunkel, *Allergy*, 1996, **51**, 540–546.
- K. M. Winningham, C. D. Fitch, M. Schmidt and D. R. Hoffman, *J. Allergy Clin. Immunol.*, 2004, **114**, 928–933.
- Z. Zou, D. L. Lopez, M. R. Kanost, J. D. Evans and H. Jiang, *Insect Mol. Biol.*, 2006, **15**, 603–614.
- S. F. Altschul, W. Gish, W. Miller, E. W. Myers and D. J. Lipman, *J. Mol. Biol.*, 1990, **215**, 403–410.
- S. Piao, Y.-L. Song, J. H. Kim, S. Y. Park, B. L. Lee, B.-H. Oh and N.-C. Ha, *EMBO J.*, 2005, **24**, 4404–4414.
- M. C. Petisch, *Biochem. Soc. Trans.*, 1996, **24**, 274–279.
- A. Romero, M. J. Romao, P. F. Varela, I. Kölln, J. M. Dias, A. L. Carvalho, L. Sanz, E. Töpfer-Petersen and J. J. Calvete, *Nat. Struct. Biol.*, 1997, **4**, 783–788.
- R. A. Laskowski, M. W. MacArthur, D. S. Moss and J. M. Thornton, *J. Appl. Crystallogr.*, 1993, **26**, 283–291.
- B. Lee and F. M. Richards, *J. Mol. Biol.*, 1971, **55**, 379–400.
- Collaborative Computational Project, Number 4. The CCP4 Suite: Program for Protein Crystallography. *Acta Crystall* 1994, **D50**, 760–763.
- W. Kabsch, *Acta Crystallogr., Sect. A: Cryst. Phys., Diffr., Theor. Gen. Crystallogr.*, 1976, **32**, 922–923.
- S. Saha and G. P. S. Raghava, *Nucleic Acids Res.*, 2006, **34**, W202–W209.
- A. S. Kolaskar and P. C. Tongaonkar, *FEBS Lett.*, 1990, **276**, 172–174.
- G. Blanc, B. Font, D. Eichenberger, C. Moreau, S. Ricard-Blum, D. J. S. Hulmes and C. Moali, *J. Biol. Chem.*, 2007, **282**, 16924–16933.
- P. Bork and G. Beckmann, *J. Mol. Biol.*, 1993, **231**, 539–545.
- K. Gehlhar, K. R. Rajashankar, E. Hofmann, Ch. Betzel, W. Weber, S. Werner and A. Bufe, *Int. Arch. Allergy Immunol.*, 2006, **140**, 285–294.
- C. Gaboriaud, N. M. Thielens, L. A. Gregory, V. Rossi, J. C. Fontecilla-Camps and G. J. Arlaud, *Trends Immunol.*, 2004, **25**, 368–373.
- H. X. Lee, A. L. Ambrosio, B. Reversade and E. M. De Robertis, *Cell*, 2006, **124**, 147–159.
- D. S. Greenspan, *Top. Curr. Chem.*, 2005, **247**, 149–183.
- Z. Tao, Y. Peng, L. Nolasco, S. Cal, C. Lopez-Otin, R. Li, J. L. Moake, J. A. Lopez and J. F. Dong, *Blood*, 2005, **106**, 4139–4145.
- C. M. Milner and A. J. Day, *J. Cell Sci.*, 2003, **116**, 1863–1873.
- L. A. Gregory, N. M. Thielens, G. J. Arlaud, J. C. Fontecilla-Camps and C. Gaboriaud, *J. Biol. Chem.*, 2003, **278**, 32157–32164.
- J. P. Gillespie, M. R. Kanost and T. Trenczek, *Annu. Rev. Entomol.*, 1997, **42**, 123–138.
- M. Zhao, I. Soderhall, J. W. Park, T. Osaki, N. C. Ha, C. F. Wu, K. Soderhall and B. L. Lee, *J. Biol. Chem.*, 2005, **280**, 24744–24751.
- N. Vidal, *J. Toxicol. Toxin Rev.*, 2002, **21**, 21–41.
- B. G. Fry and W. Wüster, *Mol. Biol. Evol.*, 2004, **21**, 870–883.