

Molecular cloning and expression of icarapin, a novel IgE-binding bee venom protein

Nico Peiren^a, Dirk C. de Graaf^{a,*}, Marleen Brunain^a, Chris H. Bridts^b, Didier G. Ebo^b, Wim J. Stevens^b, Frans J. Jacobs^a

^a Laboratory of Zoophysiology, Ghent University, K.L. Ledeganckstraat 35, B-9000 Ghent, Belgium

^b Immunology, Allergology and Rheumatology, University Antwerp, Antwerp, Belgium

Received 21 June 2006; revised 31 July 2006; accepted 3 August 2006

Available online 10 August 2006

Edited by Jesus Avila

Abstract The 1045 bp full-length cDNA sequence of a new bee venom component was obtained by rapid amplification of cDNA ends. The 672 bp coding sequence corresponds to a protein with a signal peptide and multiple carbohydrate binding sites, and it was named icarapin. It has the new consensus sequence N-[TS]-T-S-[TV]-x-K-[VI](2)-[DN]-G-H-x-V-x-I-N-[ED]-T-x-Y-x-[DHK]-x(2,6)-[STA]-[VLFI]-x-[KR]-V-R-[VLI]-[IV]-[DN]-V-x-P. At least two transcript variants were found. Recombinant icarapin was tested for recognition by IgE antibodies and gave a positive dot blot with sera from 4 out of 5 bee venom allergic patients, all beekeepers. Indirect immunofluorescent staining localized the protein in the cuticular lining of the venom duct.

© 2006 Federation of European Biochemical Societies. Published by Elsevier B.V. All rights reserved.

Keywords: Allergen; Honeybee; IgE; Hymenoptera; Icarapin; Prokaryotic expression; Venom

1. Introduction

When honeybees sting they deliver the venom sac contents into their victims. Most bee stings cause transient localized non IgE-mediated reactions in man requiring only local symptomatic treatment. In contrast, large local reactions, that affect approximately 15–20% of general population, expand from the site of the sting, consist of extensive erythematous swelling surrounding the sting site, and represent late-phase IgE-mediated hypersensitivity. More serious systemic reactions occur in 0.4–0.8% of the children and 3% of adults, particularly beekeepers. They can be mild, manifesting as a generalized cutaneous response, but generalized anaphylaxis may also occur and can be life-threatening [1,2]. Six bee venom components are known to cause allergic reactions in man: Api m 1 (phospholipase A₂), Api m 2 (hyaluronidase), Api m 3 (acid phosphatase), Api m 4 (melittin), Api m 6 and Api m 7 (CUB serine protease) [3]. With the recent finding of new bee venom components [4,5] the question raised whether these proteins could also contribute to the adverse immunological reactions evoked by bee venom.

Here we tested the IgE reactivity against “new venom protein 2”, found on 2-DE gels of bee venom and identified by mass spectrometry [4]. Because of its low abundance in bee venom we have chosen to provide sufficient amounts of the protein for immunological testing by the recombinant technology. As Mascot searching of the mass spectrometric data of venom protein 2 matched only a partial nucleotide sequence (expressed sequence tag, EST) of the honeybee brain [6], its coding sequence needed to be completed by rapid amplification of cDNA ends (RACE). The authors proposed to name this new venom protein “icarapin”.

2. Materials and methods

2.1. Patient sera

Sera of five patients from the University Hospital of Antwerp (Antwerp, Belgium) with a compelling history of bee venom allergy (urticaria, angio-edema, bronchospasm, hypotension and/or shock) documented by positive quantification of bee venom specific IgE by ImmunoCAP FEIA (Phadia AB) (respective values: 100, 12.7, 2.57, 1.17, 7.23 kUa/L for patients 1–5), skin test (ALK-Abellö) and/or basophil activation test [1], were analyzed. All patients were beekeepers, except patient 4. Negative control sera were from non-responsive stung control persons (IgE values <0.35 kUa/L and negative basophil activation test).

2.2. Bee venom gland preparation

The venom glands of 320 Carniolan honeybees (*Apis mellifera carnica*) were dissected under anesthesia by chilling. Firstly, the whole sting apparatus was removed from the abdomen and submerged in RNALater® (Ambion). Subsequently, the glands were separated from the reservoir and collected all together in 100 µl new solution. Homogenization and mRNA isolation was done using the Micro-FastTrack™ 2.0 Kit (Invitrogen) following the protocol for fresh and frozen tissue. After the yield was determined, the mRNA was stored in elution buffer at –80 °C until further use.

2.3. cDNA preparation and primer development

Venom gland cDNA was prepared using AMV reverse transcriptase and the oligo dT primer from the cDNA Cycle® Kit (Invitrogen) according to the manufacturer's instructions. Icarapin gene specific primers (GSP) for rapid amplification of cDNA ends (RACE; see further) were developed inside the gene fragments encoding the two tryptic peptides found in previous work [4]: GSP forward primer (GSP-fw) 5'-TGGACACTGTTCTCGTCTACCGTCCA-3' and GSP reverse primer (GSP-rv) 5'-ACGTGACAAAATGCCAGCCATCTGC-3'.

2.4. 5'-Rapid amplification of cDNA ends (5'-RACE) and 3'-RACE

RACE ready cDNA was prepared following the protocol described in the GeneRacer™ Kit (Invitrogen). Briefly, 100 ng of bee venom gland mRNA was treated with calf intestinal phosphatase and tobacco

*Corresponding author. Fax: +32 9 264 5242.

E-mail address: Dirk.deGraaf@UGent.be (D.C. de Graaf).

Abbreviations: EST, expressed sequence tag; HIRS, hyperimmune rabbit serum; GSP, gene specific primer; GSP-fw, gene specific forward primer; GSP-rv, gene specific reverse primer; RACE, rapid amplification of cDNA ends; UTR, untranslated region

acid pyrophosphatase, to be ligated at its 5'-end with the GeneRacer™ RNA oligo. This ligated mRNA was reverse transcribed using the SuperScript™ III RT and the GeneRacer™ oligo dT primer to create RACE-ready first-strand cDNA with known priming sites at the 5'- and 3'-ends. Generation of 5'-RACE fragment was done by PCR, using a combination of GeneRacer™ 5'-primer and GSP-rv, whereas the 3'-amplification needed a combination of GSP-fw and GeneRacer™ 3' primer. Both reactions were done in an Eppendorf Mastercycler using a touchdown protocol.

2.5. Expression of icarapin in bacteria

The complete coding sequence (stop codon excluded) of icarapin was amplified using the following primer set: 5'-ATGAAGACCTTGGCGTTCT-3' and 5'-AGCAGTTAATACATCTCCTTGGTTC-3'. The resulting amplicon had the expected length and was subsequently cloned in pBAD-TOPO-TA (Invitrogen) according to manufacturer's instructions. The retained clone was sequenced for confirmation.

Recombinant His-tagged icarapin was purified from L-arabinose stimulated cells on the ProBond Purification System (Invitrogen) under denaturing conditions as described in the manufacturer's instruction manual. The purified protein was dialysed overnight at 4 °C with Tris–buffer (pH 8.0) supplemented with 0.1% Triton X-100 and concentrated on Vivaspin 2 ml microconcentrators (cut-off: 10 kDa; Vivascience Ltd.). An irrelevant recombinant *Cryptosporidium parvum* surface protein CP15/60 expressed in the same pBAD-TOPO-TA vector [7] served as negative control.

2.6. Dot blot

Two µg recombinant icarapin was spotted on nitrocellulose membrane, air-dried and was washed twice with deionized water before incubating in milk diluent (Sigma) for 1 h at room temperature. Subsequently, the membrane was immersed with diluted human serum (1:4 in milk diluent) overnight at 4 °C. Indirect staining of specifically bound human IgE was performed in three steps: first with a monoclonal anti-human IgE antibody (1:1000, Sigma) for three hours at room temperature, thereafter with a peroxidase conjugated rabbit anti-mouse IgG (1:8000, Sigma) for one hour and finally by incubating in the chemiluminescent substrate (SuperSignal West Dura Extended Duration, Pierce) for 10 min. Blots were washed with Tris–buffered saline after each treatment. The chemiluminescent reaction was visualized and analyzed with a Chemi Doc system (Bio-Rad).

2.7. Western blot

A purified recombinant icarapin sample was separated on a 12% SDS–PAGE and blotted to polyvinylidene difluoride membrane. Icarapin-specific IgE staining was done as described above (dot blot). Anti-His (C-terminal) HRP antibodies (1:5000, Invitrogen) were used to localize the His-tagged recombinant proteins after chemiluminescent reaction. The same antibodies were also used in an immunostaining with 3,3'-diaminobenzidine tetrahydrochloride, permitting to localize the His-tagged bands on adjacent Coomassie Brilliant Blue G-250 stained strips. These Coomassie stained strips were used for N-terminal sequencing of individually cut out bands.

2.8. Nucleotide and amino acid sequencing

DNA sequencing was performed using an ABI Prism 377 automated DNA sequencer (Perkin Elmer), using the ABI Prism BigDye V 3.1 Terminator Cycle Sequencing kit. N-terminal amino acid sequencing of blotted Coomassie stained protein bands was carried out on a 476 A pulsed liquid sequenator with on-line detection of the PTH-amino acids (Applied Biosystems).

2.9. Immunolocalization

A hyperimmune rabbit serum (HIRS) was developed against recombinant icarapin by repeated immunizations (Ethics Committee file 041126-10). Disturbing anti-*Escherichia coli* antibodies were absorbed with bacterial lysates. Ten µm paraffin sections of pooled honeybee venom glands were made using a Microm HM360 microtome. Immunostaining was done in a humidified chamber by incubating glass sections first in primary antibodies (1:25 in Tris–buffered saline with 0.05% Tween-20 and 3% bovine serum albumin) overnight at 4 °C, thereafter in FITC-conjugated anti-rabbit IgG (1:150, Sigma) for 30 min at room temperature. Images were taken with a confocal

laser-scanning microscope (MRC1024 UV, Bio-Rad). Excitation was done with a Kr/Ar laser at 488 nm (dichroic filter UBHS, emission filter: O G515). The specificity of the staining was evaluated by depleting the anti-icarapin antibodies in a 4-h incubation step with affinity purified recombinant icarapin, prior to the overnight incubation.

3. Results

3.1. Nucleotide sequence

5'-Rapid amplification of cDNA ends (5'-RACE) using gene-specific primers resulted in an amplicon of approximately 500 bp in size that was cloned in the pCR®4 vector. Transformation to chemically competent TOP-10 *E. coli* yielded a large number of transformants, nine of which were retained for further analysis. The DNA sequences found were identical and represented by the insert of *clone 5.V4*. 3'-RACE resulted in a fragment of approximately 850 bp and some minor rapidly migrating bands. The 850 bp band was cut out and selectively amplified again by PCR, and thereafter cloned and sequenced as described above. These sequencing data were represented by the insert of *clone 3.A4*. Subsequently, we amplified the complete icarapin transcript from a venom gland cDNA preparation using primers designed, respectively, at the extreme 5'- and 3'-ends of *clone 5.V4* and *clone 3.A4*. This full-length cDNA of icarapin is 1045 bp long and consists of a 5' untranslated region (UTR) of 160 bp, a 672 bp coding sequence and a 3' UTR of 213 bp. Some remarkable sequence differences were found when compared to the previously cloned 3'RACE ends (see further). We named this full-length cDNA transcript variant 1 and deposited its nucleotide sequence to Genbank (accession number [DQ485318](#)). It was demonstrated that the icarapin gene has three introns located at the genome positions 6756463–6780887 (or in between transcript positions 224/225), 6781092–6795150 (428/429), and 6795279–6796248 (556/557) of chromosome 1 in the honeybee genome assembly 3.08 ([Supplementary Fig. 1](#)). Beside, the insert of *clone 3.A4* was considered to be a part of another transcript variant (variant 2; GenBank accession number [DQ485319](#)) of icarapin, characterized by an alternative splicing at the intron 2 end (genome positions 6781092–6795162) resulting in a transcript that is 12 bp shorter. Both transcript variants showed to have one single deletion mutation of 4 Ts in a 16-mer poly-T repeat when compared to the honeybee genome assembly 3.08. As the deletion mutation was situated in the 3' UTR there were no consequences for the expressed gene products.

3.2. Deduced amino acid sequence

The deduced amino acid sequences of transcript variant 1 represents a polypeptide of 223 amino acids (GenBank acc. n° [ABF21077](#)) and based on an in silico SigP-NN prediction [8], a signal peptide cleavage site was identified between position 19 and 20 ([Supplementary Fig. 2](#)). The mature protein has multiple putative N- and O-linked glycosylation sites as determined by NetOGlc 3.1 and NetNGLC 1.0 [9]. BLASTP searching revealed a strong homology (BLAST score up to 63.9; E-value: 4e-9) with a series of mostly unknown insect proteins (Genbank acc. n° [XP_975949](#), [XP_966682](#), [EAA45242](#), [ABF18040](#), [EAL33820](#), [AAQ09844](#), [AAL48770](#), [NP_995688](#) and [NP_723747](#)). They all share a consensus sequence of 41 (sometimes 37) residues: N-[TS]-T-S-[TV]-x-K-[VI](2)-[DN]-G-H-x-V-x-I-N-[ED]-T-x-Y-x-[DHK]-x(2,6)-

<i>Apis mellifera</i>	ABF21077	(223)	126	NTTSTTKI	IDGHVVT	INETTYT	DGSDDY	STLIRV	RVIDVRP	166
<i>Apis mellifera</i>	ABF21078	(175)	78	NTTSTTKI	IDGHVVT	INETTYT	DGSDDY	STLIRV	RVIDVRP	118
<i>Tribolium castaneum</i>	XP_975949	(193)	103	NTTSVTKV	IDGHKVV	INETEYK	HGDLGG	AFFKVRI	IDVLP	143
<i>Tribolium castaneum</i>	XP_966682	(249)	103	NTTSVTKV	IDGHKVV	INETEYK	HGDLGG	AFFKVRI	IDVLP	143
<i>Anopheles gambiae</i>	EAA45242	(489)	126	NTTSTVKI	IDGHKVV	INDTYYT	KKTEFG	TSIFKVR	IDVKP	166
<i>Aedes aegypti</i>	ABF18040	(268)	141	NSTSTVKI	IDGHKVI	INDTYYT	KKTEYGT	SVYKVR	IDVKP	181
<i>Drosophila pseudoobscura</i>	EAL33820	(463)	138	NTTSTVKV	VDGHKVE	INETVY	GDT---	NSLFKVR	VNV	174
<i>Drosophila yakuba</i>	AAQ09844	(198)	97	NTTSTIKV	VDGHKVE	INETVY	GDS---	NSVFKVR	LNV	133
<i>Drosophila melanogaster</i>	AAL48770	(431)	133	NTTSTIKV	VNGHKVE	INETVY	GDS---	NSVFKVR	LNV	169
<i>Drosophila melanogaster</i>	NP_995688	(234)	133	NTTSTIKV	VNGHKVE	INETVY	GDS---	NSVFKVR	LNV	169
<i>Drosophila melanogaster</i>	NP_723747	(431)	133	NTTSTIKV	VNGHKVE	INETVY	GDS---	NSVFKVR	LNV	169

***. *:::** * **:* *

. :. :***:*

Fig. 1. Alignment of the 41 (sometimes 37) residues of a consensus sequence found in a series of mostly unknown insect proteins, flanked by their start and stop positions. Species names on the left are followed by Genbank accession numbers and total length (in number of amino acids) of the protein fragment (in parentheses). “*” means that the residues in that column are identical in all sequences in the alignment; “:” means that conserved substitutions have been observed; “.” means that semi-conserved substitutions are observed.

[STA]-[VLFI]-x-[KR]-V-R-[VLI]-[IV]-[DN]-V-x-P (where x represents any amino acid, the value in parentheses indicate the number of residues and the letters between brackets represent only one of the given amino acids) (Fig. 1).

3.3. Expression of recombinant icarapin and testing of its IgE reactivity

Expression of icarapin in the pBAD-TOPO prokaryotic vector gave a major His-tagged protein band of 32.2 kDa on SDS-PAGE. After purification on a nickel-chelating affinity column a number of thin, rapidly migrating His-tagged bands (between 17.7 and 25.7 kDa) accompanied this major band.

Four out of 5 patients with varying venom specific IgE titers gave a positive dot blot against icarapin (Fig. 2; patient 1, 2, 3 and 5). Serum from non-responsive stung control persons tested negative. Whole (native) bee venom (ALK-Abellö) gave the same ratio of IgE reactivity as icarapin (4 out of 5), but the one patient that did not react was different (here patient 5). Dots with the irrelevant apicomplexan protein CP15/60 remained uncolored (not shown). Westernblot analysis revealed

that the IgE recognition of the preparation was not only caused by the major 32.2 kDa His-tagged band, but also by its accompanying His-tagged breakdown products (Fig. 3). N-terminal amino acid sequencing of a Coomassie blue stained band cut from the PVDF blot at the same height as the 32.2 kDa IgE reactive band (position determined by the anti-His staining) revealed 7 consecutive residues of the enterokinase recognition site that precedes icarapin in the (recombinant) fusion protein.

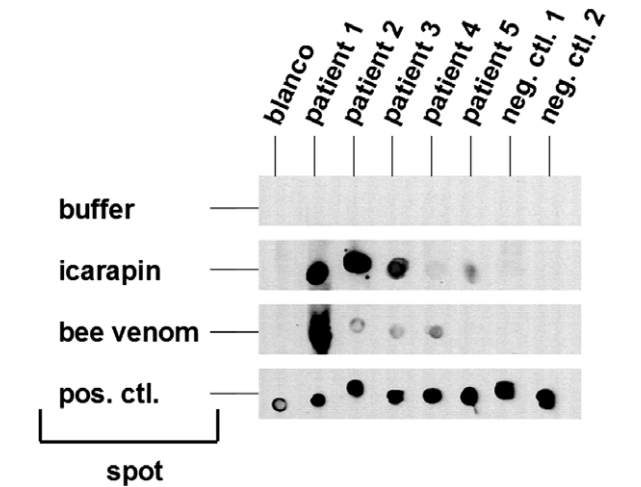


Fig. 2. IgE-specific staining of dot blots using sera from bee venom allergic patients (all were beekeepers except patient 4) and non-responsive stung control persons (neg.ctl.). In the blanco nitrocellulose strips no serum was added. Protein spots were made with recombinant icarapin, natural bee venom, a peroxidase conjugated immunoglobulin (pos.ctl.) or with buffer only.

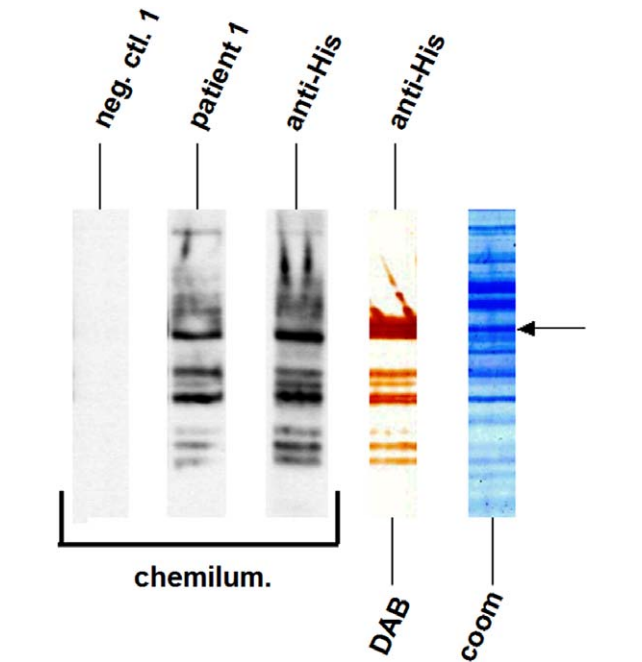


Fig. 3. Westernblots and Coomassie stained blots of affinity chromatography purified recombinant icarapin. IgE-specific staining was done with chemiluminescent substrate using sera from a bee venom allergic beekeeper (patient 1) and a non-responsive stung control person (neg.ctl.). The recombinant icarapin was localized using an anti-His monoclonal that was peroxidase conjugated and stained with the chemiluminescent substrate and with 3,3'-diaminobenzidine tetrahydrochloride (DAB), permitting to localize the His-tagged bands on the adjacent Coomassie Brilliant Blue G-250 stained strips. The arrow indicates the protein band that was cut out the blot and identified as icarapin by sequencing.

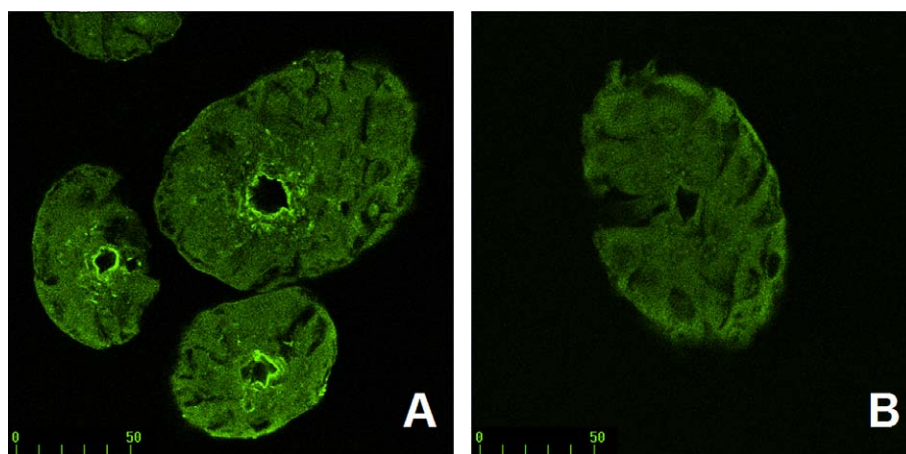


Fig. 4. Immunolocalization of icarapin using a hyperimmune rabbit serum. (A) A strong fluorescent area at the cuticular lining of venom duct was seen, with some minor staining inside the secretory cells. (B) When the serum was depleted with affinity purified recombinant icarapin the signal disappeared completely. Scale is give in µm.

3.4. Immunolocalization in the venom glands

Staining of bee venom gland sections with anti-icarapin HIRS revealed strong fluorescent area at the cuticular lining of the venom duct, with some minor staining inside the secretory cells (Fig. 4). When the HIRS was depleted with affinity purified recombinant icarapin the signal disappeared completely.

4. Discussion

Icarapin was first identified in a proteomic study of bee venom that was obtained from manually stimulated bees [4], leaving a trace of uncertainty about its origin: had this component leaked from gland tissue by squeezing out the venom sac? The finding in the present paper that it has a signal peptide, typical for secreted proteins, and that it evokes an immune response in some subjects after a bee sting, provides the required confirmation of its release during the normal sting action. In addition, this new bee venom component was independently found by another research group in reversed phase chromatography fractionated bee venom [5]. However, its occurrence in the cuticular lining of the venom duct is at the moment difficult to interpret: or icarapin penetrates from the venom into the cuticular lining that surrounds the duct and the sac, or icarapin is a component of this cuticular lining and leaks into the venom lumen. The latter seems unlikely as icarapin lacks the consensus sequences that are typical for cuticular proteins [10].

Sequencing data demonstrated that icarapin displays a protein heterogeneity that is at least partially due to alternative splicing of a single transcript, an observation also made for tropomyosins from the lobster *Homarus americanus* striated muscles [11]. Tropomyosin is an important food allergen from many other crustacean species (see the International Union of Immunology Societies' list of allergens at <http://www.allergen.org/List.htm>). It was remarkable to find homology with a set of mostly unknown insect proteins, all sharing the same consensus sequence of 37–41 residues. One of these homologues is a putative salivary secreted mucin 3 from *Aedes aegypti*, but as icarapin has no regions rich in proline, threonine and serine residues occurring in clusters (in spite of its high Pro, Thr, Ser content of 21%) and cysteine residues

are absent, it lacks some common features of mucins or mucin-like proteins [12].

Because of the failure to assign to this new venom protein a biological function, we proposed to name it icarapin, a junction of Icarus (Greek mythology) and *Apis* (genus of the honeybee) and referring to one of its most striking characteristics so far: its unstable nature. Indeed, it was previously described that on a 2-DE gel of pure venom, the full-size protein is accompanied by low molecular weight degraded forms [4]. A similar picture was seen in the affinity purified recombinant samples of the present paper, although here the degradation occurred in a more stepwise way.

The finding that 4 out of 5 patients with a compelling case history of bee venom allergy developed an icarapin-specific IgE-response suggests that it might concern a new bee venom allergen. This IgE recognition should now be further validated.

Acknowledgement: Technical assistance of Mrs. Christel Mertens was very much appreciated.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at [doi:10.1016/j.febslet.2006.08.005](https://doi.org/10.1016/j.febslet.2006.08.005).

References

- [1] Ebo, D.G., Hagendorens, M.M., Bridts, C.H., Schuerwegh, A.J., De Clerck, L.S. and Stevens, W.J. (2004) In vitro allergy diagnosis: should we follow the flow? *Clin. Exp. Allergy* 34, 332–339.
- [2] Ellis, A.K. and Day, J.H. (2005) Clinical reactivity to insect stings. *Curr. Opin. Allergy Clin. Immunol.* 5, 349–354.
- [3] Hoffman, D.R. (2006) Hymenoptera venom allergens. *Clin. Rev. Allergy Immunol.* 30, 109–128.
- [4] Peiren, N., Vanrobaeys, F., de Graaf, D.C., Devreese, B., Van Beeumen, J. and Jacobs, F.J. (2005) The protein composition of honeybee venom reconsidered by a proteomic approach. *Biochim. Biophys. Acta – Prot. Proteom.* 1752, 1–5.
- [5] Schmidt, M., Weimer, E.T., Sakell, R.H. and Hoffman, D.R. (2005) Proteins in the high molecular weight fraction of honeybee venom. *J. Allergy Clin. Immunol.* 115, S107.

- [6] Whitfield, C.W., Band, M.R., Bonaldo, M.F., Kumar, C.G., Liu, L., Pardin, J.R., Robertson, H.M., Soares, M.B. and Robinson, G.E. (2002) Annotated expressed sequence tags and cDNA microarrays for studies of brain and behavior in the honey bee. *Genome Res.* 12, 555–566.
- [7] de Graaf, D.C., De Coninck, H., De Clercq, C. and Peeters, J.E. (2002) Screening of the T- and B-cell antigenicity in neonatal calves of the His-tagged *Cryptosporidium parvum* antigens CP15, CP15/60, P23 and TRAP-C1. *Folia Parasitol.* 49, 319–322.
- [8] Bendtsen, J.D., Nielsen, H., Von Heijne, G. and Brunak, S. (2004) Improved prediction of signal peptides: SignalP 3.0. *J. Mol. Biol.* 340, 783–795.
- [9] Julenius, K., Molgaard, A., Gupta, R. and Brunak, S. (2005) Prediction, conservation analysis, and structural characterization of mammalian mucin-type O-glycosylation sites. *Glycobiology* 15, 153–164.
- [10] Willis, J.H. (1999) Cuticular proteins in insects and crustaceans. *Am. Zool.* 39, 600–609.
- [11] Mykles, D.L., Cotton, J.L.S., Taniguchi, H., Sano, K.I. and Maeda, Y. (1998) Cloning of tropomyosins from lobster (*Homarus Americanus*) striated muscles: fast and slow isoforms may be generated from the same transcript. *J. Muscle Res. Cell Motil.* 19, 105–115.
- [12] Rayms-Keller, A., McGaw, M., Oray, C., Carlson, J.O. and Beaty, B.J. (2000) Molecular cloning and characterization of a metal responsive *Aedes Aegypti* intestinal mucin cDNA. *Insect Mol. Biol.* 9, 419–426.