Available online at www.sciencedirect.com
ScienceDirect
Insect Biochemistry and Molecular Biology 38 (2008) 233–243
Insect Biochemistry and Molecular Biology
www.elsevier.com/locate/ibmb

Structure, processing and midgut secretion of putative peritrophic membrane ancillary protein (PMAP) from *Tenebrio molitor* larvae

A.H. Ferreira[a,1], P.T. Cristofoletti[a,1], D.C. Pimenta[b], A.F. Ribeiro[c], W.R. Terra[a], C. Ferreira[a,*]

[a]*Departamento de Bioquímica, Instituto de Química, Universidade de São Paulo, C.P. 26077, 05513-970 São Paulo, Brasil*
[b]*Laboratório de Espectrometria de Massa, CAT/CPID, Av. Vital Brasil 1500, 05503-900 São Paulo, Brasil*
[c]*Departamento de Genética e Biologia Evolutiva, Instituto de Biociências, Universidade de São Paulo, C.P. 11461, 05422-970 São Paulo, Brasil*

Received 10 September 2007; received in revised form 13 November 2007; accepted 15 November 2007

Abstract

A cDNA coding for a *Tenebrio molitor* midgut protein named peritrophic membrane ancillary protein (PMAP) was cloned and sequenced. The complete cDNA codes for a protein of 595 amino acids with six insect-allergen-related-repeats that may be grouped in A (predicted globular)- and B (predicted nonglobular)-types forming an ABABAB structure. The PMAP–cDNA was expressed in *Pichia pastoris* and the recombinant protein (64 kDa) was purified to homogeneity and used to raise antibodies in rabbits. The specific antibody detected PMAP peptides (22 kDa) in the anterior and middle midgut tissue, luminal contents, peritrophic membrane and feces. These peptides derive from PMAP, as supported by mass spectrometry, and resemble those formed by the *in vitro* action of trypsin on recombinant PMAP. Both *in vitro* and *in vivo* PMAP processing seem to occur by attack of trypsin to susceptible bonds in the coils predicted to link AB pairs, thus releasing the putative functional AB structures. The AB-domain structure of PMAP is found in homologous proteins from several insect orders, except lepidopterans that have the apparently derived protein known as nitrile-specifier protein. Immunocytolocalization shows that PMAP is secreted by exocytosis and becomes entrapped in the glycocalyx, before being released into midgut contents. Circumstantial evidence suggests that PMAP-like proteins have a role in peritrophic membrane type 2 formation.

Keywords: Allergen; Nitrile-specifier protein; Peritrophic membrane; Coleoptera; Lepidoptera; Digestion

1. Introduction

Several allergenic proteins have been found in insects, such as those that cause IgE-mediated allergic respiratory illnesses. Some of them are similar to well-known proteins such as tropomyosin (Asturias et al., 1999) and aspartic peptidase (Arruda et al., 1995). Others have insect-allergen-repeat-related domains and play a so far unknown role in most insects (Pomes et al., 1998; Yang et al., 2000; Shao et al., 2005; Ferreira et al., 2007), although some related lepidopteran proteins have a nitrile-specifier activity (Wittstock et al., 2004). The last mentioned proteins prevent the formation of isothiocyanates by redirecting glucosinolate hydrolysis toward nitrile formation.

The screening of an expression cDNA library of *Tenebrio molitor* using antibodies raised against midgut microvillar proteins identified several ESTs that code for proteins homologous to those having insect-allergen-repeat-related domains (Ferreira et al., 2007). The insect midgut cell microvillus is homologous to that of vertebrates (Bement and Mooseker, 1996) and midgut microvillar proteins play several roles in insects. Depending on insect taxa, microvillar proteins include: digestive enzymes, proteins involved in midgut protection, peritrophins, proteins associated with secretory mechanisms, transporters and receptors (Terra and Ferreira, 1994, 2005; Ferreira et al., 2007). The finding of those allergen proteins among midgut *T. molitor* microvillar membranes is puzzling and

*Corresponding author. Fax: +55 11 3091 2186.
E-mail address: clfterra@iq.usp.br (C. Ferreira).
[1]The contribution of these authors is equal.

0965-1748/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ibmb.2007.11.002

this work was carried out to detail their structures and secretory routes in an attempt to gather evidence to support a proposal of a biological function for those proteins. As the *T. molitor* allergen was thought to help peritrophic membrane (PM) assembly, it was named putative peritrophic membrane ancillary protein (PMAP).

A cDNA for PMAP was cloned, expressed and immunocytolocalized using antibodies that were raised against the recombinant protein. The data provided details of structure, midgut secretion and processing of PMAP molecules that are released in the feces after a putative role in PM assembly.

2. Materials and methods

2.1. Animals

Stock cultures of *T. molitor* were maintained under natural photo regime conditions on wheat bran at 24–26 °C and 70–75% r.h. Larvae of both sexes (each weighing about 0.12 g), having midguts full of food, were used.

Larvae of *T. molitor* were immobilized by placement on ice, dissected in cold 342 mM NaCl, and the midgut tissue was pulled apart. The midgut tissue, after being rinsed thoroughly with saline, was stored at −70 °C until use.

2.2. Chemicals

The plasmid Wisard Miniprep System and the pGEM-T Easy Vector plasmid kits were purchased from Promega (Madison, WI); the DNA gel extraction kit from QIAGEN Inc. (Santa Clarita, CA); the agar, agarose and oligonucleotides from Invitrogen (Carlsbad, CA) and the dNTPs and modification and restriction enzymes from New England Biolabs (Ipswich, MA). All other chemicals were purchased from Merck (Darmstadt, Germany) or Aldrich-Sigma (USA) unless stated otherwise.

2.3. Purification of microvillar membrane proteins from midgut

Microvillar membrane proteins were isolated from cytoskeleton-free microvilli prepared from midgut tissue with a procedure derived from that of Schmitz et al. (1973) and improved by Jordão et al. (1995).

2.4. Preparation of soluble midgut contents and PM from T. molitor

The PM and its contents were homogenized in double-distilled water with the aid of a Potter–Elvehjem homogenizer and centrifuged at 25,000*g* for 30 min at 4 °C. The resulting supernatant was used as source of soluble luminal contents. The pellet contains the insoluble fraction of ingested food as well as the PM that corresponds to a gel-like material seen at the top of the pellet. This gel-like material was collected with a spoon and re-suspended in 60% glycerol and centrifuged at 10,000*g* for 15 min at 4 °C. The supernatant was diluted in water to a final 10% glycerol concentration and centrifuged at 25,000*g* for 30 min at 4 °C. The resulting pellet was re-suspended in double-distilled water and labeled as PM preparation.

2.5. Preparation of wheat bran homogenates and fecal pellets of T. molitor

A sample of wheat bran was suspended in double-distilled water. The suspension was then homogenized with the aid of an Omni-Mixer (Sorvall, USA) at 6000 rpm for 3 cycles of 30 s with 30 s pause in between, followed by centrifugation at 3000*g* for 5 min at 4 °C. The supernatant was collected and re-centrifuged at 25,000*g* for 15 min at 4 °C. The new supernatant was used as source of soluble wheat bran protein.

The fecal pellets of *T. molitor* are found in the bottom of the larvae culturing vessels. They were separated manually from the wheat bran fragments. The isolated fecal pellets were homogenized in double distilled water and further processed as described for the wheat bran preparation.

2.6. Protein determination and hydrolase assays

Protein was determined according to the method of Smith et al. (1985), as modified by Morton and Evans (1992), using ovalbumin as a standard. Alternatively, the Bradford (1976) method was used.

2.7. Sodium dodecyl sulfate (SDS)–polyacrylamide gel electrophoresis (PAGE) and immunoblotting

After a desalting step in a fast desalting column, the amount of protein in the samples was determined and they were then freeze-dried. The pellets were re-suspended in SDS–PAGE sample buffer [60 mM Tris–HCl, pH 6.8, 2.5% (w/v) SDS, 0.36 mM 2-mercaptoethanol, 10% glycerol and 0.05% (w/v) bromophenol blue]. The samples were heated for 2 min at 95 °C in a water bath before being loaded onto the gels. The SDS–PAGE of samples was carried out as previously described (Cristofoletti et al., 2005). The gel was stained with a solution of 0.1% (w/v) Comassie Blue R in 10% acetic acid and 40% methanol for 30 min. Destaining was achieved with several washes in a solution containing 40% methanol and 10% acetic acid. Alternatively, the gel was silver stained following Blum et al. (1987). Molecular-mass markers used were: lysozyme (14.4 kDa), soybean trypsin inhibitor (21.5 kDa), carbonic anhydrase (31 kDa), ovoalbumin (45 kDa), bovine serum albumin (66 kDa), and phosphorilase b (97.4 kDa).

Slab gel electrophoresis was also done under native conditions, with no heating and the addition of neither mercaptoethanol nor SDS in the gel and running buffer and sample buffer, maintaining all the other conditions as for SDS–PAGE.

Immunoblotting was performed as follows: after SDS–PAGE, the proteins were electrophoretically transferred onto a nitrocellulose membrane filter (pore size 0.45 mm; BioRad, USA) (Towbin et al., 1979). The transfer efficiency was evaluated by observing the pre-stained molecular mass markers (BioRad or Sigma, USA). Filters were reacted (after a blocking step) with the antiserum diluted 200 fold in Tris-buffered solution (TBS): 50 mM Tris–HCl buffer, pH 7.4, with 0.15 M NaCl) containing 0.05% Tween 20 (TBS-T) for 2 h at room temperature (25 °C). After extensive washing with TBS-T, the filters were reacted with anti-rabbit IgG coupled with peroxidase (Sigma, USA) diluted 1:1000 in TBS-T for 2 h at room temperature. After washing extensively with the same buffer, the strips were incubated with 0.08% 4-chloro-1-naphthol in TBS containing 0.1% hydrogen peroxide until gray bands appeared where antigens were recognized. The reagent was prepared before use by the addition of one volume of 0.5% chloro-naphthol in methanol to five volumes of TBS with hydrogen peroxide.

Pre-immune serum was used in control experiments that showed no reactive bands against insect proteins.

2.8. Cloning and sequencing the cDNA coding for PMAP

The sequence of the cDNA coding for PMAP was obtained from the ESTs assembled as described by Ferreira et al. (2007) for *T. molitor* library screenings. A full clone was sequenced in an automatic DNA sequencer "ABI 3100" (Applied Biosystems) performed with the DNA kit Big Dye Terminator Cycle sequencing (Applied Biosystems) using the primers: T3 (5′ AATTAACCCTCACTAAGG 3′), T7 (5′ TAATACGACTCACTATAGGG 3′), and two specific primers for the cDNA coding for PMAP, PMAP1 (5′ CCCTCCTCGATGAAG 3′) and PMAP2 (5′ GTGGCACCTGTGATG 3′). The quality of the complete assembled sequences was above 20 as determined by the algorithm Phred-phrap (http://www.phrap.org/phredphrapconsed.html) (Ewin et al., 1998; Ewin and Green, 1998).

2.9. Recombinant expression of PMAP in Pichia pastoris

PMAP–cDNA sequence was amplified using designed primers for its amplification: forward 5′ CTC GAG AAA AGA GAG GCT GAA GCT GCC CCC AAG CAC CAC CCT AGG 3′ and reverse 5′ GCG GCC GCT TAG AAT TTC TTT GGT TCA C 3′. The PCR reaction was performed with TAQ DNA polymerase (Invitrogen, La Jolla, CA), according to the manufacturer's instructions: 30 cycles of amplification (94 °C for 30 s; 55 °C for 30 s; 72 °C for 120 s), followed by incubation at 72 °C for 10 min. The product of PCR reaction was cloned into pGEM-T Easy Vector (Promega), using the kit protocol.

The pGEM-T Easy vector containing the PCR fragment was used to transform chemically competent *Escherichia coli* SL1-blue. The resulting plasmid was purified using Wizard Miniprep Kit (Promega). The purified plasmid was digested with a combination of the restriction enzymes NotI and XhoI (New England Labs), following the supplier's instructions. The cohesive ends in the fragment and pPic9 vector were linked with T4 DNA ligase (Invitrogen), according to the manufacturer's protocol. The reaction mixture was used to transform chemically competent *E. coli* XL1-blue, in order to obtain the plasmid pPic9 with PMAP DNA fragment. The resulting plasmid was sequenced to confirm the correct insertion into the vector and the absence of changes in the coding sequence.

The transformation and expression in *P. pastoris* (strain GS115) followed the instructions of the "Pichia Expression Kit Protocol" (Invitrogen). Transformation was done with 20 mg of linearized plasmid with SacI restriction enzyme using Gene Pulser apparatus (Bio Rad). Control without the PMAP insert was included. Transformed colonies were selected in minimum dextrose (MD) medium (2% dextrose, 1.34% yeast nitrogen base with ammonium sulfate without amino acids, 4×10^{-5}% biotin) after 72 h incubation at 30 °C.

Colonies with high expression level were selected by mini-expression protocols. Each test colony was incubated in 10 mL of yeast extract–peptone–dextrose (YPD) medium for 16 h at 28 °C under agitation. Cultures with absorbance of 1.0 were centrifuged at 1900*g* at room temperature for 5 min. The cells were re-suspended in sterilized Milli *Q* water and re-centrifuged at the same conditions. The new pellets were re-suspended in 10 mL induction buffer (buffered minimal methanol yeast extract, BBMY) containing 0.5% methanol. At each 12 h interval, 50 μL methanol was added to the induction medium (to retain 0.5% methanol) and the supernatant collected to evaluate the expression level, since the expression vector is designed to secrete the protein into the medium. The best-expressing colony was used to express the recombinant protein on a large scale with the 1 L expression medium protocol with a 48 h-methanol induction time. The final level of expression was evaluated by SDS–PAGE.

2.10. Purification of PMAP

For PMAP purification, 1 L medium after 48 h methanol induction was used. The medium was made 45% in ammonium sulfate with salt addition and left standing for 30 min in ice-bath with agitation, after which the suspension was centrifuged at 27,000*g* for 10 min at 4 °C. The supernatant was set to 85% in ammonium sulfate (with the aid of a saturated solution in 50 mM MES buffer pH 6.0) and after standing and centrifuging like described before, the pellet was collected. The pellet was re-suspended in 20 mM MES buffer pH 6.0, and centrifuged at 27,000*g* for 10 min at 4 °C. The recovered protein was applied onto a HiTrap column (Pharmacia, Sweden), equilibrated with 20 mM MES buffer, pH 6.0, in order to remove the remaining ammonium sulfate. The eluate was

applied onto a 5 mL EconoPac High *Q* column (BioRad) equilibrated with the same MES buffer used in the previous step. Elution was accomplished with sequential 5-column volumes of each of the following solutions: 0, 0.4, 0.8 and 1 M NaCI in the equilibration buffer. PMAP was eluted during the 0.8 M NaCI wash. The samples containing the recombinant protein were pooled, adjusted to 1.7 M ammonium sulfate using a saturated solution and applied onto a 5 mL-EconoPac Methyl HIC (BioRad) column. The elution was attempted with different 5-column-volume washes of ammonium sulfate (1.7, 1.36, 0.68, and 0 M) in equilibration buffer. PMAP was eluted during the 0.68 M ammonium sulfate wash. Silver-staining of the gels after SDS–PAGE was used to recognize PMAP in each eluted and desalted sample.

2.11. Preparation of PMAP antibodies

Purified PMAP was used to raise antibodies in rabbits, as detailed elsewhere (Silva et al., 1995; Cristofoletti et al., 2001). For each injection, 35 μg of purified PMAP was submitted to a SDS–PAGE. The region of the gel corresponding to the 64 kDa band was excised (to be free from the PMAP 56 kDa form that contaminates the preparation), homogenized in water and mixed with Freund's adjuvant that was applied subcutaneously into rabbits. The antibody specificity was evaluated by immunoblotting. Pre-immune serum was collected as control.

2.12. PMAP peptide fingerprinting

The samples were analyzed by matrix assisted desorption ionization (MALDI)-time-of-flight (TOF) mass spectrometry using alpha-cyano-4-hydroxycinnamic acid as matrix on an Ettan MALDI-TOF/Pro instrument (Amersham Biosciences), according to Westermeier and Naven (2002). Briefly, gel bands were picked, destained, and treated with dithiotreitol and iodoacetic acid for reduction and alkylation of eventual disulfide bridges. Then, the treated spot was re-hydrated with a 40 ng/μL trypsin solution (Promega, sequencing grade) in 50 mM NH_4HCOOH and the reaction was carried out overnight at 30 °C. The tryptic peptides were extracted from the gel using an H_2O/acetonitrile/trifluoroacetic acid (1:1:0.05) solution incubated in an ultrasound bath. The peptide solution was concentrated and desalted with ZipTip C-18 (Millipore) prior to analysis. The sample was pre-mixed with the matrix solution (1:1, v:v) and approximately 0.4 μL were deposited on the sample loader and let dry over the bench. Mass spectrometry was performed on reflectron mode and internal calibrants (trypsin autolysis peptide) were used. External calibration was also performed, in case trypsin peptides were not present or clear, using $P_{14}R$ and Angiotensin-II, part of the ProteoMass kit for mass spectrometry (Sigma, St. Louis, MO, USA). The obtained mass list was analyzed by peptide mass fingerprinting algorithms (MASCOT and/or ALDENTE) (http://www.matrixscience.com/) for matches with known proteins sequences deposited on the public databases.

2.13. Confocal fluorescence microscopy and electron microscopy

For immunofluorescent visualization of PMAP along *T.molitor* midgut, the animals were dissected in their own haemolymph and midguts were fixed in Zamboni's fixative (Stefanini et al., 1967) for 16 h at 4 °C. After rising in PBS, the samples were embedded in Tissue Tek compound (Miles Inc., USA), frozen in liquid nitrogen and cut (14 μm thick) in a Leica CM 1850 cryostat. Tissue sections were collected on polylysine-coated glass slides and immunostained using an indirect immunofluorescence method. After immersion in PBS containing 0.2% Triton X-100 for 2 h at room temperature, the preparations were rinsed in PBS and incubated with the primary antibody diluted 1:300 in PBS plus 0.1% bovine serum albumin and 0.01% NaN_3 for 16 h at 4 °C. Sections were incubated with nonimmune serum in the same conditions for control experiments. The samples were then rinsed in PBS at room temperature and incubated with the secondary antibody (FITC or TRITIC, Sigma, St. Louis, USA) diluted 1:1000 in PBS plus 0.1% bovine serum albumin, for 1 h at room temperature. After rinsing in PBS at room temperature, the sections were mounted in Vectashiel (Vector Labs, Inc., USA) mounting medium and examined in a Zeiss LSM 410 confocal microscope.

For immunolabeling of PMAP at the ultrastructural level, midgut pieces were fixed (paraformaldehyde–glutaraldehyde), embedded in L.R. white acrylic resin, incubated with the primary and secondary (goat anti-rabbit IgG coupled to 15 nm gold) antibodies and examined in a Zeiss EM 900 electron microscope as detailed elsewhere (Jordão et al., 1996). As controls, sections were incubated with pre-immune serum under the same conditions.

3. Results

3.1. Cloning and characterization of T. molitor PMAP

The cDNA coding for a putative PMAP, previously obtained as EST and annotated and deposited as *T. molitor* cockroach allergen-like protein (Ferreira et al., 2007) was fully sequenced (Fig. 1). The sequence corresponds to 1900 nucleotides that codes for a protein with 595 amino acids, with molecular mass of 65.481 kDa and theoretical pI of 4.08. The translated protein has a signal peptide of 16 amino acids in its N-terminal, but no GPI anchor, according to the predictive GPI anchor software, DGPI (http://129.194.185.165/dgpi/index_en.html) nor transmembrane regions (http://www.enzim.hu/hmmtop/) and glycosylation sites (http://cbs.dtu.dk/services/). PMAP presents several motifs, identified with the software PROSITE motif search (http://expazy.org.tools), that are usual among proteins and do not deserve mentioning here.

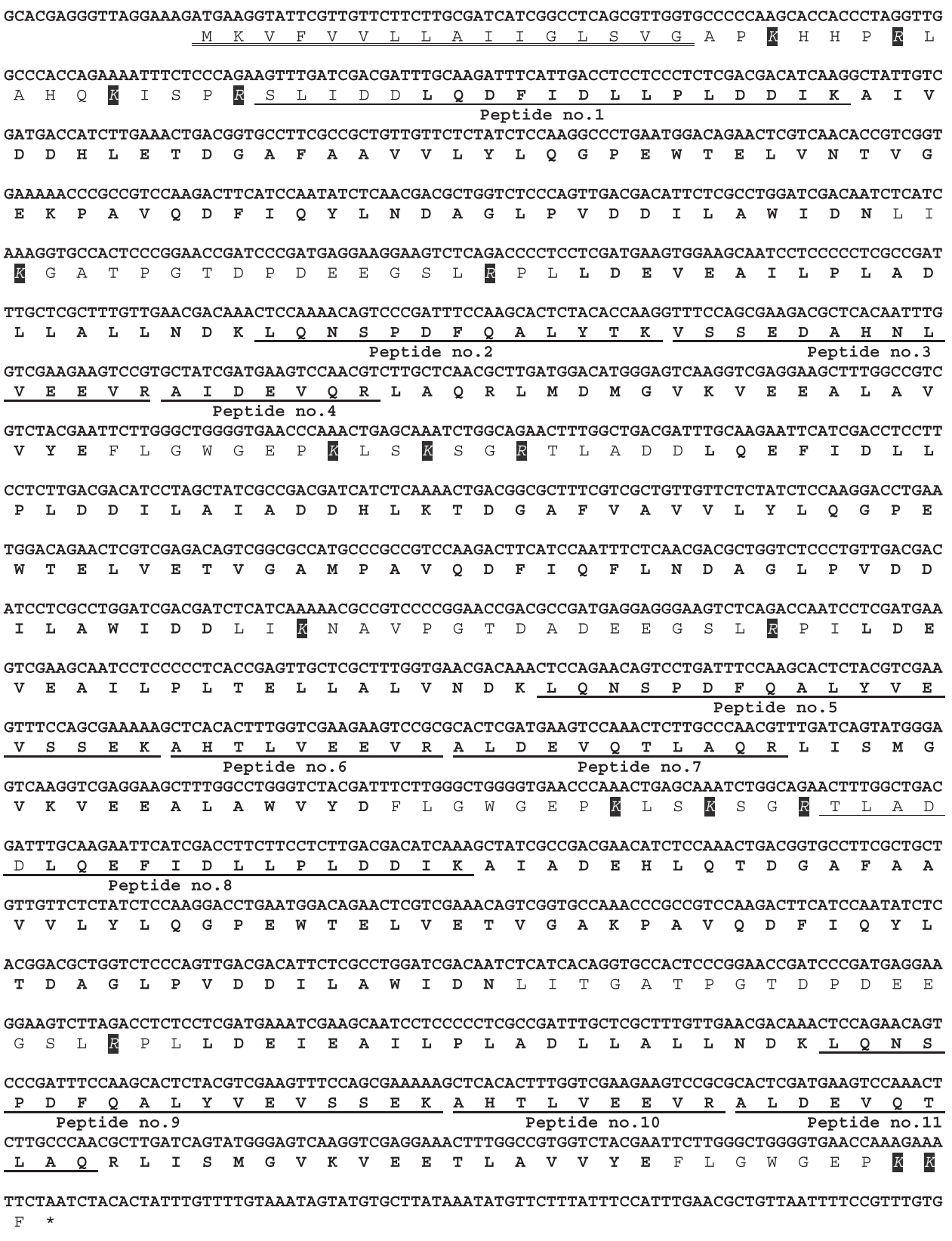

Fig. 1. Nucleotide and deduced amino acid sequence of *T.molitor* PMAP. Signal peptide is double underlined. Peptides that are underlined and numbered 1–11 correspond to sequences obtained after trypsin hydrolysis and mass spectrometric analysis. Bold amino acids correspond to the insect-allergen-related-repeat domains. Shaded K and R occur in domain linkers and are potential targets for trypsin attack. The sequence was deposited in the GenBank under the access number AAP92419.

```
17  APKHHPRLAHQKISPRSLIDDLQDFIDLLPLDDIKAIVDDHLETDGAFAA 66
224 ------------------ADDLQEFIDLLPLDDILAIADDHLKTDGAFVA 255
413 ------------------ADDLQEFIDLLPLDDIKAIADEHLQTDGAFAA 444
                       AAAAAAAAAAAAAAAAAAAAAAAAAAAAA

67  VVLYLQGPEWTELVNTVGEKPAVQDFIQYLNDAGLPVDDILAWIDNLIKG 116
256 VVLYLQGPEWTELVETVGAMPAVQDFIQFLNDAGLPVDDILAWIDDLIKN 305
445 VVLYLQGPEWTELVETVGAKPAVQDFIQYLTDAGLPVDDILAWIDNLITG 494
    AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

117 ATPGTDPDEEGSLRPLLDEVEAILPLADLLALLNDKLQNSPDFQALYTKV 166
306 AVPGTDADEEGSLRPILDEVEAILPLTELLALVNDKLQNSPDFQALYVEV 355
495 ATPGTDPDEEGSLRPLLDEIEAILPLADLLALLNDKLQNSPDFQALYVEV 544
                     BBBBBBBBBBBBBBBBBBBBBBBBBBBBBBB

167 SSEDAHNLVEEVRAIDEVQRLAQRLMDMGVKVEEALAVVYEFLGWGEPKL 216
356 SSEKAHTLVEEVRALDEVQTLAQRLISMGVKVEEALAWVYDFLGWGEPKL 405
545 SSEKAHTLVEEVRALDEVQTLAQRLISMGVKVEETLAVVYEFLGWGEPKK 594
    BBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBB

217 SKSGRTL 223
406 SKSGRTL 412
595 F------ 595
```

Fig. 2. Clustal X alignment of the insect-allergen-related-repeat domains found in the *T. molitor* PMAP sequence. Amino acids boxed in black and gray correspond to the identical and similar residues, respectively. A, amino acids pertaining to domains 1, 3, and 5 (numbered from N-terminal of the protein); B, amino acids pertaining to domains 2, 4, and 6.

PMAP has six insect-allergen-related-repeat domains (pfam06757) (Fig. 1). The first, third, and fifth ones are similar to one another and were named A1, A2, and A3, respectively (Fig. 2). The same is true for the second, fourth, and sixth domains that were labeled domain B1, B2, and B3, respectively (Fig. 2). With the aid of the software APSSP2 (www.expazy.org/tools), domains A and B are predicted to have six α-helices, all of them linked together with short coils. Domains A are globular while domains B are not, according to GLOBE (www.perdictprotein.org). The sequences between domains A and B are links having 19–20 residues (predicted to be in a coil) and were recognized as such by the software DLP as links (domain linker prediction, www.expazy.org/tools). The sequences between B and A also have 19–20 residues in a coil according to APSSP2 but were not recognized by DLP as links (Fig. 1).

The cloned PMAP sequence was expressed in *P. pastoris*. For this the sequence was cloned in pPic 9 vector without a tag sequence, which allows the expressed recombinant PMAP to be secreted to the *P. pastoris* culture medium. A highly expressing colony was selected by comparing the amounts of recombinant proteins found in the supernatant of several colonies (not shown). The selected colony was used to produce the recombinant protein in large scale and the supernatant of the culture was used as a source for crude recombinant PMAP. The protein was precipitated between 45% and 85% ammonium sulfate solution. The pellet was dialyzed and submitted to ion exchange chromatography followed by a hydrophobic chromatography, resulting in a single major band after SDS–PAGE

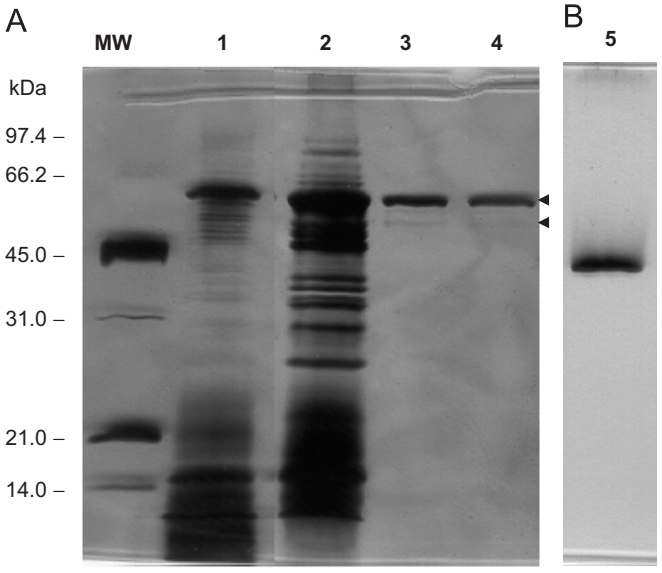

Fig. 3. Purification of the recombinant PMAP expressed in *P. pastoris*. (A) SDS–PAGE; MW, molecular-mass markers; 1–4, fractions obtained at each purification step: 1, supernatant of the culture medium of induced *P. pastoris;* 2, 85% ammonium sulfate pellet; 3, EconoPac High *Q* eluate; 4, Econo Pac Methyl HIC eluate. Arrowheads correspond to proteins with 56 and 64 kDa. (B) Native PAGE of EconoPac Methyl HIC eluate. The gels were silver stained.

(Fig. 3A). Native PAGE of the purified recombinant PMAP also results in a single major band (Fig. 3B).

The recombinant PMAP is recovered mainly as a protein of 64 kDa, although a faint band of 56 kDa is observed during

Table 1
Peptide mass fingerprinting of proteins recognized with the PMAP antiserum and isolated from different sources

Peptide no.	Mass (Da)	64 kDa protein	56 kDa protein	Luminal content	Peritrophic membrane
1	2201.1	+	–	–	–
8	2187.1	+	+	+	–
5 and/or 9	2054.0	+	+	+	+
3	1583.7	+	+	+	+
2	1524.7	+	+	–	–
7 and/or 11	1243.6	+	+	+	+
6 and/or 10	1053.5	+	+	+	+
4	830.4	+	+	–	–

The 64 kDa protein is the full size recombinant protein expressed in *P. pastoris*, whereas the 56 kDa protein is its partially degraded form. Luminal content and peritrophic membrane proteins were those resolved by SDS–PAGE and recognized with the PMAP antiserum.
Mascot analysis (http://www.matrixscience.com/) showed that all peptides are associated with PMAP. In all trials, the best hit recovered was with PMAP. Mowse score for 64 kDa protein was 144,681; for 56 kDa protein, 18,785; luminal content protein, 1701; peritrophic membrane protein, 235.

all purification steps. The 64 and 56 kDa peptides were submitted to trypsin digestion, followed by peptide mass fingerprinting. Both fingerprints match significantly ($p<0.05$) with the PMAP sequence, with probability base Mowse scores of 92 for the 64 kDa protein and 79 for the 56 kDa protein. The 64 kDa PMAP form results in 11 peptides identical with sequences present in PMAP, whereas peptide number 1 (found in the N-terminal sequence) is missing from the peptides originated from the 56 kDa PMAP form (Table 1 and Fig. 1). The 56 kDa form was considered to be an artifact caused by limited proteolysis during PMAP purification.

The recombinant 64 kDa PMAP was employed to raise antibodies in rabbit and immunoblots were used to evaluate the ability of the antibody in recognizing PMAP. The antibodies recognize the purified recombinant PMAP both in the 56 and 64 kDa forms (Fig. 4). When the PMAP antiserum is used to detect the protein in different preparations (total epithelium, midgut contents, PM and feces), the major revealed band has approximately 22 kDa, which is much smaller than the mass of the original recombinant PMAP. Nevertheless, the 22 kDa band is not recognized in the wheat bran used as food.

The 22 kDa protein bands (known to correspond to PMAP, Fig. 4) were recovered from gels after SDS–PAGE of PM and midgut contents preparations (Fig. 5A). The proteins, when submitted to peptide mass fingerprinting, result in PMAP peptides: 5 peptides from the midgut contents and 4 peptides from the PM (Table 1). Incubation of the purified 64 kDa-PMAP with trypsin led to the 22 kDa-PMAP form (Fig. 5B).

The PMAP antiserum was also used to immunolocalize the protein in the midgut. Confocal microscopy shows PMAP at the cell apex both in the anterior and middle regions of the

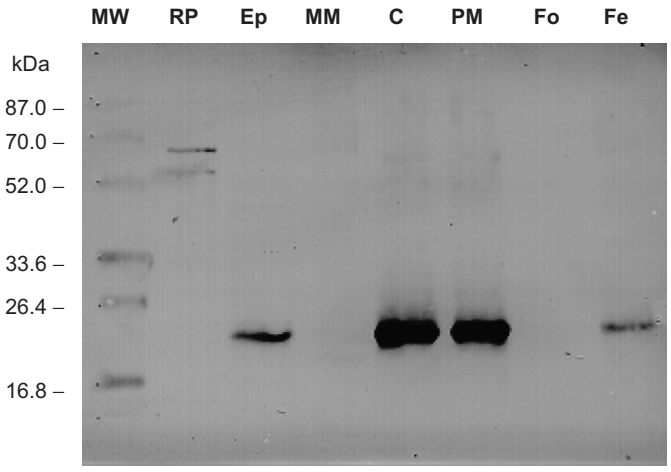

Fig. 4. Immunoblot after SDS–PAGE of samples revealed with *T. molitor* PMAP antiserum. The lanes were loaded with 0.2 (PMAP) or 10 μg (other samples) of proteins. MW, molecular mass standards; RP, recombinant PMAP; Ep, midgut of epithelium; MM, midgut microvillar membrane C, midgut contents; PM, peritrophic membrane; Fo, food (wheat bran); Fe, feces. Controls done with pre-immune serum give no reaction.

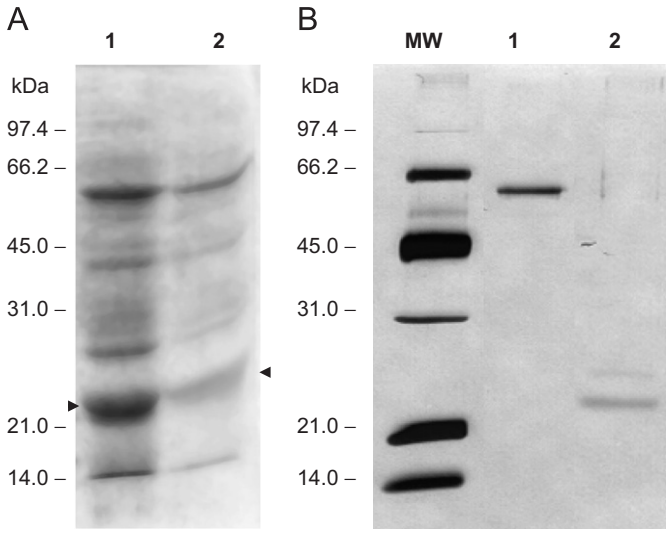

Fig. 5. SDS–PAGE of silver-stained PMAP peptides released *in vivo* and *in vitro*. (A) Midgut samples: 1, midgut contents; 2, peritrophic membrane. Arrowheads indicate peptides that were submitted to mass spectrometric fingerprinting (Table 1). (B) Treatment of 64 kDa recombinant PMAP with trypsin. 1, Control; 2, after trypsin incubation.

midgut and at the PM along the entire midgut, although at the posterior midgut the PM is less labeled (Fig. 6A–C).

Under the transmission electron microscope, PMAP was immunolocalized in the glycocalyx and inside the secretory vesicles of the anterior and middle midgut cells (Fig. 6D and E). In the posterior midgut cells, PMAP is not immunolocalized at all (Fig. 6F).

3.2. PMAP-like proteins among the insects

A search in GenBank reveals that out of 55 PMAP homologous proteins, 32 have two domains (probably one

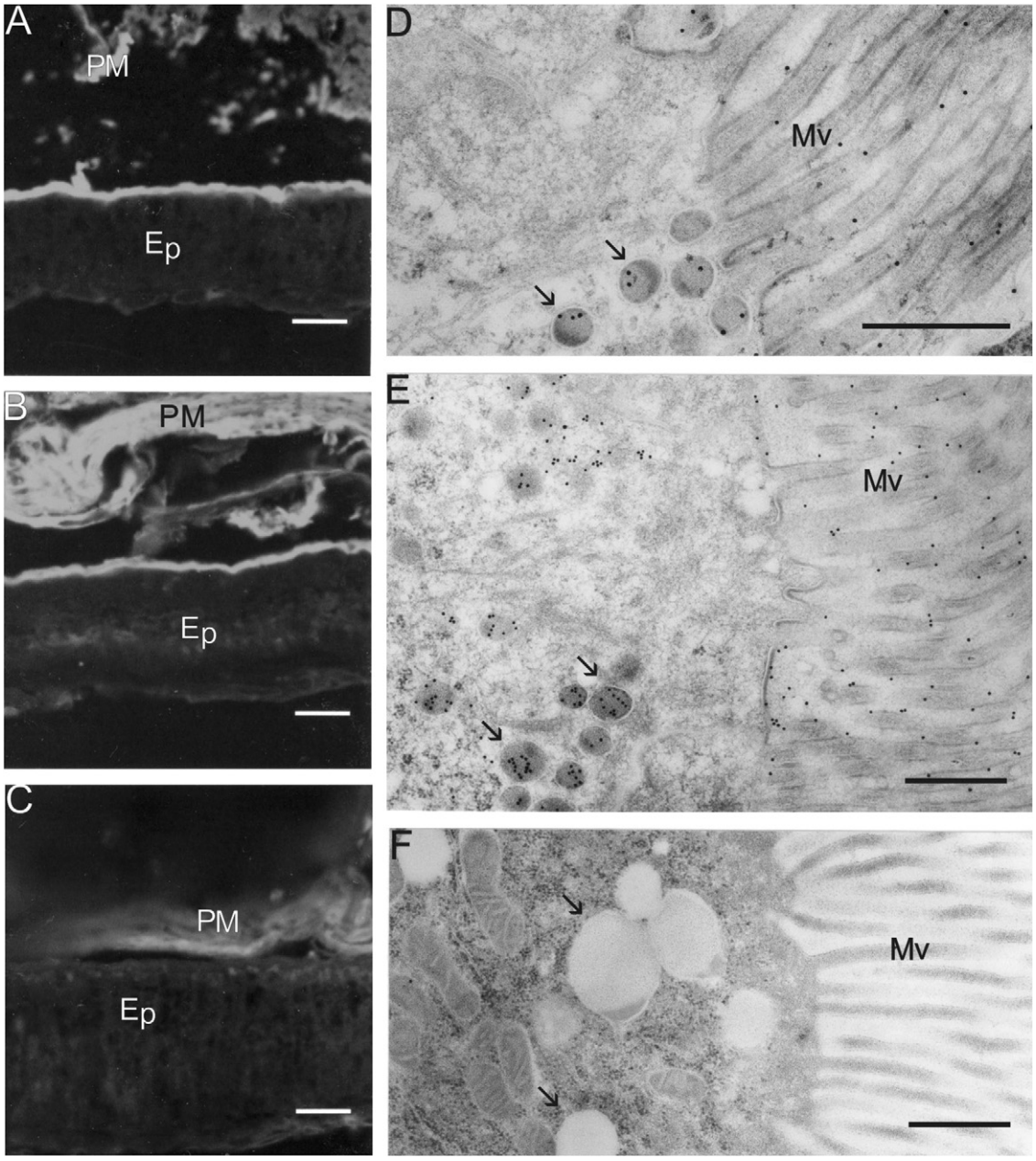

Fig. 6. Confocal immunofluorescence (A, B, C) and electron transmission immunocytochemical (D, E, F) localization of PMAP along the *T.molitor* midgut. A and D, anterior midgut; B and E, middle midgut; C and F, posterior midgut. Ep, epithelium, Mv, microvilli; PM, peritrophic membrane. No labeling was observed in controls with pre-immune serum. Arrows point secretory vesicles containing (D and E) or not (F) PMAP. Bars (μm): 25 (A, B, C), 0.5 (D, E, F).

A-like and the other B-like). The other sequences have an even (4) or an odd (19) number of domains. All the sequences with an odd number of domains are hypothetical or incomplete and thus should not be considered.

An unrooted tree of domain sequences corresponding to complete sequences that were biochemically characterized results in two major branches: one with A-like domains and the other with B-like domains (Fig. 7). It is interesting to note that both domains of the *Pieris rapae* protein are B-like (Fig. 7).

4. Discussion

4.1. Structure of PMAP

PMAP is a soluble protein with no GPI anchors, transmembrane regions or glycosylation sites and has 6 insect-allergen-related-repeat domains (pfam 06757). These domains may be grouped in types A and B occurring in pairs along the peptide chain. Domains A and B are predicted to have six α-helices linked by short coils, whereas domains are linked one another with longer coils.

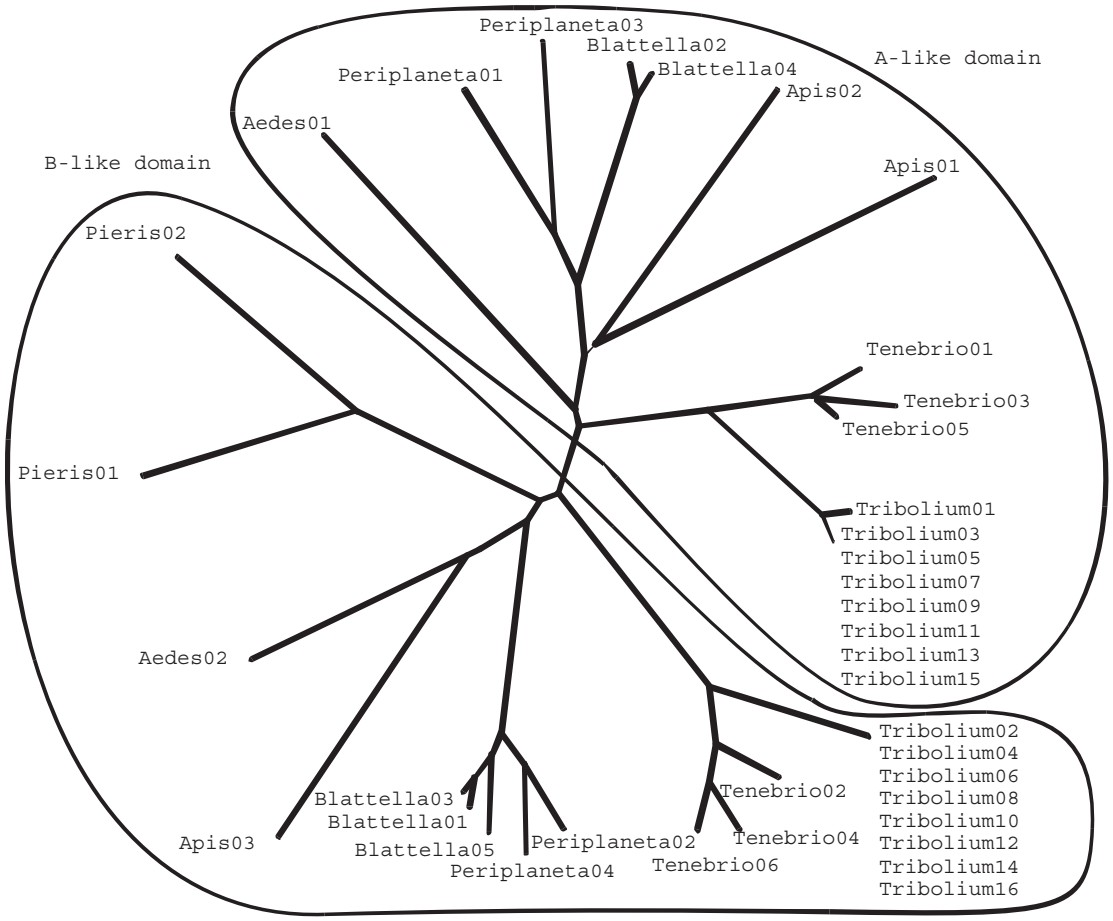

Fig. 7. Unrooted tree of insect-allergen-related-repeat domains of PMAP-like proteins from several insects. The domains were those annotated in the NCBI site, except for PMAP that were recognized in Fig. 2. The domains were classified as A- and B-like according to their sequence identities regarding PMAP domains and were numbered from the N-terminal of the protein. Access numbers (GenBank) of the sequences employed were: *Aedes aegypti* (AAL05408), *Apis mellifera* (XP_001123053), *Blatella germanica* (AAD13531), *Periplaneta americana* (AAB82404, AAP13554AAP13554, *P. rapae* (AAR84202), *Tenebrio molitor* (AAP92419), *Tribolium castaneum* (XP_971514).

Domains A are globular while domains B are not. Although several phosphorylation and myristoylation sites were predicted for PMAP, their meaning is yet unknown.

The AB-domain structure of PMAP is also found in proteins from insects pertaining to the orders Dictyoptera, Coleoptera, Hymenoptera and Diptera (adult mosquitoes). All known A-like domains branch separately from B-like domains (Fig. 7), suggesting that A and B are derived by gene duplication in an event occurring before the evolution of the neopteran orders. The A and B domains putatively derived from mitochondrial energy transfer proteins (Pomés et al., 1998).

Lepidopterans have no PMAP-like proteins. In spite of that, B-like domains are recognized in some lepidopteran proteins. Those proteins are like the nitrile-specifier protein first described in *P. rapae* where it prevents the formation of isothiocyanates by redirecting glucosinolate hydrolysis toward nitrile formation (Wittstock et al., 2004). ESTs coding to similar proteins are also found (access codes corresponding to clusters numbers in the butterfly base, BLAST PSI) in *Spodoptera frugiperda* (SF 02788_2; SFC00065_2), *Helicoverpa armigera* (HAC00198_1; HAC00198_2; HAC00176_1; HAC00022_1, *Euclidea glyphica* (EGC00004_1), *Plutella xylostella* (PXC00502_1; *Choristoneura fumigerana* (CFC0007_1); *Bombyx mori* (Bmb008629).

The two B-like domains of lepidopteran proteins derive from the B-like branch of PMAP-like proteins (Fig. 7). Thus, these proteins may have arisen from PMAP-like proteins by the loss of the A-like domain and duplication of the B-like domain.

4.2. Midgut secretion and processing of PMAP

Most anterior and middle midgut cells of *T. molitor* larva secrete PMAP molecules by exocytosis. Differently, amylase is secreted from few cells by an apocrine process (Cristofoletti et al., 2001). Once outside the cells, PMAP molecules become attached to the cell glycocalyx, where they are processed to a 22 kDa form that is found in tissue homogenates, PM, luminal contents, and feces. The 22 kDa form of PMAP is supposed to be a mixture of AB

sequences (A1B1, A2B2, and A3B3), as supported by the evidence discussed below.

Recombinant PMAP treated in native conditions with trypsin results in a 22 kDa form, like those found *in vivo*. If denaturing conditions are employed, trypsin releases the short peptides shown in Fig. 1 that include sequences inside domains. Thus, in native conditions trypsin should attack susceptible peptide bonds in the linkers between AB pairs, mirroring *in vivo* processing. If this is true, PMAP processing should result in AB units that may correspond to the functional molecules.

Chymotrypsin (Vinokurov et al., 2006), cathepsin L, and trypsin (Terra and Cristofoletti, 1996) are found in the *T. molitor* midgut lumen and may be responsible for PMAP processing. The primary specificities of those enzymes are: chymotrypsin, C-side of Tyr or Phe residues and other hydrophobic residues with much less efficiency (Barrett et al., 2004); cathepsin L, C-side of Arg residues next to Phe residues (Cristofoletti et al., 2005); trypsin, C-side of Arg or Lys residues (Lopes et al., 2006).

Fig. 1 shows (in shade) Arg and Lys residues occurring in the domain linkers. Arg or Lys residues occur close to the domains in A–B linkers and far from them in the B–A linkers. This suggests that B–A linkers are more susceptible to trypsin proteolysis because of less steric hindrance than in A–B linkers. Otherwise, an inspection in Fig. 1 also shows that Tyr residues do not occur in the AB–AB linkers and the Phe residues are found only too close to the domains. This discounted a chymotrypsin role in PMAP processing, because chymotrypsin action on hydrophobic residues other than Tyr and Phe are expected to be too slow to compete with trypsin action. Cathepsin L is also disregarded when taking into account the lack of susceptible bonds. Thus, it is highly probable that PMAP processing *in vivo* occurs due to trypsin action. One possibility is that the C-side of K218 and R410 be attacked. In this case 3 peptides will result: peptide A17 to K218 (includes A1 and B1), peptide S219 to R410 (includes A2 and B2), and peptide T411 to F595 (includes A3 and B3) with molecular masses of 22,431, 21,111, and 20,343, respectively. These peptides will migrate together in SDS–PAGE, as actually observed.

4.3. The physiological role of PMAP and homologous proteins

There is not a single hypothesis concerning the role of PMAP-like proteins. Below, a PMAP role in PM formation is proposed based on several lines of evidence.

PM is an anatomical structure that separates the food from the midgut epithelia in insects and plays several roles in the digestive process associated with the compartmentalization of the midgut lumen (Terra, 2001; Terra and Ferreira, 2005). PM is classified into two types (Peters, 1992). Type 1 PM is formed either by the whole midgut or by the anterior midgut. Type 2 PM is secreted by a few rows of cells at the entrance of the midgut. PM is composed of chitin fibrils interlocked with the chitin-binding domains of PM proteins (peritrophins) (Schorderet et al., 1998; Wang and Granados, 2001). Chitin is synthesized outside the cells by a chitin synthase bound to microvillar membranes (Zimoch and Merzendorf, 2002) and once self-organized in chitin fibers interlocks with peritrophin. The micrographs of Harper and Hopkins (1997) show that, during the formation of type 1 PM (found in most insects except larval Diptera that have type 2), a fibrous material appears first at the tips of the microvilli of anterior midgut cells and then is rapidly included into a thin PM surrounding the food bolus. Thus, the crucial events in PM formation are supposed to take place among microvilli.

Peritrophins are soluble proteins that are extensively immobilized at the surface of *T. molitor* midgut cells due to cell glycocalyx association (Ferreira et al., 2007). This arguably facilitates chitin–peritrophin association and it is conceivable that ancillary proteins help this process. It is proposed here that *T. molitor* PMAP may be one of those ancillary proteins. This hypothesis is based on the following circumstantial evidence: (a) PMAP is a soluble protein that, like peritrophins, occurs partly immobilized at the cell surface; (b) PMAP is found associated with PM, but in contrast with peritrophins, it seems to be removed from the PM as it moves along the midgut (PM confocal fluorescence fades along the midgut), appearing intact in the feces (the immediate cause of allergenic reactions commented in Introduction).

Circumstantial evidence favoring the view that PMAP homologues are ancillary proteins in PM formation is even stronger in mosquitoes. The PM of adult *Aedes aegypti* is produced *de novo* in response to blood feeding (Perrone and Spielman, 1988; Billingsley and Rudin, 1992). Accordingly, mRNAs coding for peritrophins (exemplified by Ae-Aper50) and PMAP-like protein (AEG12) are absent prior to, and rapidly accumulate after blood feeding followed by their translation. AEG12 is putatively inserted at the midgut microvillar membrane by a GPI anchor (Shao et al., 2005). Like PMAP, AEG12 is immobilized at the cell surface, although as an integral protein of the microvillar membrane and not as a glycocalyx-associated protein as PMAP. AEG12 is found in PM, thus meaning that it is processed to become soluble like PMAP.

A search in the EST data banks confirms that PMAP-like sequences are found among the *A. aegypti* midgut adult ESTs, but no sequences at all were recognized among larval midgut sequences. Dyctiopterans, coleopterans, hymenopterans and adult mosquitoes have PM type 1 and PMAP-like sequences, whereas mosquitoes larvae present PM type 2 (Peters, 1992). This leads to the suggestion that PMAP-like proteins concur in the formation of PM type 1, but not PM type 2.

As remarked above, lepidopterans do not have PMAP-like proteins. Thus, the PM type 1 of lepidopterans is formed without PMAP-like proteins, in contrast with what is supposed to occur with the other insects.

The function of PMAP needs further research concerning its putative role in PM type 1 formation and if this is proved to be correct, it will be necessary to show how lepidopteran PM is formed in its absence.

Acknowledgments

This work was supported by the Brazilian research agency FAPESP. We are indebted to W. Caldeira and M.V. Cruz for technical assistance. A.H.P. Ferreira, P.T. Cristofoletti, D.C. Pimenta are post-doc fellows of FAPESP. W.R. Terra and C. Ferreira are staff members of the Biochemistry Department and research fellows from CNPq.

References

Arruda, L.K., Vailes, L.D., Mann, B.J., Shannon, J., Fox, J.W., Vedvick, T.S., Hayden, M.L., Chapman, M.D., 1995. Molecular cloning of a major cockroach (*Blattella germanica*) allergen, Bla g 2. Sequence homology to the aspartic proteases. J. Biol. Chem. 270, 19563–19568.

Asturias, J.A., Gomez-Bayon, N., Arilla, M.C., Martinez, A., Palacios, R., Sanchez-Gascon, F., Martinez, J., 1999. Molecular characterization of American cockroach tropomyosin (*Periplaneta americana* allergen 7), a cross-reactive allergen. J. Immunol. 162, 4342–4348.

Barrett, A.J., Rawlings, N.D., Woessner, J.F., 2004. Handbook of Proteolytic Enzymes, second ed. Elsevier, London, UK.

Bement, W.M., Mooseker, M.S., 1996. The cytoskeleton of the intestinal epithelium: components, assembly, and dynamic rearrangements. In: Hesketh, J.E., Pryme, J.F. (Eds.), The Cytoskeleton: A Multi-volume Treatise, vol. 3. JAI Press, Greenwich, pp. 359–404.

Blum, H., Beier, H., Gross, H.J., 1987. Improved silver staining of plant proteins, RNA and DNA in polyacrylamide. Electrophoresis 8, 93–99.

Billingsley, P.F., Rudin, W., 1992. The role of the mosquito peritrophic membrane in bloodmeal digestion and infectivity of plasmodium species. J. Parasitol. 78, 430–440.

Bradford, M.M., 1976. A rapid and sensitive method for the quantization of microgram quantities of protein utilizing the principle of protein–dye binding. Analyt. Biochem. 72, 248–254.

Cristofoletti, P.T., Ribeiro, A.F., Terra, W.R., 2001. Apocrine secretion of amylase and exocytosis of trypsin along the midgut of *Tenebrio molitor*. J. Insect Physiol. 47, 143–155.

Cristofoletti, P.T., Ribeiro, A.F., Terra, W.R., 2005. The cathepsin L-like proteinases from the midgut of *Tenebrio molitor* larvae: sequence, properties, immunocytochemical localization and function *Insect Biochem.* Molec. Biol. 35, 883–901.

Ewin, B., Green, P., 1998. Basecalling of automated sequencer traces using phred II. Error probabilities. Genome Res. 8, 186–194.

Ewin, B., Hillier, L., Wendl, M., Green, P., 1998. Basecalling of automated sequencer traces using phred I. Accuracy assessment. Genome Res. 8, 175–185.

Ferreira, A.H.P., Cristofoletti, P.T., Lorenzini, D.M., Guerra, L.O., Paiva, P.B., Briones, M.R.S., Terra, W.R., Ferreira, C., 2007. Identification of midgut microvillar proteins from *Tenebrio molitor* and *Spodoptera frugiperda* by cDNA library screenings with antibodies. J. Insect Physiol. 53, 1112–1124.

Harper, M.S., Hopkins, T.L., 1997. Peritrophic membrane structure and secretion in European corn borer larvae (*Ostrinia nubilalis*). Tissue Cell 29, 461–475.

Jordão, B.P., Terra, W.R., Ferreira, C., 1995. Chemical determinations in microvillar membranes purified from brush borders isolated from the larval midgut from one Coleoptera and two Diptera species. Insect Biochem. Molec. Biol. 25, 417–426.

Jordão, B.P., Terra, W.R., Ribeiro, A.F., Lehane, M.J., Ferreira, C., 1996. Trypsin secretion in *Musca domestica* larval midgets. A biochemical and immunocytochemical study. Insect Biochem. Molec. Biol. 26, 337–346.

Lopes, A.R., Juliano, M.A., Marana, S.R., Juliano, L., Terra, W.R., 2006. Substrate specificity of insect trypsins and the role of their subsites in catalysis. Insect Biochem. Molec. Biol. 36, 130–140.

Morton, R.E., Evans, T.A., 1992. Modification of the bicinchoninic acid protein assay to eliminate lipid interference in determining lipoprotein protein content. Analyt. Biochem. 204, 332–334.

Perrone, J.B., Spielman, A., 1988. Time and site of the assembly of the peritrophic membrane of the mosquito *Aedes aegypti*. Cell. Tissue Res. 252, 473–478.

Peters, W., 1992. Peritrophic Membranes. Springer, Berlin, 238pp.

Pomés, A., Melen, E., Vailes, L.D., Retief, J.D., Arruda, L.K., Chapman, M.D., 1998. Novel allergen structures with tandem amino acid repeats derived from German and American cockroach. J. Biol. Chem. 273, 30801–30807.

Schorderet, S., Pearson, R.D., Vuocolo, T., Eisemann, C., Riding, G.A., Tellam, R.L., 1998. CDNA and deduced amino acid sequence of a peritrophic membrane glycoprotein, "peritrophin-48," from the larvae of *Lucilia cuprina*. Insect Biochem. Molec. Biol. 28, 99–111.

Schmitz, J., Preiser, H., Maestracci, D., Ghosh, B.K., Cerda, J., Crane, R.K., 1973. Purification of the human intestinal brush border membrane. Biochim. Biophys. Acta 323, 98–112.

Shao, L., Devenport, M., Fujioka, H., Ghosh, A., Jacobs-Lorena, M., 2005. Identification and characterization of a novel peritrophic matrix protein, Ae-Aper50, and the microvillar membrane protein, AEG12, from the mosquito, *Aedes aegypti*. Insect Biochem. Mol. Biol. 35, 947–959.

Silva, C.P., Ribeiro, A.F., Gulbenkian, S., Terra, W.R., 1995. Organization, origin and function of the outer microvillar (perimicrovillar) membranes of *Dysdercus peruvianus* (Hemiptera) midgut cells. J. Insect Physiol. 41, 1093–1103.

Smith, P.K., Krohn, R.I., Hermanson, G.T., Mallia, A.K., Gartner, F.H., Provenzano, M.D., Fujimoto, E.K., Goeke, N.M., Olson, B.J., Klenk, D.C., 1985. Measurement of protein using bicinchoninic acid. Anal. Biochem. 150, 76–85.

Stefanini, M., De Martino, C., Zamboni, L., 1967. Fixation of ejaculated spermatozoa for electron microscopy. Nature 216, 173–174.

Terra, W.R., 2001. The origin and functions of the insect peritrophic membrane and peritrophic gel. Arch. Insect Biochem. Physiol. 47, 47–61.

Terra, W.R., Cristofoletti, P.T., 1996. Midgut proteinases in three divergent species of Coleoptera. Comp. Biochem Physiol. 113B, 725–730.

Terra, W.R., Ferreira, C., 1994. Insect digestive enzymes: properties, compartmentalization and function. Comp. Biochem. Physiol. 109B, 1–62.

Terra, W.R., Ferreira, C., 2005. Biochemistry of digestion. In: Gilbert, L.I.K., Iatrou, S.S., Gill (Eds.), Comprehensive Molecular Insect Science, vol. 4. Elsevier, Oxford, UK, pp. 171–224.

Towbin, H., Staehelin, T., Gordon, J., 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: Procedure and some applications. Proc. Natl. Acad. Sci. (Wash) 76, 4350–4354.

Vinokurov, K.S., Elpidina, E.N., Oppert, B., Prabhakar, S., Zhuzhikov, D.P., Dunaevsky, Y.E., Belozersky, M.A., 2006. Diversity of digestive proteinases in *Tenebrio molitor* (Coleoptera: Tenebrionidae) larvae. Comp. Biochem. Physiol. 145B, 126–137.

Wang, P., Granados, R.R., 2001. Molecular structure of the peritrophic membrane (PM): identification of potential PM target sites for insect control. Arch. Insect Biochem. Physiol. 47, 110–118.

Westermeier, R., Naven, T., 2002. Proteomics in Practice. Wiley, Weinheim.

Wittstock, U., Agerbirk, N., Stauber, E.J., Olsen, C.E., Hippler, M., Mitchell-Olds, T., Gershenzon, J., Vogel, H., 2004. Successful herbivore attack due to metabolic diversion of a plant chemical defense. Proc. Natl. Acad. Sci. (Wash.) 101, 4859–4864.

Yang, C.Y., Wu, J.D., Wu, C.H., 2000. Sequence analysis of the first complete cDNA clone encoding an American cockroach Per a 1 allergen. Biochim. Biophys. Acta. 1517, 153–158.

Zimoch, L., Merzendorfer, H., 2002. Immunolocalization of chitin synthase in the tobacco hornworm. Cell Tissue Res. 308, 287–297.