Available online at www.sciencedirect.com
SCIENCE DIRECT®
Insect Biochemistry and Molecular Biology 35 (2005) 931–943
Insect Biochemistry and Molecular Biology
www.elsevier.com/locate/ibmb

Molecular identification of a bevy of serine proteinases in *Manduca sexta* hemolymph ☆

Haobo Jiang[a,*], Yang Wang[a], Yongli Gu[b], Xiaoping Guo[a], Zhen Zou[a], Frank Scholz[c], Tina E. Trenczek[c], Michael R. Kanost[b]

[a]*Department of Entomology and Plant Pathology, Oklahoma State University, Stillwater, OK 74078, USA*
[b]*Department of Biochemistry, Kansas State University, Manhattan, KS 66506, USA*
[c]*Institute of Zoology, University of Giessen, Germany*

Received 8 March 2005; received in revised form 24 March 2005; accepted 29 March 2005

Abstract

Extracellular serine proteinase pathways control immune and homeostatic processes in insects. Our current knowledge of their components is limited—**p**rophenoloxidase-**a**ctivating **p**roteinases (PAPs) are among the few **h**emolymph **p**roteinases (HPs) with known functions. To identify components of proteinase systems in the hemolymph of *Manduca sexta*, we amplified cDNAs from larval fat body or hemocytes using degenerate primers coding for two conserved regions in S1 family serine proteinases. PCR yielded fragments encoding seven known (HP1-HP4, PAP-1, PAP-2 and PAP-3) and 18 unknown (HP5-HP22) serine proteinases. We screened cDNA libraries and isolated clones for 17 of the newly discovered HPs (HP5-HP22 except for HP11) and prepared antibodies to 14 recombinant proteins (HP6, HP8-HP10, HP12, HP14-HP19, HP21 and HP22). Fourteen of the HPs contain regulatory clip domain(s) at their amino-terminus—HP1, HP2, HP6, HP8, HP13, HP17, HP18, HP21, HP22 and PAP-1 have one, whereas HP12, HP15, PAP-2 and PAP-3 have two clip domains. Multiple sequence alignment of catalytic domains in these and other arthropod serine proteinases provided useful clues for future functional analysis. Northern blot and reverse transcription PCR (RT-PCR) analyses showed increases in HP2, HP7, HP9, HP10, HP12–HP22 mRNA levels at 24 h after a bacterial challenge, and immunoblot analysis confirmed elevated concentrations of HP12, HP14–HP19, HP21 and HP22 proteins in plasma in response to injected bacteria. Hemocytes express HP13 and HP18; fat body produces HP12, HP20–HP22; both tissues synthesize the other HPs. These results collectively indicate the existence of a complex serine proteinase network in *M. sexta* hemolymph, predicted to mediate rapid defense responses upon wounding and/or microbial infection.

Keywords: Clip domain; Insect immunity; Phenoloxidase; Melanization; Hemolymph protein; Proteinase cascade

1. Introduction

Over 30 serine proteinases in human plasma are involved in blood coagulation, fibrinolysis and complement activation (O'Brien and McVey, 1993; Whaley and Lemercier, 1993). They are normally present in the circulation as inactive zymogens and become sequentially activated upon recognition of aberrant tissues or microbial polysaccharides. Through specific molecular interactions and limited proteolysis, a localized reaction is rapidly initiated to stop bleeding, dismantle clots, or attack invading microorganisms. After accomplishing their functions, the active enzymes are inactivated by serine proteinase inhibitors of the serpin superfamily or by α2-macroglobulin (Silverman et al., 2001; Gettins, 2002).

Complex serine proteinase systems have also evolved in invertebrates to maintain homeostasis and mediate immune/developmental signals (Jiang and Kanost,

☆ *The nucleotide sequences reported in this paper have been submitted to the GenBank™/EBI Data bank with accession numbers genbank:AY672781AY672781–genbank:AY672798AY672798.*

*Corresponding author. Tel.: +1 405 744 9400; fax: +1 405 744 6039.
E-mail address: haobo@okstate.edu (H. Jiang).

0965-1748/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ibmb.2005.03.009

2000). Those best characterized include the horseshoe crab hemolymph coagulation pathway and the *Drosophila* serine proteinase cascade for dorsal–ventral patterning (Iwanaga et al., 1998; Morisato and Anderson, 1995). Both of these cascade pathways comprise serine proteinases with an amino-terminal clip domain predicted to mediate protein–protein interactions. Accumulating evidence indicates that such enzyme systems also exist in insect hemolymph to coordinate humoral and cellular defense responses (Kanost et al., 2001, 2004). For instance, melanotic encapsulation, a resistance mechanism against parasite infection, involves a proteinase pathway that generates active phenoloxidase (PO) for melanin production (Ashida and Brey, 1998; Gorman and Pakewitz, 2001; Cerenius and Söderhäll, 2004). Proteolytic activation of plasmatocyte-spreading peptide precursor is probably mediated by a serine proteinase in lepidopteran insects (Hayakawa et al., 1995; Clark et al., 1998; Wang et al., 1999). In adult *Drosophila*, a proteinase pathway stimulated by microbial infection is responsible for processing of spätzle, which binds to the Toll receptor and induces drosomycin expression (Lemaitre et al., 1996; Hoffmann and Reichhart, 2002). *Persephone*, a clip-domain serine proteinase, is a component of this enzyme system (Ligoxygakis et al., 2002a). As in vertebrates, serpins have been demonstrated to down-regulate key proteinases in arthropods (Kanost, 1999; Levashina et al., 1999; De Gregorio et al., 2001; Ligoxygakis et al., 2002b; Tong et al., 2005; Zou and Jiang, 2005).

In previous studies on biochemistry of serine proteinases in hemolymph of the tobacco hornworm, *Manduca sexta*, we identified three prophenoloxidase-activating proteinases (PAPs) and four other hemolymph proteinases of unknown function (HP1–HP4) (Jiang et al., 1998, 1999, 2003a, b). We predicted that *M. sexta* hemolymph may contain a significantly greater number of serine proteinases participating in innate immunity. To obtain an overview of this enzyme system and develop specific probes for its components, we took a molecular approach involving PCRs to isolate cDNA for serine proteinases expressed in hemocytes and fat body. Reported here are the molecular cloning and structural features of these enzymes, as well as changes in mRNA and protein levels after a microbial challenge.

2. Materials and methods

2.1. Cloning of serine proteinase cDNA fragments

Two directional cDNA libraries in λZAP2 (Stratagene) were prepared using hemocyte or fat body mRNA from *M. sexta* larvae injected with bacteria (Jiang et al., 2003a). Naïve *M. sexta* larval hemocyte and fat body cDNA libraries were also constructed in the same vector. Bacteriophage DNA samples were isolated from these four libraries using Wizard Lambda DNA Purification System (Promega). Degenerate primers were designed from conserved regions based on analysis of the chymotrypsin (S1) family of serine proteinase genes in the *Drosophila* genome (Ross et al., 2003). Primer 659 (5′-GTATCGATACVGCSGCNCAYTG-3′) encodes TAAHC, whereas the reverse complement sequences of primers j601 (5′-ATCAACGTTGGRCCRCCRGARTCNCC-3′) and j602 (5′-CTATCTAGAGGRCCRCCRCTRTCNCC-3′) encode GDSGGP. The library DNA samples (0.1 μg) were used as templates in 25 μl PCRs containing primer pair 659-j601 or 659-j602 (10 pmol/primer) and *Taq* DNA polymerase (2.5 U). The thermal cycling conditions were: 94 °C, 3 min; 30 cycles of 94 °C, 30 s; 50 °C, 40 s, 72 °C, 40 s; 72 °C, 5 min. After 1% agarose gel electrophoresis, 0.4–0.6 kb PCR products were recovered from the gel and cloned into pGem-T vector (Promega). Plasmids were extracted by alkaline lysis from overnight cultures of the transformants.

2.2. Screening and sequence analysis of the PCR-derived clones

To avoid repeatedly isolating cDNA for known proteinases, the crude plasmid DNAs were spotted on a nitrocellulose membrane and hybridized with a probe mixture of known HP fragments. The cDNA fragments of PAP-1, PAP-2, PAP-3, HP1-HP8, and HP21 were individually labeled with [α-^{32}P]dCTP by PCR. HP5–HP8 and HP21 were isolated in the initial phase of this project (see Section 3.1). Each reaction mixture (25 μl) contained plasmid DNA (0.2 ng), primers 659 and j601 (10 pmol each), [α-^{32}P]dCTP (5 μl), dATP/dGTP/dTTP (50 μM each), *Taq* DNA polymerase (2.5 U, Promega), and 10 × buffer (2.5 μl). The cDNA inserts were amplified by 35 cycles of 94 °C, 30 s; 45 °C, 40 s; 72 °C, 40 s. Unincorporated ^{32}P-dCTP was removed from the pooled labeling mixtures by gel filtration chromatography on a PD-10 column (Amersham Biosciences). The plasmid DNA dot blot was hybridized with the probe mixture (1×10^6 cpm/ml) at 58 °C for 16 h, washed in 0.1 × SSC, 0.1% SDS, and subjected to autoradiography. The plasmid samples that did not display strong hybridization signals were treated with RNase A and purified by Wizard Minipreps DNA Purification System (Promega). Sequence analysis was performed using the BigDye Terminator Cycle Sequencing Ready Reaction Kit (PE Applied Biosystem).

2.3. Selection of serine proteinase cDNA clones by clone capturing

M. Sexta induced hemocyte and fat body λZAPII cDNA libraries were converted to the plasmid form by

mass in vivo excision of phagemids according to the instruction manual (Stratagene). The total number of plated colonies was adjusted to 10 times the number of recombinants in the original λ libraries to ensure complete coverage. The colonies were harvested from 360 LB agar plates (150 mm *i.d.*, 3×10^4 colonies/plate, 150 μg/ml ampicillin) by rinsing with LB medium (10 ml/plate). The combined bacterial suspension was centrifuged at 8000*g* for 20 min to remove the media, and the pellet was thoroughly resuspended in 120 ml of 50 mM glucose, 20 mM Tris-HCl, 10 mM EDTA, pH 8.0. For clone capturing, high-quality plasmid DNA was prepared from 4 ml of the cell suspension using NucleoBond Plasmid Kit (BD BioSciences).

For probe preparation, the cloned cDNA fragments were individually amplified by 25 cycles of 94 °C, 30 s; 54 °C, 40 s; 68 °C, 60 s. The PCR reaction mixtures (50 μl) contained plasmid DNA (20 ng), primers 659 and j601 (10 pmol each), dNTPs (50 μM each), biotin-21-dUTP (40 μM), 10 × buffer (5 μl) and Advantage PCR enzyme mix (1 μl, BD BioSciences). DNA products of expected size were recovered after agarose gel electrophoresis using a NucleoSpin Extraction Kit (BD BioSciences). The probe mixture was incubated with the plasmid library DNA in the presence of RecA protein and streptavidin magnetic beads, according to the manufacturer's instructions (BD BioSciences). Target plasmid DNAs were selected by the RecA-mediated homologous pairing, biotin–streptavidin binding, and magnetic interaction. The bound DNA was then eluted from the beads and used for transforming high-efficiency *E. coli* JM109 competent cells (Promega).

2.4. Screening for full-length serine proteinase cDNA clones

Plasmids were prepared by the alkaline lysis method from 3 ml cultures of the transformants resulting from clone capturing. The DNA samples immobilized on nitrocellulose membranes were individually hybridized with the probes labeled with ^{32}P (instead of biotin) to identify positive clones from the enriched mini-library. Plasmid DNA purification and sequencing were performed as described above. For some HPs that were not isolated by clone capturing, their cDNA fragments were labeled with α-^{32}P-dCTP by extension of random hexamers and then used as probes for screening the induced hemocyte and fat body cDNA libraries (2×10^5 λ plaques/probe) (Jiang et al., 1999). Plasmids were in vivo excised from the purified positive λ phages and sequenced using vector- and cDNA-specific primers.

2.5. Sequence analysis

Nucleotide sequences of *M. sexta* HPs were assembled using MacVector 6.5 (Genetics Computer Group). The deduced amino acid sequences were used to search the GenBank non-redundant database using the BLASTP program (Altschul et al., 1997). Clip domains were identified as described previously (Ross et al., 2003). Multiple sequence alignment and phylogenetic analysis of HP catalytic domains was performed using the CLUSTALW program (Thompson et al., 1994), along with the same region in the other invertebrate serine proteinases with known functions. The sequences start from four residues before the putative or known activation sites and end at the carboxyl termini. A Blosum 30 matrix was used for pairwise and multiple sequence alignments at a gap penalty of 10 and an extension penalty of 0.1. Construction of a phylogenetic tree was performed using the same program.

2.6. Northern blot and RT-PCR analyses of HP expression

As described previously (Wang et al., 1999), total RNAs were prepared from fat body and hemocytes of day 3, fifth instar larvae at 24 h after injection with saline (10 μl) or formalin-killed *E. coli* (1×10^8 cells/larva, 10 μl). RNA samples (20 μg each) were separated by 1% agarose gel electrophoresis and transferred to nitrocellulose membranes. After ultraviolet crosslinking, fixed RNA on the membranes was stained with 0.04% methylene blue to reveal the 28S rRNA band for loading estimation. Following prehybridization, the membranes were individually hybridized with ^{32}P-labeled, full-length HP12, HP18 or HP22 cDNA probe, and subjected to autoradiography for 60 days at −70 °C.

For reverse transcription PCR (RT-PCR), the RNA sample (2–4 μg), oligo dT (0.5 μg), and dNTPs (10 mM each, 1 μl) were mixed with DEPC-treated H_2O to a final volume of 12 μl, denatured at 65 °C for 5 min, and quickly chilled on ice for 3 min. cDNA was synthesized by MMLV reverse transcriptase (1 μl) in the presence of DTT (0.1 M, 2 μl), RNase OUT (40 U/μl, 1 μl, Ambion), buffer (4 μl), and the denatured RNA sample (12 μl). The *M. sexta* ribosomal protein S3 mRNA was used as an internal standard to normalize the cDNA pools in a preliminary PCR experiment. HP cDNA fragments were amplified using primer pairs specific for individual proteinases under the cycling conditions empirically chosen to avoid saturation. Relative HP cDNA levels in the normalized samples were determined by agarose electrophoresis (Jiang et al., 2003b).

2.7. Recombinant protein expression, antibody preparation, and immunoblot analysis

HP cDNAs were amplified by PCR using the longest clones and specific primers for individual HPs (Table 1). After digestion with *Nco*I and another restriction

enzyme (*Hind*III, *Pst*I, or *Sph*I), the PCR products were cloned into the same sites in H6pQE60 (Lee et al., 1994). Following sequence verification, the resulting plasmids were used for producing recombinant HP precursors in *E. coli* (Jiang et al., 2003b). Rabbit polyclonal antibodies were prepared against the proteins purified by Ni^{2+} affinity chromatography and preparative SDS-PAGE. To examine possible changes in HP concentrations after an immune challenge, day 2 fifth instar larvae were injected with 50 µl H_2O or a mixture of killed *M. luteus* (2×10^9 cells/ml), *E. coli* (2×10^9 cells/ml) and *S. cerevisiae* (5×10^8 cells/ml) cells. Plasma samples were obtained from the insects 24 h later, separated by 10% SDS-PAGE, and electrotransferred onto a nitrocellulose membrane. Immunoblot analysis was carried out using 1:2000 diluted HP antiserum as the first antibody (Jiang et al., 2003b).

3. Results

3.1. Isolation of HP5-HP22 cDNAs

Using degenerate primers 659 and j601, corresponding to conserved regions surrounding the active site His and Ser residues of S1 family proteinases, we amplified cDNA fragments from the hemocyte and fat body cDNA libraries. PCR products of the expected size (0.4–0.6 kb) were cloned in a plasmid vector. We randomly selected eight of the 468 transformants, determined their sequences, and identified four unknown serine proteinase cDNA fragments (HP5–HP8). We amplified the cDNA fragments from these and other known HP cDNA clones (HP1–HP8, PAP-1, PAP-2, PAP-3 and HP21)—HP21 had been fortuitously isolated from 10 randomly sequenced clones from the induced fat body cDNA library. To reduce the chances of repeatedly isolating these same sequences, we labeled these 12 cDNA fragments with ^{32}P-dCTP by PCR and used them in a probe mixture to hybridize with the plasmid DNA from the remaining transformants. In addition, we probed 392 transformants derived from PCRs using primers 659 and j602. The plasmid samples with strong hybridization to this probe mixture were eliminated from further analysis, and the remainder were purified and sequenced. Of the 287 plasmids analyzed, 12 new serine proteinase cDNA fragments were identified. Their names and the number of clones isolated for each (in parentheses) are HP9 (1), HP10 (7), HP11 (1), HP12 (1), HP13 (9), HP14 (1), HP15 (1), HP16 (64), HP17 (48), HP18 (20), HP19 (1) and HP20 (1). We also identified 33 HP1, 7 HP5, 8 HP6, 5 HP7, 11 HP21, 43 PAP-1 and 16 PAP-2 cDNA fragments in this collection of clones. Our previous analysis indicated that in the region near the active site His residue, TAGHC and SAAHC motifs are second to TAAHC in abundance among *Drosophila* serine proteinase sequences (Ross et al., 2003). Therefore, we synthesized degenerate primers encoding these two motifs and individually paired them with j601 or j602 in PCRs. From the resulting 392 plasmid samples, 255 were further purified and sequenced. However, all of the serine proteinases in this group of clones had been previously identified.

To isolate full-length cDNA corresponding to the fragments obtained by PCR, we screened the induced hemocyte and fat body libraries by clone capturing. The phage libraries were converted to the plasmid form by mass in vivo excision. Seven HP fragments (HP5–HP8,

Table 1
Oligonucleotides used in construction of recombinant plasmids for HP expression in *E. coli*

Name	Forward primer (5′–3′)	Reverse primer (5′–3′)	Cloning sites
HP6	TCA**CCATGG**AAAACGTTGGTG	CCATCTGGCAACAGTGA	*Nco*I, *Sph*I
HP8	TTG **CCATGG**AGTCTTGTAATGGTG	TGA**GCATGC**GTAATACGAC[a]	*Nco*I[d], *Sph*I
HP9	TTA**CCATGG**AATTGCGCGAAGAG	ACTCA**CTGCAG**GGCGAATTG[a]	*Nco*I, *Pst*I
HP10	SEQ CHAPTER CGG **CCATGG**CGATTTCATTCAAC	TGA**GCATGC**GTAATACGAC[a]	*Nco*I[d], *Sph*I
HP12	GCAGGCTTTG**CCATGG**GACAGTCATGTACA	CCTGCGCAAAGCT**T**GCCATTCCACAAAG[c]	*Nco*I, *Hind*III
HP13	SEQ CHAPTER TCG**CCATGG**CCCAATTCGAG	TGA**GCATGC**GTAATACGAC[a]	*Nco*I, *Sph*I
HP14	CCC**CCATGG**CAGTGTTGAAAG	CGTCCATTTTCGAGCTAG	*Nco*I
HP15	TCG**CCATGG**GGCAGAATTGCTC	ACTCA**CTGCAG**GGCGAATTG[b]	*Nco*I[d], *Pst*I
HP16	TTA**CCATGG**AAAAGCTGTCAGATG	ACTCA**CTGCAG**GGCGAATTG[b]	*Nco*I, *Pst*I[d]
HP17	TCCG**CCATGG**GACAGAATGGT	TGA**GCATGC**GTAATACGAC[b]	*Nco*I, *Sph*I
HP18	GTCCGGGA**CCATGG**AAAGCAGGATTCA	CCATATT**AAGCTT**TCCCTTGTCGCCGTAGC[c]	*Nco*I, *Hind*III
HP19	CCA**CCATGG**AACAAAGCACG	ACTCA**CTGCAG**GGCGAATTG[b]	*Nco*I[d], *Pst*I
HP21	AGATA**CCATGG**TTCGGGAAGTGTTG	GTCA**AGCTT**CTAAGGCCACAC	*Nco*I, *Hind*III
HP22	GGTGCATCG**CCATGG**GAGCTCAATACGAAG	GTGGCCA**AAGCTT**ATCTGCGTATCC[c]	*Nco*I, *Hind*III

[a,b]Modified T7 primers with an *Sph*I ([a]) or *Pst*I ([b]) site incorporated for directional cloning of digested PCR products.
[c]Approximately 65 residues were truncated at the carboxyl-terminus, including the reactive site Ser residue.
[d]Due to an internal site, the PCR product was partially digested to yield right-sized fragment which was then separated by agarose gel electrophoresis and cloned into H6pQE60.

Table 2
M. sexta HP cDNA clones[a]

Name	Accession #	Tissue source (occurrence)	cDNA clone names	Reference/cloning method	Note
HP1	AF017663	Hemocyte (4)	52, 57, 60, 66	Jiang et al., 1999	5′-RACE
HP2	AF017664	Hemocyte (3)	4, 7*, 9	Jiang et al., 1999	
HP3	AF017665	Hemocyte (3)	J2, J5, J7	Jiang et al., 1999	5′-RACE
HP4	AF017666	Hemocyte (3)	15, 19, 27	Jiang et al., 1999	5′-RACE
HP5	AY627781	Fat body (1)	C2	Library screening	5′ end incomplete
HP6	AY627782	Fat body (35)	D1-D12, D14-D18, D21-D25, D26*, D27-D38	Library screening	
HP7	AY627783	Fat body (1)	E3*	Library screening	
HP8	AY627784	Hemocyte/fat body (1)	335*	Clone capture	
HP9	AY627785	Fat body (9)	J1, J2, J3*, J4-J6, J8-J10	Library screening	
HP10	AY627786	Hemocyte (7), fat body (2)	K1, K3-K8, K10*, K23	Library screening	
HP11	AY627787	?	?	Library screening	Not obtained in screens
HP12	AY627788	Hemocyte/fat body (1)	198*	Clone capture	
HP13	AY627789	Hemocyte (15)	M1-M11, M12*, M13-M15	Library screening	
HP14	AY380790	Fat body (4)	N1*, N2, N3, N5	Library screening	A shorter 3′ UTR in N3
HP15	AY627790	Fat body (1)	O2*	Library screening	
HP16	AY627791	Hemocyte/fat body (11)	203, 253, 279, 299, 387, 396, 465, 526, 682*, 703, 730	Clone capture	
HP17	AY627792	Fat body (28[b])	P1, P2, P3*, P4-P14	Library screening	Alternative splicing in P1
HP18	AY627794	Hemocyte/fat body (4)	371, 393*, 496, 681	Clone capture	
HP19	AY627795	Fat body (8)	Q3-Q7, Q8*, Q9, Q10	Library screening	
HP20	AY627796	Fat body (~30[b])	R1-R3, R4*, R5-R10	Library screening	
HP21	AY627797	Fat body (~80[b])	S1-S12, S14*, S15-S20	Library screening	
HP22	AY627798	Hemocyte/fat body (1)	1298	Clone capture	5′ end near complete
PAP-1	AY789465	Hemocyte (7)	1, 2*, 3, 5-7, 10	Jiang et al., 1998	
PAP-2	AY077643	Hemocyte/fat body (14)	6*, 96, 290, 373, 421, 462, 508, 543, 547, 646, 689, 699, 713, 727	Clone capture	HP12 probe used
PAP-3	AY188445	Fat body (~100[b])	13, 69, 73, 76*	Jiang et al., 2003b	

[a]Each underlined clone has the longest cDNA insert in that group and "*" indicates that clone contains a complete open reading frame.
[b]Some of the positive plaques from 1° screening were not further purified.

HP12, HP16 and HP18) were amplified from their corresponding plasmids with biotin-21-dUTP incorporated into the PCR products. Through RecA-mediated homologous pairing and biotin–streptavidin interaction, cDNAs were selected from the converted library, and plasmids captured on the magnetic beads were used to transform *E. coli*. Plasmid DNAs from the resulting 735 transformants were isolated and probed individually with these seven ^{32}P-labeled HP fragments. Thirty-nine (5.3%) of them were positive, indicating that HP sequences were significantly enriched during the selection process. Four of the seven HPs (HP8, HP12, HP16 and HP18) were isolated as full-length clones (Table 2). In addition, 14 PAP-2 clones were found among the positives, probably due to the high sequence similarity between HP12 and PAP-2 (Fig. 2A and Table 3). Our second batch of clone capturing using a mixture of nine fragments (HP9–HP11, HP13–HP15, HP17, HP19 and HP20) yielded 1641 transformants. Due to poor incorporation of biotin into the probes, only one of them hybridized weakly with ^{32}P-labeled HP10 fragment. This clone (♯1298) was designated HP22, as its sequence differed from HP1 through HP21 (Tables 4 and 5).

We subsequently screened the induced hemocyte and fat body cDNA libraries by the conventional plaque hybridization using ^{32}P-labeled cDNA fragments of HP5–HP7, HP9–HP11, HP13–HP15, HP17, HP19, and HP20, separately. From 2×10^5 recombinant plaques in each screening, we obtained positive clones for HP5 (1), HP6 (35), HP7 (1), HP9 (9), HP10 (9), HP13 (15), HP14 (4), HP15 (1), HP17 (28), HP19 (8), HP20 (~30) (Table 2).

HP10 mRNA appears to be more abundant in hemocytes than fat body—we isolated seven and two positives from the same number of plaques in the hemocyte and fat body libraries, respectively (Table 2). HP16 and HP21 transcripts are apparently abundant: 64 of the 287 PCR-derived clones contained HP16 and screening of induced fat body library yielded ~80 HP21

Table 3
Structural features of *M. sexta* HPs

Name	Length (aa)	# of clip domain	Signal peptide	Putative activation site	M_r (kDa)	cDNA (bp)	Predicted specificity[a]	Glycosylation sites		p*I*
								N-linked	*O*-linked	
HP1	374	1	1-14S*E	AQGR*VFGS	42	1381	T(DGG)	3	1	5.6
HP2	388	1	1-17G*A	ADEL*IVGG	43	1571	T(DGG)	2	10	6.7
HP3	235	0	1-20G*L	NNVG*IYGG	26	931	T(DGG)	3	0	8.2
HP4	246	0	1-18A*L	PMVG*VAGG	27	944	C(GGT)	3	0	7.5
HP5	>334	0	?	ESDR*IIGG	?	1398	T(DGG)	1	0	?
HP6	337	1	1-20A*E	LDLH*ILGG	37	3405	T(DGG)	2	6	5.3
HP7	250	0	1-17A*I	GKPC*GAEA	27	1514	T(DGG)	0	0	6.0
HP8	347	1	1-24G*Q	NNDR*IVGG	38	1549	T(DGG)	1	1	5.8
HP9	375	0	1-18T*T	KPSF*AIGG	42	1441	T(DGG)	1	6	8.9
HP10	252	0	1-18A*I	NPKC*GVEA	28	1937	T(DGG)	0	0	5.1
HP11	>250	?	?	?	?	461	?	?	?	?
HP12	436	2	1-19G*Q	VSDK*IIGG	48	1428	T(DGG)	3	1	5.3
HP13	394	1	1-17A*Q	ADSL*IVGG	44	2564	T(DGG)	4	9	6.6
HP14	649	0	1-17T*S	GTEL*VLGG	72	4489	C(GAT)	7	0	5.0
HP15	424	2	1-17G*Q	VGNK*IIGG	47	1376	T(DGG)	6	1	5.0
HP16	429	0	1-15C*Q	HTGL*IVNG	47	1489	E(SSG)	3	1	7.4
HP17	586	1	1-20Q*S	SFSR*VVGG	65	2590	T(DGG)	2	22	5.1
HP18	381	1	1-18C*Q	RRFA*SYNG	42	1287	T(DGG)	3	0	5.9
HP19	528	0	1-20Q*S	PIPL*VVNG	59	2480	E(SSV)	3	14	8.1
HP20	328	0	1-17A*G	LDHF*VSGG	36	1301	E(ASA)	1	1	8.1
HP21	397	1	1-16A*A	ADDL*IIGG	44	1690	T(DGG)	4	10	5.9
HP22	398	1	1-12A*Q	ADEL*IVGG	44	2755	T(DGG)	4	10	5.7
PAP1	364	1	1-19S*Q	NGDR*IYGG	40	1512	T(DGG)	0	5	6.0
PAP2	422	2	1-19G*Q	FDNK*ILGG	46	2299	T(DGG)	2	1	6.0
PAP3	408	2	1-19G*Q	VGNK*IIGG	44	1381	T(DGG)	2	3	6.9

[a]Enzyme specificity predicted based on Perona and Craik, 1995. T: trypsin, C: chymotrypsin, E: elastase. Letters in parentheses: amino acid residues determining the primary specificity of a serine protease.

Table 4
Percentage identity (lower triangle) and similarity (upper triangle) of HPs with two clip domains

	PPAE	PAP-3	HP15	PAP-2	HP12
PPAE	—	64	64	64	63
PAP-3	51	—	60	65	59
HP15	48	48	—	72	66
PAP-2	50	52	58	—	73
HP12	47	46	54	63	—

Table 5
Percentage identity (lower triangle) and similarity (upper triangle) of HP2-related clip-domain serine proteinases

	HP18	HP2	HP13	HP21	HP22
HP18	—	51	49	55	51
HP2	34	—	68	66	67
HP13	33	52	—	68	65
HP21	37	49	52	—	75
HP22	37	50	47	62	—

clones. In contrast, we isolated one cDNA clone each for HP5, HP7, HP8, HP15 and HP22 and none for HP11.

3.2. Structural features

HP5 cDNA is 1398 nucleotides long and contains an incomplete open reading frame spanning nucleotides 1–1005 (Table 3). Its catalytic domain is most similar in sequence to single clip-domain serine proteinases (HP8, easter, PAP-1, PPAF1 and PPAF3) (Fig. 1). However, because the 5′ end of the cDNA is incomplete, we do not know whether HP5 contains a clip domain. HP5 has a typical proteolytic activation site (ESDR*IIGG) and is, therefore, likely activated by a trypsin-like serine proteinase. Cleavage at this site would yield an amino-terminal domain and a carboxyl-terminal serine proteinase domain connected by an interchain disulfide bond. Based on the residues determining the primary specificity pocket, we predict that HP5 has a trypsin-like specificity (Table 3).

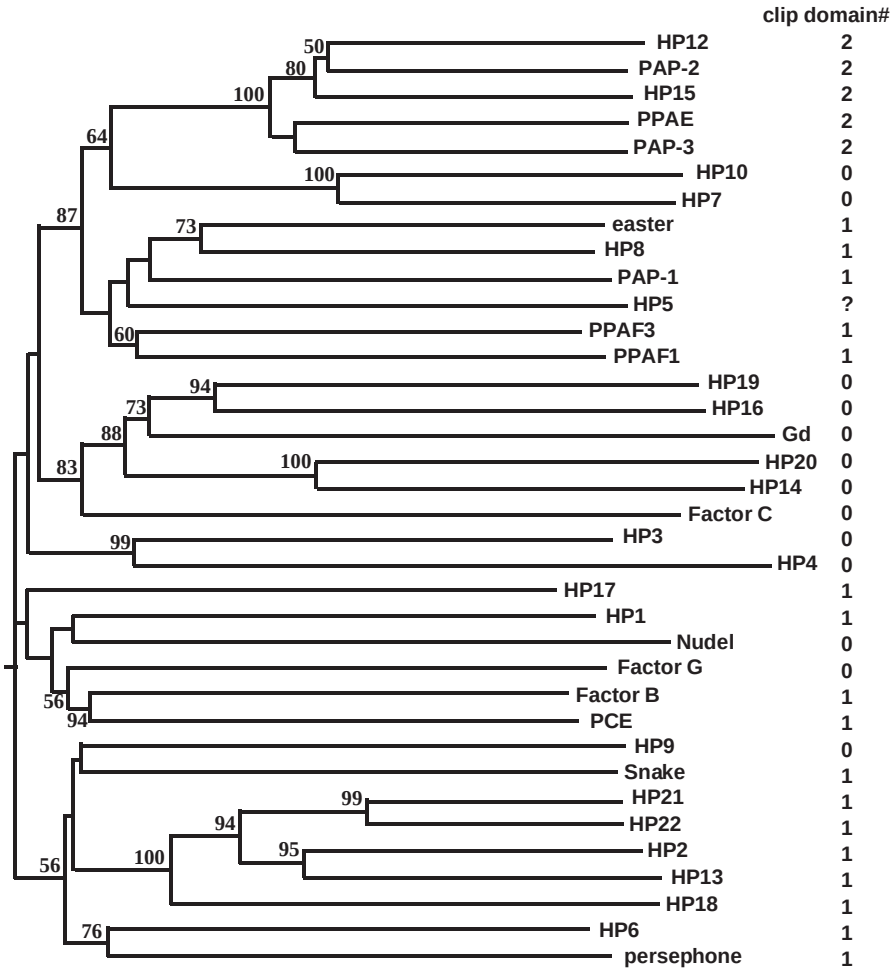

Fig. 1. Phylogenetic analysis of the catalytic domains of *M. sexta* HPs and other arthropod serine proteinases with known functions. To understand evolutionary relationships among arthropod serine proteinases, multiple sequence alignment and phylogenetic analysis were performed as described in Materials and methods. The phylogenetic tree was derived from an alignment that can be found at http://www.ento.okstate.edu/profiles/jiang.htm. Percent numbers at the nodes indicate bootstrap values (>500 of 1000 trials). Nudel, gastrulation defective (Gd), Snake and easter constitute a proteinase cascade that establishes the dorsoventral axis during *Drosophila* embryogenesis (Morisato and Anderson, 1995). Factors C, G, B, and proclotting enzyme (PCE) are members of the horseshoe crab hemolymph coagulation pathway (Iwanaga et al., 1998). *M. sexta* PAP-1, PAP-2, and PAP-3, *H. diomphalia* PPAF1 and PPAF3, as well as *B. mori* PPAE are involved in proPO activation. The numbers of clip domains in the proteinases are listed on the right. It is not yet known whether or not HP5 (marked "?") contains a clip domain.

HP6 and HP8 are most similar in their catalytic domain sequences to *Drosophila* persephone and easter, respectively (Fig. 1). HP6 has a predicted activation site of LDLH*ILGG, suggesting that its activating proteinase has an unusual specificity of cleaving after His (Table 3). This is also true of persephone whose predicted activation site is LVIH*IVGG. Like the two *Drosophila* homologs, HP6 and HP8 are expected to cleave after positively charged residues (e.g., Lys or Arg).

HP7 and HP10 are short serine proteinases about 250 residues long (Table 3). BLAST search indicated that their catalytic domain sequences are most similar to those of silkworm proPO activating enzyme (PPAE), *A. gambiae* 14D2 and other clip-domain proteinases (Fig. 1). Like HP3 and HP4 (Jiang et al., 1999), they are more similar in sequence to serine proteinases in human leucocytes than those in insect digestive tracks. They both contain a typical signal peptide at the amino terminus, whereas their predicted propeptides end with an unusual Cys residue. HP7 and HP10 do not contain any predicted sites for *N*- or *O*-linked glycosylation.

HP9 is most similar in its catalytic domain sequence with *Drosophila* snake and the other single clip-domain serine proteinases, but it does not contain a clip domain at the amino-terminus (Fig. 1). From its putative proteolytic activation site (KPSF*AIGG), we predict that HP9 is activated by a chymotrypsin-like enzyme. Since HP9 has a trypsin-like S1 specificity pocket (Table 3), it may link a chymotrypsin-like activator to a protein substrate activated by a trypsin-like enzyme in a proteinase pathway.

We identified four dual clip-domain proteinases in our collection of *M. sexta* HPs, namely HP12, HP15, PAP-2

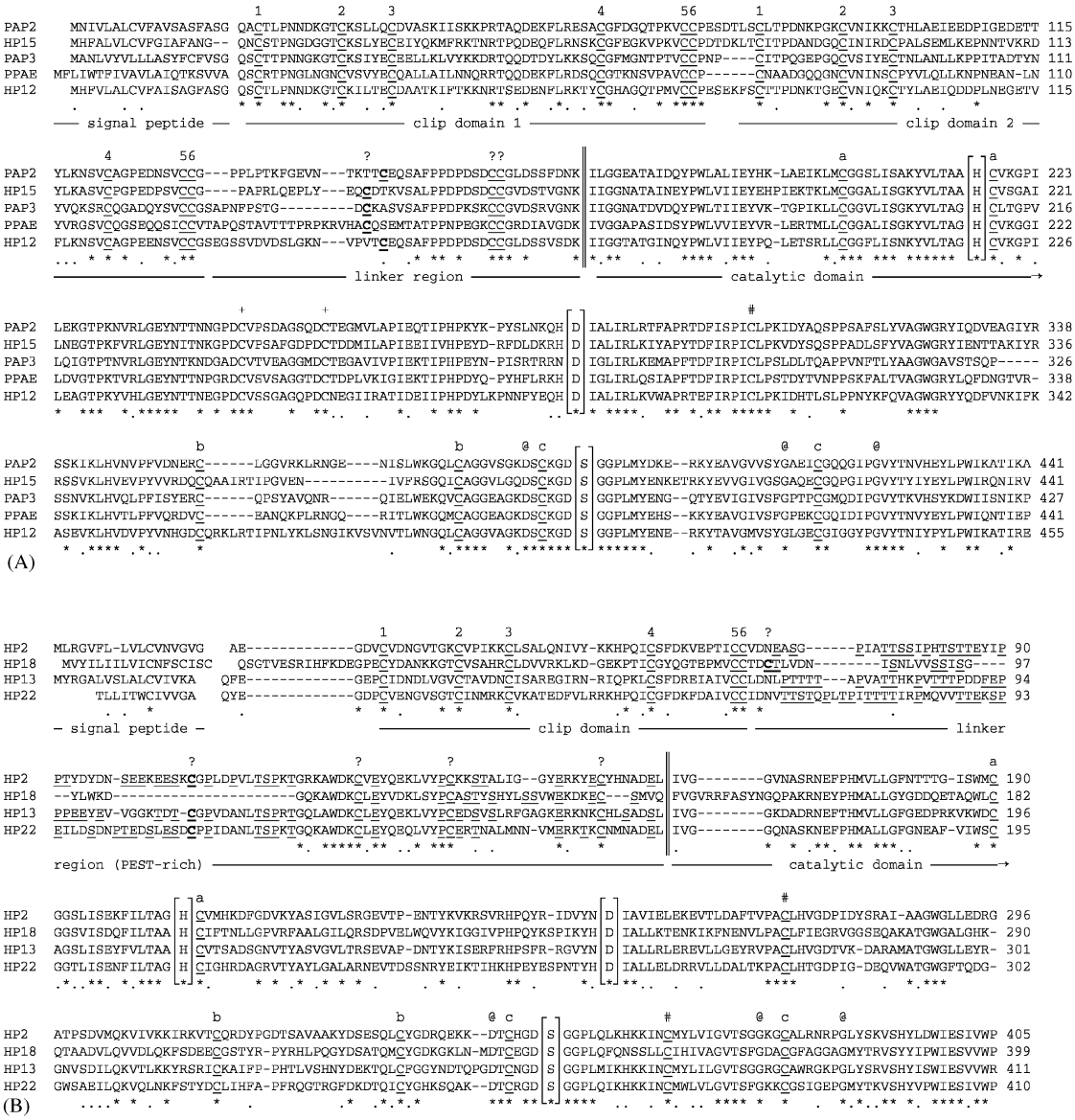

Fig. 2. Multiple sequence alignment of clip-domain serine proteinases related to PAP-2 (A) and HP2 (B). Completely conserved amino acid residues are indicated by "*", and conservative substitutions by "." underneath the sequences. Cys residues in the mature proteins are underlined, and the absolutely conserved ones in each clip domain are numbered 1–6. They form three disulfide bonds (1–5, 2–4, and 3–6) in horseshoe crab proclotting enzyme. The (predicted) cleavage activation site is marked by "‖". The residues of the catalytic triad (His, Asp, and Ser) are marked by boxes, and the important determinants of the specificity pocket by "@" on top of the sequences. The paired letters (a–a, b–b, c–c) indicate the disulfide linkage in the catalytic domain that are conserved in all S1 serine proteinases. The two unique Cys residues in most group-2 proteinases are shown by "+" (Jiang and Kanost, 2000). Cysteines possibly involved in interdomain disulfide bonds are marked with "#", and their partners in the amino-terminal regulatory domain are indicated by "?" since the linkage pattern is unknown. Pro, Glu, Ser and Thr residues in the linker regions of HP2, HP13, HP18 and HP22 are underlined.

and PAP-3 (Table 4). With highly similar sequences, these enzymes and silkworm PPAE fall on the same branch of the phylogenetic tree (Fig. 1). Although *Drosophila* SP18 has two clip domains, it differs significantly from this group of lepidopteran protein as the 71-residue linker between its clip domains is much longer than those (1–7 residues long) in HP12, HP15, PAP-2, PAP-3 and PPAE (Ross et al., 2003). An expansion of dual clip-domain serine proteinase genes may have occurred during the evolution of lepidopteran insects. Multiple sequence alignment and phylogenetic analysis indicated that *M. sexta* PAP-3 is more closely related to *B. mori* PPAE than are PAP-2, HP12, and HP15 (Figs. 1 and 2A). The chromatographic behaviors of PAP-3 and PPAE on ion exchange columns are also similar (Jiang et al., 2003b). The amino acid sequence identity and similarity between PAP-2 and HP12 are 63% and 73%, respectively (Table 4). They probably

arose from a relatively recent gene duplication and subsequent sequence divergence. Their common ancestor diverged from HP15 gene earlier—HP15 is 54% and 58% identical in amino acid sequence to HP12 and PAP-2, respectively. PAP-3 is less similar to the other three *M. sexta* dual clip-domain serine proteinases (identity: 46–52%). Nevertheless, the number and position of cysteine residues are absolutely conserved in all of these dual clip-domain serine proteinases, suggesting that their disulfide linkage patterns remain the same despite the sequence divergence (Fig. 2A) (Jiang et al., 2003b). PAP-2, PAP-3, PPAE, HP12, HP15 and their activating enzymes are anticipated to have trypsin-like specificity (Table 3).

HP2, HP13, HP18, HP21 and HP22 are closely related in sequence and domain structure, and have a conserved pattern of Cys residues that differs from other clip-domain proteinases (Fig. 2B). These single clip-domain serine proteinases constitute a small branch in the phylogenetic tree (Fig. 1). HP21 and HP22 genes probably evolved from recent gene duplication since their sequences are most similar to each other (identity: 62%; similarity: 75%) (Table 5). It is possible that HP2 and HP13 (identity: 52%; similarity: 68%) evolved by the same mechanism at an earlier time. The linker sequences in HP2, HP13, HP21 and HP22 are rich in Pro, Glu, Thr, and Ser residues, and may be heavily glycosylated (Table 3). The predicted activation sites in these four HPs are located between Leu and Ile residues. After proteolytic activation, these five serine proteinases are anticipated to cleave after Arg or Lys.

HP14 appears to be the most interesting protein in this group of serine proteinases. Its amino terminus contains a battery of Cys-rich regulatory domains, including five low-density lipoprotein receptor class A repeats, one Sushi domain, and one 7C region that links to the proteinase catalytic domain in the carboxyl terminus. The recombinant HP14 precursor interacts with peptidoglycan, autoactivates, and initiates the proPO activation cascade (Ji et al., 2004).

The 1.5 kb HP16 cDNA includes an open reading frame ranging from nucleotide 62 to 1396 (Table 3). BLASTP search indicated that the amino-terminal 170 residues of the protein had no significant sequence similarity with proteins in the public databases. Its catalytic domain is most similar to *M. sexta* HP19 as well as *D. melanogaster* gastrulation defective (Fig. 1). The latter is the second component of the serine proteinase cascade that establishes the dorsoventral axis of *Drosophila* embryo (LeMosy et al., 2001). HP16 has a putative proteolytic activation site (HTGL*IVNG), suggestive of a chymotrypsin/elastase-like activating proteinase. Once activated, HP16 may cleave after small- to medium-sized hydrophobic residues (Table 3).

Different from all the other *M. sexta* HPs, HP20 contains a carboxyl-terminal extension of ~50 residues after the catalytic domain. There are four cysteine residues and many hydrophilic/charged residues in this region. While carboxyl-terminal extensions are present in *Drosophila* Nudel, SP18, SP33 and SP36 (Ross et al., 2003), we did not identify any sequence similarity with HP20 among these extended regions. HP20 contains a 15-residue predicted pro-region not expected to form a domain with regulatory function. Its removal may involve a chymotrypsin-like enzyme (Table 3).

3.3. Expression analysis

To investigate the transcriptional regulation of HP genes upon microbial infection, we first examined the mRNA levels of HP12, HP18 and HP22 by Northern blot analysis (Fig. 3, panels A–C). HP12 transcripts (1.4 kb) were detected at a low level in fat body from larvae injected with saline, and after a bacterial challenge, the mRNA level strongly increased. We did not detect HP12 or HP22 transcripts in the control or induced hemocytes. HP22 transcripts were only present in induced fat body: the 3 kb minor band is the size expected for the cloned cDNA with a 1.5 kb 3′ untranslated region. The 1.5 kb major transcript has not been cloned but might represent an alternate polyadenylation site. HP18 expression was detected only in induced hemocytes, indicating that it is also an acute-phase gene but with differing tissue specificity. Because the Northern blot results for HP12, HP18, and HP22 were quite consistent with the results from RT-PCR analysis (Fig. 3, panels D–F), we then performed RT-PCR to determine relative mRNA levels for the other HPs.

RT-PCR analyses showed basal levels of HP1–HP7, HP9–HP11, HP14, HP16 and HP19 transcripts in naïve hemocytes and fat body (Fig. 3, panel G). After a bacterial challenge, HP2, HP5, HP7, HP9, HP10, HP14, HP16 and HP19 transcripts increased, whereas HP3, HP4 and HP11 mRNAs decreased in both tissues. HP6 mRNA levels remained unchanged in hemocytes and fat body. Other bacteria-inducible messages include HP13, HP15, HP17, HP20 and HP21. HP13 and HP15 were specifically expressed in hemocytes, while HP20 and HP21 mRNAs were detected in fat body only. Constitutive expression of HP17 was not observed in hemocytes or fat body. HP17 was produced in both tissues after the microbial injection. In contrast, the level of HP8 mRNA in fat body decreased upon immune challenge.

HP6 and HP8 antibodies each recognized a 37 kDa band (Fig. 4), similar in size to the calculated M_r's of the mature proenzymes (Table 3). While intensities of the immunoreactive bands did not change much after a microbial challenge, HP8 antibody also weakly recognized a 39 kDa band in the induced hemolymph.

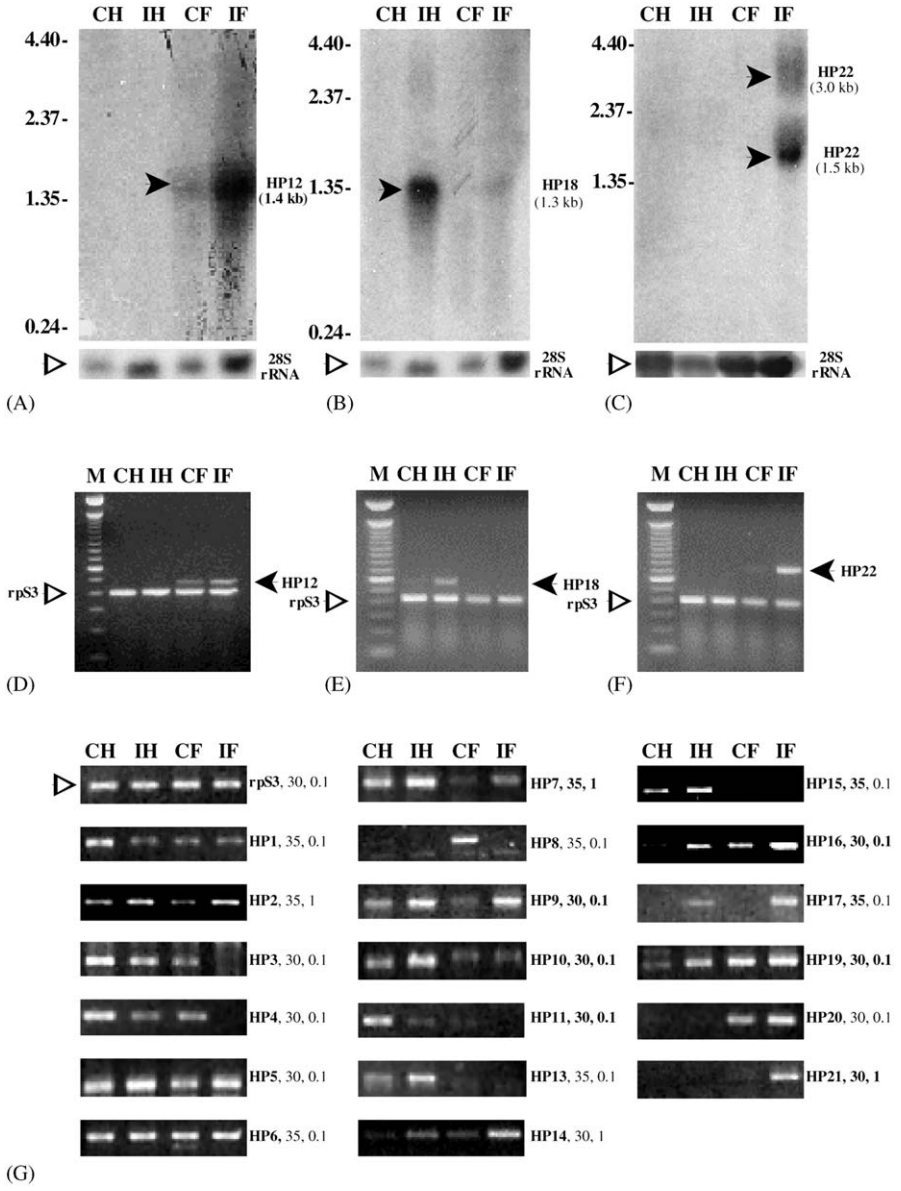

Fig. 3. Northern blot and RT-PCR analyses of *M. sexta* HP transcripts. Panels A–C: total RNA samples (20 μg) from hemocytes (H) or fat body (F) collected 24 h after injection of saline (C, control) or *E. coli* (I, induced) were separated by 1% agarose gel electrophoresis as described in *Materials and methods*. Northern blot analysis was performed using ^{32}P-labeled *M. sexta* HP12 (A), HP18 (B), and HP22 (C) cDNA probes. The positions of RNA size standards are shown on the left, and hybridization signals are marked with arrows. The band of 28S ribosomal RNA, indicated by an open arrowhead, was used as a loading control. Panels D–F: cDNA samples from control and induced hemocytes (CH, IH) and fat body (CF, IF) were normalized with *M. sexta* ribosomal protein S3 (rpS3, open arrowhead). The duplex PCR was performed using primers specific for rpS3 and HP12 (D)/HP18 (E)/HP22 (F). The expected sizes for HP PCR products are indicated by arrows. M, 100 bp DNA size markers. Panel G: After method validation, the normalized cDNA samples were diluted and analyzed by PCRs using gene-specific primers. The product name, cycle number, and amount of templates used are marked on the right.

Consistent with their transcriptional up-regulation (Fig. 3), there were elevations in HP9 and HP12 protein levels. Both antibodies recognized a broad band at 45–49 kDa, suggesting that HP9 is posttranslationally modified—the calculated M_r's of proHP9 and proHP12 are 42 and 48 kDa, respectively (Table 3). A 31 kDa immunoreactive band is close to the theoretical M_r of proHP10 (28 kDa), and its low intensity remains unchanged after microbial infection (Fig. 4). HP13 antibody did not detect any protein in the plasma samples (data not shown).

Due to the interference from hemolymph storage proteins, we were unable to accurately measure the apparent mobility of HP15, HP17, HP18, HP19, and HP22 precursors. However, we observed an increase in their protein levels after the immune challenge (Fig. 4).

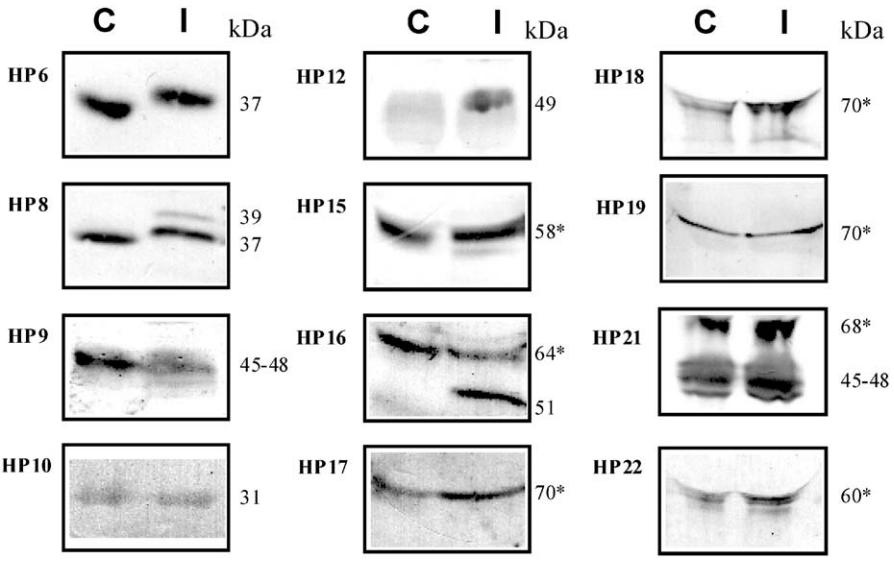

Fig. 4. Examination of *M. sexta* HP protein levels in plasma by immunoblot analysis. Hemolymph was collected from *M. sexta* larvae at 24 h after injection with H_2O or *E. coli*. The plasma samples (1.0 μl/lane) were subjected to 10% SDS-PAGE and immunoblot analysis using 1:2000 diluted HP antiserum as the first antibody. Left lane, control (C) plasma; right lane, induced (I) plasma. The estimated sizes of immunoreactive bands are indicated in kiloDalton. "*" denotes the band whose mobility was severely affected by the storage proteins in the plasma.

HP16 antibody recognized two bands at about 64 and 51 kDa. The 51 kDa one was only detected in induced hemolymph. With a calculated M_r of 47 kDa, HP16 precursor is probably a glycoprotein. HP21 antibody detected multiple bands around 68 and 45–48 kDa. Intensities of these bands increased after the immune challenge.

4. Discussion

For the past 10 years, we have investigated the biochemical roles of serine proteinases in prophenoloxidase (proPO) activation, a component of the host defense system in *M. sexta*. The purification and characterization of PAP-1, PAP-2 and PAP-3 were carried out with an approach which utilized purified proPO and a large amount of starting materials (Jiang et al., 1998, 2003a, b). However, low abundance, instability, stickiness, cofactor requirement, and lack of specific detection methods made it quite difficult to isolate other HPs from the hemolymph or cuticular extract. Our initial PCR experiments yielded cDNA clones for four additional serine proteinases (Jiang et al., 1999). Considering that the *M. sexta* genome sequence will not be available in the near future, we have continued this approach and isolated 18 new serine proteinase cDNAs from fat body and hemocytes. Identification of these HPs is a key step toward elucidating the pathways and functions of serine proteinases in this surprisingly complex system, which serves as a model for understanding the roles of proteinases in insect immune responses. Acquisition of specific cDNA probes, sequence information, and polyclonal antibodies has laid a foundation for our future biochemical investigation of *M. sexta* HPs.

Serine proteinases have been identified by a similar approach from several other insect species including *D. melanogaster*, *A. gambiae*, and *C. felis* (Coustau et al., 1996; Gorman and Pakewitz, 2001; Gaines et al., 1999). There are good indications that many such proteins are involved in immune responses rather than functioning as digestive enzymes. Recently, genome analysis revealed that serine proteinases and their homologs constitute one of the largest protein families in *D. melanogaster* and *A. gambiae* (Adams et al., 2000; Christophides et al., 2002). Transcriptome analysis further demonstrated changes in some of their mRNA levels in response to microbial infection (Irving et al., 2001; De Gregorio et al., 2001). Biochemical analysis of HP systems provides a valuable alternative that complements the genetic and genomic analyses. For instance, we hypothesized that HP14, with a large size and complex domain structure, could be the first enzyme in a serine proteinase pathway. Then, we confirmed that recombinant proHP14 indeed autoactivates in the presence of peptidoglycan and triggers the proPO activation pathway (Ji et al., 2004). It would be interesting to test whether proPO activation is also affected by *Drosophila* AY118964 or *Anopheles gambiae* CP12488, orthologs of HP14.

The clip domain is a structural module commonly existing in arthropod serine proteinases and serine

proteinase homologs that lack proteolytic activity (Jiang and Kanost, 2000). About 40 serine proteinase-related proteins from *D. melanogaster* contain at least one clip domain (Ross et al., 2003). Such domains are present in 14 of the 25 *M. sexta* HPs (Table 3), four of which have two clip domains at their amino termini. Although the prototype of clip domains was identified some time ago in the horseshoe crab proclotting enzyme (Muta et al., 1990), the structures and functions of clip domains are still largely unknown. We proposed earlier that clip domains could be the site for interactions of a serine proteinase with its upstream activating enzyme, its downstream protein substrate, or a cofactor that regulates the enzyme's activity (Jiang and Kanost, 2000). Now, having a sizable collection of single/dual clip domains will allow exploration of their structure–function relationships using proteins and domains isolated from a recombinant expression system or hemolymph of *M. sexta* larvae.

Immunoblot analyses, using antisera raised to the recombinant HPs made in *E. coli*, indicated that most of the HPs expressed in fat body or hemocytes can be detected in larval plasma. (The exception is HP13.) These polyclonal antibodies are proving to be useful tools for monitoring the purification of corresponding HPs without prior knowledge of their functions (data not shown). By detecting the proteolytic processing of HP zymogens, we can use these antibodies as reagents to study the activation process and assay for the upstream activating proteinases. Another application of these antibodies is to detect covalent complexes of HPs and serpins known to regulate proPO activation and other defense responses. Immunoblot and peptide mass fingerprint analyses identified HP1, HP6, HP8, HP21, PAP-3, serpin-4, serpin-5, and serpin-6 in high M_r complexes formed during immune responses (Tong et al., 2005; Zou and Jiang, 2005).

Sequence, size, inducibility, activation site, substrate specificity, and domain structures of the HPs provided useful clues in predicting organization of the serine proteinase network in *M. sexta* plasma. We successfully confirmed the prediction of HP14 being a first component of the proPO activation cascade (Ji et al., 2004). We hypothesize that H16 or HP19 could be the second component of this pathway, because their catalytic domains are similar in sequence to that of *Drosophila* gastrulation defective (Fig. 1) and because their predicted activation sites fit well with the specificity of HP14 (Table 3). The identification of these HPs has also led to development of a model for proPO activation pathways proposed based on our study of serpin–HP interactions (Tong et al., 2005). This sizable collection of information and molecular probes, along with the availability of large amounts of plasma for biochemical analysis provides a unique opportunity to formulate hypotheses and examine the roles of HPs in insect defense responses.

Acknowledgments

This work was supported by National Institutes of Health Grants GM58634 (to H. Jiang) and GM41247 (to M. Kanost). This article was approved for publication by the Director of the Oklahoma Agricultural Experiment Station and supported in part under project OKLO2450. We thank Drs. Dillwith and Essenberg for their helpful discussion and comments on the manuscript.

References

Adams, M.D., Celniker, S.E., Holt, R.A., Evans, C.A., Gocayne, J.D., et al., 2000. The genome sequence of *Drosophila melanogaster*. Science 287, 2185–2195.

Altschul, S.F., Madden, T., Schaffer, A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J., 1997. Gapped BLAST and PSI BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25, 3389–3402.

Ashida, M., Brey, P.T., 1998. Recent advances on the research of the insect prophenoloxidase cascade. In: Brey, P.T., Hultmark, D. (Eds.), Molecular Mechanisms of Immune Responses in Insects. Chapman & Hall, London, pp. 135–172.

Christophides, G.K., Zdobnov, E., Barillas-Mury, C., Birney, E., Blandin, S., et al., 2002. Immunity-related genes and gene families in *Anopheles gambiae*. Science 298, 159–165.

Clark, K.D., Witherell, A., Strand, M.R., 1998. Plasmatocyte spreading peptide is encoded by an mRNA differentially expressed in tissues of the moth *Pseudoplusia includens*. Biochem. Biophys. Res. Commun. 250, 479–485.

Coustau, C., Rocheleau, T., Carton, Y., Nappi, A.J., ffrench-Constant, R.H., 1996. Induction of a putative serine protease transcript in immune challenged *Drosophila*. Dev. Comp. Immunol. 20, 265–272.

Cerenius, L., Söderhäll, K., 2004. The prophenoloxidase-activating system in invertebrates. Immunol. Rev. 198, 116–126.

De Gregorio, E., Spellman, P.T., Rubin, G.M., Lemaitre, B., 2001. Genome-wide analysis of the *Drosophila* immune response by using oligonucleotide microarrays. Proc. Natl. Acad. Sci. USA 98, 12590–12595.

Gaines, P.J., Sampson, C.M., Rushlow, K.E., Stiegler, G.L., 1999. Cloning of a family of serine protease genes from the cat flea *Ctenocephalides felis*. Insect Mol. Biol. 8, 11–22.

Gettins, P.G.W., 2002. Serpin structure, mechanism, and function. Chem. Rev. 102, 4751–4803.

Gorman, M.J., Pakewitz, S.M., 2001. Serine proteases as mediators of mosquito immune responses. Insect Biochem. Mol. Biol. 31, 257–262.

Hayakawa, Y., Ohnishi, A., Yamanaka, A., Izumi, S., Tomino, S., 1995. Molecular cloning and characterization of cDNA for insect biogenic peptide, growth blocking peptide. FEBS Lett. 376, 185–189.

Hoffmann, J.A., Reichhart, J.M., 2002. *Drosophila* innate immunity: an evolutionary perspective. Nature Immunol. 3, 121–126.

Irving, P., Troxler, L., Heuer, T.S., Belvin, M., Kopczynski, C., Reichhart, J.M., Hoffmann, J.A., Hetru, C., 2001. A genome-wide analysis of immune responses in *Drosophila*. Proc. Natl. Acad. Sci. USA 98, 15119–15124.

Iwanaga, S., Kawabata, S., Muta, T., 1998. New types of clotting factors and defense molecules found in horseshoe crab hemolymph: their structures and functions. J. Biochem. 123, 1–15.

Ji, C., Wang, Y., Guo, X., Hartson, S., Jiang, H., 2004. A pattern recognition serine proteinase triggers the prophenoloxidase

activation cascade in the tobacco hornworm, *Manduca sexta*. J. Biol. Chem. 279, 34101–34106.

Jiang, H., Kanost, M.R., 2000. The clip-domain family of serine proteinases in arthropods. Insect Biochem. Mol. Biol. 30, 95–105.

Jiang, H., Wang, Y., Kanost, M.R., 1998. Pro-phenol oxidase activating proteinase from an insect, *Manduca sexta*: a bacteria-inducible protein similar to *Drosophila easter*. Proc. Natl. Acad. Sci. USA 95, 12220–12225.

Jiang, H., Wang, Y., Kanost, M.R., 1999. Four serine proteinases expressed in *Manduca sexta* haemocytes. Insect Mol. Biol. 8, 39–53.

Jiang, H., Wang, Y., Yu, X.-Q., Kanost, M.R., 2003a. Prophenoloxidase-activating proteinase-2 (PAP-2) from hemolymph of *Manduca sexta*: a bacteria-inducible serine proteinase containing two clip domains. J. Biol. Chem. 278, 3552–3561.

Jiang, H., Wang, Y., Yu, X.-Q., Zhu, Y., Kanost, M.R., 2003b. Prophenoloxidase-activating proteinase-3 (PAP-3) from *Manduca sexta* hemolymph: a clip-domain serine proteinase regulated by serpin-1J and serine proteinase homologs. Insect Biochem. Mol. Biol. 33, 1049–1060.

Kanost, M.R., 1999. Serine proteinase inhibitors in arthropod immunity. Dev. Comp. Immunol. 23, 291–301.

Kanost, M.R., Jiang, H., Wang, Y., Yu, X., Ma, C., Zhu, Y., 2001. Hemolymph proteinases in immune responses of *Manduca sexta*. Adv. Exp. Med. Biol. 484, 319–328.

Kanost, M.R., Jiang, H., Yu, X., 2004. Innate immune responses of a lepidopteran insect, *Manduca sexta*. Immunol. Rev. 198, 97–105.

Lee, E., Linder, M., Gilman, A.G., 1994. Expression of G-protein alpha subunits in *Escherichia coli*. Methods Enzymol. 237, 146–163.

Lemaitre, B., Nicolas, E., Michaut, L., Reichhart, J.M., Hoffmann, J.A., 1996. The dorsoventral regulatory gene cassette spätzle/Toll/cactus controls the potent antifungal response in *Drosophila* adults. Cell 86, 973–983.

LeMosy, E.K., Tan, Y.Q., Hashimoto, C., 2001. Activation of a protease cascade involved in patterning the *Drosophila* embryo. Proc. Natl. Acad. Sci. USA 98, 5055–5060.

Levashina, E.A., Langley, E., Green, C., Gubb, D., Ashburner, M., Hoffmann, J.A., Reichhart, J.M., 1999. Constitutive activation of toll-mediated antifungal defense in serpin-deficient *Drosophila*. Science 285, 1917–1919.

Ligoxygakis, P., Pelte, N., Hoffmann, J.A., Reichhart, J.M., 2002a. Activation of *Drosophila* Toll during fungal infection by a blood serine protease. Science 297, 114–116.

Ligoxygakis, P., Pelte, N., Ji, C., Leclerc, V., Duvic, B., Belvin, M., Jiang, H., Hoffmann, J.A., Reichhart, J.M., 2002b. A serpin mutant links Toll activation to melanization in the host defence of *Drosophila*. EMBO J. 21, 6330–6337.

Morisato, D., Anderson, K.V., 1995. Signal pathways that establish the dorsal–ventral pattern of the *Drosophila* embryo. Ann. Rev. Genet. 29, 371–399.

Muta, T., Hashimoto, R., Miyata, T., Nishimura, H., Toh, Y., Iwanaga, S., 1990. Proclotting enzyme from horseshoe crab hemocytes: cDNA cloning, disulfide locations, and subcellular localization. J. Biol. Chem. 265, 22426–22433.

O'Brien, D., McVey, J., 1993. Blood coagulation, inflammation, and defense. In: Sim, E. (Ed.), The Natural Immune System, Humoral Factors. IRL Press, New York, pp. 257–280.

Perona, J.J., Craik, C.S., 1995. Structural basis of substrate specificity in the serine proteases. Protein Sci. 4, 337–360.

Ross, J., Jiang, H., Kanost, M.R., Wang, Y., 2003. Serine proteases and their homologs in the *Drosophila melanogaster* genome: an initial analysis of sequence conservation and phylogenetic relationship. Gene 304, 117–131.

Silverman, G.A., Bird, P.I., Carrell, R.W., Church, F.C., Coughlin, P.B., Gettins, P.G., Irving, J.A., Lomas, D.A., Luke, C.J., Moyer, R.W., Pemberton, P.A., Remold-O'Donnell, E., Salvesen, G.S., Travis, J., Whisstock, J.C., 2001. The serpins are an expanding superfamily of structurally similar but functionally diverse proteins. Evolution, mechanism of inhibition, novel functions, and a revised nomenclature. J. Biol. Chem. 276, 33293–33296.

Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. ClustalW: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22, 4673–4680.

Tong, Y., Jiang, H., Kanost, M.R., 2005. Identification of plasma proteases inhibited by *Manduca sexta* serpin-4 and serpin-5 and their association with components of the prophenoloxidase activation pathway. J. Biol. Chem. 280, 14932–14942.

Wang, Y., Jiang, H., Kanost, M.R., 1999. Biological activity of *Manduca sexta* paralytic and plasmatocyte spreading peptide and primary structure of its hemolymph precursor. Insect Biochem. Mol. Biol. 29, 1075–1086.

Whaley, K., Lemercier, C., 1993. The complement system. In: Sim, E. (Ed.), The Natural Immune System, Humoral Factors. IRL Press, New York, pp. 121–150.

Zou, Z., Jiang, H., 2005. *Manduca sexta* serpin-6 regulates immune serine proteinases PAP-3 and HP8: cDNA cloning, protein expression, inhibition kinetics, and function elucidation. J. Biol. Chem. 280, 14341–14348.