Published in final edited form as:

Insect Biochem Mol Biol. 2020 January ; 116: 103261. doi:10.1016/j.ibmb.2019.103261.

Inhibition of immune pathway-initiating hemolymph protease-14 by *Manduca sexta* serpin-12, a conserved mechanism for the regulation of melanization and Toll activation in insects

Yang Wang¹, Fan Yang¹, Xiaolong Cao¹, Rudan Huang¹, Susan Paskewitz², Steve D. Hartson³, Michael R. Kanost⁴, Haobo Jiang¹

¹Department of Entomology and Plant Pathology, Oklahoma State University, Stillwater, OK 74078, USA

²Department of Entomology, University of Wisconsin-Madison, Madison, WI 53706, USA

³Department of Biochemistry and Molecular Biology, Oklahoma State University, Stillwater, OK 74078, USA

⁴Department of Biochemistry and Molecular Biophysics, Kansas State University, Manhattan, KS 66506, USA

Abstract

A network of serine proteases (SPs) and their non-catalytic homologs (SPHs) activates prophenoloxidase (proPO), Toll pathway, and other insect immune responses. However, integration and conservation of the network and its control mechanisms have not yet been fully understood. Here we present evidence that these responses are initiated through a conserved serine protease and negatively regulated by serpins in two species, *Manduca sexta* and *Anopheles gambiae*. We have shown that *M. sexta* serpin-12 reduces the proteolytic activation of HP6, HP8, proPO activating proteases (PAPs), SPHs, and POs in larval hemolymph, and we hypothesized that these effects are due to the inhibition of the immune pathway-initiating protease HP14. To test whether these changes are due to HP14 inhibition, we isolated a covalent complex of HP14 with serpin-12 from plasma using polyclonal antibodies against the HP14 protease domain or against serpin-12, and confirmed formation of the complex by 2D-electrophoresis, immunoblotting, and mass spectrometry. Upon recognition of bacterial peptidoglycans or fungal β -1,3-glucan, the zymogen proHP14 became active HP14, which formed an SDS-stable complex with serpin-12 *in vitro*. Activation of proHP21 by HP14 was suppressed by serpin-12, consistent with the decrease in steps downstream of HP21, proteolytic activation of proPAP3, proSPH1/2 and proPO in hemolymph. Guided by the results of phylogenetic analysis, we cloned and expressed *A. gambiae* proSP217 (an ortholog of HP14) and core domains of *A. gambiae* serpin-11 and -17. The

Send correspondence to: Haobo Jiang, Department of Entomology and Plant Pathology, Oklahoma State University, Stillwater, OK, 74078, Telephone: (405)-744-9400, Fax: (405)-744-6039, haobo.jiang@okstate.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

The sequence of *Manduca sexta* serpin-12 has been submitted to the GenBank™/EBI Data bank with accession number [MG732913](#).

recombinant SP217 zymogen became active during expression, with cleavage between Tyr³⁹⁴ and Ile³⁹⁵. Both MsHP14 and AgSP217 cleaved MsSerp12 and AgSRPN11 at Leu*Ser (P1*P1') and formed complexes *in vitro*. ProPO activation in *M. sexta* plasma increased after recombinant AgSP217 had been added, indicating that it may function in a similar manner as the endogenous initiating protease HP14. Based on these data, we propose that inhibition of an initiating modular protease by a serpin may be a common mechanism in holometabolous insects to regulate proPO activation and other protease-induced immune responses.

Graphical Abstract

Keywords

insect immunity; hemolymph protein; serine protease cascade; melanization; serpin-protease complex

1. Introduction

Extracellular serine protease (SP) systems mediate rapid defense responses to tissue damage and pathogen attack in vertebrates and invertebrates (Krem and Di Cera, 2002; Kanost and Jiang, 2015). Upon recognition of aberrant or microbial surfaces, proteolytic processing sequentially activates serine proteases and their homologs to finally convert prophenoloxidases (proPOs) to POs in insects. POs catalyze the formation of reactive compounds and melanin to kill and sequester pathogens. Cleavage of pro-Spätzle (a cytokine precursor) at a specific site yields Spätzle that binds the Toll receptor to induce antimicrobial protein synthesis (Veillard et al., . Specific proteolysis generates another group of cytokines, known as stress responsive peptides, to induce some cellular immune responses (Schrag et al., 2017). Non-catalytic serine protease homologs (SPHs) are also activated by specific proteolysis, converting them to active regulators of melanization (Kwon et al., 2000; Yu et al., 2003). For instance, proPO activating proteases (PAPs) and a high M_r complex of SPH1 and SPH2 together generate active POs in the tobacco hornworm, *Manduca sexta* (Wang and Jiang, 2004)

The immune SP-SPH systems and their regulation by serpins are best studied in *Drosophila melanogaster*, *Tenebrio molitor* and *Manduca sexta* (Fig. 1) (Veillard et al., 2016; Park et al., 2010; Kanost and Jiang, 2015). Genetic analyses have revealed a serine protease cascade including the serine proteases ModSP, Grass, and SPE that induces the Toll pathway

(Buchon et al., 2009) in *D. melanogaster*. Proteases MP1, MP2 and Haysan participate in melanization but their connections to ModSP, Grass or SPE are unclear. Serpins 43Ac (*i.e.* Necrotic) and 28Da may inhibit related target SPs to affect Toll and proPO activation (Levashina et al., 1999; Scherfer et al., 2008). Biochemical studies in *T. molitor* unveiled a cascade pathway of MSP-SAE-SPE leading to proSPH1, proPOs and proSpätzle activation (Kim et al., 2008). *T. molitor* Spn40 and Spn93C form covalent complexes with MSP (Jiang et al., 2009 and 2011). In *M. sexta*, the pathway of proteases HP14-HP21-PAP2/3 generates active SPH1, SPH2 and POs (Wang and Jiang, 2007; Wang et al., 2014; Gorman et al., 2007). HP6 activates PAP1 and HP8. PAP1 activates proPO, whereas HP8 activates proSpätzle-1 (An et al., 2009 and 2010). It is unclear how proHP6 activation may be linked to HP14 or HP21. We recently reported that *M. sexta* serpin-12 down-regulates the proPO activation system (Yang et al., 2018). Serpin-12 has an amino-terminal extension rich in Gln, Ser, and Pro. The scissile bond in the reactive center loop of serpin-12 is after the P1 residue Leu⁴²⁹, and the serpin-12 core domain inhibits cathepsin G, chymotrypsin, and pancreatic elastase, which cleave after residues with hydrophobic side chains. A phylogenetic analysis of serpins from *Drosophila*, *Tenebrio*, *Manduca* and other insects and alignment of their reactive suite loop sequences indicate that *Drosophila* Spn28Da and Spn43Ac, *Tenebrio* Spn40 and Spn93C, and *Manduca* serpin-12 are related, but no clear orthologous relationship can be established with serpin-12. The detection of a 72 kDa band by antibodies to serpin-12 and the HP14 protease domain (HP14PD) suggests that a complex of HP14 and serpin-12 formed after an immune response was elicited in *M. sexta* larval hemolymph by exposure to peptidoglycan (Yang et al., 2018).

Our understanding of the insect SP-SPH-serpin system is still fragmentary. The protease and serpin network that functions in immune responses and its extent of conservation among major insect groups remain unclear, which is in sharp contrast with the extensive sequences and expression patterns of insect SPs, SPHs, and serpins (Cao et al., 2015 and 2017; Cao and Jiang, 2018; Reichhart, 2005; Zou et al., 2009; Suwanchaichinda and Kanost, 2009; Meekins et al., 2017; Li et al., 2018). In this study, we first examined the formation of a complex between HP14 and serpin-12 isolated from plasma after microbial elicitation of protease pathways. We then used purified recombinant proteins to test if the HP14-HP21-PAP2/3-PO pathway is regulated by serpin-12, which is known to decrease proPO activation in larval plasma (Yang et al., 2018). To find out if a similar reaction occurs in other insects, we cloned and expressed *A. gambiae* modular protease SP217 (orthologous to *M. sexta* HP14) and serpins SRPN11 N and SRPN17 N to examine their biochemical activities. Finally, we discuss the inhibition of HP14 and other homologous modular proteases by serpins as a conserved mechanism for modulating melanization, Toll signaling, and cytokine responses in holometabolous insects.

2. Materials and methods

2.1. Microbial elicitors, insect rearing, immune challenge, and hemolymph collection

Micrococcus luteus and curdland of *Alcaligenes faecalis* were purchased from Sigma-Aldrich. High purity Lys-peptidoglycan of *Staphylococcus aureus* and *m*-diaminopimelate-peptidoglycan of *Escherichia coli* K12 were ordered from InvivoGen. Human neutrophil

cathepsin G and porcine pancreatic elastase (PPE) were purchased from Athens Research & Technology. *M. sexta* eggs were ordered from Carolina Biological Supply and larvae were reared on an artificial diet (Dunn and Drake, 1983). Each fifth instar larva (day 2) was injected with a mixture of killed *E. coli* (2×10^7 cells), *M. luteus* (20 μ g), and curdlan (20 μ g) suspended in 30 μ l of 0.85% NaCl. At 24 h post injection, induced hemolymph was individually collected from a cut proleg into a microfuge tube. After hemocyte removal by centrifugation at $5,000 \times g$ for 15 min at 4 °C, the induced plasma (IP) samples were pooled and stored at -80 °C.

2.2 Preparation of proteins

M. sexta peptidoglycan recognition protein-1 (PGRP1) (Sumathipala and Jiang, 2010), microbe binding protein (MBP) (Wang et al., 2011), β -1,3-glucan recognition protein-2 (β GRP2) (Jiang et al., 2004), proHP21 (Wang and Jiang, 2007), and serpin-12 N (Yang et al., 2018) were produced in Sf9 cells infected with respective baculoviruses. Serpin-12 N was also expressed in *E. coli* as a soluble protein. Unless specified, the serpin-12 N from *E. coli* was used in the assays. *M. sexta* proHP14 and proPOs were purified from hemolymph as described in previous studies (Wang and Jiang, 2006; Jiang et al., 1997) and were stored at -80 °C for *in vitro* assays.

2.3. Immunoaffinity purification and identification of serpin-12-associated proteins from induced larval plasma

Coupling of the serpin-12 and HP14PD antibodies to protein A-Sepharose (GE Healthcare) was performed according to the manufacturer's instructions. To activate the plasma SP-SPH system and suppress protein crosslinking and melanization, 20 ml of induced plasma was incubated with 5.0 mg insoluble Lys-peptidoglycan of *S. aureus*, 10 mM diethylthiocarbonate and 1 mM 1-phenyl-2-thiourea for 30 min at room temperature (Yang et al., 2018). A separate control sample from the same batch of induced plasma was not treated with peptidoglycan but was incubated with diethylthiocarbonate and 1-phenyl-2-thiourea as described. Then, fresh phenylmethanesulfonyl fluoride at 1 mM final concentration and 0.5 ml of a cocktail of protease inhibitors (P-8849, Sigma-Aldrich) were added to the mixture to inactivate hemolymph proteases for 10 min prior to centrifugation at $5000 \times g$ at 4 °C for 15 min. Aliquots (10 ml) of the supernatant were separately mixed with beads conjugated with antibodies to serpin-12 or HP14PD (1.0 ml) at 4 °C for 8 h before loading into an empty Poly-Prep column (Bio-Rad). Unbound proteins were removed from the column as flow-through and washing fractions using 20 ml of 1 M NaCl and 20 ml of 10 mM sodium phosphate (pH 6.8). The bound proteins were eluted with 50 mM glycine-HCl (pH 2.5) and 0.5 ml fractions were collected into tubes containing 50 μ l of 1.0 M sodium phosphate (pH 8.0). Similarly, the induced plasma sample that was not activated by peptidoglycan, as a negative control, was loaded onto the two antibody columns and separated as described above. Fractions of the eluted proteins from both types of samples were subjected to 7.5% SDS-PAGE, Coomassie blue staining, and immunoblot analysis to identify peak fractions.

Three aliquots of the peak fraction (200 μ g protein each, *S. aureus* peptidoglycan-treated) from the serpin-12 antibody column were simultaneously separated under identical

conditions on three gels by two-dimensional gel electrophoresis (2DE) at Nevada Proteomics Center. Proteins on gel-1 and -2 were electrotransferred onto nitrocellulose membranes, imaged after Ponceau S staining, and subjected to immunoblot analysis using diluted serpin-12 or HP14PD antiserum (Yang et al., 2018; Wang and Jiang, 2010) as the first antibody. After gel-3 stained with SYPRO® Ruby was aligned with blot-1 and blot-2 via image superimposition, gel pieces corresponding to spots recognized by both antibodies were cut for in-gel trypsinolysis, peptide extraction, and LC-MS/MS analysis.

2.4. Formation of HP14-serpin-12 N complex in vitro and its effect on proHP21 activation

As controls, the purified serpin-12 N alone, curdlan-activated HP14 (a mixture of curdlan, β GBP2, and proHP14), and peptidoglycan-activated HP14 (a mixture of *E. coli* peptidoglycan, PGRP1, MBP, and proHP14) were incubated in buffer A (20 mM Tris-HCl, pH 7.5, 0.001% Tween-20, 5 mM CaCl₂) at 37 °C for 90 min. In tests, serpin-12 N was incubated with HP14 activated by curdlan or peptidoglycan under the same conditions. The mixtures were subjected to SDS-PAGE followed by immunoblot analysis using the diluted antisera to serpin-12 and HP14PD. To test whether formation of serpin-12-HP14 complex affects proHP21 cleavage activation, buffer A, peptidoglycan-activated HP14, and a mixture of HP14 and serpin-12 N were pre-incubated for 1 h at 37 °C. Aliquots of proHP21 were then added to the tubes and incubated for 1 h at 37 °C before treatment with 1×SDS sample buffer. The samples were separated by SDS-PAGE, transferred onto a nitrocellulose membrane, and detected using diluted HP21 antiserum as the primary antibody (Wang and Jiang, 2007).

2.5. Cloning and expression of *A. gambiae* SP217 in *E. coli* for antibody preparation

A fragment of SP217 was amplified from a cDNA pool of *A. gambiae* larvae using primers F6 (5'-GCGCGATCGTCGAAAAGTG) and R7 (5'-TTATATCACGGGATACCGAC) and cloned into pCR2.1 (Thermo Fisher). The cDNA was amplified again from the TOPO clone using j605 (5'-TGCGCCATGGGCGATCGTCGAA) and j611 (5'-CCTCAGTCGACCACGGGATACCGACT), digested with *Nco*I and *Hind*III, and inserted to the same sites in H6pQE60 (Lee et al., 1994). *E. coli* JM109 cells carrying the recombinant plasmid were induced with 1 mM IPTG for 5 h at 37 °C to produce proSP217 (Yang et al., 2018). The 75 kDa insoluble protein was purified from the bacteria by nickel affinity chromatography under denaturing condition and preparative SDS-PAGE (Wang and Jiang, 2010). The band was cut from the gel and used as an antigen for production of a rabbit polyclonal antiserum against *A. gambiae* SP217 at Cocalico Biologicals.

2.6. Expression of *A. gambiae* proSP217 in insect cells

To produce native proSP217 for functional tests, its 5' and 3' cDNA ends were amplified from *A. gambiae* larvae using primers j626 (5'-TGAATTCACAGGCGCGATCGT) and j608 (5'-CCAGTATGGTGCCACTGTTGC) and primers j625 (5'-GGTGTACTACCTGTACGGGCT) and j611, respectively. The PCR products were cloned into pGem-T and completely sequenced to ensure no missense mutation. The validated plasmids were digested with *Eco*RI and *Sph*I (5' end in pGem-T), *Sph*I and *Sac*II (central part in pCR2.1), *Sac*II and *Sal*I (3' end in pGem-T), and *Eco*RI and *Xho*I (pMFH6 vector). The four fragments at the expected sizes were cut from the gel after electrophoresis,

recovered, and ligated, as *Sa*I and *Xho*I are compatible. The recombinant plasmid proSP217/pMFH6 was used to generate a bacmid and high-titer baculovirus for expression in a suspension culture (Sumathipala and Jiang, 2010). According to the design, the mature proSP217 is GIHR⁵RDRR... ESRYPV⁶⁴⁴VEHHHHHH, where the underlined region is encoded by cDNA. The recombinant protein was purified from 300 ml of the culture medium by cation exchange and Ni²⁺ affinity chromatography (Sumathipala and Jiang, 2010) and concentrated using Centricon-30 (Amicon). Following buffer exchange to 20 mM Tris-HCl, pH 7.5, 50 mM NaCl, the protein was stored at -80 °C. An aliquot of the induced plasma was incubated with active SP217 to test if the mosquito protease could elicit melanization in *M. sexta*. IP alone was used as a negative control.

2.7. Determination of the cleavage activation site in *A. gambiae* SP217 by LC-MS/MS mass analysis and Edman degradation.

The purified SP217 was denatured and reduced for 30 min in 8 M urea, 100 mM Tris-HCl, pH 8.5, 10 mM tris(2-carboxyethyl)phosphine, alkylated for 20 min with 10 mM iodoacetamide, and digested overnight at 37 °C with 4 µg/ml Glu-C endopeptidase (Worthington). The desalted peptides were analyzed by LC-MS/MS on an Orbitrap Fusion mass spectrometer (Thermo) in the DNA/Protein Resource Facility at Oklahoma State University. The spectrometer was programmed to perform “top speed” (5 sec) data-dependent MS/MS fragmentations, wherein precursor ions were analyzed in the Orbitrap sector and fragment ions in the dual-stage ion trap sector. Precursor ions were selected by data-dependent quadrupole filtering and two MS/MS spectra were collected each scan cycle, with the first scan utilizing HCD activation (30% energy) in the ion routing multipole and the second utilizing CID activation (35% energy) in the ion trap sector. Peptide-spectrum matches were generated by utilizing MaxQuant v1.5.3.8 to search a database comprised of 13,515 *A. gambiae* sequences downloaded from UniProt on 11-21-2018 and supplemented with (E)TLSCVDGSDGVPVFRCEPVCCTPTDPAEAY, (D)GPVFRCEPVCCTPTDPAEAY, (E)PVCCTPTDPAEAY and IIGGRNVISIAE, putative peptides from SP217 that might be obtained after proteolysis in the reaction cleavage step, for protein identification. The residues in bold font are Glu-C sites, and the underlined residues are adjacent to the predicted activation site. To verify results of the mass analysis, the SP217 sample was separated by SDS-PAGE and transferred onto a PVDF membrane. After Ponceau S staining, a 30 kDa major band expected to be the catalytic domain was sequenced by automated Edman degradation at UC Davis Proteomics Core.

2.8. cDNA cloning and expression of *A. gambiae* SRPN11 N and SRPN17 N as soluble proteins in *E. coli*

To express the core domain of *A. gambiae* serpin-11, its cDNA was amplified from larvae using primers j1486 (5'-GCCATGGACTTTGCGGTCAAAC) and j1487 (5'-GCATGCTTAACCTCGTCTGAATGGATCT). The PCR product was cloned into pGem-T and five plasmids were completely sequenced. The cDNA with the fewest (likely allelic) differences in sequence from a previously sequenced SRPN11 cDNA clone (Suwanchaichinda and Kanost, 2009) was inserted into H6pQE60 by *Nco*I-*Sph*I cloning. Similarly, the *A. gambiae* serpin-17 cDNA was amplified using primers j1484 (5'-GCCATGGAGTTTCGCCTGGAACATG) and j1485 (5'-

CAAGCTTAGTTCAAAGTCGGGTTCAC). The product was inserted to pGem-T and six plasmids were completely sequenced. The *NcoI*-*HindIII* fragment with a single nonsynonymous substitution was subcloned into H6pQE60. *E. coli* JM109 cells carrying the recombinant plasmid were grown at 37 °C in 100 ml 2×YT medium with 100 µg/ml ampicillin. When OD₆₀₀ of the culture reached 0.7, IPTG at a final concentration of 0.1 mM was used to induce synthesis of the soluble SRPN11 N at 16°C for 18 h. The expression of SRPN17 N was performed under same conditions except that induction was at 37 °C for 4–6 h. Affinity purification of the serpins was performed as described previously (Yang et al., 2018).

2.9. Determination of cleavage sites in the serpins by mass spectrometry

The C-terminal peptide released in a serpin-protease reaction was examined by MALDI-TOF spectrometry in the linear mode (Jiang et al., 2003). The controls were HP14 (0.1 µg), SP217 (0.05 µg), cathepsin G (0.5 µg), PPE (0.35 µg), and 1.0 µg each of serpin-12 N, SRPN11 N or SRPN17 N together in a total volume of 20 µl buffer A. HP14 was generated in a mixture of *E. coli* peptidoglycan (1 µg), PGRP1 (0.5 µg), MBP (0.5 µg) and proHP14 (0.1 µg) incubated for 60 min at room temperature. The test reactions contained the same amounts of protease and serpin incubated for 30 min at room temperature in 20 µl buffer A. These reactions included HP14, SP217, cathepsin or PPE with each of the three serpins. The molecular mass of a peak in a test reaction but not present in its controls was compared with calculated masses of the serpin's C-terminal peptides to detect a match within the error range (500 ppm) of the method. Such a match indicated a cleavage occurred at the preceding peptide bond to release the observed peptide.

2.10. Formation of the complexes of *M. sexta* HP14 or *A. gambiae* SP217 with serpin-12 N, SRPN11 N and SRPN17 N

To test complex formation with the mosquito serpins, HP14 precursor was incubated with *E. coli* peptidoglycan, PGRP1 and MBP for activation in the presence of SRPN11 N or SRPN17 N. The control and test samples were subjected to SDS-PAGE and immunoblot analysis using 1:1000 diluted HP14PD antiserum as the primary antibody (Wang and Jiang, 2010). In the test of SP217 inhibition, the recombinant protease (50 ng) was separately incubated for 15 min at room temperature with 1 µg of serpin-12 N, SRPN11 N, or SRPN17 N. The reactions and controls were resolved by 10% SDS-PAGE and analyzed by immunoblotting using 1:1000 diluted anti-pentahistidine monoclonal antibody (Bio-Rad) as the first antibody.

3. Results and discussion

3.1. Isolation and characterization of hemolymph protease-14 associated with serpin-12

Our previous experiment indicated that a 71 kDa band in *M. sexta* larval plasma recognized by the serpin-12 and HP14 antibodies may represent a complex of HP14 protease domain (PD) and serpin-12, formed in response to added *S. aureus* peptidoglycan (Yang et al., 2018), but we considered the evidence insufficient for making an unequivocal conclusion. Most serpin-protease complexes migrate to 70–75 kDa positions during SDS-PAGE under reducing condition due to typical sizes of the serpin fragments (*c.a.* 40 kDa) and SP catalytic

domains (*c.a.* 30 kDa). In other words, HP14PD complexed with another serpin could have a mobility close to that of serpin-12 complexed some other SPs in the plasma. In addition, abundant storage proteins in hemolymph can impact apparent M_r 's of the complexes and lead to erroneous size estimation (Yang et al., 2018). Therefore, we carried out a more thorough characterization of the band by immunoaffinity chromatography using antibodies to HP14PD or serpin-12 to isolate proteins from induced plasma (IP) that had been activated by exposure to *S. aureus* peptidoglycan. Proteins eluted from the serpin-12 antibody column included the 52 kDa serpin-12, its cleavage products, a 71 kDa band, and a fainter band at about 80 kDa. The latter two bands were recognized by antibodies to serpin-12 and HP14PD (Fig. 2A, lane 2) but absent in the plasma not treated with the peptidoglycan (lane 1). The elution fraction from the HP14PD antibody column contained the 71 and 80 kDa bands (Fig. 2B, lane 2). Their absence in the control (untreated IP, lane 1) suggested that the added peptidoglycan had led to their formation. We also tested the elution fraction from the serpin-12 antibody column by immunoblot analyses using polyclonal antibodies against proteases HP2, HP9, HP13, HP16, HP19, HP21, HP22 and HP25 but did not detect any band at 70–80 kDa (data not shown). Neither was there any band in the same range recognized by antibodies against *M. sexta* serpin-1, 3–6 or 13 in the fraction of peptidoglycan-treated plasma eluted from the HP14PD antibody column (data not shown). Together, these data support the hypothesis that serpin-12 inhibits HP14 and rule out some other HPs or serpins as partners of serpin-12 or HP14PD, respectively. The isolation of the complexes using antibodies to both the protease and serpin further strengthen the conclusion that these bands contain complexes of HP14 and serpin-12.

3.2. 2DE separation and LC-MS/MS analysis of the serpin-12-complexed protein

For validating the HP14-serpin-12 complexes, aliquots of the proteins eluted from the serpin-12 antibody column were resolved by 2DE under identical conditions. After electrotransfer onto two membranes and staining with Ponceau S, immunoblot analysis was performed using antibodies against HP14PD and serpin-12 (Fig. 3, A and B). A majority of the serpin-12 signal was detected at 52, 50, 48, 40, and 38 kDa positions, consistent with the 1DE data (Fig. 2A). Interestingly, instead of being electrofocused into a single spot at a specific size, a series of spots with different pI's were revealed using antibody to serpin-12 (Fig. 3B). While proteolysis and glycosylation may have caused the size variations, charge differences are responsible for the pI changes at a particular M_r value. Superimposition with the SYPRO® Ruby stained gel (Fig. 3, C and D) indicated, for instance, the 52 kDa serpin was separated into five major spots with estimated pI's of 6.5, 6.8, 7.2, 8.0, and 8.8. The theoretical pI (8.51) of the mature serpin-12 is closest to 8.8, where the largest spot is located. Since the amino-terminal extension of *M. sexta* serpin-12 is rich in Gln and Asn (Yang et al., 2018), we suggest that the more acidic isoforms were products of deamination in some of these residues. Mass spectrometric analysis indicated that deamination occurred in the serpin domain as well (Table 1). Elevated pH and temperature during isoelectric focusing may have increased the reaction rate and caused steaking in pH 8.0–8.8 (Robinson and Robinson, 2001). While we detected the complexes at 71 kDa (spot-2) and around 80 kDa (spot-1) by the serpin antibody (Fig. 3B), signals from the HP14PD antibody were visible but much weaker in the same locations (Fig. 3A).

Through image superimposition (Fig. 3D), the gel pieces containing spot-1 and spot-2 were excised, reduced, alkylated, and digested with trypsin prior to LC-MS/MS analysis. Twelve unique peptides were identified in spot-1, eight corresponding to the N-terminal large fragment of serpin-12 and four to the HP14 catalytic domain (Table 1). In spot-2 (71 kDa), nineteen peptides were detected for serpin-12 and six peptides for HP14, respectively. Detection of the serpin-12 peptide with C-terminal Leu⁴²⁹ (Table 1, Fig. S1) is consistent with the cleavage of serpin-12 at Leu⁴²⁹ (the predicted P1 residue) by HP14 during the inhibition reaction, since this peptide would not be generated by trypsin (cleavage after Arg or Lys). This interpretation is in line with HP14's enzyme specificity and serpin-12's inhibitory selectivity (Wang and Jiang, 2006; Yang et al., 2018).

3.3. Formation of a serpin-12 N-HP14 complex and its effect on proHP21 activation in vitro

Another line of evidence suggests that *M. sexta* serpin-12 is a physiological regulator of HP14. After HP14 was activated by incubating purified proHP14 with curdlan and β GRP2 or with *E. coli* peptidoglycan, PGRP1 and MBP, the active HP14 formed a covalent complex with serpin-12 N (a functional serpin-12 core domain construct expressed without its amino-terminal extension) (Fig. 4A). The complex migrated to the 71 kDa position during SDS-PAGE under reducing condition and was recognized by antibodies to both serpin-12 and HP14PD. No immunoreactivity was found at this position in the negative controls of serpin-12 N and HP14 alone.

We further examined the HP14 activity using its substrate proHP21 (Wang and Jiang, 2007) to test the effect of HP14 inhibition by serpin-12 on proHP21 activation. In the assay, HP14, generated by pre-incubating *E. coli* peptidoglycan, PGRP1, MBP and proHP14, was reacted with proHP21 in the absence or presence of serpin-12 N (Fig. 4B). SDS-PAGE and immunoblot analysis indicated that most of proHP21 was converted to HP21 by HP14. However, serpin-12 N inhibited this reaction. This inhibition likely caused the previously observed changes in proteolytic processing of SP-related proteins downstream from HP14 in immune cascades after adding serpin-12 N to hemolymph (Yang et al., 2018).

3.4. Properties, expression, activation, and possible function of *A. gambiae* SP217

The SP-SPH-serpin system is conserved to some extent in different insects to generate POs and cytokines (*e.g.* Spätzle-1, stress responsive peptides) (Kanost and Jiang, 2015; Schrag et al., 2017; Cao and Jiang, 2018). *A. gambiae* SP217, an ortholog of *M. sexta* HP14 and *D. melanogaster* ModSP (Cao and Jiang, 2018), contains a signal peptide for secretion, four LDLa repeats, a Sushi domain, a Wonton domain, and a catalytic domain (Cao et al., 2017). Its predicted activation site is AEAY³⁹³*I³⁹⁴IGG, located before the C-terminal SP domain. After cleavage activation, an interchain disulfide bond is anticipated to tether the regulatory and catalytic chains. Intrigued by a possible role for proSP217 as an initiating enzyme that activates itself and downstream proteases in response to bacterial or fungal infection, we cloned the proSP217 cDNA, expressed it in *E. coli* for use as an antigen to generate a rabbit antibody, and then constructed a baculovirus to produce proSP217 in Sf9 cells for functional assays. Most of the proSP217 zymogen was insoluble and associated with cell pellet, and a small amount of activated SP217 in the medium was poorly recognized by the polyclonal

antibody at 30 kDa (data not shown). We isolated from 300 ml of the cell culture medium 7.0 µg *A. gambiae* SP217 at an estimated purity of 40%. While its heavy chain was hardly stained by Coomassie brilliant blue, the 30 kDa light chain was stained by the dye and recognized by antibody against the C-terminal hexahistidine tag (Fig. S2A).

Several factors seem to contribute to the proteolysis of recombinant SP precursors produced in baculovirus-infected insect cells. Interaction with a hydrophobic surface during concentration caused partial autoactivation of proPAP1 (Yu et al., 2003). Cleavage activation of proHP6 and proHP8 (Wang and Jiang, 2008) could be triggered by a trypsin-like SP secreted by Sf9 cells. The presence of a considerable amount of SP217 in the culture medium may be related to its role as a self-activating pathway initiator. Besides, having four instead of five LDLa repeats might make *A. gambiae* proSP217 easier to activate than *M. sexta* HP14 (data not shown).

To test if the 30 kDa band resulted from proSP217 autoactivation, we treated the protein with Glu-C endopeptidase and analyzed the products by LC-MS/MS. (Trypsin was not selected because the expected peptide I³⁹⁴IGGR may be too hydrophilic to bind to the column.) The experimental MS/MS spectra matched the hypothetical peptides that covered 91% of the mature protein, which included (E)TLSCVDGWDGPVFRCEPVCCTPTDAEAY³⁹³ and (D)GPVFRCEPVCCTPTDAEAY³⁹³ (Fig. S3, A and B). Then, we sequenced the 30 kDa band by automated Edman degradation and detected Ile (26.7), Ile (23.1), Gly (20.9), Gly (18.2), Arg (11.6), Xaa, Val (17.2), Ser (8.4), Ile (14.4), Ala (14.6), and Glu (11.8) in cycles 1–11, where the numbers in parentheses denote amounts of PTH derivatives in picomoles. This sequence matched the predicted amino-terminus of the catalytic domain I³⁹⁴IGGRN³⁹⁹VSIAE even though we did not detect PTH-Asn in cycle 6. Instead, there was an unknown peak before PTH-Asp (Fig. S3C), whose intensity decreased in cycle 7 and disappeared in the subsequent cycles. Since N³⁹⁹X(S/T) is an N-linked glycosylation site, we interpret this peak as modified PTH-Asn³⁹⁹, which is more hydrophilic than non-glycosylated. In summary, these data demonstrated that proSP217 was cleaved after AEAY³⁹³ by an enzyme with chymotrypsin-like specificity (Cao et al., 2017). With a primary specificity pocket made up of Ser⁶⁰², Ala⁶³⁵ and Val⁶⁴⁹, SP217 is anticipated to cleave after Leu, Tyr, or other hydrophobic residues in its protein substrates. Thus, we hypothesize that auto-cleavage between Tyr³⁹³ and Ile³⁹⁴ is likely responsible for proSP217 activation.

Due to their small sizes, we did not collect hemolymph from *A. gambiae* larvae to test the biochemical function of SP217. Instead, we added the purified SP217 to plasma from *M. sexta* larvae and found that this protease stimulated a major increase in PO activity (Fig. 5) in the absence of microbial elicitors. This result is consistent with a conserved function of SP217 as an initiating protease for the proPO activation cascade and demonstrates that plasma from *M. sexta* larvae can be used to test function of this enzyme. This approach of using *M. sexta* plasma has been used successfully to identify *A. gambiae* CLIPB9 as an activator of proPO (An et al., 2011).

3.5. Structural features of *A. gambiae* SRPN3, SRPN11 and SRPN17

To identify candidate physiological regulator(s) of the SP217, we examined the sequences of all 18 serpin genes in the *A. gambiae* genome (Suwanchaichinda and Kanost, 2009). Among them, SRPN3, SRPN11 and SRPN17 were predicted to be secreted, inhibitory serpins with a large, hydrophobic residue (Phe or Leu) at the putative P1 site in the reactive center loop (Fig. S4). *A. gambiae* SRPN1–3, *D. melanogaster* Spn27A, and *M. sexta* serpin-3 belong to an orthologous group (Fig. S5). Mature SRPN1–3 have a 17–27-residue region before the serpin domain, much shorter than the N-terminal extensions in *A. gambiae* SRPN11 (118), SRPN17 (75), *M. sexta* serpin-12 (106), and *Drosophila* Necrotic (89) (Yang et al., 2018). This extension is a characteristic feature of these inhibitors, but its purpose remains unclear. Immune challenge causes an amino-terminal truncation of Necrotic at an unknown site. There is no major difference in association rate constant (k_{ass}) or stoichiometry of inhibition (molar ratio serpin behaving as inhibitor vs. substrate) between full-length Necrotic (Nec-fl) and Nec N (Pelte et al., 2006), except that the later has a 13-fold increase in k_{ass} and 5-fold decrease in SI to porcine pancreatic elastase (PPE). Thus, removing the amino-terminal extension does not have a major effect on the biochemical activity of this serpin.

We detected a possible folding unit, approximately 62 residues long, with six Cys residues that may form three stabilizing disulfide bonds, located in the first part of SRPN11's 118-residue amino-terminal extension (Fig. S4). No protein with significant similarity to this region of SRPN11 was found in a BLAST search of the non-dipteran reference sequences in GenBank. In comparison, the 75-residue extension in *A. gambiae* SRPN17 is more like that in *Drosophila* Necrotic or *Manduca* serpin-12: rich in Gln, Pro and Ser/Thr, hydrophilic, with a GRAVY (grand average of hydropathy) score of -1.125 , and so far only found in a few lepidopteran and dipteran serpins (Yang et al., 2018).

Following the amino-terminal extension is a 367-residue serpin domain in both SRPN11 and SRPN17, which are closely similar to each other but less so to Necrotic or Spn28Da (Waterhouse et al., 2007). However, in their reactive center loops, both of these mosquito serpins have Leu-Ser at the putative P1 and P1' positions (Fig. S6), identical to that in *M. sexta* serpin-12, *D. melanogaster* Necrotic and Spn28Da. *A. gambiae* SRPN11 has a Pro at P2 and P17, SRPN17 has a stretch of four Pro from P4' to P7', and Necrotic has a Pro at P2, P3', P4' and P6'. These residues may provide SRPN11, SRPN17 and Necrotic unusual structural properties.

3.6. Expression and functional analysis of the core domains of *A. gambiae* SRPN11 and SRPN17

While a role of SRPN3 in SP217 regulation has not yet been ruled out, we focused on SRPN11 and SRPN17, as they group in phylogenetic analysis with *M. sexta* serpin-12 (Fig. S5) and share several structural features as described above. We amplified and cloned into H6pQE60 the cDNA for the core serpin domains of SRPN11 and SRPN17. According to the primer design, the recombinant proteins have the following sequences: MHHHHHHAMDFA...FRRG (SRPN11 N, 43,932.13 Da, pI = 6.43) and MHHHHHHAMEFAWN...PTLN (SRPN17 N, 42,441.95 Da, pI = 8.95), where the underlined parts are the ends of the sequences in the constructs encoded by the respective

cDNAs (Fig. S4). Sequencing multiple clones revealed the same nonsynonymous changes in SRPN11 (T331S, T402S, I417V) and SRPN17 (A452T) compared with the previously determined sequence (Suwanchaichinda and Kanost, 2009), suggesting that these substitutions are due to allelic variations in the mosquito colony used for total RNA and cDNA preparation. The differing residues in SRPN11 are similar in chemical property to the original ones whereas A452T at P12' of SRPN17 is in the end of strand-1 in β -sheet C. Thus, we did not expect a major change in conformation or function in the domain level, and the serpins were indeed soluble and active protease inhibitors. From 100 ml culture, we obtained 4.2 mg SRPN11 N and 3.6 mg SRPN17 N at an estimated purity of 95% (Fig. S2B). SRPN11 N inhibited 80% of bovine chymotrypsin and PPE activities at molar ratios of 5:1 and 4:1, respectively; SRPN17 N inhibited 80% of the PPE activity at a ratio of 2:1 but did not inhibit chymotrypsin or cathepsin G (data not shown).

To further characterize these related serpins, we examined by MALDI-TOF mass spectrometry their cleavage by *M. sexta* HP14, *A. gambiae* SP217, cathepsin G and PPE during inhibition reactions. Incubation of serpin-12 N with HP14 generated a 5323.30 Da peptide identical in size to a peptide produced in the cathepsin G-serpin-12 N reaction (Fig. 6A, Table 2, Yang et al., 2018). Their close match with the calculated M_r of S⁴³⁰YDE...AAEA (5326.09 Da) indicated that the proteolysis occurred between L⁴²⁹ and S⁴³⁰ of serpin-12 N. Analysis of mixtures of SRPN11 N with HP14 or cathepsin G (Fig. 6, B and J) showed that the protease inhibition was accompanied by cleavage at L⁴³⁷*S⁴³⁸. These data agreed perfectly with the P1 site predictions for serpin-12 N and SRPN11 N (Fig. S6). However, HP14 cleaved SRPN17 N at A*F⁴³⁷ (P5*P4) and S*A⁴⁴² (P1'*P2') (Fig. 6C) and yielded 5006.40 and 4489.18 Da peaks matching the calculated M_r 's of F⁴³⁷VAL...PTLN⁴⁸¹ (5009.00) and A⁴⁴²APP...PTLN⁴⁸¹ (4491.38 Da). We interpreted this deviation from the predicted L*S⁴⁴¹ (P1*P1') cleavage site as result of a suboptimal interaction between *M. sexta* HP14 and *A. gambiae* SRPN17 N. With a broader specificity, PPE cleaved SRPN17 N at A*L⁴⁴⁰ (P2*P1), L*S⁴⁴¹ (P1*P1'), and S*A⁴⁴² (P1'*P2') (Fig. 6K, Table 2). We also examined the reaction mixtures of *A. gambiae* SP217 with serpin-12 N, SRPN11 N, and SRPN17 N (Fig. 6, E–G, Table 2). The results indicated that the proteolysis occurred at the predicted P1*P1' site in all these cases including L*S⁴³⁸ and L*S⁴⁴¹ in *A. gambiae* SRPN11 N and SRPN17 N, respectively. Less-than-perfect interaction between SP217 and *M. sexta* serpin-12 N may have caused the secondary cleavage at F*A⁴²⁵ (P6*P5, Fig. 6E). In summary, these three serpins are in most cases cleaved at L*S (P1*P1') by HP14, SP217 and cathepsin G during inhibition reactions.

We then examined the reaction mixtures of *M. sexta* HP14 and *A. gambiae* SRPN11 N or SRPN17 N to detect serpin-protease complex formation (Fig. 7). After proHP14 was incubated with *E. coli* peptidoglycan, PGRP1 and MBP, the level of the 75 kDa zymogen decreased whereas the 30 kDa band corresponding to the active protease domain appeared and was recognized by the protease domain antibody (lane 3). When SRPN11 N and SRPN17 N were included in the reactions, intensity of the 30 kDa band decreased and the 75 kDa band had a small intensity increase (lane 5). Since proHP14 and the serpin-protease complex have the same gel mobility, we interpret the increase in the 75 kDa band as a result of the complex formation. Similarly, SP217 formed 75 kDa complexes with serpin-12 N, SRPN11 N and SRPN17 N, which were recognized by the pentahistidine antibody,

indicating labeling of the hexahistidine tag present in the serpins (Fig. 8). These results are consistent with inhibition of *M. sexta* HP14 by *A. gambiae* SRPN11 N and SRPN17 N, suggesting that they may have functions similar to *M. sexta* serpin-12.

3.7. Concluding remarks

Elucidation of protein function can be challenging. In the case of finding a physiological pair of serpin and SP, it often requires testing a panel of inhibitors and proteases that are present in the same location *in vivo*, such as the complex mixture of protease and serpins in insect plasma. Analysis of the inhibitory kinetics can eliminate certain combinations whose association rates are too low to be biologically relevant, but association rate constants alone may not reveal physiological regulation of proteases. During the investigation of serpins in hemolymph of *M. sexta*, we have developed a systematic approach to address this problem. At first, selective inhibition of commercially available SPs may show an agreement between P1 prediction in the serpin and known specificity of the proteases. Then, determination by MALDI-TOF mass spectrometry of the C-terminal serpin peptide(s) released during the serpin-protease reaction is convenient for assessing the cleavage site in the serpin and specificity of the serpin-protease interaction. Demonstrated inhibition of an endogenous hemolymph protease by the serpin indicates a potential biological target for the serpin *in vivo*. Further experiments, such as measurement of a serpin's effect on proPO activation *ex vivo* or antimicrobial peptide induction *in vivo*, allow initial functional correlation of the serpin. While SDS-PAGE separation and immunoblot identification of serpin-protease complex is a step in identifying serpin-protease pairs, the isolation of serpin-protease complex formed in plasma by immunoaffinity chromatography followed by separation by two-dimensional electrophoresis and characterization of the composition of the complex by LC-MS/MS analysis provides the strongest evidence for the existence of a physiological pair of SP and serpin formed in hemolymph.

In this paper, we have presented strong evidence that serpin-12 is a physiological regulator of HP14 in *M. sexta*. While the roles of HP14 and serpin-12 in melanization has been studied and established in *M. sexta*, it is not known if this type of protein pair is conserved in other insects. We showed that *A. gambiae* modular protease SP217, an ortholog of *M. sexta* HP14, can cause activation of the proPO pathway, and we found that the mosquito serpins SRPN11 and SRPN17 can form complexes with SP217. These results suggest that inhibition of orthologous modular initiating proteases such as HP14 by serpins may be a key and conserved step in regulating activation of protease pathways leading to melanization and activation of the Toll pathway in insects.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

This work was supported by National Institutes of Health Grants GM58634 and AI112662 (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 OKL03054.

Abbreviations:

PO and proPO	phenoloxidase and its proenzyme
PAP	proPO activating protease
PGRP	peptidoglycan recognition protein
MBP	microbe binding protein
βGRP	β-1,3-glucan recognition protein
HP	hemolymph protease
SP and SPH	serine protease and noncatalytic serine protease homolog
ModSP/MSP	modular serine protease
PPE	porcine pancreatic elastase
Spn/SRPN	serine protease inhibitor in the serpin superfamily
IP	induced plasma from day 3, 5 th instar larvae challenged with a mixture of bacteria

References

- An C, Budd A, Kanost MR, Michel K, 2011 Characterization of a regulatory unit that controls melanization and affects longevity of mosquitoes. *Cell Mol. Life Sci* 68, 1929–1939. [PubMed: 20953892]
- An C, Ishibashi J, Ragan EJ, Jiang H, Kanost MR, 2009 Functions of *Manduca sexta* hemolymph proteinases HP6 and HP8 in two innate immune pathways. *J. Biol. Chem* 284, 19716–19726. [PubMed: 19487692]
- An C, Jiang H, Kanost MR, 2010 Proteolytic activation and function of the cytokine Spätzle in innate immune response of a lepidopteran insect, *Manduca sexta*. *FEBS J.* 277, 148–162. [PubMed: 19968713]
- Cao X, He Y, Hu Y, Zhang X, Wang Y, Zou Z, Chen Y, Blissard GW, Kanost MR, Jiang H, 2015 Sequence conservation, phylogenetic relationships, and expression profiles of nondigestive serine proteases and serine protease homologs in *Manduca sexta*. *Insect Biochem. Mol. Biol* 62, 51–63. [PubMed: 25530503]
- Cao X, Gulati M, Jiang H, 2017 Serine protease-related proteins in the malaria mosquito, *Anopheles gambiae*. *Insect Biochem. Mol. Biol* 88, 48–62. [PubMed: 28780069]
- Cao X, Jiang H, 2018 Building a platform for predicting functions of serine protease-related proteins in *Drosophila melanogaster* and other insects. *Insect Biochem. Mol. Biol* 103, 53–69. [PubMed: 30367934]
- Buchon N, Poidevin M, Kwon HM, Guillou A, Sottas V, Lee BK, Lemaitre B, 2009 A single modular serine protease integrates signals from pattern-recognition receptors upstream of the *Drosophila* Toll pathway. *Proc. Natl. Acad. Sci. USA* 106, 12442–12447 [PubMed: 19590012]
- Dunn PE, Drake DR, 1983 Fate of bacteria injected into naive and immunized larvae of the tobacco hornworm *Manduca sexta*. *J. Invertebr. Pathol* 41, 77–85.
- Gorman MJ, Wang Y, Jiang H, Kanost MR, 2007 *Manduca sexta* hemolymph proteinase 21 activates prophenoloxidase-activating proteinase-3 in an insect innate immune response proteinase cascade. *J. Biol. Chem* 282, 11742–11749. [PubMed: 17317663]
- Jiang H, Ma C, Lu Z, Kanost MR, 2004 β-1,3-glucan recognition protein-2 (βGRP-2) from *Manduca sexta*: an acute-phase protein that binds β-1,3-glucan and lipoteichoic acid to aggregate fungi and

bacteria and stimulate prophenoloxidase activation. *Insect Biochem. Mol. Biol* 34, 89–100. [PubMed: 14976985]

Jiang H, Wang Y, Ma C, Kanost MR, 1997 Subunit composition of prophenoloxidase from *Manduca sexta*: molecular cloning of subunit proPO-p1. *Insect Biochem. Mol. Biol* 27, 835–850. [PubMed: 9474780]

Jiang H, Wang Y, Yu XQ, Kanost MR, 2003 Prophenoloxidase-activating proteinase-2 from hemolymph of *Manduca sexta*. A bacteria-inducible serine proteinase containing two clip domains. *J. Biol. Chem* 278, 3552–3561. [PubMed: 12456683]

Jiang R, Kim EH, Gong JH, Kwon HM, Kim CH, Ryu KH, Park JW, Kurokawa K, Zhang J, Gubb D, Lee BL, 2009 Three pairs of protease-serpin complexes cooperatively regulate the insect innate immune responses. *J. Biol. Chem* 284, 35652–35658. [PubMed: 19858208]

Jiang R, Zhang B, Kurokawa K, So YI, Kim EH, Hwang HO, Lee JH, Shiratsuchi A, Zhang J, Nakanishi Y, Lee HS, Lee BL, 2011 93-kDa twin-domain serine protease inhibitor (Serpin) has a regulatory function on the beetle Toll proteolytic signaling cascade. *J Biol Chem*. 286, 35087–35095. [PubMed: 21862574]

Kanost MR, Jiang H, 2015 Clip-domain serine proteases as immune factors in insect hemolymph. *Curr. Opin. Insect Sci* 11, 47–55. [PubMed: 26688791]

Kim CH, Kim SJ, Kan H, Kwon HM, Roh KB, Jiang R, Yang Y, Park JW, Lee HH, Ha NC, Kang HJ, Nonaka M, Söderhäll K, Lee BL, 2008 A three-step proteolytic cascade mediates the activation of the peptidoglycan-induced toll pathway in an insect. *J Biol Chem*. 283, 7599–7607. [PubMed: 18195005]

Krem MM, Di Cera E, 2002 Evolution of enzyme cascades from embryonic development to blood coagulation. *Trends Biochem Sci*. 27, 67–74. [PubMed: 11852243]

Kwon TH, Kim MS, Choi HW, Joo CH, Cho MY, Lee BL, 2000 A masquerade-like serine proteinase homologue is necessary for phenoloxidase activity in the coleopteran insect, *Holotrichia diomphalia* larvae. *Eur. J. Biochem* 267, 6188–6196. [PubMed: 11012672]

Lee E, Linder ME, Gilman AG, 1994 Expression of G-protein α -subunits in *Escherichia coli*. *Methods Enzymol*. 237, 146–164. [PubMed: 7934993]

Levashina EA, Langley E, Green C, Gubb D, Ashburner M, Hoffmann JA, Reichhart JM, 1999 Constitutive activation of toll-mediated antifungal defense in serpin-deficient *Drosophila*. *Science* 285(5435), 1917–1919. [PubMed: 10489372]

Li M, Christen JM, Dittmer NT, Cao X, Zhang X, Jiang H, Kanost MR, 2018 The *Manduca sexta* serpinome: analysis of serpin genes and proteins in the tobacco hornworm. *Insect Biochem. Mol. Biol* 102, 21–30. [PubMed: 30237077]

Meekins DA, Kanost MR, Michel K, 2017 Serpins in arthropod biology. *Semin. Cell Dev. Biol* 62, 105–119. [PubMed: 27603121]

Park JW, Kim CH, Rui J, Park KH, Ryu KH, Chai JH, Hwang HO, Kurokawa K, Ha NC, Söderhäll I, Söderhäll K, Lee BL, 2010 Beetle immunity. *Adv. Exp. Med. Biol* 708, 163–180. [PubMed: 21528698]

Pelte N, Robertson AS, Zou Z, Belorgey D, Dafforn TR, Jiang H, Lomas D, Reichhart JM, Gubb D, 2006 Immune challenge induces N-terminal cleavage of the *Drosophila* serpin Necrotic. *Insect. Biochem. Mol. Biol* 36, 37–46. [PubMed: 16360948]

Reichhart JM, 2005 Tip of another iceberg: *Drosophila* serpins. *Trends Cell Biol*. 15, 659–665. [PubMed: 16260136]

Robinson NE, Robinson AB, 2001 Molecular clocks. *Proc. Natl. Acad. Sci. USA* 98, 944–949. [PubMed: 11158575]

Scherfer C, Tang H, Kambris Z, Lhocine N, Hashimoto C, Lemaitre B, 2008 *Drosophila* Serpin-28D regulates hemolymph phenoloxidase activity and adult pigmentation. *Dev. Biol* 323, 189–196. [PubMed: 18801354]

Schrag LG, Herrera AI, Cao X, Prakash O, Jiang H, 2017 Structure and function of stress responsive peptides in insects *In: Srivastava VP (Ed.), Peptide-based drug discovery: challenges and new therapeutics*. Royal Society of Chemistry, London, UK, pp. 438–451.

- Sumathipala N, Jiang H, 2010 Involvement of *Manduca sexta* peptidoglycan recognition protein-1 in the recognition of bacteria and activation of prophenoloxidase system. *Insect Biochem. Mol. Biol.* 40, 485–495.
- Suwanchaichinda C, Kanost MR, 2009 The serpin gene family in *Anopheles gambiae*. *Gene* 442, 47–54. [PubMed: 19394412]
- Veillard F, Troxler L, Reichhart JM, 2016 *Drosophila melanogaster* clip-domain serine proteases: structure, function and regulation. *Biochimie* 122, 255–269. [PubMed: 26453810]
- Wang Y, Jiang H, 2004 Prophenoloxidase (PPO) activation in *Manduca sexta*: an initial analysis of molecular interactions among PPO, PPO-activating proteinase-3 (PAP-3), and a cofactor. *Insect Biochem. Mol. Biol.* 34, 731–742. [PubMed: 15262278]
- Wang Y, Jiang H, 2006 Interaction of β -1,3-glucan with its recognition protein activates hemolymph proteinase 14, an initiation enzyme of the prophenoloxidase activation system in *Manduca sexta*. *J. Biol. Chem.* 281, 9271–9278 [PubMed: 16461344]
- Wang Y, Jiang H, 2007 Reconstitution of a branch of *Manduca sexta* prophenoloxidase activation cascade *in vitro*: Snake-like hemolymph proteinase 21 cleaved by HP14 activates prophenoloxidase-activating proteinase-2 precursor. *Insect Biochem. Mol. Biol.* 37, 1015–1025. [PubMed: 17785189]
- Wang Y, Jiang H, 2008 A positive feedback mechanism in *Manduca sexta* prophenoloxidase activation. *Insect Biochem. Mol. Biol.* 38, 763–769. [PubMed: 18625399]
- Wang Y, Jiang H, 2010 Binding properties of the regulatory domains in *Manduca sexta* hemolymph proteinase-14, an initiation enzyme of the prophenoloxidase activation system. *Dev. Comp. Immunol.* 34, 316–322. [PubMed: 19900472]
- Wang Y, Lu Z, Jiang H, 2014 *Manduca sexta* prophenoloxidase activating proteinase-3 (PAP3) stimulates melanization by activating proPAP3, proSPHs, and proPOs. *Insect Biochem. Mol. Biol.* 50, 82–91. [PubMed: 24768974]
- Wang Y, Sumathipala N, Rayaprolu S, Jiang H, 2011 Recognition of microbial molecular patterns and stimulation of prophenoloxidase activation by a β -1,3-glucanase-related protein in *Manduca sexta* larval plasma. *Insect Biochem. Mol. Biol.* 41, 322–331. [PubMed: 21296155]
- Waterhouse RM, Kriventseva EV, Meister S, Xi Z, Alvarez KS, Bartholomay LC, Barillas-Mury C, Bian G, Blandin S, Christensen BM, Dong Y, Jiang H, Kanost MR, Koutsos AC, Levashina EA, Li J, Ligoxygakis P, Maccallum RM, Mayhew GF, Mendes A, Michel K, Osta MA, Paskewitz S, Shin SW, Vlachou D, Wang L, Wei W, Zheng L, Zou Z, Severson DW, Raikhel AS, Kafatos FC, Dimopoulos G, Zdobnov EM, Christophides GK, 2007 Evolutionary dynamics of immune-related genes and pathways in disease-vector mosquitoes. *Science* 316, 1738–1743. [PubMed: 17588928]
- Yang F, Wang Y, Sumathipala N, Cao X, Kanost MR, Jiang H, 2018 *Manduca sexta* serpin-12 controls the prophenoloxidase activation system in larval hemolymph. *Insect Biochem. Mol. Biol.* 99, 27–36. [PubMed: 29800677]
- Yu XQ, Jiang H, Wang Y, Kanost MR, 2003 Nonproteolytic serine proteinase homologs are involved in prophenoloxidase activation in the tobacco hornworm, *Manduca sexta*. *Insect Biochem. Mol. Biol.* 33, 197–208. [PubMed: 12535678]
- Zou Z, Zhao P, Weng H, Mita K, Jiang H, 2009 A comparative analysis of serpin genes in the silkworm genome. *Genomics* 93, 367–375. [PubMed: 19150649]

Highlights

- Isolation and identification of the complex of serpin-12 and HP14 formed in *M. sexta* hemolymph during immune responses
- Inhibition of the pathway-initiating protease and its effects on downstream cascade members
- Conservation of the control mechanism in *A. gambiae* involving SRPN11/17 and SP217, the HP14 ortholog

Fig. 1.

The immune serine protease pathways and their regulation by serpins in *D. melanogaster* (A), *T. molitor* (B), and *M. sexta* (C). Some steps (marked by dashed arrows) in the *Drosophila* system may represent indirect proteolytic activation. Names of the serpins (in red bold font) are further abbreviated by removing “serpin or Spn”.

Fig. 2.

SDS-PAGE separation and immunoblot analysis of the hemolymph proteins eluted from the serpin-12 (**A**) or HP14PD (**B**) antibody columns. As described in Section 2.3, plasma proteins associated with the antibodies were eluted from the column. Aliquots of the peak fractions were separated by 7.5% SDS-PAGE, stained with Coomassie brilliant blue (CBB) (*left* panel), and analyzed using diluted serpin-12 or HP14PD antiserum (*middle* and *right* panels), as indicated at the bottom. Lane 1, eluted proteins from induced plasma not treated with *S. aureus* peptidoglycan; lane 2, eluted proteins from the induced plasma elicited with peptidoglycan. Positions and sizes (in kDa) of the pre-stained M_r standards are marked on the *left*, with the 75 kDa marker highlighted *red*. *, putative serpin-protease complexes; ▶, uncoupled IgG heavy chain.

Fig. 3.

Separation of the hemolymph proteins eluted from the serpin-12 antibody column by 2DE for immunoblot and LC-MS/MS analysis. As described in Section 2.3, an induced plasma sample was treated with *S. aureus* peptidoglycan and then separated by serpin-12 immunoaffinity chromatography. The eluted proteins (200 μ g) were resolved by 2DE followed by immunoblot analysis using diluted antiserum to HP14PD (**A**, 1:1000) or serpin-12 (**B**, 1:2000), by staining with SYPRO Ruby (**C**), or by superposition (**D**) of panels B and C. After image alignment, spot-1 and spot-2, marked by *dashed ovals*, were excised from the stained gel for trypsin digestion and mass spectrometric analysis. Positions and sizes (in kDa) of the pre-stained M_r standards are marked on the *left*, with the 75 kDa marker highlighted *red*.

Fig. 4.

Inhibition of *M. sexta* HP14 by serpin-12 N *in vitro* and its effect on proHP21 activation.

(A) Formation of the serpin-12-HP14 complex. As controls, the purified serpin-12 N (lane 1, 0.5 μ g), curdlan-activated HP14 (lane 2, a mixture of 10 μ g curdlan, 40 ng β GBP2, and 0.1 μ g proHP14), and peptidoglycan-activated HP14 (lane 3, a mixture of 1 μ g *E. coli* peptidoglycan, 0.3 μ g PGRP1, 0.3 μ g MBP, and 0.1 μ g proHP14) were incubated in a total volume of 24 μ l at 37 $^{\circ}$ C for 90 min. The test reactions were curdlan-activated HP14 (0.1 μ g, lane 4) or peptidoglycan-activated HP14 (0.1 μ g, lane 5) incubated with 0.5 μ g serpin-12 N under the same conditions. Two halves of the mixtures were subjected to 10% SDS-PAGE under reducing condition and immunoblot analysis using polyclonal antisera against serpin-12 (1:2000 dilution, *left* panel) or HP14PD (1:1000 dilution, *right* panel). *, serpin-protease complex. (B) Suppression of proHP21 activation. Buffer A (lane 1), peptidoglycan-activated HP14 (lane 2, 0.1 μ g), and a mixture of 0.5 μ g serpin-12 N and 0.1 μ g HP14 (lane 3) were incubated at 37 $^{\circ}$ C for 1 h in a total volume of 23 μ l. Aliquots of proHP21 (50 ng in 1 μ l) were added to the tubes and incubated for another 1 h at 37 $^{\circ}$ C. The samples were separated by 10% SDS-PAGE followed by immunoblot analysis using 1:2000 diluted antisera against HP21 as the primary antibody. Positions and sizes (in kDa) of the M_r standards are marked on the *left*, with the 75 kDa marker highlighted *red*.

Fig. 5. Stimulation of proPO activation in induced plasma from *M. sexta* larvae by *A. gambiae* SP217. Aliquots of IP (1.0 μ l) were incubated with SP217 (0.1 μ g) in 18 μ l of buffer A for 10 min at room temperature before PO activity measurement. As negative controls, IP or SP217 alone were incubated with buffer A under the same conditions. PO activities in the reaction and controls were measured and plotted as mean $\pm$ range (n = 2) on the bar graph.

Fig. 6. MALDI-TOF mass spectrometric analysis of the serpin C-terminal peptides released from the SP-serpin reactions. A representative single accumulation spectrum is presented for each of the sixteen samples. Protein compositions of the samples are indicated in panels A—P. CathG, human cathepsin G; PPE, porcine pancreatic elastase. The mass spectrometer was calibrated with a mixture of standards and its performance was monitored using the external standards between sample runs. Positive identifications were made for peaks whose observed masses match the calculated mass within an error range of 500 ppm. The 4975.21 Da peak in panel I may not be T⁴⁶⁶HTL...FRRG (4986.95 Da), as the mass difference was 2360 ppm. The predicted P1 residue (Leu) is in *red* font; the peptide released upon cleavage (*) in most cases occurs between predicted P1 and P1' sites in the serpin reactive center loop.

Fig. 7.

Complex formation between *M. sexta* HP14 and *A. gambiae* SRPN11 N or SRPN17 N *in vitro*. The controls of proHP14 (lane 1, 0.1 μg), recognition complex (lane 2, 1 μg *E. coli* peptidoglycan, 0.4 μg PGRP1, 0.4 μg MBP), HP14 (lane 3, activation mixture of peptidoglycan, PGRP1, MBP and proHP14), serpin (lane 4, 1 μg, SRPN11 N on the *left* panel; SRPN17 N on the *right*), and test reaction (lane 5, HP14 and serpin) were incubated at 37 °C for 90 min. Half of the samples were subjected to SDS-PAGE and immunoblot analysis using diluted polyclonal antiserum against HP14PD. *, 75 kDa complex of HP14 protease domain and SRPN11 N or SRPN17 N; ■, 75 kDa proHP14; ►, 30 kDa catalytic domain of HP14; ?, a nonspecific protein from *E. coli*.

Fig. 8.

Covalent complexes of *A. gambiae* SP217 with serpin-12 N (**A**), SRPN11 N (**B**) or SRPN17 N (**C**). To test SP217 inhibition, aliquots of SP217 (50 ng) were separately reacted with serpin-12 N (0.15 μg), SRPN11 N (1 μg), or SRPN17 N (1 μg) in buffer A (12 μl) for 30 min at room temperature. For use as negative controls, the same amounts of SP217 and serpins alone were incubated under the same conditions. The samples were separated by 10% SDS-PAGE under reducing condition and analyzed using the antibody to pentahistidine. *, serpin-protease complex; •, serpin; ▶, 30 kDa protease domain of SP217.

Table 1.

A list of trypsinolytic peptides identified by LC-MS/MS in the two protein spots on the 2D gel

Identified peptide sequence ^a	Start ^b	Stop ^b	Probability (%)	Modification ^c	Variants ^d	PPM ^e	Σ Intensity ^f
Spot 1 (upper), Serpin-12, 8 unique peptides:							
(R)EQQPQInGGVHSPTR(S)	61	75	100	DeAm (+1)	1	0.6	4420
(R)ALQMTPEK(S)	150	157	99		1	-0.6	10956
(R)IAVNnFSDSLTPTYFGK(P)	199	215	100	DeAm (+1)	2	-1.2	15456
(K)PALAAQNInSWIASK(T)	216	230	93	DeAm (+1)	1	0.2	3261
(K)GLWEIPFR(E)	260	267	99		1	-0.3	9401
(K)DVALELPK(F)	352	359	99		1	0.5	16814
(K)ADINLEPVLNK(M)	364	374	100		2	0.2	36875
(R)GGSAAAATSFAAVAL(-)	415	429	100		1	-0.5	6296
Spot 1 (upper), HP14, 4 unique peptides:							
(K)EEYAVALK(L)	444	452	96		1	-1.0	25175
(K)EQLYVGSLGK(V)	520	529	100		1	-0.5	14845
(K)DEAGNPSQVLK(V)	537	547	100		2	-0.6	206664
(K)GDSGGGLSFPVNR(L)	589	602	100		1	-0.8	47887
Spot 2 (lower), Serpin-12, 19 unique peptides:							
(R)SSGFNQNmQTNTR(G)	28	40	100	Oxid (+16)	2	-2.3	5903
(R)GPFESQFDIQASLQTTK(S)	41	57	100		2	-1.5	8342
(R)SPTQDTSFK(S)	76	84	100		1	-2.5	5008
(K)SLSPmIPTTSSFHSGPASTSFGVNVFK(Q)	85	111	100	Oxid (+16)	2	-1.7	27163
(K)QmATEQSGNLAAPFSITILLAmLQQGAAGNTLDEITR(A)	112	149	100	2x Oxid (+16)	2	0.1	1634
(R)ALQmTPEK(S)	150	157	99	Oxid (+16)	1	-2.1	11091
(K)SAEIFKK(V)	158	164	90		2	-2.5	7718
(K)VNEEIQR(R)	165	171	90		3	-2.7	16416
(K)TANVFLSENFNLNPQFK(R)	180	197	100	DeAm (+1)	6	-2.6	32845
(R)IAVNnFSDSLTPTYFGK(P)	199	215	100		4	-1.7	57506
(K)PALAAQNINSWIASK(T)	216	230	100		2	-0.7	12663
(K)LVSPDDLSGNTQmVmVNAVYFK(G)	238	259	100	2x Oxid (+16)	5	-1.0	45840
(K)GLWEIPFR(E)	260	267	99		1	-1.6	10928
(K)VASFmQTR(R)	285	292	99	Oxid (+16)	2	-2.9	15915
(K)VVVLPEYNEYSLIVVLPLK(S)	307	326	100		5	-1.5	25433
(K)SSNVDALLSSLSmEDVASFLDLPPK(D)	327	351	100	Oxid (+16)	5	0.0	25581
(K)SSNVDALLSSLSmEDVASFLDLPPKDVALELPK(F)	327	359	100	Oxid (+16)	8	0.1	50171
(K)ADInLEPVLNK(M)	364	374	100	DeAm (+1)	2	-2.0	13029
(R)GGSAAAATSFAAVAL(-)	415	429	98		2	-1.7	6138
Spot 2 (lower), HP14, 6 unique peptides:							
(R)AQFGELPWQAGIYTK(N)	394	408	100		2	-1.8	8008

Identified peptide sequence ^a	Start ^b	Stop ^b	Probability (%)	Modification ^c	Variants ^d	PPM ^e	Σ Intensity ^f
(R)DIHISPYFLGR(T)	473	483	100		2	-1.8	22071
(K)EQLYVGSLGK(V)	520	529	100		1	-2.5	5244
(K)DEAGNPSQVLK(V)	537	547	100		2	-1.7	7913
(K)LPYVDVLQcISQSPQAFR(P)	551	568	100	CAM (+57)	2	-2.3	5602
(K)GDSGGGLSFPAVNR(L)	589	602	100		2	-2.1	8944

^a: peptide sequence with the modified residue(s) in lower case;

^b: positions of the 1st and last residues of the peptide in the protein;

^c: types of peptide modification: CAM, carbamidomethyl, DeAm, deamidation, Oxid, oxidation;

^d: number of variants in charge and modification;

^e: actual (*i.e.* observed $\times$ charge) minus calculated mass expressed as parts per million;

^f: sum of precursor intensity of all the variants. Highlighted in *bold* are the peptides whose 2nd mass spectra (Fig. S1) indicate they may be dissociation products of the trypsin-digested serpin-protease complexes.

Table 2.

A summary of the MALDI-TOF results used to locate the proteolytic cleavage site

serpin name	C-terminal peptides ^a	calc. M_r (Da)	+ cathepsin G	+ PPE	+ HP14	+ SP217	serpin + buffer
serpin-12 N	LSYDE...RAAEA	5439.25				5748.01 ^e	5747.82 ^e
	SYDE...RAAEA	5326.09^b	5323.30^{b, c}	nd ^d	5323.30^b	5322.69^b	5676.76 ^e
	YDE...RAAEA	5239.01					5605.51 ^e
	DE...RAAEA	5075.84					
SRPN11 N	LSITH...PFRRG	5300.30					
	SITH...PFRRG	5187.14^b	5186.76^b		5183.86^b	5184.44^b	no ^f
	ITH...PFRRG	5100.06					
	TH...PFRRG	4986.90		4975.21 ^g			
SRPN17 N	LSAAP...NPTLN	4691.61		4693.55	5006.40 ^e		
	SAAP...NPTLN	4578.45^b	nd ^d	4580.38^b		4578.82^b	no ^f
	AAP...NPTLN	4491.38		4493.34	4489.18		
	AP...NPTLN	4420.24					

^a: The C-terminal sequences are: FAVALSYDEPSLYFRANKPFLAILWDNRSSIPLFMARIMDPTLEHARAAEA in serpin-12 N, PLSITHTLDFKADQPFALIIM DKQNKLPFFAKISKPSKPKDPFRRG in SRPN11 N, and AFVALSAAPPPPIINF TVNEPFLMMIVDKIHEYPLFVGKIVNPTLN in SRPN17 N, where the P1 residues are shown in *bold*. As described in Section 3.6, Thr⁴⁵² (in *italic*) of SRPN17 N may be an allelic variation of Ala⁴⁵².

^b: Calculated and observed M_r 's of the peptide released from P1-P1' cleavage are shown in *bold*.

^c: Data from Yang et al., 2018.

^d: not determined;

^e: The 5,748.01/5747.82, 5676.76, 5605.51, and 5,006.40 Da peaks correspond to serpin-12 N's AAVA...AEA (5751.62 Da), AVAL...AEA (5680.54 Da), VALS...AEA (5609.46 Da), and SRPN17 N's FVAL...TLN (5009.00 Da), respectively. Numbers in parentheses are calculated monoisotopic masses of the peptides.

^f: no peak in the range of 4,000–6,000 Da in the control samples.

^g: unexplained mass peak.