CHARACTERIZATION OF CO-FACTOR PROTEINS FOR PROPHENOLOXIDASE ACTIVATION IN INSECTS

By

QIAO JIN

Bachelor of Ecology
Agriculture University of Hebei Province
Baoding, Hebei
2013

Master of Biochemistry and molecular biology
Chinese academy of sciences
Beijing
2017

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
DOCTOR OF PHILOSOPHY
Dec, 2022

CHARACTERIZATION OF CO-FACTOR PROTEINS FOR PROPHENOLOXIDASE ACTIVATION IN INSECTS

Dissertation Approved:

Dr. Haobo Jiang
Dissertation Adviser

Dr. Bruce H. Noden

Dr. Junpeng Deng

Dr. Jose Soulages

ACKNOWLEDEMENTS

First and foremost, I want to thank my primary supervisor, Dr. Jiang, Haobo, who guides me on my research path, he gives me a constant support and courage when I have setbacks. He is really a scholarly mentor and beneficial friend.There are many things Haobo taught me, but nothing was more precious than his enthusiasm to push the boundary on insect immunity. Not only did I learn from him how to be a qualified scientist, but also an honest person. He is also a very good friend to me, when I'm confused in my life, he always helps me.

I would like to thank my committee members, Dr. Deng, Junpeng, Dr. Soulages, Jose L, and Dr Noden, Bruce. They always provide help when I'm confused and when I need professional guidance. Their work ethic taught not to settle for mediocrity, which has deeply changed the way I think. Similar gratitude to Dr. Hartson, Steven, he taught me mounts of knowledge of proteomics, which lead me to explore more about Mass Spectrometry and Proteomics and broaden my eyes for both my research and my career.

I want to thank my current and previous lab member, including Yang Wang, Xiaolong Cao, Yan He, Udeshika Kariyawasam, Zelong Miao, Tisheng Shan, Chunxiang Hou, Chao Xiong, and Haibo Bao. Previous students provided me many precious experiences for my research projects and my career I also would like to thank Yang, who teach me almost everything I need during my Ph.D, she is a good friend in daily life, who usually give many pertinent advice.

Thanks to all people in in Entomology and Plant Pathology department, including our previous and current department head, Dr. Phillip Mulder and Dr. Justin Talley, Graduate Program Coordinator, Dr. Nathan Walker and Dr. George Opit, the office staffs including Barbara Brown, Sharon Hillock, Lafonda Barrera, Noah Dunagan, etc. They were extremely patient and kind, and helped me in every critical academic timepoint during my Ph.D period. I really appreciate the support from members of Core Facility in OSU regarding the friendly support and training.

I'm also greatly indebted to my master's advisor, Dr. Zhen Zou, who gives me a lot of helpful and positive advice, his continuous support makes me keep going without fear. I want to thank all my friends in my motherland, their endless caring keeps me optimistic about everything around me.

Of course, thanks must go to my parents, my parents don't know the details of what I'm doing, but they tried their best to support me to do what I want. Without their constant support, I would never have had a chance to start and finish my Ph.D. They always be my motivation.

Finally, I want to thank my husband, Fei Liang. There is so much I want to say, but little can be expressed. Without his help and support, it would have been impossible for me to go through my Ph.D period without any regret.

What I experienced in OSU makes me to be a better scientist and better person. Thanks to my alma mater, I will commit to the principles of truth and honesty

Name: QIAO JIN

Date of Degree: DECEMBER 2022

Title of Study: SERINE PROTEASE HOMOLOGS WORKING AS COFACTORS FOR PROPHENOLOXIDASE ACTIVATION IN MANDUCA SEXTA, ANPHOPHELES GAMBIAE, AND DROSOPHILA MELANOGASTER

Major Field: ENTOMOLOY

Abstract:

Melanization is a key part of insect innate immunity, which is mainly induced in host after pathogen invasion. Phenoloxidase (PO) is a crucial component to catalyze melanization. As PO is not stable in insects, it is always stored as prophenoloxidase (PPO), thus the activation of PPO is an important process for melanization induction. PO is generated from its precursor proPO by prophenoloxidase activating proteases (PAPs) while PO activity can't be highly induced, at least in some insects. It has been demonstrated that proPO activation requires additional proteins to serve as a cofactor for PAPs Serine protease homolog (SPH) in Manduca sexta and some other insects. However, it is not fully known if PAPs and cofactors are required for PPO activation in mosquito and *Drosophila melanogaster*. Moreover, formerly demonstrated SPH1 (1a) in *M. sexta* seems having a series orthologs, which need to be explored.

Here we analyzed four proSPHI proteins (proSPH4, proSPH1a, proSPH1b, proSPH101) in *M. sexta* including their abundance and cleavage activation and functional comparison. Based on the mRNA and protein data, we established SPH1b-2 pair as the major cofactor for proPO activation by PAPs in M. sexta. The recombinant proSPHs were tested with PAP3 and proPO to reveal the reaction mechanisms. We also identified four CLIPAs (ClipA4, ClipA6, ClipA7, and ClipA12) in An. gambiae and tested their potential to act as cofactors of CLIPB9, a PAP in An. gambiae, moreover, those ClipAs can also been cleaved by PAP3 from M. sexta. After the proteolytic activation, A. gambiae CLIPA4 formed complexes with CLIPA6, A7 or A12 to assist PPO2 or PPO7 activation. Moreover, cSPH242 and cSPH35 in drosophila were found to be cofactor proteins and work in pairs to enhance DmPPO1 PO activity. Unprecedented high levels of PO activity were achieved in the in vitro conditions, suggesting that cofactor-assisted PPO activation reaction is conserved in holometabolous insects.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIRGURES

Figure Page

Chapter I

Chapter II

Figure Page

Chapter III

Figure Page

CHAPTER I

CLEAVAGE ACTIVATION AND FUNCTIONAL COMPARISION OF MANDUCA SEXTA SERINE PROTEASE HOMOLOGS SPH1A, SPH1B, SPH4, AND SPH101 IN CONJUNCTION WITH SPH2

Qiao Jin [a], Yang Wang [a], Steven D. Hartson [b], Haobo Jiang [a]

[a] Department of Entomology and Plant Pathology, Oklahoma State University,

Stillwater, OK 74078, USA

[b] Department of Biochemistry and Molecular Biology, Oklahoma State University,

Stillwater, OK, 74078, USA

Running title: *Manduca sexta* serine proteinase homologs

Send correspondence to:

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

Key words: clip domain, insect immunity, phenoloxidase, melanization, hemolymph protein

Abstract

Phenoloxidase (PO) is a crucial component of the insect immune response against microbial infection. In the tobacco hornworm *Manduca sexta*, PO is generated from its precursor proPO by prophenoloxidase activating proteases (PAPs) in the presence of two noncatalytic serine protease homologs (SPHs). cDNA cloning and genome analysis indicate that SPH1a (formerly known as SPH1), SPH1b, SPH4, SPH101, and SPH2 contain a clip domain, a linker, and a protease-like domain (PLD). The first 22 residues of the SPH1b, SPH4, and SPH101 PLDs are identical, and differ from SPH1a only at position 4, Thr^{154} substituted with Asn^{154} in SPH1a. While the sequence from Edman degradation was used to establish PAP cofactor as a high M_r complex of SPH1a and SPH2, this assignment needed further validation, especially because *SPH1b* mRNA levels are much higher than *SPH1a*'s and better correlate with *SPH2* transcription. Thus, here we determined expression profiles of these SPH genes in different tissues from various developmental stages using highly specific primers. High levels of SPH1b and SPH2 proteins, low SPH4, and no SPH1a or SPH101 were detected in hemolymph from larvae in the feeding, wandering and bar stages, pupae, and adults by targeted LC-MS/MS analysis, based on unique peptides from the trypsin-treated SPHs. We expressed the five proSPHs in baculovirus-infected *Sf*9 cells for use as standards to identify and quantify their counterparts in plasma samples. Moreover, we tested their cleavage by PAP3 and efficacy of the SPH1a, 1b, 4, and 101 as SPH2 partners in PAP3-mediated proPO activation. PAP3 processed proSPH1b and 101 more readily than proSPH1a and 4; PAP3 activated proPO more efficiently in the presence of SPH2 with SPH101 or SPH1b than with SPH1a or SPH4. These results generally agree with their order of appearance or sequence similarity: SPH101 > SPH1b (98%) > SPH1a (90%) > SPH4 (83%). In

other words, likely due to positive selection, products of the newly duplicated genes (*SPH1b* and *SPH101*) are more favorable substrates of PAP3 and better SPH2 partners in forming a high M_r cofactor than *SPH1a* or *SPH4* is. Electrophoresis on native gel and immunoblot analysis further indicated that SPH101 or 1b form high M_r complexes more readily than SPH1a or 4 does. In comparison, SPH2 showed a small mobility decrease on native gel after PAP3 cleavage at the first site. Since the natural cofactor in bar-stage hemolymph is complexes of SPH1 and 2 with an average M_r of 790 kDa, PAP3-activated SPH2 may associate with the higher M_r SPH1b scaffolds to form super-complexes. Their structures and formation in relation to cleavage of SPH1b at different sites await further exploration.

1. Introduction

Insect phenoloxidases (POs) catalyze multiple steps of a chemical reaction series to produce quinones and other reactive compounds that kill invading microorganisms [1, 2, 3]. The damaged pathogens or parasites are often encapsulated by melanin, inert polymers of the reactive intermediates. As such, PO-mediated melanization is a critical immune response in insects [4, 5, 6]. Since the reactive compounds are also toxic to host cells, insect POs are produced as inactive proPOs and later activated via limited proteolysis when and where needed. Specific enzymes, designated proPO activating proteases or PAPs, are responsible for the cleavage activation [7]. At least in some insects, proPO activation requires additional proteins to serve as a cofactor for PAPs [8, 9, 10, 11]. The auxiliary proteins have a clip domain at the amino terminus followed by a linker and a noncatalytic serine protease homolog (SPH) domain. Similarly, PAPs are composed of one or two clip domains, a linker, and a serine protease (SP) domain. Genomic analyses revealed many genes encoding clip-domain SPs and SPHs, abbreviated as CLIPs, in

holometabolous insects [12, 13, 14]. Alignment of the CLIP sequences resulted in phylogenetic relationships useful for devising hypotheses. For instance, *M. sexta* SPH1 (likely SPH1b) and SPH2, components of a high M_r cofactor of PAPs [15], are located on two branches in a clade of CLIPAs [11]. The clade includes one honeybee, two *Drosophila*, five *Manduca*, and eleven *Anopheles* CLIPAs, suggesting that a single ancestral gene acted as a PAP cofactor (as in *Apis mellifera*), duplicated in the *Drosophila*, *Anopheles* and *Manduca* lineages, and, while DmcSPH242, AgCLIPA4 and MsSPH2 form one branch, the other includes DmcSPH35, 10 AgCLIPAs, and 4 MsSPHs (*i.e.*, 1a, 1b, 4, and 101) [11]. Note that the term SPH1 was used in the past to represent SPH1a, the first SPH whose cDNA was isolated from *M. sexta* [9] and expressed in *Sf*21 cells [16]. With new appreciation that SPHs are a large family of proteins, we use SHPI as a broad term to describe the branch including *M. sexta* SPH1a, 1b, 4, and 101, *Drosophila melanogaster* cSPH35, and so on; SPHII is used for the other branch including *M. sexta* SPH2, *D. melanogaster* cSPH242, and their orthologs in other insects.

Several lines of evidence indicate that gene duplication and adaptation gave rise to the five functionally related SPHs in *M. sexta*. The order of sequence identity (SPH101 > SPH1b (96%) > SPH1a (89%) > SPH4 (77%) > SPH2 (42%)) suggests, after split of the SPHI and SPHII branches, SPH4 emerged first, SPH1a arose later, and SPH1b or SPH101 duplicated the last as *M. sexta* evolved. Since lineage-specific expansion of the SPHI branch may have experienced functional adaptation, we propose that the SPH4, SPH1a, SPH1b, and SPH101 are increasingly fit as a PAP3 substrate, SPH2 partner, and cofactor for proPO activation and there may have been a transition in the past from SPH4 to SPH1a and then to SPH101/SPH1b usage in forming complex with SPH2. The earlier forms may have developed new functions, subfunctionalized, or lost function. The SPH1b and SPH2 transcripts are widely detected in the

52 tissue samples, higher in head, fat body, and muscle than in midgut, Malpighian tubules, testis and ovary [9]. In contrast to a close correlation in SPH1b and SPH2 expression profiles, the SPH4, SPH1a, and SPH101 mRNA levels differ drastically. At the protein level, we detected peptides derived from SPH101, SPH1b, SPH4, SPH2, but not from SPH1a [17]. Due to the high sequence similarity, especially between SPH101 and SPH1b, we suspect that the initial assignment of SPH1 as SPH1a was incorrect and quantitation of the variants was imperfect due to the principle of parsimony that is typically used in the analysis of bottom-up proteomic data. Therefore, we examined in this study mRNA levels of the five SPHs in various tissues and developmental stages by quantitative real-time PCR. We also measured their relative protein abundances in cell-free hemolymph samples in various developmental stages using targeted mass spectrometric monitoring of isoform-specific peptides. We expressed the five proSPHs, tested their cleavage by PAP3 and formation of high M_r complexes, measured cofactor activities of the four SPH pairs (101-2, 1b-2, 1a-2 and 4-2), and obtained evidence for the functional adaptation during evolution. Their cleavages and behaviors on native gels provided new insights into the mechanisms of high M_r complex formation and proPO activation.

2. Methods and materials

2.1. Insect rearing, hemolymph preparation, and SDS-PAGE analysis for sample selection

M. sexta eggs were purchased from Carolina Biological Supply, and larvae were reared on an artificial diet [18]. Fifth instar larvae at feeding (F, day 4) and wandering (W, day 2) stages, pharate pupae with metathoracic brown bars (B), day 14 pupae (P), and day 2 adults (A) were chilled on ice for 10 min and dissected to collect hemolymph into separate tubes. Each tube contained a few crystals of *p*-aminobenzamidine and 1-phenyl-2-thiourea to inhibit serine

protease and PO activities, respectively. After centrifugation at 4000×*g* for 5 min at 4 °C, the supernatants were transferred to clean tubes, and equal volumes of the plasma samples from three insects at the same stage were pooled and stored at -80 °C. Three additional biological replicates at each stage were similarly prepared from nine insects. The cell-free hemolymph samples (5 μl each, 1:5 diluted, F, W, B, P and A) were mixed with 6× SDS sample buffer (1 μl), heated at 95 °C for 5 min, and resolved by 10% SDS-PAGE. After brief staining with Coomassie brilliant blue (CBB) and destaining, the gels were imaged (Fig. S4) for selecting samples with most consistent banding patterns in each stage. While three samples in the F, W, B, and P stages were easily chosen, only two of the adult samples were selected for further analysis, due to variations and difficulties in hemolymph collection from adults.

2.2. Multiple sequence alignment and analysis of phylogenetic relationships

To identify orthologs of the *M. sexta* SPHs in other lepidopteran insects, the SPH1a, 1b, 2, 4, and 101 were used as queries in five BLASTP searches of the non-redundant sequence database at GenBank under default conditions. These searches, limited to Lepidoptera, resulted in a total of 114 hits with similarity higher than 50%. After removal of the sequences lacking a signal peptide or an intact clip domain, 88 sequences were aligned with *Holotrichia diomphalia* PPAF2 [19] to build a phylogenetic tree (Fig. S1), which was used to select the five SPHs and their putative orthologs or paralogs from representative families for a focused study. Multiple sequence alignments of the entire proteins were performed using MUSCLE [20], one module of MEGA X [21]. The aligned sequences were used to construct neighbor-joining trees using MEGA with bootstrap method for the phylogeny test (1000 replications, Poisson model, uniform rates, and complete deletion of gaps or missing data). FigTree 1.4.3 (http://tree.bio.ed.ac.uk/software/figtree) was used to display the phylogenetic trees.

2.3. cDNA cloning and recombinant production of proSPH4, 1b, and 101

SPH4 cDNA was amplified from plasmid #504 [22] by PCR using primers j1346 (5'-GAATTCATATGCAAACAATAAATGTGGATG-3') and j1347 (5'-AAGCTTACTCGAGTTCGTAAACCGAAACGTC-3'). The product was cloned into pGEM-T vector (Promega) and, after sequence validation, the *Eco*RI-*Xho*I and *Nde*I-*Hin*dIII fragments were inserted into the same sites in pMFH6 and pSKB3 (a pET28a derivative), respectively. SPH4/pMFH6 was used to generate a baculovirus to express proSPH4 (GIHMQTIN…VYELEHHHHHH) in insect cells [23], and the underlined part is encoded by the cDNA. SPH4/pSKB3 was used to produce another form of proSPH4 (MGSSHHHHHHDYDIPTTENLYFQ*GHMQTIN…VYELE) in *E. coli* BL21 (DE3). The insoluble protein (1.0 mg) was extracted from the bacteria dissolved in 8 M urea, isolated by Ni^{2+} affinity chromatography, 1:16 diluted in 20 mM Tris-HCl, pH 7.5, 300 mM NaCl, 10 mM imidazole, 0.005% Tween-20, and treated with TEV protease (0.1 mg) (New England BioLabs) overnight at room temperature. The cleaved protein (GHMQ…LE) was separated from the affinity tag and fusion protein using Ni^{2+}-nitrilotriacetic acid agarose beads, concentrated on a Centricon-10 centrifugal filter device (Millipore), and subjected to preparative SDS-PAGE. After light staining, the proSPH4 band (0.5 mg) was cut into gel pieces for use as antigen to raise a rabbit antiserum at Cocalico Biologicals.

A SPH1b fragment was amplified from cDNA of fat body from day 3 pupae using primers j1671 (5'-CTGAATTCAGTCCGGAGATCTGGAGG) and j1672 (5'-GTCCTCGAGTTCGTAAACCGTAATG). After cloning and sequence validation, the *Eco*RI-*Xho*I fragment was retrieved and inserted into the same sites in pMFH6 to generate a baculovirus for producing proSPH1b (GIHMQSGD…VYELEHHHHHH) in insect cells. The underlined

region is identical to the mature precursor, which is recognized by the SPH1a [9] and SPH101 antibodies.

A SPH101 fragment was amplified from a cDNA pool of the induced larval fat body using primers j1336 (5'-GAATTCATATGCAGTCCGGAGATTTGGAGTC) and j1337 (5'-ACTCGAGTTCGTAAACCGAAACGTCGT). After TA cloning and sequencing, the cDNA was found to have twelve nonsynonymous mutations and the *Nde*I-*Xho*I region was inserted into pSKB3 for antibody preparation. SPH101/pSKB3 was used to produce proSPH101 (MGSSHHHHHHDYDI PTTENLYFQ*GHMQSG…YELEHHHHHH) in *E. coli*. As described above, the fusion protein was purified and used as antigen to raise a rabbit antibody against proSPH101. A different SPH101 fragment was amplified from cDNA of induced *M. sexta* eggs using primers j440 (5'-CATATGCAGTCCGGAGATTTGG) and j441 (5'-AAGCTTACTCGAGGAGTTCGTAAACCG). The PCR product was TA cloned, confirmed to be error-free, retrieved by *Nde*I-*Xho*I digestion, and inserted into the same sites in pMFFMH6 [unpublished data]. Like SPH1b/pMFH6, SPH101/pMFFMH6 was used to generate a baculovirus that produces proSPH101 (GIHDYKDD DDKHMQSG…YELLEQKLISEEDLHHHHHH) in insect cells, and the underlined part is encoded by SPH101 cDNA.

2.4. Expression and purification of the recombinant proSPHs

*Sf*9 cells at 2.4×10^6 cells/ml in 1.4 L of insect serum-free medium (Invitrogen Life Technologies) were separately infected with the baculovirus stocks at a multiplicity of infection of 10 and grown at 27 °C for 96 h with gentle agitation at 100 rpm. After the cells were removed by centrifugation at 5,000×*g* for 10 min, the pH of the conditioned medium was adjusted to 8.0 using 1.0 M Tris base. Cell debris and fine particles were spun down by centrifugation at

10,000×*g*, and the supernatant was diluted with 20 mM Tris-HCl, pH 8.0 (buffer A) to a final volume of 4.2 L. The solution was applied to a Q-Sepharose FF column (20 ml bed volume) at a flow rate of 5.0 ml/min and, following a washing step with 100 ml buffer A, bound proteins were eluted from the column with a linear gradient of 0−1.0 M NaCl in 240 ml of buffer A at a flow rate of 1.0 ml/min. The proSPH fractions were combined and loaded onto a 10 ml Ni^{2+}-nitrilotriacetic acid agarose column. After washing with 50 ml of 50 mM sodium phosphate, pH 8.0, the bound proteins were eluted with a gradient of 0−0.3 M imidazole in 90 ml of the same buffer. Fractions containing proSPHs were combined, dialyzed against 20 mM Tris-HCl, pH 7.6, and concentrated on Amicon Ultra-30 centrifugal filter devices (Millipore). The protein aliquots were rapidly frozen in liquid nitrogen prior to storage at -80 °C.

2.5. Profiling of the SPH1a, 1b, 2, 4, and 101 mRNA levels in tissues from various stages

To compare the SPH RNA levels in different stages and tissues, cDNA samples (each equivalent to starting with 50 ng total RNA) were incubated with 1× iTag Universal SYBR Green Supermix (Bio-Rad) and specific primers (0.5 mM each) in triplicate in each reaction (10 µl). The ribosomal protein S3 primers were j055 (5'-CTCAGGCCGAGTCTTTGAGATACA) and j056 (5'-ACTTCATGGACTTGGCTCTCTGAC). Specific qRT-PCR primers were synthesized for SPH2 (j1565, 5'-GGGAGAGTCAATTCGGTGAG, j1566, 5'-TAACCACTTGCGGATGAATG), SPH1a (j443, 5'-CTACCATTGTTCGCCTTGCT, j444 5'-AAGTCGGTCGTATTCGCTTG), SPH1b (j445, 5'-CTGCCATTGTTCGCCTTGTT, j446 5'-TCGGTCGCATTCGCTTGTGG), SPH4 (j447, 5'-GCGCTACCTGACAAGAGACC, j448 5'-TGTAGACCGTTTGGATCGTG), and SPH101 (j449 5'-GGCGAGTGGGACACACAGCA, j450, 5'-AGGCACGCTACTCCCACATT). Each primer pair was tested to ensure amplification of the target gene only and not close homologs. Efficiency of the amplification was determined

individually and confirmed to be 90−110%. Thermal cycling conditions were 95 °C for 2 min and 40 cycles of 95 °C for 10 s and 60 °C for 30 s. After PCR was complete on a CFX Connect Real-Time PCR Detection System (Bio-Rad), melting curves of the products in all reactions were examined to ensure proper shape and Tm values. The mRNA levels were normalized against the internal control of rpS3 (set at 1.00) using corresponding Ct values for the same cDNA samples and the relative mRNA levels were calculated as $2^{-\Delta Ct}$, where $\Delta Ct = Ct_{SPH} - Ct_{rpS3}$.

2.6. Selection, validation, and quantification of specific peptide reporters from the recombinant proSPHs

Based on the sequence alignment of SPHIs and SPH2, theoretical trypsinolytic peptides were calculated for each isoform, and hypothetical peptides unique to individual isoforms were selected. To determine which ones of the peptides were detectable by LC-MS/MS, the purified proSPH1a, 1b, 2, 4 and 101 from insect cells were separately denatured in urea and digested with trypsin [24]. Resulting peptides were desalted using C18 affinity media, dried, and redissolved in mobile phase A (0.1% formic acid in H_2O). The samples were loaded onto an Acclaim PepMap RSLC C18 column (75 μm × 50 cm, Thermo Fisher) for data-dependent LC-MS/MS analysis as described previously [25]. RAW instrument files were searched against a database of sequences from *M. sexta* OGS 2.0 [26] and MCOT 1.0 [9] using the peptide identification software Byonic (v3.10-4×64). Chromatographic peaks corresponding to isoform-specific peptides were examined by manual queries of the instrument RAW data files using Xcalibur Qual Browser (version 2.2.0.23; TFS, San Jose, CA). Peptide ions were advanced as potential parallel reaction monitoring (PRM) reporters if they represented isoform-specific sequences and yielded unambiguous peaks on extracted ion chromatograms.

To validate the peptide reporters, trypsinolytic peptides from the recombinant proSPH1a, 1b, 2, 4, and 101 were diluted 1:3, 1:9, 1:27, and 1:81. Each diluted peptide sample was separated on the Acclaim column via a gradient of 0–35% mobile phase B (0.1% HCOOH, 80% AcCN, 20% H_2O) developed over 120 min. During elution, peptides were analyzed using an Orbitrap Fusion quadrupole mass spectrometer equipped with a Nanospray Flex ion source and stainless steel emitters. Peptides were analyzed using parallel reaction monitoring, or PRM [27], for which a full mass spectrum was acquired at 60,000 resolution in the Orbitrap sector using an m/z range of 350–2000. The survey scan was followed by up to 25 high-energy collisional MS/MS events triggered from an inclusion list representing the hypothetical peptide ions (see Table 1), with scanning of the collisional fragments at 15,000 resolution in the Orbitrap sector.

To quantify each isoform-specific peptide, peptide spectrum matches were reviewed in Byonic to identify the three most intense MS/MS fragment ions. These fragment ions were used to extract fragment-ion-specific chromatograms, using the parent ion scan range as an initial filter, and each of its top three fragments as a secondary filter (tolerance of mass error: 20 ppm). For each chromatogram thus created, the peaks were reviewed for specificity and elution times, the peak area of each fragment ion was calculated, and the areas of individual peptides' reporter fragments were summed to represent relative abundance of the parent peptide. Measured peak areas were compared to assess concomitant variation between dilution and peak area.

2.7. Quantification of the five SPHs in hemolymph by PRM of the isoform-specific reporters

After method validation using the purified proSPHs, aliquots (9.0 µl) of each plasma sample were analyzed by PRM. Values from three biological replicates at each developmental stage (two for adult) were averaged to represent levels of the SPHs. For comparing relative abundances of a specific SPH in different stages of *M. sexta*, the peak areas for all its PRM reporters in a bar-

stage (B) plasma sample and its biological replicates were used to normalize the corresponding peak areas of that SPH in the F, W, P, and A samples. The bar stage was selected since levels of the five SPHs were moderate when compared with those in the other four. Sums of the normalized peak areas from individual samples for the specific SPHs were used to calculate their relative levels in the five stages and to validate significant pairwise differences by Student's t-tests.

2.8. Cleavage of proSPH1a, 1b, 2, 4, and 101 by PAP3 and electrophoretic mobility changes

M. sexta PAP3 was isolated from pharate pupal hemolymph [28] and used to activate recombinant proPAP3 [29]. The proSPH1a and proSPH2 were expressed in insect cells, purified from the conditioned media [16], and used with proSPH1b, 4, and 101 as PAP3 substrates. The cleavage reactions, controls, and M_r markers were separated by 10% SDS and native PAGE, followed by electrotransfer and immunoblotting using antibody against the hexahistidine tag. To better understand the process of PAP3 cleavage and high M_r complex formation, aliquots of the proSPHs were incubated with different amounts of PAP3 for 1 h at 37 °C. The mixtures and PAP3 control were resolved by 10% SDS and native PAGE, transferred onto nitrocellulose membranes, and detected using SPH and $(His)_6$ antisera as primary antibody and goat-anti-rabbit/mouse IgG conjugated to alkaline phosphatase (Bio-Rad) as the secondary antibody, and a BCIP-NBT substrate kit (Bio-Rad) for color development.

2.9. ProPO activation and PO activity assay

M. sexta proPO was purified from hemolymph of naïve larvae and stored at -80 °C [30]. To activate proPO under different conditions, proPO (1 µl, 270 µg/ml) and PAP3 (1 µl, 40 ng) were incubated on ice with one of proSPH1a, 2b, 4 and 101 (1 µl, 200 ng), proSPH2 (1 µl, 200 ng), or both (1 µl + 1 µl) in a total volume of 20 µl buffer B (0.001% Tween-20, pH 7.5, 20 mM Tris-

HCl, 5 mM $CaCl_2$). Sixty minutes later, PO activity was determined on a microplate reader immediately after 150 µl of 2.0 mM dopamine in 50 mM sodium phosphate, pH 6.5, had been added to each sample well [31].

3. Results

3.1. Evolution of the SPHI and SPHII genes in lepidopteran insects

The genome of *M. sexta* contains 10 genes encoding clip-domain SPHs (1a, 1b, 2, 4, 42, 53, 92, 101, 221, 223), fewer than 18 in *D. melanogaster*, 19 in *Tribolium castaneum*, and 55 in *Anopheles gambiae* [12, 13, 14, 32]. To study the evolution of putative PAP cofactors, a subgroup of clip-domain SPHs, we identified the close homologs of *M. sexta* SPH1a and SPH2 in lepidopteran species, performed a phylogenetic analysis, and uncovered two main branches in the neighbor-joining tree, with *H. diomphalia* PPAF2 located near a hypothetical root (Fig. 1). Branch I is divided into two subgroups: SPHIA includes SPH1a, 1b, 4, and 101 in *M. sexta* (4) and their orthologs in *P. rapae* (3), *B. mori* (1), and *S. frugiperda* (1); SPHIB is a more ancient group consisting of SPHIA paralogs in *P. rapae* (1) and *S. frugiperda* (2). A more extended study of phylogenetic relationships revealed other SPHIB members in the butterflies (*B. anynana*, *P. machaon*, and *P. xuthus*) and moths (*A. plantaginis* and *H. armigera*) (Fig. S1). Function of the SPHIB members is unclear but can be similar to the SPHIAs' and lineage-specific gene amplifications gave rise to various members in the two subgroups. *A. plantaginis* has 3 SPHIBs and 0 SPHIA, *P. xuthus* has 4 SPHIBs and 3 SPHIAs, but all the other species have more SPHIAs than SPHIBs. There are 4 SPHIAs and 0 SPHIB in *M. sexta*. SPH4 emerged the first, SPH1a arose later, and duplication of SPH1b and SPH101 happened the last. Branch II diverged at an ancient time, likely earlier than the split of SPHI subgroups (Fig. S1). While most

of the species contain a single SPHII, the other lepidopterans have 2−8 members. For instance, *S. frugiperda* has 8. *A. plantaginis* has 6 and they belong to two subgroups, suggestive of another complex evolutionary process. In general, the splitting orders of branches I and II agree with the phylogeny of Lepidoptera and expansion of these SPH branches may be related to their putative functional importance as regulators of proPO activation.

3.2. Profiles of M. sexta SPH4, SPH1a, SPH1b, SPH101, and SPH2 transcript levels

Does *M. sexta* SPH2 have a function different from the four SPHIAs and is it possible to distinguish the roles of SPH4, SPH1a, SPH1b, and SPH101? An indirect way to address the questions is to compare levels of the SPHI and SPHII transcripts in multiple tissues at various developmental stages [12]. We first examined the RNA-seq data for each isoform (Fig. S2), finding that SPH2 and SPH1b are abundantly expressed in head, muscle, and fat body, and that their expression patterns are closely similar in the 52 tissue samples. SPH1a mRNA levels are on average 27-fold lower than SPH1b's. SPH1a expression is limited to muscle of pre-wandering and wandering larvae and testis of pupae; scarcer SPH101 transcripts are present in head; moderate levels of SPH4 mRNA are detected in some fat body samples. To confirm and extend the expression profiles, we determined the mRNA levels by qRT-PCR using highly specific primers for SPH4, SPH1a, SPH1b, SPH101, and SPH2. While a 10.3-fold increase in the SPH2 mRNA level was verified in fat body after an immune challenge [9], low but significant increases in SPH1b, 1a, and 101 were also detected (Fig. 2A). Patterns of the SPH1b and SPH2 expression in fat body of the feeding (F), wandering (W), and bar (B) stage larvae and pupae (P) are similar (Fig. 2B): lowest in the bar stage, higher in feeding and pupal stages, and highest in wandering stage. mRNA levels of SPH101, 4, and 1a were low in these four stages, except for SPH1a in pupal fat body. SPH1b mRNA was abundant in trachea and nervous tissue from

feeding larvae (Fig. 2C), at least 12-fold higher than the SPH1a mRNA levels in the same tissues. Little SPH4 and SPH101 transcripts were present in the five tissues. We detected high SPH2 expression concurring with the SPH1b expression in these tissue samples.

3.3. Levels of M. sexta SPH1a, 1b, 2, 4, and 101 proteins in hemolymph in different stages

While transcript profiling provided useful clues for functional analysis, mRNA levels do not correlate well with protein abundances [17]. On the other hand, because of the high sequence similarity (Fig. S3), it is difficult to determine the SPHIA protein levels in hemolymph by the shotgun approach. To solve this problem, we compared their amino acid sequences and identified 15, 3, 11, 4, and 35 theoretical unique peptides from SPH1a, 1b, 4, 101 and SPH2, respectively. Due to incomplete digestion or protein modification, only 13 of these 68 hypothetical peptides were detectable as peptides via trypsinolysis and LC-MS/MS using the purified proSPHs (see below): 2 for SPH1a, 1 for SPH1b, 4 for SPH4, 2 for SPH101, and 5 for SPH2 (Figs. 3A and S5). We then validated 9 of these 13 peptides as reporters of their parent proteins using a parallel reaction monitoring (PRM) method. Linear relationships between PRM peak areas and peptide amounts (Fig. S6) suggested that most of the unique peptides were specific and reliable quantifiers of their parent proteins within the ion intensity ranges assessed.

We then used these LC-MS/MS PRM assays to examine cell-free hemolymph samples for each SPH isoform. Only SPH1b, 2, and 4 were reliably detected in the plasma samples in the five stages (Figs. 3B and S7). Relative SPH1b protein abundance (8.8×10^7) was significantly higher in the feeding stage than in the bar (5.6×10^7), pupal (3.3×10^7), and adult (3.4×10^7) stages. A similar pattern was observed for SPH2 ($1.7–7.4\times10^7$), consistent with the finding that SPH1b and 2 mRNA levels in the 52 tissue samples were highly similar (Fig. S2). The abundances of SPH4 were relatively low and varied greatly ($7.0\times10^4–3.8\times10^6$). Its protein increase in W, B and P

hemolymph was consistent with the elevation of SPH4 transcripts in fat body and muscles of wandering larvae and bar-stage pharate pupae.

3.4. Recombinant expression, purification, and characterization of the proSPHs from Sf9 cells

Beginning to study functions of the five closely related SPHs, we amplified the full-length cDNAs of *M. sexta* SPH4, 1b, 101, and subcloned them into pMFH6 or pMFFH6. Upon transposition, the bacmids were isolated for transfecting insect cells and producing high-titer viral stocks through serial infection. Led by the honeybee mellitin signal peptide, the precursor proteins were efficiently secreted into media with the hexahistidine tag fused to their carboxyl-terminus. As proSPH1a and 2 were successfully prepared [16], we adopted the same protocol to isolate the other three proSPHs from the conditioned media by cationic and nickel affinity chromatography. From 100 ml of the media, 0.1, 0.8, and 0.3 mg proSPH4, 1b, and 101 were obtained, respectively. The proteins migrated to 49 (proSPH4), 60 (proSPH1a), 56 (proSPH1b), 59 (proSPH101) and 50 (proSPH2) kDa positions on a 10% SDS polyacrylamide gel (Fig. 4A, *left*). As judged based on the staining patterns, the proteins are near homogeneous, except for proSPH101. In addition to the precursor, a small amount of 39 kDa fragment of SPH101 contained the hexahistidine tag recognized by its antibody (Fig. 4A, *right*). Size of this cleavage product was close to the 36 kDa SPH1 isolated from hemolymph of bar-stage *M. sexta* pharate pupae [15], which was cleaved between Arg^{82} and Phe^{83} [26]. To test if the five proSPHs are glycosylated, we incubated them with *N*-glycosidase and observed mobility increases in SPH4, 1a, 1b, and 2 (Fig. 4B). Likewise, treatment of SPH1a, 1b, and 101 with *O*-glycosidases reduced their apparent M_r. These data are consistent with the predicted *N*- and *O*-glycosylation sites in the corresponding sequences (Fig. S3).

3.5. Sequential proteolytic processing of the five proSPHs by M. sexta PAP3

M. sexta SPH1 and SPH2 isolated from induced hemolymph of feeding larvae were cleaved next to R^{133} and R^{77}, respectively [9]. However, nearly half of the SPH1 in plasma of pharate pupae was cleaved after R^{82} (and the remaining still at R^{133}) [15, 26]. Predominant cleavage of the recombinant proSPH1a at R^{82}-F^{83} and proSPH2 at R^{77}-F^{78} by PAP3 yielded high M_r complexes acting as cofactor for proPO cleavage activation at R^{51}-F^{52} by a PAP. The first site in the SPHIAs is conserved in proSPH1a, 1b, and 101 all at R^{82}-F^{83}, but less so in proSPH4 at R^{63}-T^{64} (Fig. S3). The second site in SPH1a, 1b, and 101 are identical to each other at VR^{133}-TTGD and to SPH4 (VR^{107}-TTGD). As such, we suggest that proteolytic processing of the four SPHIAs may occur at the same sites.

Complete proteolysis of the proSPH4, 1b, and 101 yielded 31 kDa bands, indicating that PAP3 may have cut the substrate proteins eventually at the second site (Fig. 5A, *blue circle*). In contrast, a majority of the proSPH1a remained intact – PAP3 cleaved less than 1/5 of the precursor at the first site and yielded a 37 kDa C-terminal fragment co-migrating with the PAP3 catalytic chain (*red arrowhead*). Similarly, complete processing of proSPH2 by PAP3 at the first site gave rise to a 37 kDa SPH2 that migrated along with the PAP3 catalytic domain. In addition to these major products, minor bands were detected at 26 and 22 kDa positions, suggestive of cleavages at the 1st and additional sites. Roles of these cleavages in generating functional cofactor remained unclear. Native PAGE and immunoblot analysis suggested that processing of proSPH1b, 101, and 4 produced a high M_r smear extending from the stacking gel (Fig. 5B, *blue bar*). Partial digestion of proSPH1a yielded a shorter smear overlapping with the PAP3 (*red bar*). Size and intensity of the SPH4 and 1a smears in the stacking gel were lower than those of the SPH1b and 101. To our surprise, complete digestion of proSPH2 yielded a smear (*blue bar*) that migrated faster than its precursor (*green bar*).

To examine dynamic processes of the proSPH cleavage, we incubated aliquots of the five SPH precursors with decreasing amounts of PAP3, separated the reaction mixtures by 10% SDS-PAGE, and examined the products by immunoblot analysis using hexahistidine and specific antibodies (Fig. 6). As the lowest PAP3 level, the 49 kDa proSPH4 and 37 kDa SPH4 were detected as major bands (Fig. 6A, blot 1, lane 8), suggesting that cleavage at the first site occurred first. Intensity of the 31 kDa SPH4 increased as the PAP3 concentration increased and cleavage at the second site occurred to the 37 kDa SPH4 or 49 kDa proSPH4 directly but to a lesser extent. Antibodies specific to SPH4 yielded a similar result (Fig. 6B, blot 1). Precursor of the SPH4 migrated mainly as two broad bands on the native PAGE (Fig. 6C, blot 1, lane 2) and cleavage at the first and then second site yielded a smear in the stacking gel and a protein ladder in the separating gel. Ratios of the smear to ladder decreased as more PAP3 were added. While sequential proteolysis at the first and second sites seemed to also occur in proSPH1b and 101 (Fig. 6, A and B, blots 3 and 4), ratios of the smear to ladder were much higher in SPH1b and 101 than in SPH4. In other words, SPH1b and 101 have a stronger tendency to form high M_r complexes and stay in the stacking gel. Processing of proSPH1a was incomplete and complex – a majority of the precursor was not cleaved and some proteolysis occurred in an order of first, second, and then third sites (Fig. 6A, blot 2). The 26 kDa product was hardly recognized by the SPH1-specific antibodies (Fig. 6B, blot 2). Banding patterns of the reaction mixtures on native gel seemed to go through a decrease and then an increase in the amounts of high M_r smear in the stacking gel (Fig. 6C, blot 2), as the PAP3 level increased. Interestingly, SPH2 went through a decrease in gel mobility first, which is followed by a substantial increase in migration toward the anode (Fig. 6, A−C, blots 5). At the lowest level of PAP3 (lane 9), the 37 kDa product was at the peak level and, as PAP3 increased and proSPH2 diminished (lanes 3−8), the band intensity

decreased. This indicated that the first site was easily processed to form high M_r complexes and, after it was cleaved at additional sites (Fig. 5A), the complexes dissociated into pieces that migrated faster than proSPH2 on the native gel.

3.6. Effectiveness of the proteolytically processed SPHs as cofactors for proPO activation

We then tested if PAP3 processing of the four SPHI-II precursor pairs could yield active cofactors for proPO activation and how different amounts of PAP3 may affect the final PO activity levels. As shown in Fig. 7, all the SPHI-SPHII pairs generated PO activities much higher than those in the controls [16], indicating that SPH4-2, 1a-2, 1b-2, and 101-2 pairs were all active as cofactors. Along with the increases in PAP3 levels, the final PO activities increased correspondingly in the mixtures containing the same amounts of proPO, proSPHI, and proSPHII. In the presence of PAP3 at a particular level, the PO activities were highest in the reactions containing SPH101 and SPH1b, followed by SPH1a and then SPH4. This result agrees with our prediction based on evolutionary history of the SPHIAs, the newer ones are better partners of SPH2 than those emerged earlier (Fig. 1). The highest PO activity reached 522−527 U after proSPH1b or 101 had been activated by 40 ng PAP3 along with proSPH2 and 1.0 μg proPO. PO activities were negligible (≤4 U) in the control reactions (Fig. S8).

4. Discussion

4.1. Known functions of the clip-domain SPHs in holometabolous insects

Like S1A serine proteases, their noncatalytic homologs also play important roles in insect physiological processes, including digestion, development, and defense against pathogen infection [11]. While the role of serine proteases in digestion, for example, is easy to understand, it is more difficult to rationalize the function of SPHs in gut juice due to their lack of enzymatic

activity [29]. In this work, we took advantage of the close phylogenetic relationships of five clip-domain SPHs in *M. sexta* and demonstrated the role of SPH1a, 1b, 2, 4, and 101 as cofactors for the PAP-mediated proPO activation. The extension from SPH1a-SPH2 cofactor strengthens the previous finding that proPO activation is a complex reaction [33, 8, 9, 16] and further highlights the importance of PO, as five of the ten clip-domain SPHs in *M. sexta* are devoted to production of active PO. Moreover, expansion of the subgroup is not limited to *M. sexta*, Sphingidae, or Lepidoptera (Fig. S1). An extreme case was uncovered in the dipteran insect *A. gambiae*, which has as many as twelve close homologs of *M. sexta* SPH1a, 1b, 2, 4, and 101 [10, 11]. Further research is needed to validate the putative role of SPHIs and SPHIIs in proPO activation and test how widely this phenomenon spreads in various insects.

Clip-domain SPHs have functions other than the proPO activation reaction. In *M. sexta*, the other five clip-domain SPHs (42, 53, 92, 221, and 223) are likely involved in the regulation of development and immunity in *M. sexta*. The SPH53 is orthologous to *Drosophila* Masquerade, a regulator of somatic muscle attachment and axonal guidance [34]. Both proteins contain five clip domains followed by a protease-like domain. Scarface, the ortholog of *M. sexta* SPH221, regulates JNK-controlled morphogenic events such as embryonic dorsal closure [35]. In *A. gambiae*, SPCLIP1 (*i.e.*, CLIPA30), CLIPA8, and CLIPA28 form an SPH pathway to support localized CLIPC9 activation for melanization of *E. coli* and *Plasmodium berghei* [36, 37, 38]. In contrast, *A. gambiae* CLIPs A2, A5, and A14 down-regulate melanization [39, 40, 41].

4.2. Establishment of SPH1b as a major partner of SPH2 in M. sexta

We demonstrated herein that SPH1b, rather than SPH1a, is the primary partner of SPH2. The seemingly minor change from 1a to 1b was actually quite dramatic, accompanied by a series of observations worth discussing. At first, the fact that *M. sexta* SPH1a gene emerged before

SPH1b and 101 (Fig. 1) was somewhat surprising to us because, if positive selection played a major role in their evolution, SPH1b or 101 should be a better SPH2 partner than SPH1a or 4 is. Then, expression profiling by RNA-seq and qRT-PCR analyses raised more concerns (Fig. S2 and Fig. 2). While SPH1b and 2 transcription displayed a good agreement, SPH1a mRNA levels were much lower and limited to a few tissues at certain stages. As we did observe negative correlation of transcripts and translates for certain genes [17], is it possible that the SPH1a protein is more abundant than SPH1b in hemolymph? Therefore, to better understand the SPH functions in different life stages, we developed the LC-MS/MS method but did not detect SPH1a or 101 in hemolymph from larvae, pupae, or adults (Fig. 3B). Consistent with the transcriptome data, SPH1b and 2 protein levels were comparable in all five life stages when the plasma samples were taken. Finally, we decided to directly analyze the cofactor purified from bar-stage plasma [15] by trypsin digestion and mass spectrometry. The result clearly displayed that the high M_r complexes are mainly composed of SPH1b and 2 (Fig. 3B). The relative peak areas of SPH1b, 4, and 101 were 4.0×10^8, 2.6×10^7, and 2.4×10^3, respectively. And, since no SPH1a was identified, the original assignment of SPH1 to SPH1a was a mistake, partly because the first 22 residues of the SPH1b, 4, and 101 protease-like domains are only different from SPH1a at position 4, Thr^{154} substituted with Asn^{154} in SPH1a (Fig. S1). Signal of PTH-Thr (or Ser) is much lower than that of PTH-Asn at 1:1 molar ratio and signal-to-noise ratios largely decrease in later cycles of Edman degradation. Besides, the *M. sexta* genome sequence was not available until a decade later, allowing us to discern the small differences among SPH1a, 1b, 4, and 101 in this region or simply attribute the discrepancy to heterogeneity of our colony of insects.

The establishment of SPH1b-2 complex as a chief cofactor for proPO activation in *M. sexta* also disclosed a limitation of the shotgun approach we took to study the hemolymph proteomes,

that is, problematic assignment of peptides shared by two or more proteins in complex mixtures such as plasma. For instance, SPH1b levels were reported to be lower than those of SPH1a, SPH4, and SPH101 [23]. Inconsistency with the current data was not unexpected, because MaxQuant [43] at the default settings processed raw mass data not solely based on unique and, thus, rarer peptides to increase the hit numbers or sensitivity during database search. Protein groups were assembled using the principle of parsimony, meaning intensities of common peptides are assigned to the protein with the first unique peptide and other proteins were mostly disregarded as sources of the commons.

4.3. A parallel reaction monitoring method for quantifying closely related proteins in plasma

To overcome the limitation of shotgun technique, we adopted a targeted approach of proteomics analysis, known as parallel reaction monitoring (PRM) [24]. A major advantage of the method is its high specificity, which efficiently distinguishes multiple proteins with highly similar sequences (*e.g.*, SPH1b and SPH101: 96% identical) by detecting their unique peptides. With the five purified proSPHs available as references, we evaluated the degrees of difficulty in generation and detection of the predicted unique peptides and optimized the conditions of trypsin digestion based on preliminary experiments, such as mixing the proSPHs with plasma samples. While the relative levels of SPH1b and SPH2 were determined, our efforts to quantify SPH1a, 4, and 101 were less successful. Low protein levels and differences in degree of trypsin cleavage led to the failure of detection and large data variations, which is not uncommon for targeted proteomic studies [24]. In fact, the observed differences of unique peptides from the same protein were acceptable, and the magnitudes of unique peptides remained the same within individual samples and developmental stages.

4.4. Mechanism for proPO activation in insects

Generation of highly active PO in insects appears to be a conserved process more than is currently appreciated. In the silkworm, proPO activating enzyme alone generates active PO by cleaving proPO at Arg^{51}-Phe^{52} [43], although *B. mori* SPHI and SPHII both exist (Fig. 1) and may enhance proPO activation. In the beetle *H. diomphalia* or *T. molitor*, proPO is activated by *Hd*PPAF1 or *Tm*SPE, a clip-domain SP that cleaved proPO at Arg^{51}-Phe^{52} but does not yield active PO [5]. In this system, co-presence of *Hd*PPAF2/*Tm*SPH1, an SPHI activated by *Hd*PPAF3/*Tm*SPE, yielded active PO with one subunit cleaved at Arg^{163}-Ala^{164} in *Hd*PPO [30, 44, 45]. We have identified a unique SPHII in *T. molitor* (data not shown) and wondered if it could activate proPO to a higher level by working with *Tm*SPH1. In another beetle, *T. castaneum*, there are two SPHIs (cSPH3, 4) and one SPHII (cSPH2) [11] and, in light of this work, we predict that RNA interference of cSPH2 alone, 3 and 4 both, and 2 through 4 may negatively impact proPO activation and host survival after pathogen infection. In *M. sexta*, the four SPHIs, each pairing with the SPHII, formed PAP cofactors that yielded active PO at different levels (Fig. 7).

As reported before [46], extent of proPO cleavage at Arg^{51}-Phe^{52} increased after the cofactor (SPH1b-2) was included in the mixture of proPO and PAP1, and PO activity arose disproportionally. Such increases were also observed in *H. diomphalia* [30, 19]. By integrating the observation of a ternary complex of *M. sexta* PAP3, cofactor, and proPO [15], we propose that the high M_r complex of SPH1b and 2 may interact with both PAP and proPO to generate and stabilize highly active PO.

Characterization of PAP3-mediated proSPH1a, 1b, 2, 4, and 101 cleavage reactions provided new insights into the proPO activation in *M. sexta* and this understanding may be applicable to other insect systems. In *H. diomphalia*, PPAF2 forms a two-hexameric-toroidal structure to bind

and activate the 76 kDa PO (cleaved but inactive) [19]. Activation of proPO is more tightly controlled in *M. sexta*. The SPH1b-2 complex did not reactivate PO unless proPO was cleaved by PAP when the cofactor was present simultaneously [42]. In this study, we showed 1–3 cleavage sites in SPH1a, 1b, 2, 4 and 101, orders of these proteolytic events, and their electrophoretic behaviors on native gels (Fig. 5 and Fig. 6). Tendency of the four SPHIs to form high M_r complexes appeared to have a positive correlation with the PO activity level generated (Fig. 7). It would be interesting to find out how 37 kDa cleaved SPH2, which experienced a mobility decrease first and then an increase due to cleavage at additional sites, may associate with the SPHI complexes, interact with PAP, and stabilize active PO. We are exploring roles of the cleavage events in SPH1b and SPH2 in the formation of active cofactors and how activation of the two proSPHs together by PAP3 may differ from those in the separate reactions (Fig. 6C). In our current model, SPH2 oligomers act as adapters that associate with the SPHI scaffolds, orient PAP and proPO for cleavage under optimal conditions, and perhaps exchange with the leaving PAP to maintain PO in the active conformation.

4.5. Concluding remarks

In summary, we analyzed the evolutionary relationships of *M. sexta* SPH2, 4, 1a, 1b, and 101 in the context of their orthologs in lepidopteran insects. Expression profiling provided us useful clues to understand their functional differences. The five clip-domain SPH precursors were used as standards in the development of a PRM method to distinguish the highly similar SPHs and measure their relative abundances in complex mixtures of hemolymph proteins. Based on the mRNA and protein data, we established SPH1b-2 pair as the major cofactor for proPO activation by PAPs in *M. sexta*. The recombinant proSPHs were tested with PAP3 and proPO to reveal the reaction mechanisms. We identified a positive correlation among order of SPHI appearance in

evolution, formation of high M_r complexes of SPHI, and level of the cofactor activity, suggestive of a positive selection for a more efficient SPHI-II pair as the cofactor. Future research in *B. mori*, *T. castaneum*, *A. mellifera*, *A. gambiae*, and *D. melanogaster* may provide support for the conserved requirement of an SPHI-II as cofactor for generating highly active PO in insects.

Figures

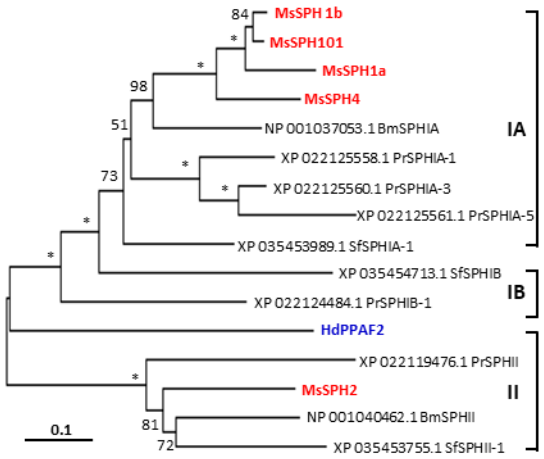


Fig. 1. Phylogenetic relationships of the representative members of SPHI and SPHII in some lepidopteran species in relation to the beetle clip-domain SPH. The neighbor-joining tree was constructed as described in *Section 2.2*. While topology of the specific SPHs in *Bombyx mori* (Bm), *Spodoptera frugiperda* (Sf), *Pieris rapae* (Pr), and *M. sexta* (Ms) is outlined here, results of a more extensive analysis are shown in Fig. S2, supporting the claims made in *Section 3.1*. The *Manduca* SPHIAs (1a, 1b, 4, and 101) and SPHII (2) are in *red* font. *H. diomphalia* proPO activating factor-2 (HdPPAF2, *blue* font) [19] is used as an outgroup for the tree. Percentage bootstrap values greater than 50 are marked at the corresponding nodes, with values ≥99 simplified as "*". Brackets and names (IA, IB, II) represent the SPH subgroups.

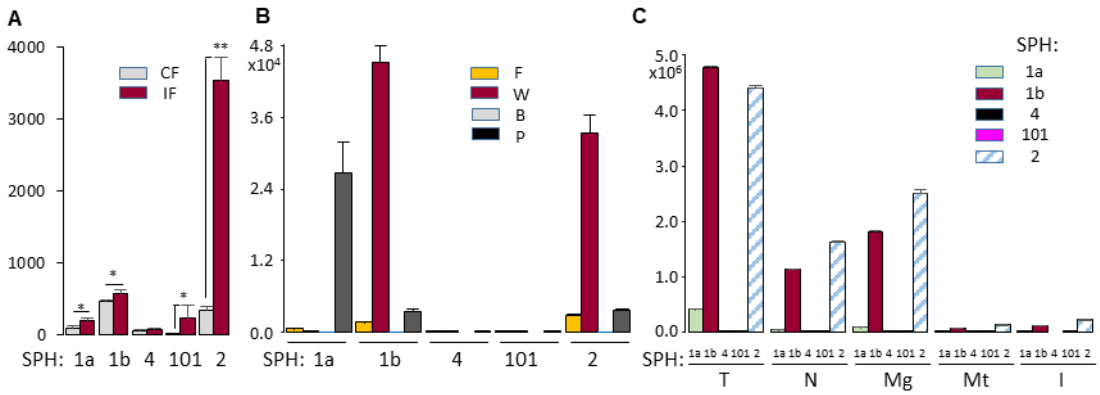
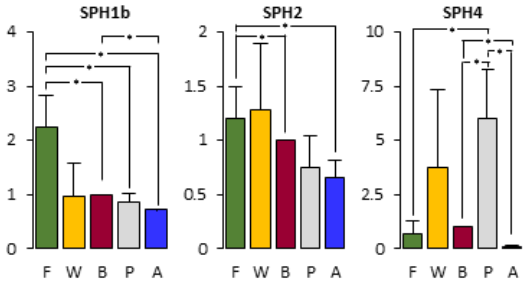


Fig. 2. *M. sexta* SPHI and SPHII mRNA immune inducibility and expression profiles in various tissues and developmental stages. The mRNA levels in **c**ontrol and **i**nduced **f**at body (CF and IF) **(A)**, fat body in four life stages **(B)**, and five tissues in day 3, 5th instar larvae **(C)** are calculated based on the C_t values from three biological replicates and plotted as bar graphs (mean ± SEM, n = 3). For panel A, pairwise comparisons were performed between control and induced samples using Student's t-test (*, $p < 0.05$). For panel B, fat body RNA samples from the **f**eeding (F) and **w**andering (W) larvae, **b**ar-stage (B) prepupae, and **p**upae (P) were prepared for qPCR analysis (*Section 2.5*). For panel C, profiles of SPHI and SPHII mRNA in **i**ntegument (I), **m**id**g**ut (Mg), **M**alpighian **t**ubules (Mt), **n**ervous tissue (N), and **t**rachea (T) were determined by the same method.

A

Protein	Predicted unique peptide	Identified unique peptide by Byonic	m/z	Scan time (min) Byonic	Scan time (min) PRM
SPH101	(R)APSTVVPSVPQGNAADLSCK*	APSTVVPSVPQGNAADLSCK	666.67	66	63
SPH101	(R)YFELHSSFMCAGGESGR*	YFELHSSFMCAGGESGR	813.67	61	61
SPH4	(R)IFGFPFTSK*	IFGFPFTSK	522.28	94.3	93
SPH4	(R)PEVTPK*	RPEVTPKPELK	431.92	24.5	n.d.
SPH4	(R)AGEWDTQHTK*	AGEWDTQHTKEPYPHQDR	732.33	31	n.d.
SPH4	(K)ITNTIINDGTDIIDIR*	ITNTIINDGTDIIDIR	893.98	94	92
SPH1a	(R)KTQLGQFFELHSSFMCAGGEPGR*	KTQLGQFFELHSSFMCAGGEPGR	646.81	56	54
SPH1a	(K)ICEGDGGSPLVCPNEYEK*	ICEGDGGSPLVCPNEYEK	765.67	84.8	n.d.
SPH1b	(R)PETDPIIPK*	PETDPIIPKPDPVKPER	521.79	38	39
SPH2	(K)DGLTINDPTLDGNGLLDIR*	DGLTINDPTLDGNGLLDIR	1006.52	109	105
SPH2	(R)LCQESVER*	FGEDENNDRLCQESVER	669.64	46	n.d.
SPH2	(K)DPSEVVK*, (K)PDPSK*	VLKDPSEVVKPKPDPSK	621.69	30	29

B



	F	W	B	P	A	purified
SPH1a	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
SPH1b	8.53×10^{7}	3.46×10^{7}	5.63×10^{7}	3.34×10^{7}	3.41×10^{7}	4.00×10^{8}
SPH2	7.36×10^{7}	7.17×10^{7}	4.67×10^{7}	2.84×10^{7}	1.71×10^{7}	4.86×10^{7}
SPH4	3.30×10^{5}	2.54×10^{6}	1.89×10^{6}	3.77×10^{6}	7.04×10^{4}	2.59×10^{7}
SPH101	n.d.	n.d.	n.d.	n.d.	n.d.	2.43×10^{3}

Fig. 3. Determination of the *M. sexta* SPHI and SPHII protein levels in cell-free hemolymph samples from feeding and wandering larvae, bar-stage prepupae, pupae and adults by targeted

LC-MS/MS analysis. **(A)** A list of identified unique trypsinolytic peptides from the recombinant proSPH1a, 1b, 4, 101, and 2 as individuals or in mixtures. **(B)** Relative abundances of the selected hemolymph proteins in the five life stages. In relative to bar (B) stage, peak areas (mean ± SEM, n = 3) of the isoform-specific peptide reporters (Fig. S7) in feeding (F), wandering (W), pupal (P), and adult (A) stages for SPH1b, 2, and 4 are shown in the bar graphs. Significant differences are indicated by "*" ($p < 0.05$) between stages. Table in the bottom shows relative abundances (*i.e.*, average peak areas in chromatograms) of *M. sexta* SPHIs and SPHII in hemolymph from insects in the five developmental stages and the cofactor purified from bar-stage hemolymph [15]. n.d., not detected.

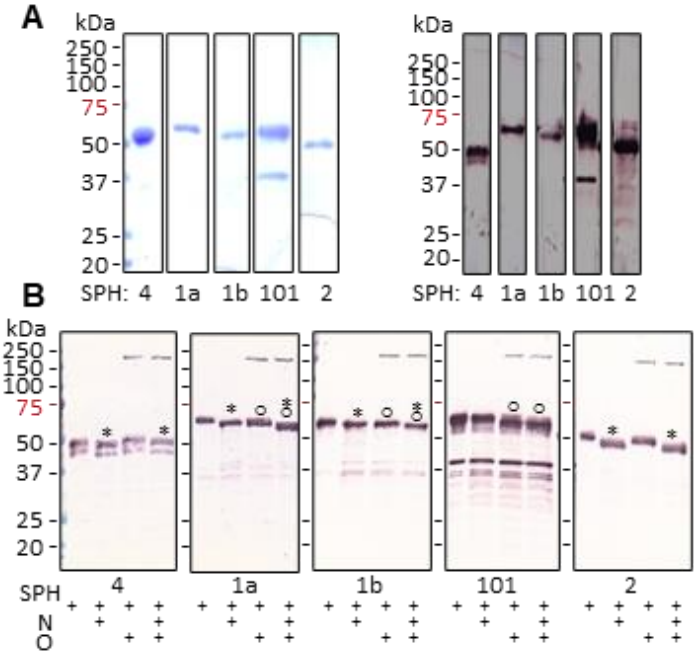


Fig. 4. SDS-PAGE and immunoblot analysis of the purified *M. sexta* proSPH4, 1a, 1b, 101, and 2 **(A)** and their deglycosylation products **(B)**. For panel A, the purified proSPHs were resolved by 10% SDS-PAGE followed by Coomassie Brilliant Blue (CBB) staining (*left*, 1 μg) or immunoblotting (*right*, 200 ng) using 1:1000 diluted antibody against 6×histidine affinity tag. For panel B, the purified proSPHs (200 ng) were separately treated with buffer, PNGase F (N),

neuraminidase A and O-glycosidase (O), or all three enzymes (N+O) prior to 10% SDS-PAGE, electrotransfer, and immunoblot analysis using the antibody to the affinity tag. The small mobility increases caused by *N*- and *O*-deglycosylation are indicated by "*" and "○", respectively. Positions and sizes (in kDa) of the prestained M_r standards are marked on the *left*, with the 75 kDa marker highlighted *red*.

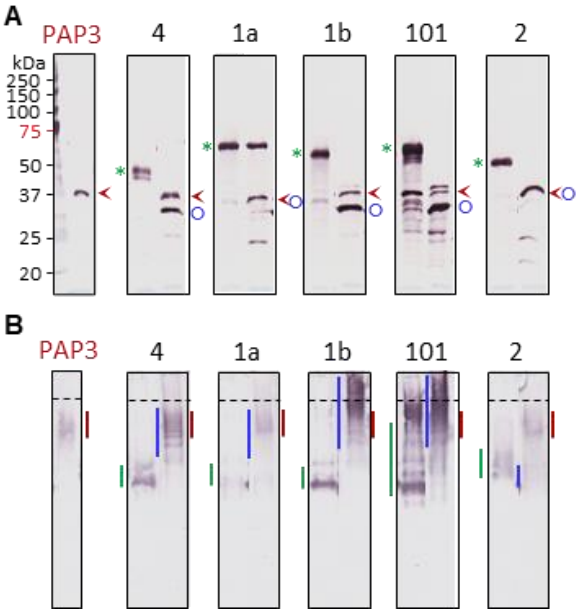


Fig. 5. Immunoblot analysis of the PAP3 proteolytic products of *M. sexta* SPH4, 1a, 1b, 101 and 2 precursors following 10% SDS **(A)** and native **(B)** polyacrylamide gel electrophoresis (PAGE). The purified proSPHs (200 ng/µl, 1 µl) were separately incubated with PAP3 (100 ng/µl, 1 µl) in 10 µl buffer B (0.001% Tween-20, pH 7.5, 20 mM Tris-HCl, 5 mM $CaCl_2$) at 37 °C for 1 h. The reaction mixtures and controls (40 ng PAP3 and 200 ng proSPH) were treated with 1×SDS sample buffer at 95 °C for 5 min or the same buffer lacking SDS and DTT at 25 °C for 5 min prior to 10% SDS-PAGE and 10% native PAGE. After electrotransfer, immunoblot analysis was performed using 1:1000 diluted antibody against the hexahistidine tag. In panel A, the catalytic domain of PAP3, proSPHs (*left* lanes), and major cleavage products (*right* lanes) are marked by *red arrowheads*, *green asterisks*, and *blue circles*, respectively. Positions and sizes of the

prestained M_r standards are indicated, with the 75 kDa marker highlighted *red*. In panel B, the dashed line divides the stacking and separating gels. The smeared bands of PAP3, proSPHs (*left* lanes), and cleavage products (*right* lanes) are marked by *red*, *green*, and *blue vertical bars*, respectively. The *blue* line in the middle points to the processed SPHs on the *right*.

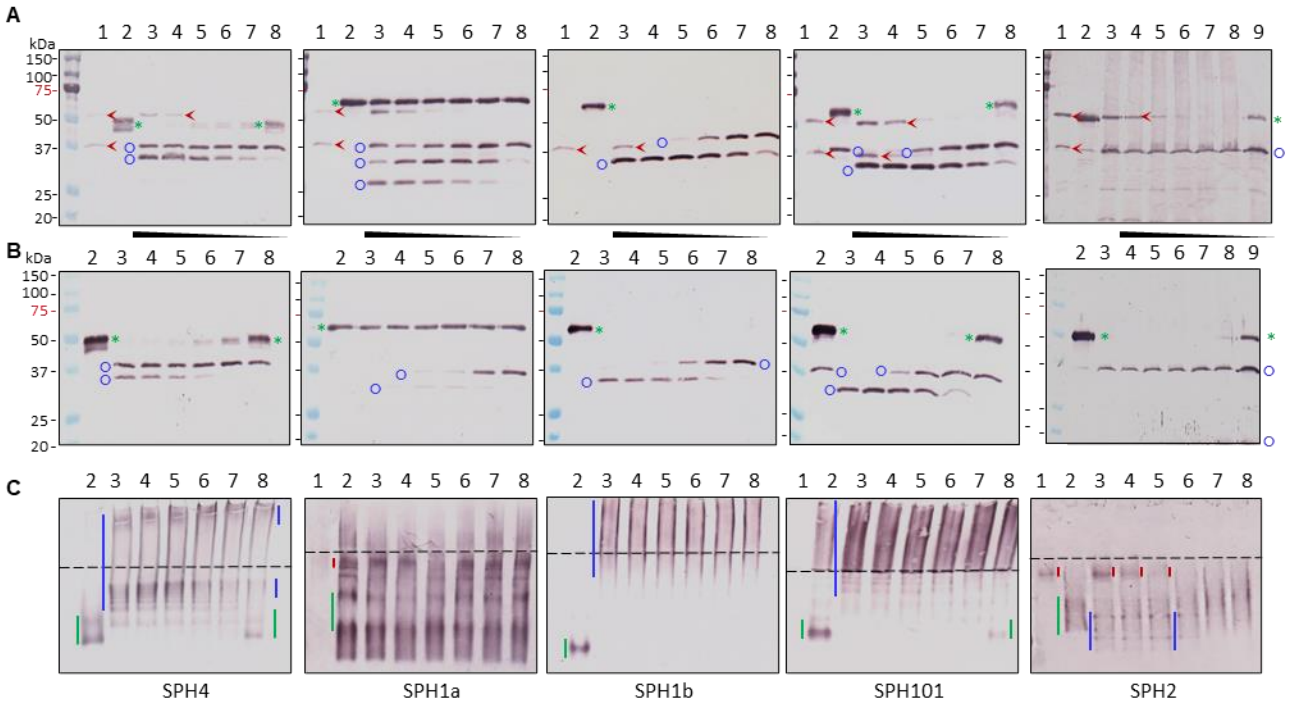


Fig. 6. PAP3 processing of *M. sexta* proSPH4, 1a, 1b, 101, and 2 analyzed by 10% SDS and native PAGE followed by immunoblotting. Aliquots of the purified proSPHs (200 ng) were incubated with 80 (lane 3), 40 (lane 4), 20 (lane 5), 10 (lane 6), 5 (lane 7), 2.5 (lane 8), 1.3 (lane 9 for proSPH2 only), and 0 (lane 2) ng of PAP3 in 10 µl buffer B at 37 °C for 1 h. The mixtures and PAP3 control (40 ng, lane 1) were subjected to 10% SDS-PAGE (**A** and **B**) and 10% native (**C**) PAGE, as described in the legend to Fig. 5. After electrotransfer, immunoblot analysis was performed using antibody against the hexahistidine tag (**A** and **C**) or specific SPHs (**B**) as the first antibody. The SPH1a, 1b, and 101 were all recognized by antibodies to SPH1a (residues 136–396) [9]; the SPH2 was recognized by antibodies to SPH2 (residues 79–385) [9]; SPH4 by antibodies to the entire protein (this work). In panels A and B, PAP3, proSPHs, and major

cleavage products are marked by *red arrowheads*, *green asterisks*, and *blue circles*, respectively. Positions and sizes of the prestained M_r standards are indicated, with the 75 kDa marker highlighted *red*. In panel C, the dashed line divides the stacking and separating gels. The smeared bands of PAP3, proSPHs, and cleavage products are marked by *red*, *green*, and *blue vertical bars*, respectively.

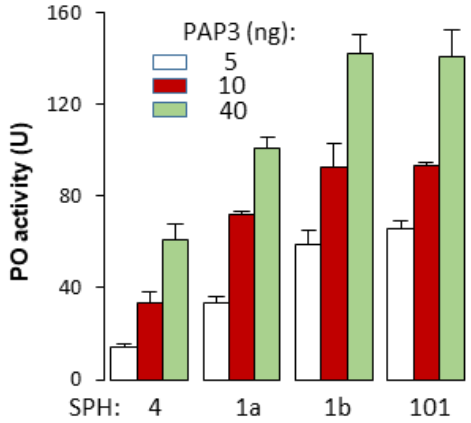


Fig. 7. Concentration-dependent activation of *M. sexta* proPO by PAP3 in the presence of SPHI-II pairs. The purified proPO (0.27 μg) and PAP3 (5, 10, and 10 ng) were incubated for 1 h on ice with proSPHI (200 ng 1a, 1b, 101 or 4), proSPH2 (200 ng), or both SPHI and SPH2 in 20 μl buffer B. PO activity was measured using dopamine (2 mM, 150 μl/well) and plotted in the bar graph (mean ± SEM, n = 3). The control reactions and their results are shown in Fig. S8.

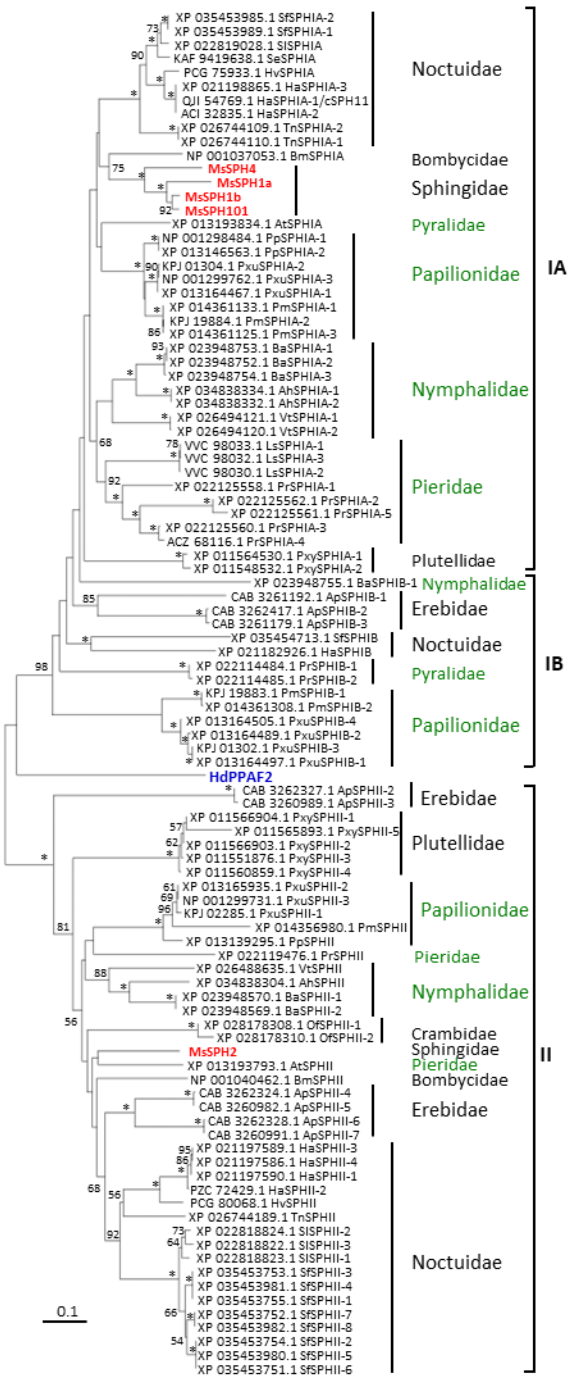


Fig. S1. Phylogenetic relationships of the SPHI and SPHII subfamily members in lepidopteran insects. Species names and their abbreviations in parenthesis are *Aphantopus hyperantus* (Ah), *Arctia plantaginis* (Ap), *Amyelois transitella* (At), *Bicyclus anynana* (Ba), *Bombyx mori* (Bm), *Helicoverpa armigera* (Ha), *Heliothis virescens* (Hv),

Holotrichia diomphalia (Hd), *Leptidea sinapis* (Ls), *Manduca sexta* (Ms), *Spodoptera exigua* (Se), *Spodoptera frugiperda* (Sf), *Spodoptera litura* (Sl), *Trichoplusia ni* (Tn), *Papilio machaon* (Pm), *Papilio polytes* (Pp), *Papilio xuthus* (Pxu), *Pieris rapae* (Pr), *Plutella xylostella* (Pxy), *Ostrinia furnacalis* (Of), and *Vanessa tameamea* (Vt). Their family and superfamily names are listed on the *right*, with the butterflies' in *green* font. Suggested protein names consist of an abbreviated species name, a group name (SPHIA, SPHIB, or SPHII), and a serial number (-1, -2, -3, …). The *M. sexta* SPHIAs *(i.e.*, 1a, 1b, 4, and 101) and SPHII (*i.e.*, 2) are in *red bold* font. *H. diomphalia* proPO activating factor-2 (*Hd*PPAF2, in *blue bold* font) [19] is included as an outgroup for the phylogenetic tree.

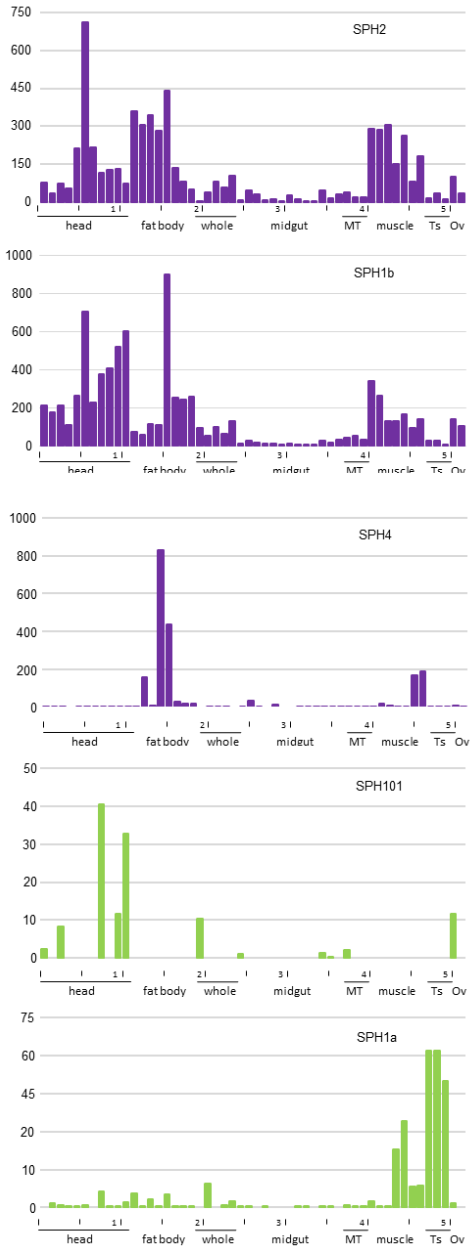


Fig. S2. Expression of *M. sexta* SPH2, 1a, 1b, 4, and 101 in 52 tissue samples. The relative mRNA levels, as represented by FPKM values, are shown in bar graphs. The 52 cDNA libraries **1** through **52** were constructed from the following tissues and stages: head [**1**, 2nd (instar) L (larvae), d1 (day 1); **2**, 3rd L, d1; **3**, 4th L, d0.5; **4**. 4th L, late; **5**. 5th L, d0.5; **6**. 5th L, d2; **7**. 5th L, pre-W (pre-wandering); **8**. P (pupae), late; **9**. A (adults), d1; **10**. A, d3; **11**. A, d7], fat body (**12**. 4th L, late; **13**. 5th L, d1; **14**. 5th L, pre-W; **15**. 5th L, W; **16**. P, d1-3; **17**. P, d15-18; **18**. A, d1-3;

19. A, d7-9), whole animals [**20**. E (embryos), 3h; **21**. E, late; **22**. 1st L; **23**. 2nd L; **24**. 3rd L), midgut (**25**. 2nd L; **26**. 3rd L; **27**. 4th L, 12h; **28**. 4th L, late; **29**. 5th L, 1-3h; **30**. 5th L, 24h; **31**. 5th L, pre-W; **32**–**33**. 5th L, W; **34**. P, d1; **35**. P, d15-18; **36**. A, d3-5; **37**. 4th L, 0h), Malpighian tubules (MT) (**38**. 5th L, pre-W; **39**. A, d1; **40**. A, d3), muscle (**41**. 4th L, late; **42**–**43**. 5th L, 12h; **44**–**45**. 5th L, pre-W; **46**–**47**. 5th L, W), testis (Ts, **48**. P, d3; **49**. P, d15-18; **50**. A, d1-3), and ovary (Ov, **51**. P, d15-18; **52**. A, d1).

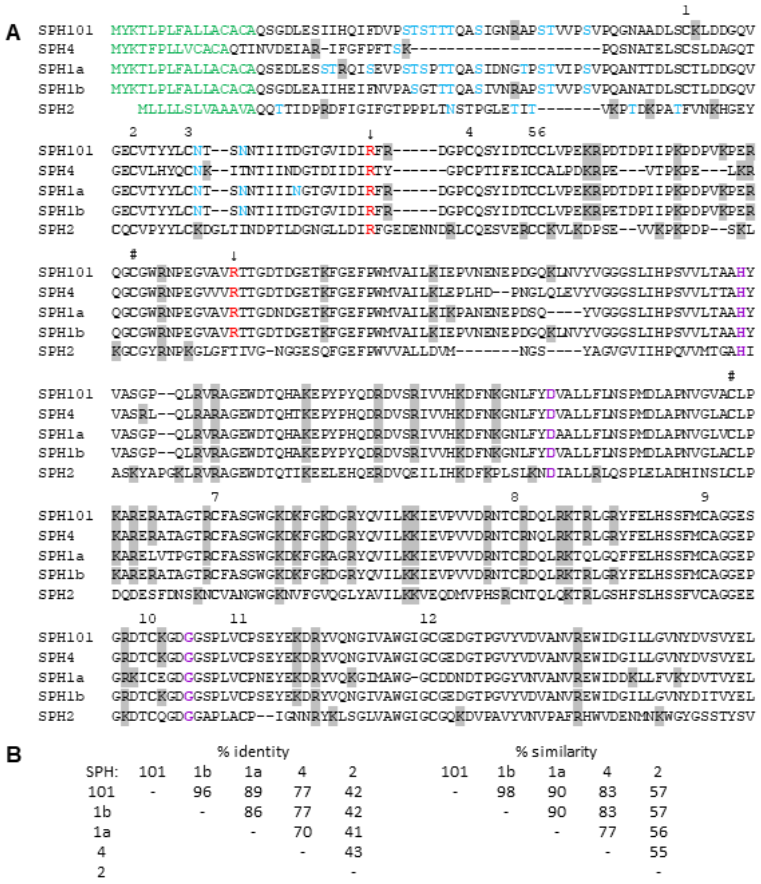


Fig. S3. Aligned sequences of *M. sexta* SPH1a, 1b, 4, 101, and SPH2. **(A)** Multiple sequence alignment was performed using MUSCLE, a module of MEGA-X. The signal peptide for secretion is shown in *green* font. Arg and Lys residues are *shaded grey* to show that trypsin cleaves next to these positions. The cleavage sites in SPH1b [9, 15] and their corresponding Arg residues in the SPHs are indicated by ↓ and in *red bold* font. Residues in the altered catalytic

triad (His-Asp-Gly) are in *purple bold* font. Predicted *N*- and *O*-linked glycosylation sites are in *blue* font. Disulfide bonds in the clip domain (Cys-1 and Cys-5, Cys-2 and Cys-4; Cys-3 and Cys-6), in the protease-like domain (Cys-7 and Cys-11, Cys-8 and Cys-9, Cys-10 and Cys-12), and between the two domains (# and #) are highlighted in *bold* font (19 Piao et al., 2005). **(B)** Percentage identity and similarity of amino acid sequence pairs.

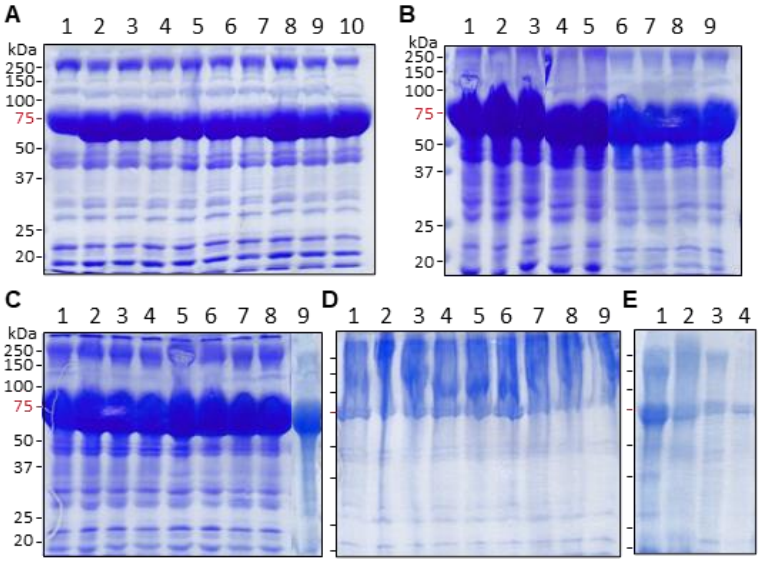
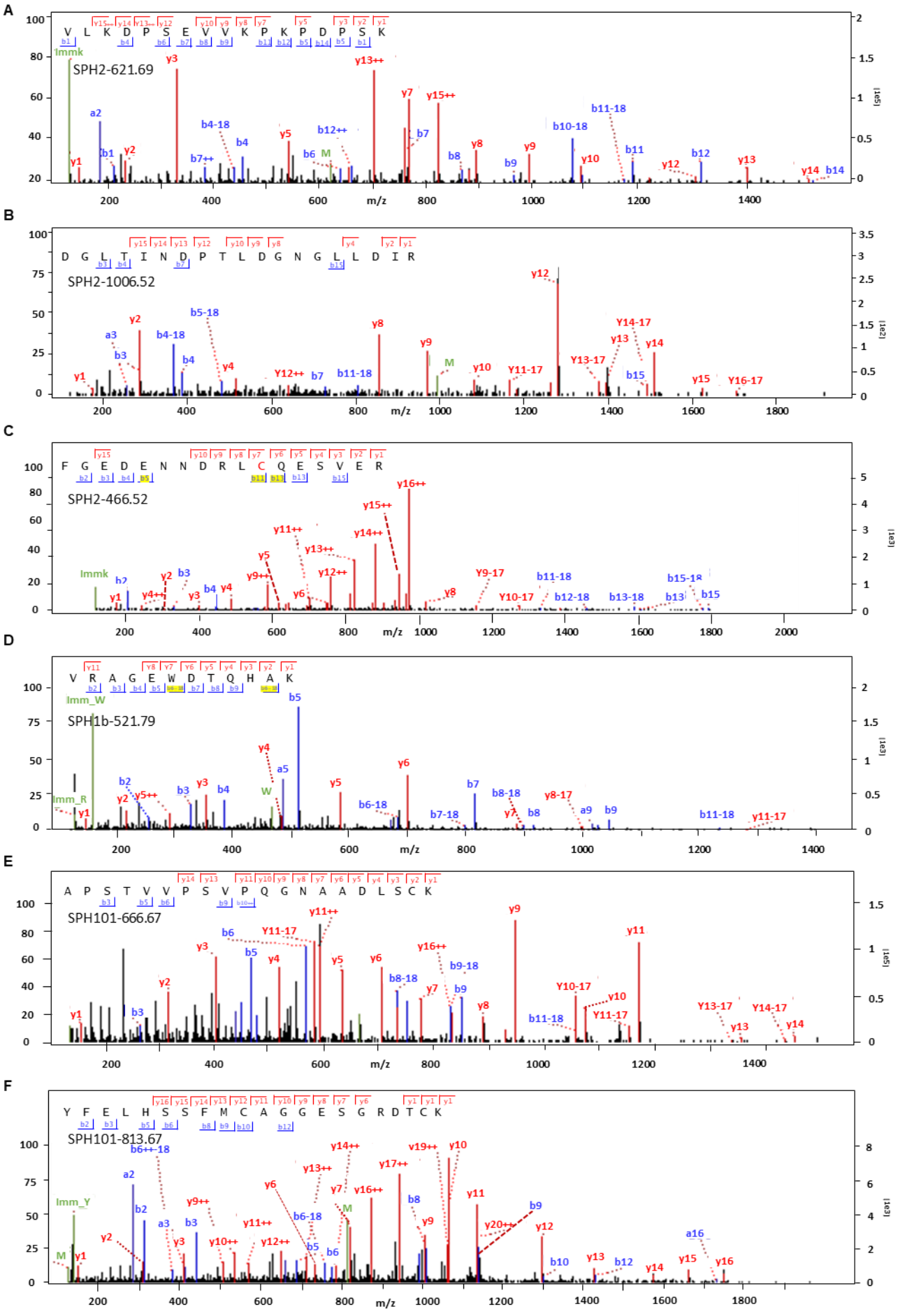


Fig. S4. SDS-PAGE analysis of cell-free hemolymph from feeding larvae **(A)**, wandering larvae **(B)**, bar-stage pharate pupae **(C)**, pupae **(D)**, and adults **(E)**. One microliter of each plasma sample was diluted with 4 µl ddH_2O, treated with 5 µl 2×SDS sample buffer in the presence of DTT, separated on 10% SDS polyacrylamide gels, and stained by Coomassie Brilliant Blue.


A
V L K D P S E V V K P K P D P S K
SPH2-621.69
B
D G L T I N D P T L D G N G L L D I R
SPH2-1006.52
C
F G E D E N N D R L C Q E S V E R
SPH2-466.52
D
V R A G E W D T Q H A K
SPH1b-521.79
E
A P S T V V P S V P Q G N A A D L S C K
SPH101-666.67
F
Y F E L H S S F M C A G G E S G R D T C K
SPH101-813.67
m/z
[as]

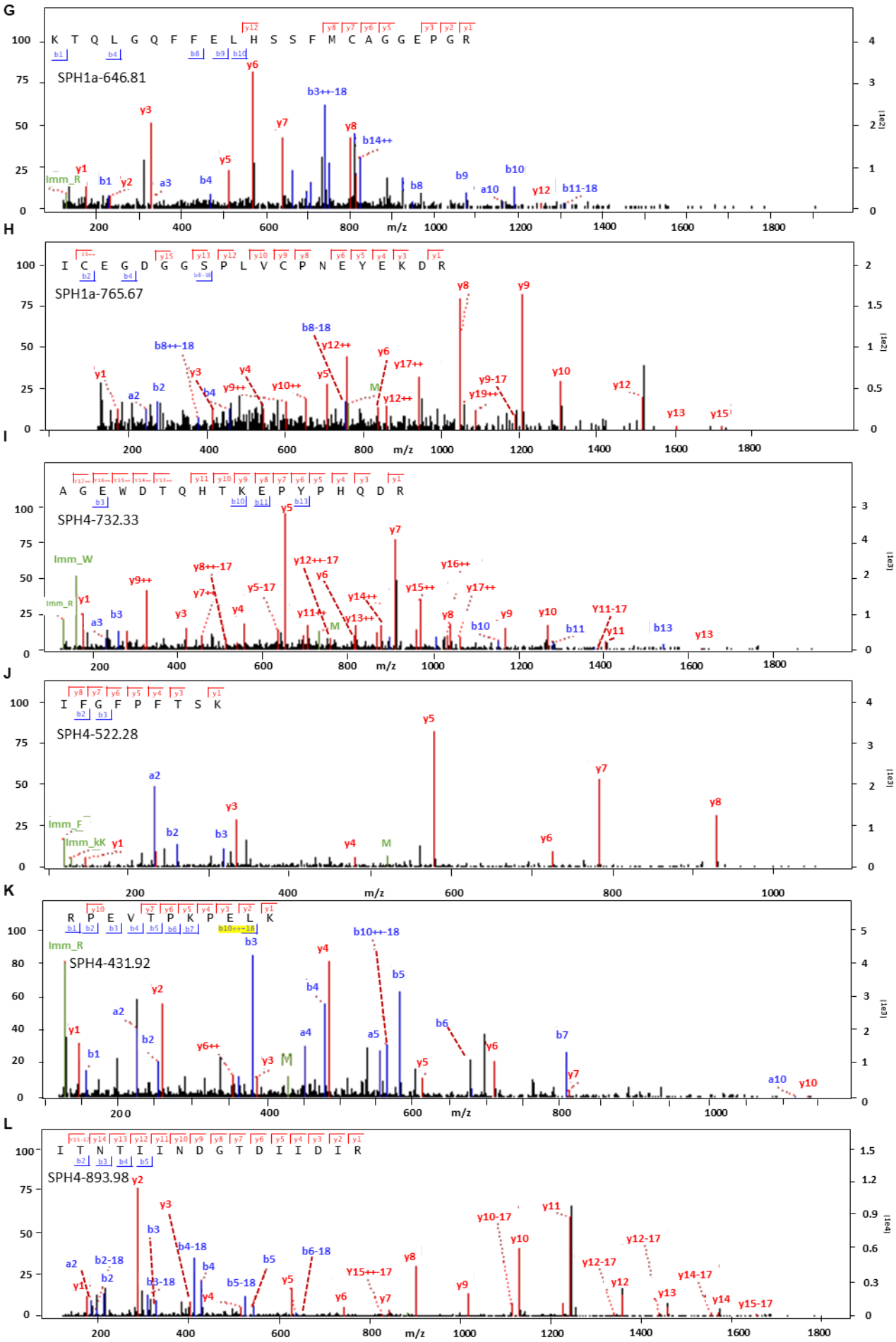

G
K T Q L G Q F F E L H S S F M C A G G E P G R
SPH1a-646.81
Imm_R
H
I C E G D G G S P L V C P N E Y E K D R
SPH1a-765.67
I
A G E W D T Q H T K E P Y P H Q D R
SPH4-732.33
Imm_W
Imm_R
J
I F G F P F T S K
SPH4-522.28
Imm_F
Imm_kK
K
R P E V T P K P E L K
SPH4-431.92
Imm_R
L
I T N T I I N D G T D I I D I R
SPH4-893.98
m/z
[a.u.]

Fig. S5. Second MS spectra of trypsinolytic peptides from the purified proSPH2 (**A**–**C**), 1b (**D**), 101 (**E**, **F**), 1a (**G**, **H**), and 4 (**I**–**L**). **(A)** VLKDPSEVVKPKPDPSK, **(B)** DGLTINDPTLDGNGLLDIR, **(C)** FGEDENNDRLCQESVER, **(D)** VRAGEWDTQHAK, **(E)** APSTVVPSVPQGNAADLSCK, **(F)** YFELHSSFMCAGGESGRDTCK, **(G)** KTQLGQFFELHSSFMCAGGEPGR, **(H)** ICEGDGGSPLVCPNEYEKDR, **(I)** AGEWDTQHTKEPYPHQDR, **(J)** IFGFPFTSK, **(K)** PEVTPKPELK, **(L)** ITNTIINDGTDIIDIR.

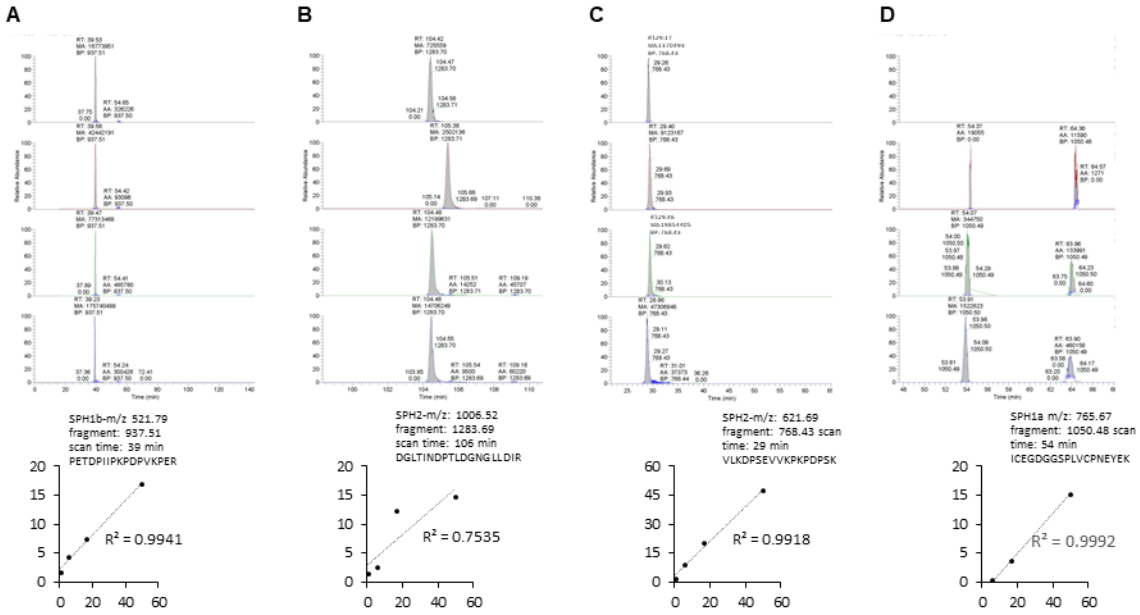


Fig. S6. Validation of the eight PRM reporters and their dilution responses. for detected reporter of SPH1 and SPH2, purified protein was diluted as supplement shown, the different dilutions of SPH1 isoforms and SPH2 samples were detected by HPLC using PRM method. The trendlines using peak areas for different dilutions of SPH1s and SPH2 were shown.

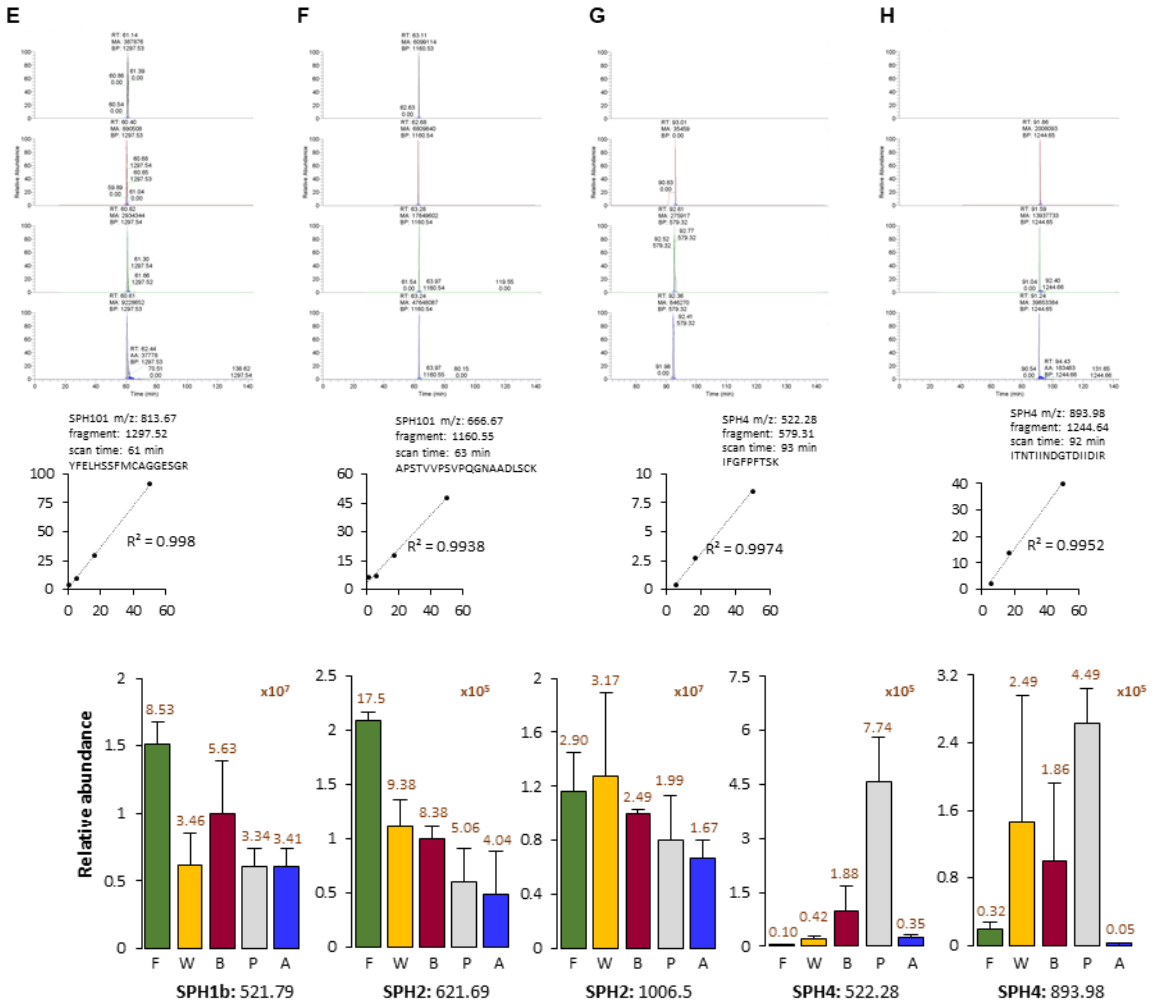


Fig. S7. Relative abundances of the five PRM peptide reporters for the hemolymph SPHs in the five developmental stages. Relative and absolute peak areas (mean ± SEM, n = 3) in F (feeding stage), W (wandering stage), B (bar stage), P (pupal stage) and A (Adult stage) for SPH1b, 2, and 4 are shown on y-axis and labeled in *maroon* font, respectively. m/z of the peptide ion was labeled in the bottom.

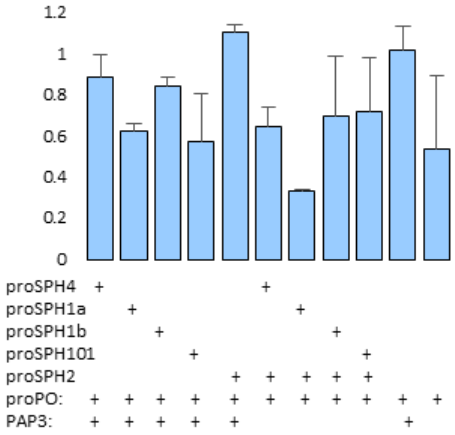


Fig. S8. Control reactions for *M. sexta* proPO activation by SPHI, SPH2, and PAP3. The purified proPO (0.27 μg) alone, its mixture with PAP3 (40 ng), proPO mixed with PAP3 and a single proSPH (200 ng 1a, 1b, 101, 4, or 2), and proPO mixed with a pair of proSPHI and II (200 ng each, 1a-2, 1b-2, 4-2, or 101-2) in 20 μl buffer B were incubated on ice for 1 h in 20 μl buffer B. PO activity was measured using dopamine (2 mM, 150 μl/well) and plotted in the bar graph (mean ± SEM, n = 3).

Acknowledgments

This work was supported by National Institutes of Health Grants GM58634 and AI112662. We thank Drs. Michael Kanost, Xiaoqiang Yu, Zhiqiang Lu, and the anonymous reviewers for their critical comments which helped us to improve the manuscript. Dr. Xiaoqiang Yu at University of Missouri-Kansas City kindly provided polyclonal antisera to *M. sexta* SPH1a and SPH2. The mass spectrometry assays were developed and performed in the Center for Genomics and Proteomics at Oklahoma State University. This article was approved for publication by the Director of Oklahoma Agricultural Experimental Station and supported in part under project OKL03054.

References

1. Zhao P, Li J, Wang Y, Jiang H. Broad-spectrum antimicrobial activity of the reactive compounds generated in vitro by Manduca sexta phenoloxidase. Insect Biochem Mol Biol. 2007 Sep;37(9):952-9.
2. Zhao P, Lu Z, Strand MR, Jiang H. Antiviral, anti-parasitic, and cytotoxic effects of 5,6-dihydroxyindole (DHI), a reactive compound generated by phenoloxidase during insect immune response. Insect Biochem Mol Biol. 2011 Sep;41(9):645-52.
3. Nappi AJ, Christensen BM. Melanogenesis and associated cytotoxic reactions: applications to insect innate immunity. Insect Biochem Mol Biol. 2005 May;35(5):443-59.
4. Jiang H, Vilcinskas A, Kanost MR. Immunity in lepidopteran insects. Adv Exp Med Biol. 2010;708:181-204.
5. Park JW, Kim CH, Rui J, Park KH, Ryu KH, Chai JH, Hwang HO, Kurokawa K, Ha NC, Söderhill I, Söderhill K, Lee BL. Beetle immunity. Adv Exp Med Biol. 2010;708:163-80.
6. Kanost MR, Jiang H, Yu XQ. Innate immune responses of a lepidopteran insect, Manduca sexta. Immunol Rev. 2004 Apr;198:97-105.
7. Kanost MR, Jiang H. Clip-domain serine proteases as immune factors in insect hemolymph. Curr Opin Insect Sci. 2015 Oct 1;11:47-55.
8. Kwon TH, Kim MS, Choi HW, Joo CH, Cho MY, Lee BL. A masquerade-like serine proteinase homologue is necessary for phenoloxidase activity in the coleopteran insect, Holotrichia diomphalia larvae. Eur J Biochem. 2000 Oct;267(20):6188-96.
9. Yu XQ, Jiang H, Wang Y, Kanost MR. Nonproteolytic serine proteinase homologs are involved in prophenoloxidase activation in the tobacco hornworm, Manduca sexta. Insect Biochem Mol Biol. 2003 Feb;33(2):197-208.
10. Jiang H, Wang Y, Kanost MR. Pro-phenol oxidase activating proteinase from an insect, Manduca sexta: a bacteria-inducible protein similar to Drosophila easter. Proc Natl Acad Sci U S A. 1998 Oct 13;95(21):12220-5.
11. Kim MS, Baek MJ, Lee MH, Park JW, Lee SY, Soderhall K, et al. A new easter-type serine protease cleaves a masquerade-like protein during prophenoloxidase activation in Holotrichia diomphalia larvae. J Biol Chem. 2002 Oct 18;277(42):39999-40004.
12. Cao X, Jiang H. Integrated modeling of protein-coding genes in the Manduca sexta genome

using RNA-Seq data from the biochemical model insect. Insect Biochem Mol Biol. 2015 Jul;62:2-10.

13. Cao X, Gulati M, Jiang H. Serine protease-related proteins in the malaria mosquito, *Anopheles gambiae*. Insect Biochem Mol Biol. 2017 Sep;88:48-62.
14. Cao X, Jiang H. Building a platform for predicting functions of serine protease-related proteins in Drosophila melanogaster and other insects. Insect Biochem Mol Biol. 2018 Dec;103:53-69.
15. Wang Y, Jiang H. Prophenoloxidase (proPO) activation in Manduca sexta: an analysis of molecular interactions among proPO, proPO-activating proteinase-3, and a cofactor. Insect Biochem Mol Biol. 2004 Aug;34(8):731-42.
16. Lu Z, Jiang H. Expression of Manduca sexta serine proteinase homolog precursors in insect cells and their proteolytic activation. Insect Biochem Mol Biol. 2008 Jan;38(1):89-98.
17. He Y, Cao X, Zhang S, Rogers J, Hartson S, Jiang H. Changes in the Plasma Proteome of Manduca sexta Larvae in Relation to the Transcriptome Variations after an Immune Challenge: Evidence for High Molecular Weight Immune Complex Formation. Mol Cell Proteomics. 2016 Apr;15(4):1176-87.
18. Ross DR, Dunn PE. Effect of parasitism by Cotesia congregata on the susceptibility of Manduca sexta larvae to bacterial infection. Dev Comp Immunol. 1989 Summer;13(3):205-16.
19. Piao S, Song YL, Kim JH, Park SY, Park JW, Lee BL, et al. Crystal structure of a clip-domain serine protease and functional roles of the clip domains. EMBO J. 2005 Dec 21;24(24):4404-14.
20. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 2004 Mar 19;32(5):1792-7.
21. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. Mol Biol Evol. 2018 Jun 1;35(6):1547-1549.
22. Jiang H, Wang Y, Gu Y, Guo X, Zou Z, Scholz F, et al. Molecular identification of a bevy of serine proteinases in Manduca sexta hemolymph. Insect Biochem Mol Biol. 2005 Aug;35(8):931-43.
23. Sumathipala N, Jiang H. Involvement of Manduca sexta peptidoglycan recognition protein-1 in the recognition of bacteria and activation of prophenoloxidase system. Insect Biochem

Mol Biol. 2010 Jun;40(6):487-95.
24. Zhang S, Cao X, He Y, Hartson S, Jiang H. Semi-quantitative analysis of changes in the plasma peptidome of Manduca sexta larvae and their correlation with the transcriptome variations upon immune challenge. Insect Biochem Mol Biol. 2014 Apr;47:46-54.
25. Cao X, Wang Y, Rogers J, Hartson S, Kanost MR, Jiang H. Changes in composition and levels of hemolymph proteins during metamorphosis of Manduca sexta. Insect Biochem Mol Biol. 2020 Dec;127:103489.
26. Kanost MR, Arrese EL, Cao X, Chen YR, Chellapilla S, Goldsmith MR, et al. Multifaceted biological insights from a draft genome sequence of the tobacco hornworm moth, Manduca sexta. Insect Biochem Mol Biol. 2016 Sep;76:118-147.
27. Peterson AC, Russell JD, Bailey DJ, Westphall MS, Coon JJ. Parallel reaction monitoring for high resolution and high mass accuracy quantitative, targeted proteomics. Mol Cell Proteomics. 2012 Nov;11(11):1475-88.
28. Jiang H, Wang Y, Yu XQ, Zhu Y, Kanost M. 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. 2003 Oct;33(10):1049-60.
29. Wang Y, Lu Z, Jiang H. Manduca sexta proprophenoloxidase activating proteinase-3 (PAP3) stimulates melanization by activating proPAP3, proSPHs, and proPOs. Insect Biochem Mol Biol. 2014 Jul;50:82-91.
30. Jiang H, Wang Y, Ma C, Kanost MR. Subunit composition of pro-phenol oxidase from Manduca sexta: molecular cloning of subunit ProPO-P1. Insect Biochem Mol Biol. 1997 Oct;27(10):835-50.
31. Jiang H, Wang Y, Yu XQ, Kanost MR. Prophenoloxidase-activating proteinase-2 from hemolymph of Manduca sexta. A bacteria-inducible serine proteinase containing two clip domains. J Biol Chem. 2003 Feb 7;278(6):3552-61.
32. Miao Z, Cao X, Jiang H. Digestion-related proteins in the tobacco hornworm, Manduca sexta. Insect Biochem Mol Biol. 2020 Nov;126:103457.
33. Lee SY, Kwon TH, Hyun JH, Choi JS, Kawabata SI, Iwanaga S, et al. In vitro activation of pro-phenol-oxidase by two kinds of pro-phenol-oxidase-activating factors isolated from hemolymph of coleopteran, Holotrichia diomphalia larvae. Eur J Biochem. 1998 May

15;254(1):50-7.

34. Murugasu-Oei B, Balakrishnan R, Yang X, Chia W, Rodrigues V. Mutations in masquerade, a novel serine-protease-like molecule, affect axonal guidance and taste behavior in Drosophila. Mech Dev. 1996 Jun;57(1):91-101.
35. Rousset R, Bono-Lauriol S, Gettings M, Suzanne M, Spéder P, Noselli S. The Drosophila serine protease homologue Scarface regulates JNK signalling in a negative-feedback loop during epithelial morphogenesis. Development. 2010 Jul;137(13):2177-86.
36. Povelones M, Bhagavatula L, Yassine H, Tan LA, Upton LM, Osta MA, Christophides GK. The CLIP-domain serine protease homolog SPCLIP1 regulates complement recruitment to microbial surfaces in the malaria mosquito *Anopheles gambiae*. PLoS Pathog. 2013;9(9):e1003623.
37. El Moussawi L, Nakhleh J, Kamareddine L, Osta MA. The mosquito melanization response requires hierarchical activation of non-catalytic clip domain serine protease homologs. PLoS Pathog. 2019 Nov 25;15(11):e1008194.
38. Sousa, G.L., Bishnoi, R., Baxter, RHG, Povelones, M., 2020. The clip-domain serine protease CLIPC9 regulates melanization downstream of SPCLIP1, CLIPA8, and CLIPA28 in the malaria vector *Anopheles gambiae*. PLoS Pathog., 16, e1008985.
39. Sousa GL, Bishnoi R, Baxter RHG, Povelones M. The CLIP-domain serine protease CLIPC9 regulates melanization downstream of SPCLIP1, CLIPA8, and CLIPA28 in the malaria vector *Anopheles gambiae*. PLoS Pathog. 2020 Oct 12;16(10):e1008985.
40. Yassine H, Kamareddine L, Chamat S, Christophides GK, Osta MA. A serine protease homolog negatively regulates TEP1 consumption in systemic infections of the malaria vector *Anopheles gambiae*. J Innate Immun. 2014;6(6):806-18.
41. Nakhleh J, Christophides GK, Osta MA. The serine protease homolog CLIPA14 modulates the intensity of the immune response in the mosquito *Anopheles gambiae*. J Biol Chem. 2017 Nov 3;292(44):18217-18226.
42. Cox J, Mann M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat Biotechnol. 2008 Dec;26(12):1367-72.
43. Satoh D, Horii A, Ochiai M, Ashida M. Prophenoloxidase-activating enzyme of the silkworm, Bombyx mori. Purification, characterization, and cDNA cloning. J Biol Chem.

1999 Mar 12;274(11):7441-53.

44. Gupta S, Wang Y, Jiang H. Manduca sexta prophenoloxidase (proPO) activation requires proPO-activating proteinase (PAP) and serine proteinase homologs (SPHs) simultaneously. Insect Biochem Mol Biol. 2005 Mar;35(3):241-8.
45. Kan H, Kim CH, Kwon HM, Park JW, Roh KB, Lee H, Park BJ, Zhang R, Zhang J, Söderhäll K, Ha NC, Lee BL. Molecular control of phenoloxidase-induced melanin synthesis in an insect. J Biol Chem. 2008 Sep 12;283(37):25316-25323.
46. Lee KY, Zhang R, Kim MS, Park JW, Park HY, Kawabata S, Lee BL. A zymogen form of masquerade-like serine proteinase homologue is cleaved during pro-phenoloxidase activation by Ca2+ in coleopteran and Tenebrio molitor larvae. Eur J Biochem. 2002 Sep;269(17):4375-83.

CHAPTER II

SERINE PROTEASE HOMOLOGS CLIPA4, A6, A7, AND A12 ACT AS COFACTORS FOR PROPHENOLOXIDASE ACTIVATION IN ANOPHELES GAMBIAE

Qiao Jin[a], Yang Wang[a], Yingxia Hu[a], Yan He[a], Steven D. Hartson[b], Haobo Jiang[a]

[a] Department of Entomology and Plant Pathology, Oklahoma State University,

Stillwater, OK 74078, USA

[b] Department of Biochemistry and Molecular Biology, Oklahoma State University,

Stillwater, OK, 74078, USA

Key words: clip domain, insect immunity, melanization, hemolymph protein.

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
E-mail: haobo.jiang@okstate.edu

The abbreviations used are: PO and PPO, phenoloxidase and its precursor; PAP, PPO activating protease; SP and SPH, serine protease and its noncatalytic homolog; CLIP, clip-domain SP(H), PGRP, peptidoglycan recognition protein; SDS-PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis.

Abstract

Phenoloxidase (PO) catalyzed melanization in insects is mediated and modulated by extracellular serine proteases (SPs) and their noncatalytic homologs (SPHs) that form cascade pathways. Most of the pathway members contain a regulatory clip domain and, thus, are also designated CLIPs. At least in some insects, POs are activated from their precursors PPOs by prophenoloxidase activating proteases (PAPs) in the presence of clip-domain SPHs that serve as PAP cofactors. However, it is not fully known if PAPs and cofactors are required for PPO activation in mosquitoes, and melanotic encapsulation is a key defense mechanism against pathogens and parasites. In this study, we identified four CLIPAs in the African malaria mosquito *Anopheles gambiae* and tested their potential to act as cofactors of CLIPB9, a PAP. CLIPA4 is the ortholog of *M. sexta* SPH2 whereas CLIPA6, A7 and A12 are close homologs of SPH1b. The SPH1a and 2 form a high M_r complex required for PPO activation by PAP1−3 in the tobacco hornworm. CLIPA4, A6, A7 and A12 precursors were processed by *A. gambiae* CLIPB9$_{Xa}$ and *M. sexta* PAP3. After the proteolytic activation, *A. gambiae* CLIPA4 formed complexes with CLIPA6, A7 or A12 to assist PPO2 or PPO7 activation. Unprecedented high levels of PO activity were achieved in the *in vitro* conditions, suggesting that cofactor-assisted PPO activation reaction is conserved in holometabolous insects.

1. Introduction

Melanization is an indispensable defense response of insects in the resistance to pathogen and parasite infection (Nappi and Christensen, 2005). As a critical enzyme in the pigmentation process, phenoloxidase (PO) catalyzes multiple steps of a chemical reaction series which converts monophenols to diphenols, quinones, other reactive intermediates, and then to stable polymers at last (Zhao et al., 2007 and 2011). While lightly colored intermediates can kill viruses, bacteria, fungi, parasites, and parasitoid wasp eggs, melanin polymers are inert and

protective against damage of free radicals and ultraviolet lights. To prevent damage of the reactive compounds to host tissues and cells, POs are synthesized as inactive zymogens PPOs and activated by proteolytic cleavage when and where needed. And, to ensure a potent, local, and transient reaction against non-self or altered self, a system of serine proteases (SPs), noncatalytic serine protease homologs (SPHs), serpins, and specific recognition proteins is employed to activate PPOs and other defense mechanisms at the site of infection or wounding in each insect (Kanost and Jiang, 2015; ref). Among the known system components, SPs and SPHs mostly contain a clip domain and, hence, are called CLIPs. CLIPs in the A subfamily are catalytically inactive (*i.e*. cSPHs, c for clip domain), whereas a majority of CLIPBs, CLIPCs and CLIPDs are cSPs. Recognition of pathogens or parasites can lead to autoactivation of a modular serine protease zymogen, which orderly activates CLIPs in the system. Although exceptions exist, CLIPCs are located upstream of CLIPBs, whereas CLIPAs are activated by SPs or other proteases (Zhang et al., 2020; An et al., 2011; Wang and Jiang, 2007; Wang et al., 2014). Substitution of the catalytic residues (His, Asp and Ser) in S1A family of SPs abolishes CLIPAs' amidase activities but they can still function as auxiliary factors after proteolytic activation. In some insects at least, the reaction of PPO activation is mediated by a prophenoloxidase activating protease (PAP) in the presence of a high M_r complex of cSPHs in the CLIPA subfamily (Kwon et al., 2000; Yu et al., 2003; Kan et al., 2008; Wang et al., 2020). Recently, in *Aedes aegpti*, ClipA14 was identified as a cofactor of ClipB9 to enhance PPO3 activation (Ji et al, 2022).

There are 21, 45, 52, 55 and 110 CLIP genes in the genomes of *Apis mellifera*, *Manduca sexta*, *Tribolium castaneum*, *Drosophila melanogaster* and *Anopheles gambiae*, 7, 10, 18, 19 and 55 of which encode cSPHs, respectively (Cao et al., 2015 and 2017; Cao and Jiang, 2018). While

some of the CLIPBs, CLIPCs and CLIPDs are established as serine protease system components for melanization and other processes, biochemical roles of the cSPHs have not yet been defined in nearly all cases. For instance, among the 22 *A. gambiae* CLIPAs, only A8, A28 and A30 (*i.e.* SPCLIP1) are known to form a cSPH pathway for TEP1-localized melanization of *E. coli* and *Plasmodium berghei* (Povelones et al., 2013; El Moussawi et al., 2019; Sousa et al., 2020). *A. gambiae* CLIPA2, A5, A7, A8, and A14 affect melanization (Volz et al., 2006; Yassine et al., 2014; Nakhleh et al., 2017). Bovine coagulation factor Xa-activated *A. gambiae* CLIPB9$_{Xa}$ and B10$_{Xa}$ cleaved the purified *M. sexta* PPOs and yielded low levels of PO activity (An et al., 2011; Zhang et al., 2021). Silencing CLIPC9, CLIPs B1, B3, B4, B8–10, B14, B15, and B17 impacted melanization of *Plasmodium berghei* ookinetes and oocysts (Sousa et al., 2020; Paskewitz et al., 2006; Volz et al., 2005 and 2006; Zhang et al., 2016). Nevertheless, it is unclear whether or not functional redundancy of the CLIPs had affected the results of RNAi screenings. Neither is it understood if PPO activation by PAP (*e.g.* CLIPB9 and B10) needs one or more CLIPAs as cofactor to generate fully active POs. Furthermore, since there are nine PPO genes in *A. gambiae*, it is not known if different CLIPs may play distinct roles in the activation of PPO1−9.

We and our colleagues have discovered and step-by-step reconstituted an SP-SPH-serpin network in *M. sexta* hemolymph for melanization and Toll pathway activation (Kanost and Jiang, 2015; Wang et al., 2020). We also built a platform for predicting functions of SPs and SPHs in holometabolous insects based on the genetic and biochemical data obtained from different species (Cao and Jiang, 2018). Recently, five of the ten CLIPAs (SPHI: 1a, 1b, 4 and 101; SPHII: 2) in *M. sexta* were found to form high M_r complexes that assist PPO activation by PAP3 (Jin et al., 2022). Likely due to positive selection, products of the newly duplicated genes

(*SPH1b* and *SPH101*) are more favorable substrates of PAP3 and better SPH2 partners than those of *SPH1a* and *SPH4*. Building on the same phylogenetic relationships and results of the transcriptome and proteome analyses, we selected four CLIPAs as candidates of PAP cofactors in *A. gambiae*. After being processed by Factor Xa-treated CLIPB9$_{Xa}$ or *M. sexta* PAP3, purified CLIPA6, A7, and A12 precursors from *Sf*9 cells formed SPHI complexes with decreasing oligomerization states, while cleaved CLIPA4, an SPHII, remained as a monomer. CLIPB9$_{Xa}$ and PAP3 alone also cleaved *A. gambiae* PPO2 and PPO7 but generated low PO activities. However, in the presence of SPHI-II pairs or mixture, much higher PO activities were yielded in the reactions, indicating that an SPHI-II complex is required for PPO activation by a PAP in *An. gambiae*. Such auxiliary factors can be prepared *in vitro* to probe the anti-parasitic mechanism biochemically. Curiously, except for CLIPA7, roles of the other three CLIPAs were not reported after the extensive RNAi screenings. Direct biochemical investigations can provide key insights into molecular mechanisms of melanization in mosquitoes as well.

2. Methods and materials

2.1. cDNA cloning and recombinant production of ClipA4, ClipA6, ClipA7, ClipA12, and ClipB9xa

ClipA4 was amplified from cDNA of whole body from adult (day ?-not labled) *An. gambiae* using primer:j784(5'-CAGATGCAGCAACCGATC) and primer j785(5'-CTCGAGAGCAAACAGGCT). The product was cloned into pGEM-T vector (Promega) and, after sequence validation, the *Nde*I-*Xho*I fragments were inserted into the same sites in pMFFMH6(He et al., 2022). ClipA4/pMFFMH6 was used to generate a baculovirus to express proClipA4(HMQQPI… SLFALEQKLISEEDLHHHHHH) in insect cells (Sumathipala and Jiang, 2010), and the underlined part is encoded by the cDNA.

ClipA6 was amplified from the cDNA of whole body from adult *An. gambiae* using primer: j786(5'-CATATGGATGATCTCAGCCTGG) and j787(5'-CTCGAGGGGCGTGTAGCTG). After cloning and

sequence validation, the *Nde*I-*Xho*I fragment was retrieved and inserted into the same sites in pMFFMH6 to generate a baculovirus for producing proClipA6 (HMDDLS…YTPLEQKLISEEDLHHHHHH) in insect cell. The underlined region is identical to the mature precursor.

ClipA12 was also amplified from the cDNA of whole body from adult *An. gambiae* using primer: j792(5'-CATATGCAGACTTGTGAAGGGA) and primer j793(5'-CTCGAGGGCTGGCGTGTAG). Confirmed to be error-free, retrieved by *Nde*I-*Xho*I digestion, and inserted into the same sites in pMFFMH6. ClipA12/pMFFMH6 was used to generate a baculovirus that produces proClipA12 (HMQTCEG…YTPALEQKLISEEDLHHHHHH) in insect cells, and the underlined part is encoded by ClipA12 cDNA.

The full length of ClipA7 is 821aa, which includes the low complexity region (LC region), Clip domain and serine protease like domain. To make the cloning and expression easy, we skipped the LC region, and keep the Clip domain and chymotrypsin-like domain. The first fragment of ClipA7 was amplified using the plasmid of ClipA7/? using primer: j788(5'-CATATGGAAGATGAGGAGGTTATC) and j789(5'-CTCCGGATATGGGCAGCA), the second fragment was amplified using primer: j790(5'-TCCGGAGAAGACGGTTAGCGTG) and j791(5'-CTCGAGGTAGAAGCTGTCAGT). The PCR products were cloned into pGem-T (Promega) and verified by DNA sequencing. The fragments retrieved by NdeI-BspEI/BspEI-XhoI digestion and inserted into pMFFMH6 (NdeI/XhoI). ClipA7/pMFFMH6 was used to generate a baculovirus that produces proClipA7(HMEDE…PYPE-KTVS…FYLEQKLISEEDLHHHHHH), the underlined part is encoded by ClipA7.

ClipB9 factorXa form (ClipB9xa) was amplified from plasmid (ClipB9/pMFFMHaH6) using j1254(5'-CATATGCAGCAACAGCAATGCACGA), j1255(5'-CTCGAGCTTGATATTGCTTCTCAG), j198(5'-ATCGACGGACGGATCTACGGT), and (5'- AGATCCGTCCGTCGATGCTAA). After TA cloning and sequencing, the cDNA was found to have twelve nonsynonymous mutations and the *Nde*I-*Xho*I region was retrieved

and inserted into P6.9MFHaH6(*Nde*I/*Xho*I) to generate a baculovirus that produces proClipB9xa(HMQQ…IDGR…KEGKPIPNPLLGLDSTHHHHHH) fragment was retrieved and inserted into the same sites in P6.9MFHaH6 (from Ethan) to generate a baculovirus for producing proClipB9xa. The underlined region is identical to the mature precursor.

2.2. Expression and purification of the recombinant proClipA4, proClipA6, proClipA7, proClipA12 and proClipB9XA

*Sf*9 cells at 2.4×10^6 cells/ml in 300ml of insect serum-free medium (Invitrogen Life Technologies) were separately infected with the baculovirus stocks at a multiplicity of infection of 10 and grown at 27 °C for 96 h with gentle agitation at 150 rpm. After the cells were removed by centrifugation at 5,000×*g* for 20 min, the supernatant was mixed with the same volume of H2O containing 1mM benzamidine, the pH of the conditioned medium PH was adjusted to 6.4 using 12N HCl. After mixed at 4℃ for 10 min, the fine particles were spun down by centrifugation at 8,000×*g* for 30 min, and the supernatant was applied to a DS-column (20 ml bed volume) at a flow rate of 1.5 ml/min and, following a washing step with 100 ml buffer A, bound proteins were eluted from the column with a linear gradient of 0−1.0 M NaCl in 250 ml of buffer A at a flow rate of 1.5 ml/min. The proClipA or proClipB9xa fractions were combined and loaded onto a 1 ml Ni^{2+}-nitrilotriacetic acid agarose column. After washing with 15 ml of 50 mM sodium phosphate, pH 7.5 (buffer B), the bound proteins were eluted with a gradient of 0−0.25 M imidazole in 25 ml of buffer B. Fractions containing proClipA or proClipB9xa were combined, dialyzed against 20 mM Tris-HCl (pH 7.5), and concentrated on Amicon Ultra-30 centrifugal filter devices (Millipore). The protein aliquots were rapidly frozen in liquid nitrogen prior to storage at -80 °C.

2.3. Activation of ClipB9xa and PAP3 following by the IEARase amidase detection

To get the active ClipB9xa, 0.8ug of Factorxa were mixed with 2.5 ug proClipB9xa to the final volume of 25µl. The reaction mixture and control were incubated at 37℃ for 6 hours. 2.5 µl of the

active mixture were incubated with 150 µl 25um acetyl-Ile-Glu-Ala-Arg-p-nitroanilide (A0180; Sigma) as a chromogenic substrate to test the IEAR amidase. One unit of activity is defined as ΔA405/min=0.001.

In order to get the active PAP3, PAP3(80 ng) mixed with proPAP3(400ng) to the final volume of 20 µl. After incubation at 37℃ for 1 hour, 200ng of PAP3 were used to test IEAR activity. The method was same as it's described above.

2.4. Cleavage of proClipA4, A6, A7, and A12 by PAP3 and electrophoretic mobility changes

M. sexta PAP3 was isolated from pharate pupal hemolymph (Jiang et al., 2003) and used to activate recombinant proPAP3 (Wang et al., 2014). The proClipA4 and proClipA6, A7 and A12 were expressed in insect cells, purified from condition media, and used as PAP3 substrates. To better understand the process of PAP3 cleavage and high M_r complex formation, aliquots of the proClipAs (1µl, 300ng) were incubated with 40ng PAP3 for 1 h at 37 °C. The cleavage reactions, controls, and M_r markers were separated by 10% SDS and native PAGE, transferred onto nitrocellulose membranes, and detected using MsSPH1(ClipA6, ClipA7 and ClipA12), anti-FLAG (ClipA4) and $(His)_6$ antisera as primary antibody and goat-anti-rabbit/mouse IgG conjugated to alkaline phosphatase (Bio-Rad) as the secondary antibody, and a BCIP-NBT substrate kit (Bio-Rad) for color development.

2.5. Cleavage of proClipA4, A6, A7, and A12 by ClipB9xa and electrophoretic mobility changes

proClipB9xa was expressed from insect cell and was purified from the condition media. 2.5 ug proClipB9xa was activated by 0.8ug Factorxa at 37℃ for 4 hours. To explore the cleavage and high Mr complex formation, aliquot proClipAs were mixed 100ng ClipB9xa in 37℃ for 1 hour. The cleavage reactions, controls, and M_r markers were separated by 10% SDS and native PAGE, followed by electro transfer and immunoblotting using antibody against the hexahistidine tag. Mixtures and ClipB9xa control were separated by 10% SDS and native PAGE. The bands were transferred using semi-dry

transfer machine and were analyzed by immunoblotting using antibody against the hexahistidine tag or antiMsSPH1 antibody as primary antibody and goat-anti-rabbit/mouse IgG conjugated to alkaline phosphatase as the secondary antibody, and a BCIP-NBT substrate kit for color development.

2.6 Cleavage of proPO by ClipB9xa and PAP3

To identify which proPO can be cleaved by ClipB9xa, 100ng ClipB9xa were incubated with purified AgproPO (1 μl, 300ng). Samples or controls were subjected to Western blot analysis using antibodies against *Aedes aegypti* PPOs. For measuring the processing AgPPO1-9 by MsPAP3. AgPPO (1μl, 300ng) were mixed with PAP3(1μl, 40ng) at 37℃ for 1 hour, samples and controls resolved by 10% SDS-gel, followed by electro transfer and immunoblotting using antibody against *Ae. aegypti* as primary antibody and goat-anti-rabbit IgG conjugated to alkaline phosphatase as the secondary antibody.

2.7 Identification and quantification of cleavage sites of ClipAs using LC/MS/MS

Around 5-10 μg the purified proClipA4, A6, A7 and A12 from insect cells were limited processed using ClipB9Xa (1μg/reaction) and PAP3(1 μg/reaction) separately for 1h at 37℃, the mixtures or the pro-form of ClipAs were separately denatured in urea and digested with chymotrypsin, LysC, and V8 proteinase (Zhang et al., 2014). Resulting peptides were desalted using C18 affinity media, dried, and redissolved in

mobile phase A (0.1% formic acid in H2O). The samples were loaded onto an Acclaim PepMap RSLC C18 column (75 μm × 50 cm, Thermo Fisher) for data-dependent LC-MS/MS analysis as described previously

(Cao et al., 2020). Each sample was subjected to the Acclaim column via a gradient of 0–35% mobile phase B (0.1% HCOOH, 80%AcCN, 20% H2O) developed over 120 min as described before (Jin et al., 2022). The survey scans were performed followed by both HCD and CID collisional MS/MS events

triggering from parent ions, with scanning of collisional fragment at 15,000 resolutions in the Orbitrap Fusion.

To identify and quantify the limited processed cleavage sites of ClipAs, a database combined A. gambiae protein database (PEST Peptides_AgamP4.2) and background proteins from M. sexta, human, and insect cell were constrcuted, peptide spectrum matches were reviewed in Byonic to check peptides not cut by nonspecific processing enzymes, the details of peak area calculation were same as described previously (Jin et al, 2022).

2.8 ProPO activation and PO activity assay

A. gambiae proPO1-proPO9 was purified from *E. coli* and stored at -80 °C. To activate AgproPO, AgproPO (1µl, 500 ng/µl) and PAP3 (1µl, 40 ng) were incubated on ice together with proClipA4 (1µl, 300ng/µl), ClipA6 (1µl, 300ng/µl), ClipA7(1µl, 300ng/µl), and ClipA12(1µl, 300ng/µl) in a total volume of 20 µl Tris-HCl, 5mM Cacl2, pH 7.5. Sixty minutes later, PO activity was determined on a microplate reader immediately after 150 µl of 2.0 mM dopamine in 50 mM sodium phosphate, pH 6.5, which had been added to each sample well (Jiang et al., 2003).

To activate AgproPO using ClipB9xa and ClipAs, proPO (1 µl, 500 µg/ml) and ClipB9xa (1µl, 100ng) were incubated with proClipA4(1µl, 300ng/µl), ClipA6(1µl, 300ng/µl), ClipA7(1µl, 300ng/µl), and ClipA12(1µl, 300ng/µl) on ice in a total volume of of 20 µl Tris-HCl, 5mM Cacl2, pH 7.5. The method for PO activity testing was the same as described above.

To determine the efficiency for forming “cofactor” of different combinations, ClipA4, were paired with ClipA6, ClipA7 and ClipA12 separately and mixed with PAP3 or ClipB9xa and AgproPO. PO activity was measured after incubation on ice for 1 hour.

3. Result

3.1 Selection of An. gambiae ClipA4, ClipA6, ClipA7, and ClipA12 as Cofactor candidates

Serine protease homologues (SPHs) have been identified as indispensable components for PPO activation in many insects, (Jiang et al., 1998; Kwon et al.,2000; Yu et al., 2003, Jin et al., 2022; Ji et al., 2022). While studies in the malaria vector *A. gambiae* revealed that SPHs have other roles in the regulation of immune responses. Among the 21 SPHs clustered as ClipAs in *An. gambiae*, several of them have been studied, such as ClipA30(CSCLIP1), ClipA28, ClipA8, ClipA14, which were demonstrated as the positive regulators of melanization (Schnitger et al., 2007; Nakhleh et al.,2017; Moussawi et al., 2019; Povelones et al., 2013); ClipA2, ClipA7 resulted in a more pronounced melanization response against ookinetes (Yassine et al., 2014; Awada., 2018). Despite numerous of transgenetic evidence for involvement of AgClipAs in melanization, it is still unclear that ClipAs can serve as a cofactor for PPO activation in *An. gambiae*.

The genome of *An. gambiae* contains 110 clip domain serine protease related genes, 55 genes were encoding clip-domain SPHs, which are more than 18 cSPHs in *D. melanogaster*, 19 cSPHs in *Tribolium castaneum*, and 10 cSPHs in *M. sexta*(Cao et al., 2015 and 2017; Cao and Jiang, 2018; Miao et al., 2020). Based on the phylogenetic tree, 8 ClipA genes (ClipA12, ClipA13, ClipA14, ClipA5, ClipA6, ClipA26, ClipA31, ClipA7) were clustered with *M. sexta* SPH1, and ClipA4 was clustered with *M. sexta* SPH2(Cao et al., 2017). By analyzing the transcriptome data, it has been found that the levels of ClipA4, ClipA12, ClipA14, ClipA6 and ClipA7 were relatively high, while ClipA5, ClipA13, ClipA26 and ClipA31 can't be detected at mRNA level (Cao et al., 2017). Besides, among those candidates ClipA14 has been demonstrated that can regulate *Plasmodium* melanization in a TEP1-dependent manner (Nakhleh., 2017). Therefore, we screened ClipA4, ClipA6, ClipA7 and ClipA12 as PO activation cofactors.

3.3 Recombinant expression, purification, and characterization of the ClipA4, ClipA6, ClipA7, ClipA12 and ClipB9xa from Sf9 cells

Beginning to study the function of ClipA4, ClipA6, ClipA7 and ClipA12, we amplified the full-length cDNAs of *An. gambiae* and subcloned them into pMFFMH6. Upon transformation, the bacmids were isolated for transfecting insect cells and producing high-titer viral stocks through serial infection. Led by the honeybee mellitin signal peptide, the precursor proteins were efficiently secreted into media with the hexahistidine tag fused to their carboxyl-terminus. We isolated the proClipAs from the conditioned media by cationic and nickel affinity chromatography. From 100 ml of the media, 0.5mg/proClipA4, 1mg/proClipA6, 0.3mg/proClipA7 and 0.2mg/proClipA12 and 0.4mg/proClipB9xa were obtained, respectively. The proteins migrated to 55kD(proClipA4), 60kD(proClipA6), 42kD(proClipA7), and 55kD (proClipA12) positions on a 10% SDS polyacrylamide gel (Fig. 1B).

We also designed the FactorXa activating site of ClipB9 to get the activating form of ClipB9 in vitro. proClipB9xa/P6.9FMHaH6 was constructed, and baculovirus for infecting Sf9 cells was made. 150ml media was used to do the purification by nickel affinity chromatography. 0.8mg of proClipB9xa was harvested. The recombinant proClipB9xa was migrated at 55kD on 10% SDS-PAGE (Fig. 1B). As judged based on the staining patterns, the proteins are near homogeneous, except for ClipA7 and proClipA12 (Fig. 1A). In addition to the precursors, a small amount of 34kDa of ClipA7, 41 kDa and 31 kDa fragments of ClipA12 contained the hexahistidine tag recognized by its antibody (Fig. 1A). LC-MS/MS data (Fig7) revealed that the cleaved forms of proClipA7 after “LHLCPNGELNTDGANIIDIR*” and “CGLRNVDGVGFR*” can also be detected in untreated samples of proClipA7. Same situation of proClipA12, both active forms of proClipA12 “TAEGEDDDAPAPEVDLR*” and “*IGAGKVEEAEFGEFPW” can be detected in proClipA12 sample. Compared with Fig 5A and Fig 6A, the size of 34kD band in ClipA7 was same as its cleaved form start from “IDIR*FNPDNEC……” thus it seems to be that small amount of cleaved ClipA7 between C3 and C4 in Clip domain was detected; similar situation in proClipA12: the sizes of 41 kDa and 31 kDa fragments of ClipA12 were same as the sizes of processed CliPA12 between C3 and

C4 in Clip domain and inside of serine protease like domain, therefore, both fragments start from "VDLR*IGQENSNV……", "VGFR*IGAGKVEEAEFGEFPW……" can be recognized in proClipA12.

3.4 Cloning construction, expression, and purification of recombinant AgPPO2 and PPO7(not write)

Proteolytic cleavage of An. gambiae PPO7 by ClipB9xa

ClipB9 was reported as a PAP for PPO activation in *An. gamibiae*(An, Budd et al. 2011), however, no evidence was provided for exact substrate(AgPPOs) of ClipB9. To further explore which AgPPO can be cleaved by ClipB9, the recombinant ClipB9 was expressed in vitro (sf9 cell). As the endogenous activating enzyme of proClipB9 is unknown, we prepared the FactorXa mutated form of proClipB9(ClipB9xa), which permits its activation by commercially available bovine Factor Xa. SDS-PAGE analysis of purified proCLIPB9Xa suggests that the recombinant protein has a mass of around 55kD (Fig.1A-1B). Incubation of Factor Xa with proClipB9xa led to the cleavage of proClipB9xa and a new band appeared around 36kD (Fig.S1A). Activated ClipB9xa resulted in a significant increase in activity (3.2U) as measured by cleavage of the artificial acetyl-Ile-Glu-Ala-Arg-p-nitroanilide (IEARpNA) (Fig.S1B). There are 9 PPOs in *An. gambiae* (Cao et al., 2017; He et al., 2017), to explore which AgPPO can be processed by ClipB9Xa, we incubated ClipB9xa with purified AgPPOs (AgPPO1, AgPPO2, AgPPO3, AgPPO4, AgPPO7, AgPPO8, AgPPO9). Western blot analysis of recombinant purified AgPPO2(79kD) recognized one thick new band around 75kD (Fig. 2A), which is closely below the pro-form of PPO2. Taken together, the evidence showed that ClipB9 had the characteristics of PAP and it can cleave AgPPO2 and other AgPPOs in vitro.

3.5 Proteolytic cleavage of An. gambiae PPO7 by Manduca PAP3

It has been demonstrated ClipB9 can cleave MsPPO (An, Budd et al. 2011), which shows the functional conservatism of PAP in insects. *M. sexta* PAP3 is reported as a trypsin-like serine proteinase and it can cleave after positively charged amino acid residues (Lys or Arg) efficiently in *M. sexta*(Wang, Lu et al. 2014), its substrates such as MsSPHI(4, 1a, 1b, 101)/2, HP5, MsPPO, and Spätzle are widely

ranged. Compared with MsPAP1 and MsPAP2, MsPAP3 can process SPHI/2 better, and incubating PAP3 with SPHI/SPH2 and MsPPO can achieve high PO activity. So next we want to see whether PAP3 has the conserved function across species, to explore whether MsPAP3 can activate AgPPOs or not, the zymogen of PAP3 was purified from sf9-cell media. Immunoblot analysis showed that the molecular weight of proPAP3 was around 50 kDa (Fig. 2C). To get a large amount of PAP3, recombinant proPAP3 was incubated with PAP3, which was purified from Manduca hemolymph. The active form of PAP3 migrated to 36kDa (FisS1C). IEARase activity of active PAP3(1.55U) was increased significantly compared with its zymogen (FisS1D). To determine which AgPPOs can be cleaved by PAP3, active PAP3 was incubated with AgPPOs (AgPPO1, AgPPO2, AgPPO3, AgPPO4, AgPPO7, AgPPO8, AgPPO9) separately. Western blot analysis showed that the purified AgPPO7 was around 75kD (Fig. 3B). After being cleaved by PAP3, a new band around 72kD was shown (Fig. 3B), which was supposed to be a similar position of Arg^{51} with Manduca. In conclusion, MsPAP3 can cleave PPOs in *An. gambiae*.

3.6 Sequential proteolytic processing of the four proClipAs by An. gambiae ClipB9Xa and M. sexta PAP3

It has been found in *M. sexta*, PAP3 can process both PPO, SPH1(SPH4, 1a, 1b, and 101) and SPH2(Jiang, Wang et al. 1998, Gupta, Wang et al. 2005, Wang, Lu et al. 2014., Jin et al., 2022). And *Holotrichia diomphalia* PPAF-I (PAP) can activate *H. diomphalia* PPO instead of PPAF-II(SPH)(Lee, Kwon et al. 1998). Our data showed ClipB9 and MsPAP3 can activate AgPPOs (Fig. 2), while the question that in *An. gambiae* the cofactor candidates: proClipA6, ClipA7, ClipA12 (SPH1) and proClipA4 (SPH2), can be activated by ClipB9 or MsPAP3 or not is still unclear. In previous work of Manduca, MsPAP3 was found to be a strong proteinase that can process MsSPHI(SPH4, 1a, 1b, and 101) and MsSPH2 efficiently (Jin et al., 2022). To identify the functional conservation of MsPAP3 in *An. gambiae* and study ClipA4, ClipA6, ClipA7 and ClipA12 further, ClipA4, ClipA6, ClipA7 and ClipA12 were incubated with MsPAP3 separately. Complete cleavage of ClipA4 yielded two new bands: the primary band was around 50kD, the minor

cleavage products were shown at the position of 42kD (Fig 3A). All proClipA6, A7 and A12 can be processed by PAP3 and the predominate bands occurred at the same site of 30kD, minor cleavage bands at 36kD were also shown ClipA7 being cleaved by PAP3 (Fig. 3A). Native PAGE and immunoblot analysis showed that processing of ClipA6 and A7 by PAP3 produced a high M_r smear extending from the stacking gel (Fig. 3B, *orange bar*). Digestion of ClipA12 yielded a shorter and ladder-form smear (Fig. 3B). Complete digestion of ClipA4 yielded a smear (*orange bar*) moved a little slower than its precursor (*blue bar*). To explore whether ClipA4, ClipA6, ClipA7 and ClipA12 are the substrates of ClipB9, activated ClipB9xa was incubated with proClipA4, A6, A7 and A12 separately. Small amounts of proClipA4 were processed (Fig.4A), predominant cleavage of the recombinant proClipA4 was around 50kD (Fig.4A). All proClipA6 can be processed by ClipB9xa, cleaved bands of ClipA6 around 39kD and 32kD can be recognized by specific antibody (Fig.4A). proClipA7 were cleaved into two new bands at 35kD and 30kD (Fig.4A). After being cleaved by ClipB9xa, ClipA12 mainly migrated at 30kD (Fig.4A). Native PAGE analysis indicated that the full process of ClipA6 and ClipA7 by ClipB9xa led to a high *Mr* smear expanding to stacking gel (Fig. 4B, *orange bar*). While complete digestion of ClipA12 yielded a shorter smear (Fig. 4B, *orange bar*). Due to incomplete digestion of ClipA4 by ClipB9xa, migration of its cleaved form on Native gel was not obvious, it seemed to move slower than its precursor (Fig. 4B). In summary, the results indicated both AgClipB9 and MsPAP3 can process ClipA4, ClipA6, ClipA7 and ClipA12.

To examine dynamic processes of the proClipAs cleavage, we incubated aliquots of the four ClipA precursors with decreasing amounts of PAP3 or ClipB9Xa, separated the reaction mixtures by 10% SDS-PAGE, and examined the products by immunoblot analysis using hexahistidine antibodies (Fig5 and Fig6). At the lowest level of PAP3, 30%-40% of proClipA4 was migrated to 50kD (lane7-8), as PAP3 increased and proClipA4 diminished (lane3-6), the intensity of the cleaved form increased and there is an extra band around 42kD, this indicated that as the increase of PAP3, high Mr complexes is easier to form on the native gel (Fig. 5B), compared with the pro-form, cleaved ClipA4 moved slower while the

smear still in the separating gel, which seemed different from the behavior of MsSPH2. In terms of ClipA6, as the lowest PAP3 level, pro-form was fully digested, the 34kDa ClipA6 were detected as major bands, while the first cleavage band around 49kDa also can be detected (lane7-8), suggesting that cleavage at the first site occurred first. Intensity of the 31 kDa ClipA6 increased as the PAP3 concentration increased and cleavage at the second site occurred to the 42 kDa ClipA6 decreased (Fig5A). Precursor of ClipA6 migrated as one major band on native PAGE (lane2), and cleavage at the first and then second site yielded a smear in the stacking gel and a protein ladder in the separating gel (lane3-8). Similar situation also happened in the sequential proteolysis of proClipA7 and proClipA12 (Fig5A and 5B).

The dynamic process of ClipA4, ClipA6, ClipA7 and ClipA12 also have been analyzed by incubating same amount of the four pro-ClipAs with ClipB9Xa, samples were treated with 1 × SDS sample buffer at 95 ℃ for 5 min or the same buffer lacking SDS and DTT at 25 ℃ for 5 min prior to 10% SDS-PAGE and 10% native PAGE. After electrotransfer, immunoblot analysis was performed using 1:1000 diluted antibody against the hexahistidine tag. At the lowest level of ClipB9Xa, a little proClipA4 was cleaved and 50kD cleaved form was recognized, as the increase of ClipB9Xa, active form of ClipA4 become darker, the processing of proClipA4 was incompletely even at the highest concentration of ClipB9Xa. The smear of cleaved ClipA4 moved a bit slower than it pro-form on native PAGE, and the conformation change of ClipA4 was not obvious. Concerning ClipA6, at the lowest level of ClipB9Xa, a 42kDa band was recognized weakly, as the increase of ClipB9Xa, most of proClipA6 was digested, the cleaved forms at 42kDa, and 34kDa were getting intense. On native gel, cleaved ClipA6 was migrated both at the top of separating gel and the stacking gel. The same situation was also observed in ClipA7 process. A little bit different of ClipA12 digestion, at low concentration of ClipB9Xa, little of bands at 41kDa and 31kD were detected, as the concentration of ClipB9Xa getting higher, the precursor was fully processed, at the same time the first cleavage site decreased, and the second cleavage sites is increased.

Banding patterns of the reaction mixtures on native gel seemed to go through a decrease and then an increase in the amounts of high Mr smear.

3.8 Identification of the cleavage sites of ClipAs' precursors by M. sexta PAP3 and An. gambiae ClipB9Xa using LC-MS/MS

Our previous work has identified that *M. sexta* SPH1 and SPH2 isolated from induced hemolymph of feeding larvae were cleaved next to R^{133} and R^{77}(automated Edman degradation) respectively (Yu et al.,2003) at first, then by peptide mass fingerprint analysis, SPH1 in plasma of pharate pupae was also detected being cleaved after R^{82}(Wang and Jiang, 2004; Wang et al., 2014). Sequential cleavage from result 3.6 indicated one or multiple PAP specific cleavage sites was present in ClipA4 (SPH2) or ClipA6, A7, and A12 (SPH1). To identify the cleavage sites of ClipAs in *An. gambiae*, various digested enzymes (Chymo, V8, and LysC) were used to process their precursors or active forms of ClipAs separately, then we then used LC-MS/MS to examine the mixtures. A database including all sequences for detecting and background in the cocktails was constructed for sequence searching, softwares of PMi-Byonic-Viewer (https://proteinmetrics.com/) and Xcalibur (https://www.thermofisher.com/order/catalog/product/OPTON-30965) were used to identify and quantify targeted peptides. MaxQuant version 1.5.2.28 (Cox and Mann, 2008) and Perseus (https://maxquant.net/perseus/) 1.6.14.0 was used to normalize the detected peptides in total between different treatments. The Second MS spectra were listed in supplement (Fig. S4). The peak areas were calculated as the relative abundances of detected peptides. Theoretically we thought ClipB9Xa or PAP3 cleaved proClipA4 after ……EQNK*, after the treatment by Chymotrysin (cleaved after Y, W, F, M, L), there is no peptides processed after EQNK* was detected in proClipA4, while great amount peptides of "FTCQPPPEFAEQNK" were detected in ClipB9Xa (1.5×10^8) and PAP3(1.55×10^9) processed samples, and 10 times higher in PAP3 treated samples compared with ClipB9Xa plus proClipA4, which indicated that more efficient limited process was observed in *M. sexta* PAP3 cleaved samples. Interestingly, from the

sequential processing western blot (Fig 5A), there was a minor cleavage band around 43kDa in PAP3 processed proClipA4 sample was detected, this light chain was around 7kD less than major cleaved product. Thus, we check the possible cleaved position and found ……PVGR was exposed in cleaved ClipA4 samples. The abundance of this peptide is 1535 times more intense in PAP3(1.52×10^9) processed treatment compared with that using ClipB9Xa (9.9×10^5), which explained the band was detected in PAP3 cleaved ClipA4 instead of process using ClipB9Xa (Fig 5A and Fig 6A). As it's shown (Table 1), certain amounts of cleaved products of ……IDIR* and ……LGFR* can be detected in the precursor sample, after being further cleaved by B9Xa or PAP3, more peptides were detected at both cleavage sites. The results above indicated that the limited cleavage preference of ClipB9Xa and PAP3 are the same although they derived from different species. We also identified the cleavage sites of ClipA7, ClipA12, similar with the situation in ClipA6, two hypothetical cleavage sites of each ClipAs were detected in their pro-form samples, more products were quantified in specific-cleaved samples (Table 1). According to these results, we identified and quantified the cleavage sites of *An. gambiae* ClipAs.

3.9 Effectiveness of the proteolytically processed ClipA4, ClipA6, ClipA7 and ClipA12 as cofactors for proPO activation

From our previous work, we know serine protease homologues can enhance the PAP3 ability to cleave PPO, and there is a positive nonlinear relationship between the cleavage and PO activation (Gupta et al., 2005). To investigate whether the ClipAs in *An. gambiae* can enhance the cleavage and activation of AgPPOs as well as the relation between cleavage efficiency and PO activation, AgPPOs were incubated with PAP, ProClipA, PAP and all proClipAs, or paired proClipAs on ice for 1 hour, then the mixtures were used to do immunoblot analysis or PO activity detection (Fig8 and Fig9). The PPOs and the cleaved forms' darkness were measured using Image J 1.53t (https://imagej.nih.gov/ij/download.html), the cleavage efficiency (percentage) was measured using intensity of PO band plus uncleaved PPO band divide PO band intensity, the final number was got using

the ratio times 100%. About 62% of proPO2 can be cleaved by ClipB9xa (lane 8, Fig 9A), yielding 16 units (per microgram of PPO2) of PO activity (Fig 9B). When proClipA6, A7, A12, which are homologues of *M. sexta* SPH1, were incubated with AgPPO2 and ClipB9Xa leading to an unobvious enhance process of PPO2 (68%) (lane 10, Fig 9A), and the PO activity was 10 unit (per microgram of PPO2) (Fig 9B). While proClipA4, the homologue of M. sexta SPH2, was added together with other ClipAs, ClipB9Xa, and AgPPO2, almost 90% of PPO2 was cleaved (lane11, Fig 9A), which is 28% enhanced compared with PPO2 cleaved only using ClipB9Xa, and interestingly, the PO activity was about 99 units (Fig 9B), it's almost 13 time higher than ClipB9Xa processing AgPPO2 sample, 13 times higher than the value of ClipB9Xa with ClipA4 and ClipB9Xa(Fig 9B), and 20 times higher than ClipB9Xa with ClipA6, A7, A12, and PPO2. In conclusion, cofactors for PAP were essential for PPO2 activation in *An. gambiae*; as an SPH2, ClipA4 is indispensable for cofactor formation, playing a key role for PPO2 activation. Next, we tested which SPH1 (A6, A7, A12) can pair with ClipA4 and form a cofactor, leading to an efficient activation of AgPPO2. We paired ClipA4 with ClipA6, A7, A12 separately, and incubated it with ClipB9Xa and PPO2. About 80% of PPO2 was processed if ClipA4 and ClipA6 were paired (lane 12, Fig 9A), 50 unites of PO activity was harvest, which is not statistically changed compared with four ClipAs mixed with ClipB9Xa (positive control, Fig 9B), and PPO2. 91% and 78% PPO2 was cleaved when ClipA4 paired with ClipA7 or ClipA12 separately (lane13 and lane 14, Fig 9A), triggering 60 unites and 56 unites PO activity in corresponding, which are not significantly change compared with western compared with positive control (Fig 9B). Taken together, there is no preference of ClipA4 pairing with ClipA6, A7, A12 for forming cofactors to activate PPO2.

Similarly, we then tested if PAP3 processing of ClipA4, ClipA6, ClipA7 and ClipA12 precursors could yield active cofactors for proPO7 activation. The same process and software as above paragraph were used for the samples' treatment. The cleavage of PPO7 by PAP3 could not be enough to be detected by western blot (lane 8, Fig 8A), there was about 7.6 units of PO activity (per microgram of

PPO7) was achieved (Fig 8B). Mixing ClipAs with PAP3 and PPO7 except ClipA4 neither enhanced the PPO7 (lane9, Fig 8A) process nor it's activation (16 units) (Fig 8B). Four ClipAs were added together with PAP3 and PPO7, 14% of PPO7 were cleaved into PO (lane 10, Fig 8A); on the other hand, 70 units PO activity was yielded under this condition, which is 7.3 time higher than the treatment just with PAP3, 5 time increased compared the value of PAP3 with ClipA4 and PPO7 (Fig 8B), 4.5 times higher than sample in absence of ClipA4 (Fig 8B). These data revealed that external PAP from *M. sexta* also needs the assistance from cofactors to activate AgPPO7. Besides, SPH1(A6/A7/A12) seems to be necessary for the formation of cofactors. Then we also wonder which SPH1(A6/A7/A12) pairing with SPH2(A4) can be the best cofactor for MsPAP3 to activate AgPPO7. 4%, 5%, and 2% PPO7 were cleaved when ClipA4 were paired with A6, A7, A12 separately (lane 11, 12, 13, Fig 8A). Corresponding with the process, 62 units and 65 units of PO activity were achieved if ClipA4 paired with ClipA6 or ClipA7(Fig. 8C), which has no statistical change compared with positive control (four ClipAs mixed with PAP3 and PPO7). While 25 units PO activity was harvested when ClipA4 paired with ClipA12, 2.5 times lower than positive control, suggesting that ClipA4 paired either with ClipA6 or ClipA7 can form an efficient cofactor complex to activate AgPPO7.

4. Discussion

The arthropod melanization needs the specific proteolytic cleavage of PPO into PO, this process precisely regulates by upstream serine protease cascade, which mainly composes serine proteases (SP) and serine protease homologs (SPH). Biochemical studies indicated that SPHs function as cofactors of PAPs to regulate PPO activation in Lepidoptera and Coleoptera (Lee et al., 1998; Kwon et al., 2000; Yu et al., 2003). Previous work in *An. gambiae* revealed that several ClipC, ClipB, and ClipA were required for microbiol melanization (L. Sousa et al., 2020; An et al., 2011; Moussawi et al., 2019). Evidence showed SPCLIP1(ClipA30), CLIPA8, and CLIPA28 were in the upstream of ClipC9, which located on the upper layer of ClipB, then certain PPOs were activated by ClipB and melanization occurred (Zhang et al., 2021; L.

Sousa et al., 2020; An et al., 2011; Moussawi et al., 2019; Povelones et al., 2013). In *Aedes aegypti*, ClipA14 was identified as cofactor of ClipB9 to activate AAPPO3(Ji et al., 2022). However, the mechanism of PPO activation in mosquitoes are still ambiguous. Here, we identified four cofactors of ClipB: ClipA4, ClipA6, ClipA7, and ClipA12, in the presence of ClipB and its cofactor, AgPPO2 and AgPPO7 can achieve high PO activities. It is worth to mention that high AgPO2 activities triggered by ClipB and ClipB cofactors indicated ClipA4, CllipA6, ClipA7 and ClipA12 help AgPPO2 participate in melanization killing *P. berghei* oocysts.

Our previous work showed ClipA4, ClipA6, ClipA7, and ClipA12 were in same clade, indicating high and similar expression profiles of these four genes in whole development of *An. gambiae,* all these four genes can be detected in mosquito hemolyph (Xiaolong Cao et al., 2017; Xuesong He et al., 2017). In case of the transcription levels of four cofactor genes, there is no big difference between male and female in multiple development stages (Xiaolong Cao et al., 2017). Interestingly, there were obviously low expression of ClipA4 and ClipA12 during egg (early mid) and pupa (mid-late) stages, while ClipA6 kept high level during in whole life cycle; besides, four ClipAs clustered with some ClipBs such as ClipB5, ClipB1, and ClipB4, suggesting they may participate in same immune response. Here we selected biochemically identified AgClipB9 as endogenous PAP. We have determined that the cleavages and conformation changes of proClipAs are important for playing roles as cofactors (Fig), proClipA6, proClipA7, and ClipA12 can be cut efficiently, however, the processing proClipA4 by ClipB9xa is not efficient, one possible reason is that the modification of factorXa site will influence the cleavage efficiency of ClipB9xa; another reason is that unlike ClipA6/A7/A12(GFR*IGA), the predicted cleavage site of ClipA4 is EQNK*FNEI, which was difficult for ClipB9xa to process, we believe there are more PAPs needs to be identified, thus it was guessed that some unidentified ClipB proteins could cleaved ClipA4 efficiently. Cleavage sites identification demonstrated ClipA4 can be cleaved at EQNK*FNEI and PVGR*GCGL by PAP3 or ClipB9Xa, the sequential processing Native gel (Fig 5B and 6B) of proClipA4 by

ClipB9Xa and PAP3 showed different patterns, because the lower band, triggering the smear move slower on native PAGE, can form easier in PAP3 treated ClipA4 compared with processing using ClipB9Xa (Fig 5B and 6B).

Different functions and action mechanisms of PPOs seem to take place in *An. gambiae*, in addition to catalyzing the process of melanin production, PPOs also participate in wound healing and coagulation reactions. RNA interference (RNAi) of PPO2, PPO3, or PPO9 increased the intensity of malaria parasite infection for the late phase killing of *P. berghei* oocysts, while loss of PPO4, PPO5, and PPO6 have no positive effects on *P. berghei* invasion, suggesting diverse roles of PPOs invading microorganisms in *An. gambiae*. Two thirds of AgPPOs (PPO1-PPO4, and PPO7-PPO8) can be cleaved by AgClipB9xa and MsPAP3 (data not shown). However, only PPO2 and PPO7 achieve high PO activities in the presents of PAPs (ClipB9 and PAP3) and ClipAs. We guessed 9 PPO proteins involved in immune response against pathogens in multiple ways. It was reported that in *Drosophila*, unlike DmPPO1 and PPO2, DmPPO3 is not required for Toll and Imd pathway and it contributes to melanization during the encapsulation process. Therefore, PPOs that couldn't achieve high PO activity using ClipAs and ClipB may be involved in coagulation and encapsulation. Another possibility is the activation of other PPOs may need the presence of unidentified cofactors.

Although four ClipA proteins (ClipA4, ClipA6, ClipA7, and ClipA12) contain one clip domain and one serine protease like domain, molecular weight of ClipA7 is much larger than other three ClipAs because of a 437 aa unknown region between Clip domain (28-71) and serine protease like domain (593-763) (from 77 to 514). The function of this region may contribute to it works as a negative regulator in melanizating *P. berghei* oocysts, which was identified by silencing ClipA7 gene using RNAi (Awada et al., 2018). Additionally, we couldn't exclude the compensation effects of other SPH1 (ClipA6, ClipA12) after knocking down one single SPH1. The SP-SPH network seems much more complicated in managing the immune reactions in mosquitoes, and further explorations need to be devoted.

In summary, this report describes ClipA4, ClipA6, ClipA7, and ClipA12 as ClipB cofactor units regulating melanization and demonstrated their essential roles in increasing PPO activation in *An. gambiae*. Our biochemical experiments demonstrated that both *M. sexta* PAP3 and *An. gambiae* ClipB9xa can process ClipA4, ClipA6, ClipA7, and ClipA12, which process is indispensable for them to play cofactor function, these results indicate functional conservation of PAPs in insects in the process of evolution. Our data additionally revealed the ClipA4 needs to work with ClipA6/ClipA7/ClipA12 to form cofactor, those results highlight the importance of ClipAs function as cofactors and reveal the complexity of melanization regulatory in *An. gambiae*.

Figures

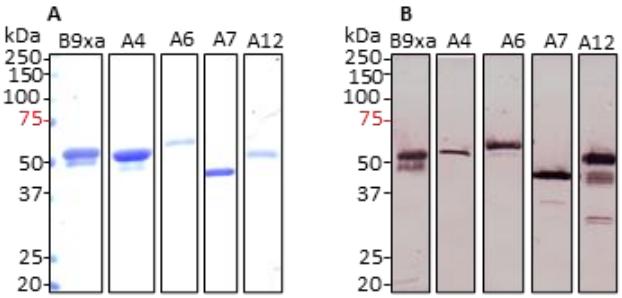


Fig 1. SDS-PAGE and immunoblot analysis of purified *A.gambiae* ProClipB9xa, proClipA4, proClipA6, proClipA7 and ClipA12. The purified proteins were resolved by 10% SDS-PAGE followed by Coomassie Brilliant Blue (CBB) staining (left, 1μg) or immunoblotting (right) using 1:1000 diluted antibody against 6×histidine affinity tag. Positions and sizes (in kDa) of the prestained Mr standards are marked on the left, with the 75 kDa marker highlighted red.

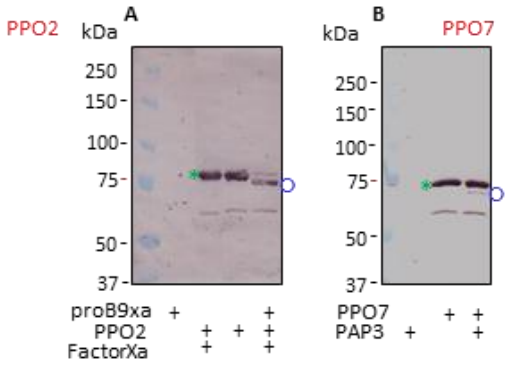


Fig 2. Limited proteolysis of *A. gambiae* PPO2(AgPPO2) and PPO7(AgPPO7) by ClipB9xa and PAP3.

Purified PPO2(500ng/ul, 1ul) and PPO7 (500 ng/µl, 1 µl) were reacted with ClipB9xa (100ng/ul, 1ul) (**A**) or PAP3 (40 ng/µl, 1 µl) (**B**) separately, the mixture was incubated with buffer (20 mM Tris-HCl, 5 mM CaCl2, 0.001% Tween 20, pH 7.5) at 37 °C for 1 h. The reaction mixtures were treated with SDS-sample buffer, separated by 7% SDS-PAGE, and subjected to immunoblot analysis using 1:1000 diluted antibody against PPO as the primary antibody. The red triangle represented the cleavage site around 70kD, green triangle indicated the non-classical cleavage site of PO. The number indicated the maker of protein molecular weight(kD) in the left lane of blot.

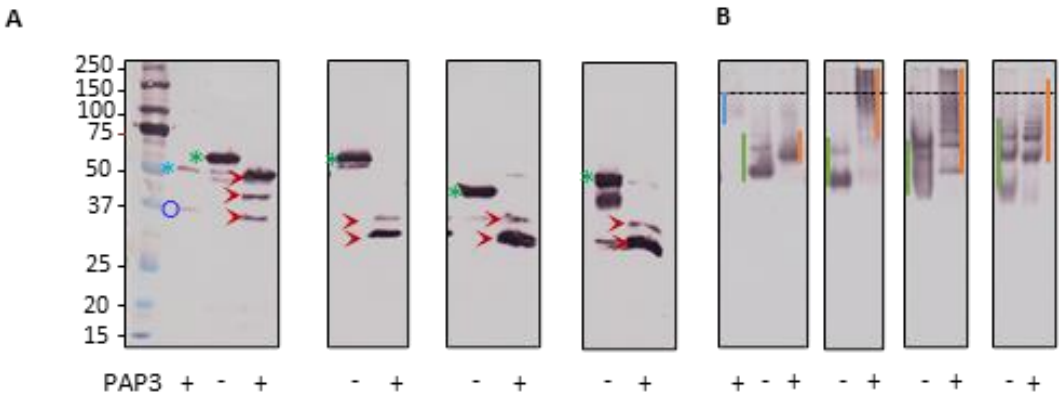


Fig 3. Processing ClipA4, ClipA6, ClipA7 and ClipA12 by PAP3 using DTT gel and native gel. (A)

Recombinant ClipA4 (220ng ng/µl, 1 µl), A6(300ng/ul, 1ul), A7(300ng/ul, 1ul) and A12(200) were reacted with PAP3 (40 ng/µl, 1 µl), the mixture was incubated in 18 ul buffer (20 mM Tris-HCl, 5 mM CaCl2, 0.001% Tween 20, pH 7.5) at 37 °C for 1 h. The reaction mixtures were treated with 1 ×SDS-sample

buffer or native sample buffer at 95 ◦C for 5 min or the same buffer lacking SDS and DTT at 25 ◦C for 5 min prior to 10% SDS-PAGE(A) and 10% native PAGE (**B**). The blue marker "*" represented proPAP3, "○" indicated cleaved PAP3, green marker "*" indicated the proClipA4, A6, A7 and A12." ">" indicated the cleaved bands of ClipA4, A6, A7 and A12. The marker "|" showed the proClipA4, A6, A7 and A12 on Native-Page, "|" indicated the cleaved bands of ClipA4, A6, A7 and A12 on Native-Page, PAP3 were marked as "|".

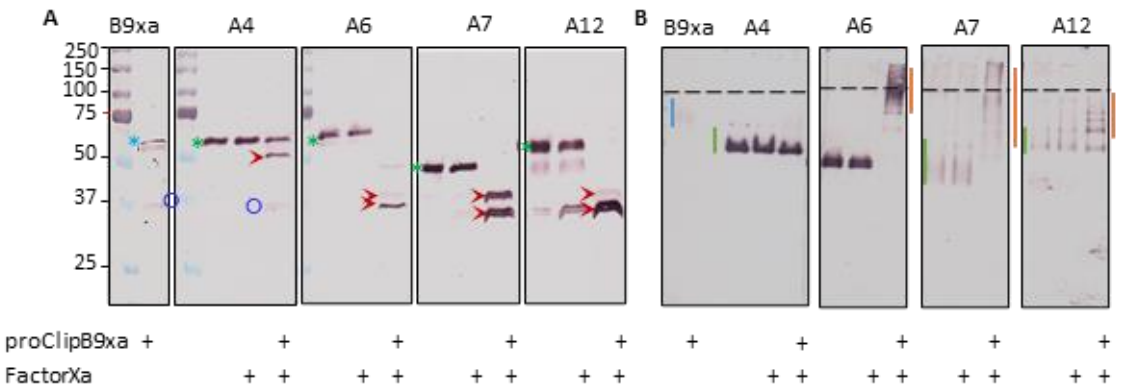


Fig 4. Processing ClipA4, ClipA6, ClipA7 and ClipA12 by ClipB9xa using DTT gel and native gel. (A) Recombinant ClipA4 (220ng ng/µl, 1 µl), A6(300ng/ul, 1ul), A7(300ng/ul, 1ul) and A12(200) were reacted with ClipB9xa (100 ng/µl, 1 µl), the mixture was incubated in 18 ul buffer (20 mM Tris-HCl, 5 mM CaCl2, 0.001% Tween 20, pH 7.5) at 37 °C for 1 h. The reaction mixtures were treated with SDS-sample buffer or native sample buffer, separated by 10% SDS-PAGE (**A**) or 10% Native-PAGE (**B**) and the membranes were analyzed using 1:1000 diluted His antiserum (**A**, **B**) as the primary antibody. The blue marker" *" represented proClipB9xa, "○" indicated cleaved ClipB9xa by factorxa, green marker "*" indicated the proClipA4, A6, A7 and A12. ">" indicated the cleaved bands of ClipA4, A6, A7 and A12. The marker "|" showed the proClipA4, A6, A7 and A12 on Native-Page, "|" indicated the cleaved bands of ClipA4, A6, A7 and A12 on Native-Page.

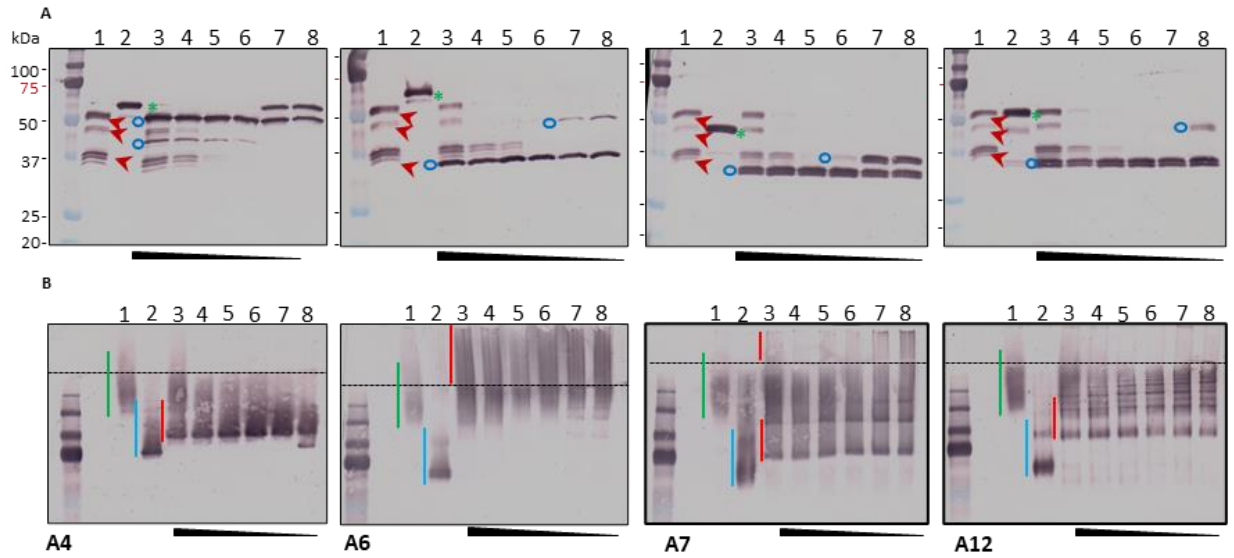


Fig 5. PAP3 processing of An. gambiae proClipA4, A6, A7, and A12 analyzed by 10% SDS and native PAGE followed by immunoblotting. Aliquots of the purified proClipAs (200 ng) were incubated with 100 (lane 3), 50 (lane 4), 25 (lane 5), 12.5 (lane 6), 6.25 (lane 7), 3.1 (lane 8), and 0 (lane 2) ng of PAP3 in 10 µl buffer B at 37 ◦C for 1 h. The mixtures and PAP3 control (100 ng, lane 1) were subjected to 10% SDS-PAGE (A) and 10% native (B) PAGE, as described in the legend to Fig. 3. After electrotransfer, immunoblot analysis was performed using antibody against the hexahistidine tag (A and B). In panels A, PAP3, proClipAs, and major cleavage products are marked by red arrowheads, green asterisks, and blue circles, respectively. Positions and sizes of the prestained Mr standards are indicated, with the 75 kDa marker highlighted red. In panel B, the dashed line divides the stacking and separating gels. The smeared bands of PAP3, proClipAs, and cleavage products are marked by green, blue, and red vertical bars, respectively.

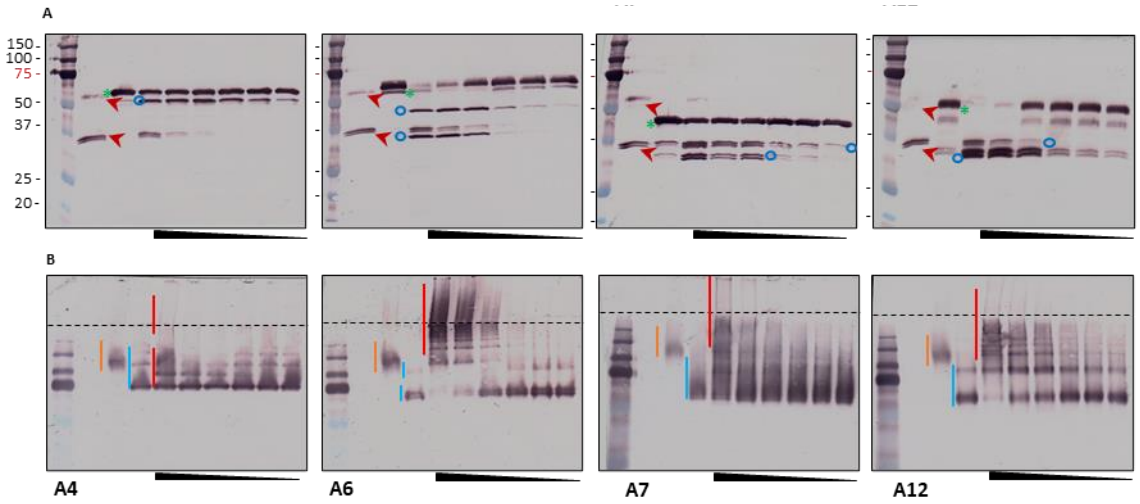


Fig 6. ClipB9xa processing of An. gambiae proClipA4, A6, A7, and A12 analyzed by 10% SDS and native PAGE followed by immunoblotting. Aliquots of the purified proClipAs (200 ng) were incubated with 100 (lane 3), 50 (lane 4), 25 (lane 5), 12.5 (lane 6), 6.25 (lane 7), 3.1 (lane 8), and 0 (lane 2) ng of actived ClipB9Xa in 10 µl buffer B at 37 ◦C for 1 h. The mixtures and ClipB9Xa control (100 ng, lane 1) were subjected to 10% SDS-PAGE (A) and 10% native (B) PAGE, as described in the legend to Fig. 3. The immunoblot analysis was performed using antibody against the hexahistidine tag (A and B). In panels A, ClipB9Xa, proClipAs, and major cleavage products are marked by red arrowheads, green asterisks, and blue circles, respectively. Positions and sizes of the prestained Mr standards are indicated, with the 75 kDa marker highlighted red. In panel B, the dashed line divides the stacking and separating gels. The smeared bands of PAP3, proClipAs, and cleavage products are marked by green, blue, and red vertical bars, respectively.

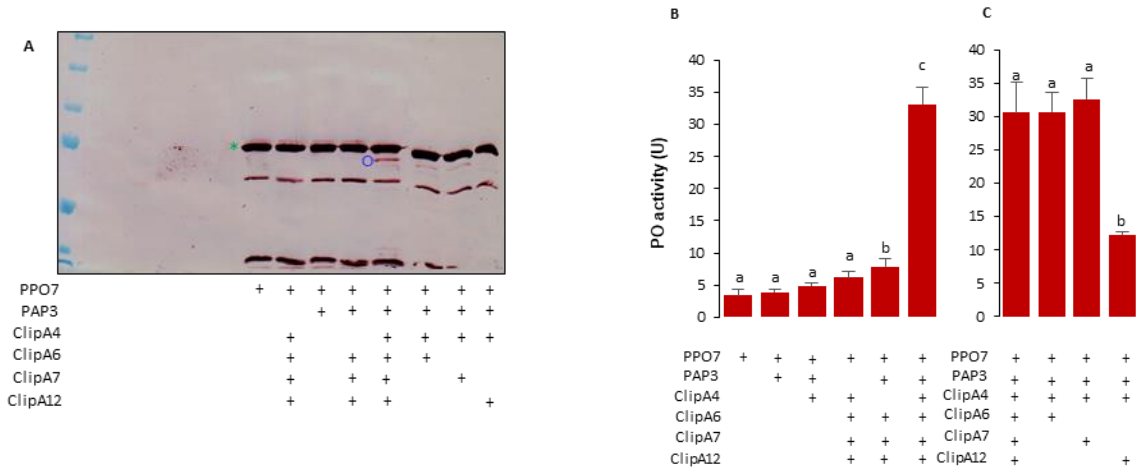


Fig 7. *An. gamibiae* proPO7 activation using Manduca PAP3 and ClipA4, ClipA6, ClipA7 and ClipA12 and its cleavage detection using Western blot. (A) Recombinant *An. gamibiae* proPO7 (1 µl, 0.49 µg/ul) and Manduca PAP3 (1µl, 40ng) were incubated on ice with ClipA4 (1µl, 200-300ng), ClipA6/A7/A12 (1µl, 200-300ng), or both (1µl + 1µl) in a total volume of 20 µl Tris-HCl, pH=7.5, 5mM Cacl2. Dopamine (150 µl, 2.0 mM) was added to the reaction mixtures 60 min, the samples were analyzed using immunoblotting(**A**) or PO activity was measured immediately (**B**). **(C)** Same amount as described above of *An. gambiae* proPO7 and *M. sexta* PAP3 (1µl, 40ng) were incubated with ClipA4 paired with ClipA6, ClipA7 and ClipA12 on ice, respectively. All test conditions were same as described in **(A),** the mixtures were used to do Western blot analysis **(A)** or PO activity detection**(C)**. One way ANOVA analysis was used to do the statistical analysis ($p<0.05$). Three biological replicates were performed and plotted in the bar graphs (mean ± SE, n = 3).

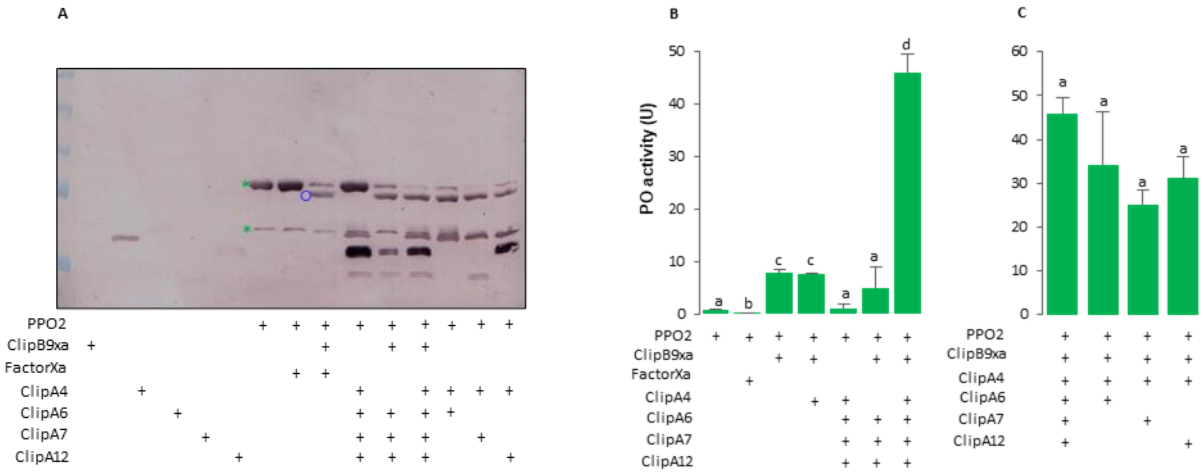
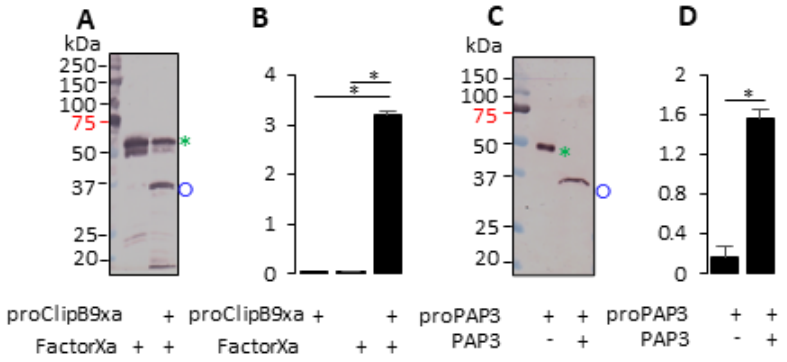
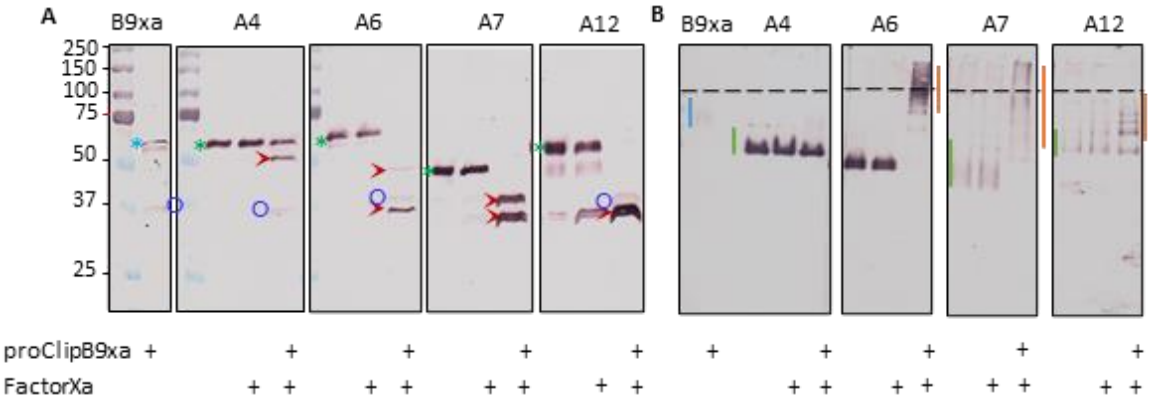


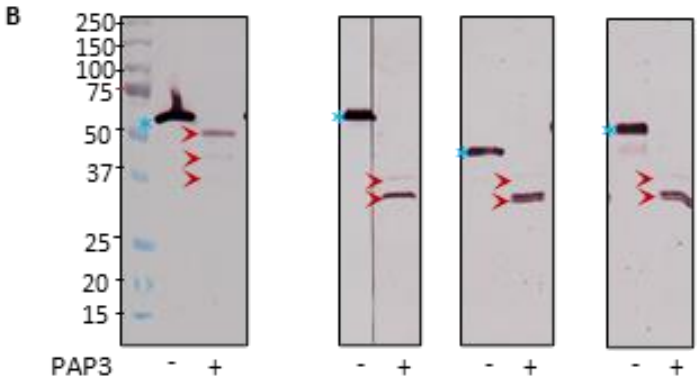
Fig 9. *An. gamibiae* proPO2 activation using ClipB9xa and ClipA4, ClipA6, ClipA7 and ClipA12 and the detection of proPO2 processing using immunoblotting parallel with PO activation. (A) Recombinant *An. gamibiae* proPO2 (1 µl, 0.47 µg/ul) and ClipB9xa (1µl, 100ng/ul) cleaved by Factorxa (1ul, 40ng/ul) were incubated with ClipA4 (1µl, 200-300ng), ClipA6/A7/A12 (1µl, 200-300ng), or both (1µl + 1µl) on ice, in a total volume of 20 µl Tris-HCl, pH=7.5, 5mM Cacl2. Dopamine (150 µl, 2.0 mM) was added to the reaction mixtures for 1 hour and western blot **(A)** or PO activity **(B)** was measured immediately. **(C)** Same amount as described above of *An. gambiae* proPO2 and ClipB9xa were incubated with ClipA4 paired with ClipA6, ClipA7 and ClipA12 on ice, respectively. The mixtures were used to do immunoblotting **(A)** and PO activity testing**(C)**. Same test condition was used as described in **(A)**. One way ANOVA analysis was used to do the statistical analysis ($p<0.05$). Three biological replicates were performed and plotted in the bar graphs (mean ± SE, n = 3).



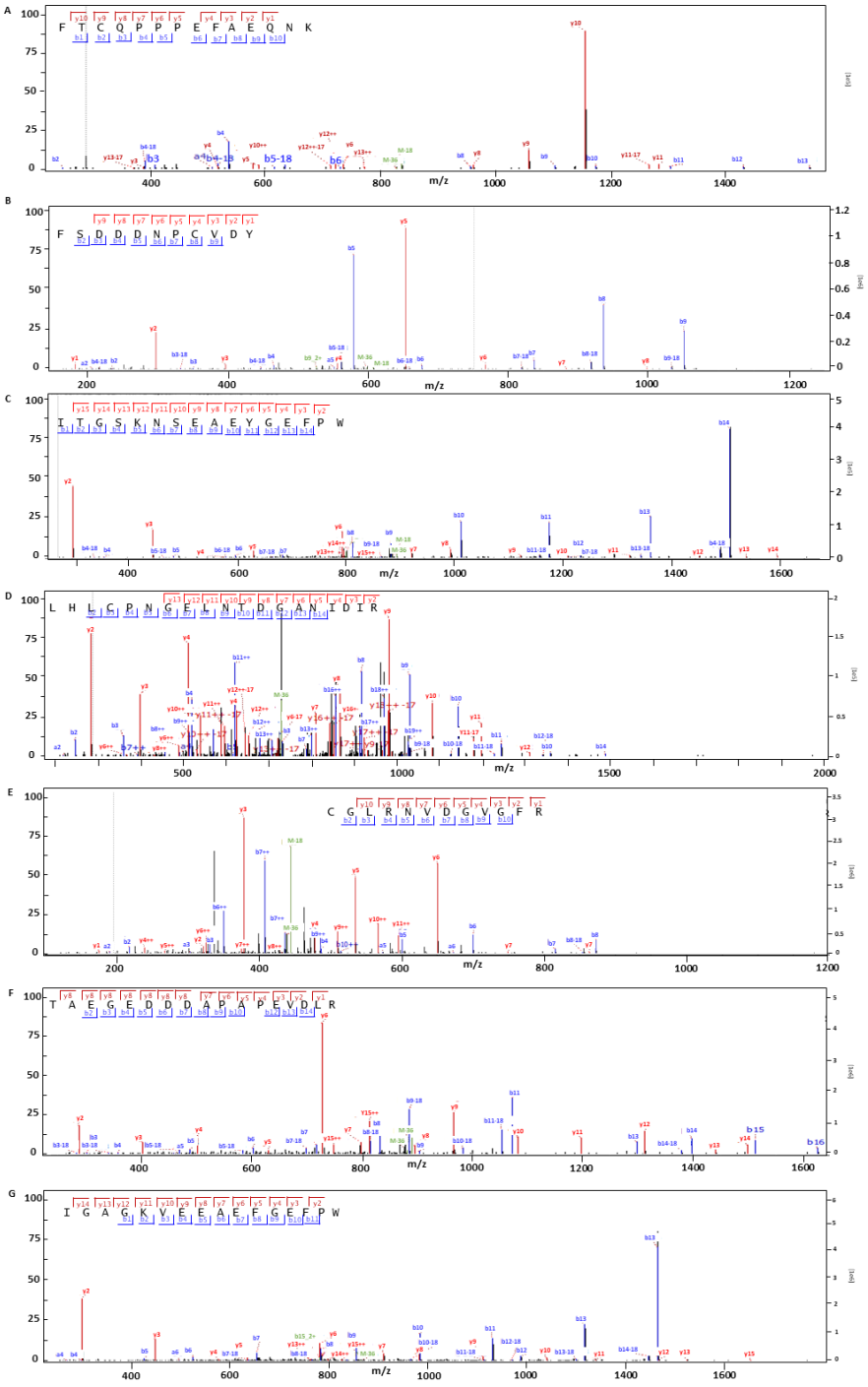
FigS1: **The activation of ClipB9xa and PAP3 and the IEARase activity of ClipB9xa and PAP3.** The processing of ClipB9xa using Factorxa and PAP3. 2.5ug of proClipB9xa was incubated with 0.8ug of Factoxa in the total volume is 25ul, after incubation in 37℃ for 3 hours, 2.5ul of total volume to do the western(**A**). (**B**) IEARase activity of ClipB9xa detection. 2.5 μl cleaved ClipB9Xa were mixed with 150 μl IEAR*p*Na (25μM) and tested its IEARase activity under 405nm absorbtion. The cleavage of ProPAP3(0.4ug) using PAP3(40ng) (C), the IEARase activity of PAP3. 150 μl of 25 μM IEARase mixed with 0.2ug PAP3. Three biological replicates were performed and plotted in the bar graphs (mean ± SE, n = 3).



FigS2. Processing of ClipA4, ClipA6, ClipA7 and ClipA12 by B9Xa using DTT gel and native gel. (A) Recombinant ClipA4 (200ng ng/μl, 1 μl), A6(200ng/ μl, 1 μl), A7(200ng/ul, 1 μl) and A12(200) were reacted with ClipB9Xa (100 ng/μl, 1 μl), the mixture was incubated in 10 μl buffer (20 mM Tris-HCl, 5 mM CaCl2, 0.001% Tween 20, pH 7.5) at 37 °C for 1 h. The reaction mixtures were treated with SDS-sample buffer or native sample buffer, separated by 10% SDS-PAGE (**A**) or 10% Native-PAGE (**B**)and the membranes were analyzed using 1:1000 diluted Anti *M. sexta* SPH1(ClipA6, A7, and A12) or *M. sexta* SPH2 (ClipA4) as the primary antibody. Green "*" represented proClipB9Xa, "○" indicated cleaved ClipB9Xa, blue "*" indicated the proClipA4, A6, A7 and A12. Red arrow indicated the cleaved bands of ClipA4, A6, A7 and A12. Blue, green, and orange vertical bar showed the ClipB9Xa; proClipA4, A6, A7 and A12; and cleaved proClipA4, A6, A7 and A12 on native gel.



FigS3. Cleavge of ClipA4, ClipA6, ClipA7 and ClipA12 by PAP3 using DTT gel. (A) Recombinant ClipA4 (200ng ng/µl, 1 µl), A6(200ng/ul, 1ul), A7(200ng/ul, 1ul) and A12(200) were reacted with PAP3 (40 ng/µl, 1 µl), the mixture was incubated in 18 ul buffer (20 mM Tris-HCl, 5 mM CaCl2, 0.001% Tween 20, pH 7.5) at 37 °C for 1 h. The reaction mixtures were treated with SDS-sample buffer or native sample buffer, separated by 10% SDS-PAGE using 1:1000 diluted Anti *M. sexta* SPH1(ClipA6, A7, and A12) or *M. sexta* SPH2 (ClipA4) as the primary antibody. Blue "*" indicated the proClipA4, A6, A7 and A12. Red arrow indicated the cleaved bands of ClipA4, A6, A7 and A12.



FigS4. Second MS spectra of Chymotryspin/V8 protease/ LysC processed peptides, which exposed the cleavage sites by ClipB9Xa or PAP3, from the purified proClipA4 (A), A6 (B, C), A7 (D, E), A12 (F, G). (A) FTCQPPPEFAEQNK, (B) FSDDNPCVDY, (C) ITGSKNSEAEYGEFPW, (D) LHLCPNGELNTDGANIIDIR, (E) CGLRNVDGVGFR, (F) TAEGEDDDAPAPEVDLR, (G) IGAGKVEEAEFGEFPW.

Table

Table 1

Sample name	1st cutting site	m/z	Scan time (min)	Intensity of 1st cutting site(Peak area)	2nd cutting site	m/z	Scan time	Intensity of 2nd cutting site(Peak area)
ProClipA4	FTCQPPPEFAEQNK*	846.9	64	nd	NONE	NONE	NONE	NONE
ProClipA4+PAP3	FTCQPPPEFAEQNK*	846.9	64	1.55×10^{9} (10 times higher than B9xa processed)	PVGR*GCGLRN	639.82	91	1.52×10^{9}
ProClipA4+ClipB9xa(Cleaved by factorxa)	FTCQPPPEFAEQNK*	846.9	64	1.5×10^{8}	PVGR*GCGLRN	639.82	91	9.9×10^{5}
ProClipA7	LHLCPNGELNTDGANIIDIR*	745.7	89	1.93×10^{8}	CGLRNVDGVGFR*	450.6	54	3.77×10^{8}
ProClipA7+PAP3	LHLCPNGELNTDGANIIDIR*	745.7	89	1.25×10^{10} (64 times higher than pro)	CGLRNVDGVGFR*	450.6	54	8.13×10^{10} (277 times higher than pro)
ProClipA12	TAEGEDDDAPAPEVDLR*	900.4	63	7×10^{6}	GVGFR*IGAGKVEEAEFGEFPW	883.4	106	7.8×10^{8}
ProClipA12+PAP3	TAEGEDDDAPAPEVDLR*	900.4	63	1.95×10^{9} (3 times higher than pro)	GVGFR*IGAGKVEEAEFGEFPW	883.4	106	3.71×10^{9}(5 times higher than pro)
ProClipA6	IDIR*FSDDNPCVDY	616.2	76	6.99×10^{5}	LGFR*ITGSKNSEAEYGEFPW	923.9	86	1.65×10^{6}
ProClipA6+ClipB9xa(Cleaved by factorxa)	IDIR*FSDDNPCVDY	616.2	76	2.71×10^{8} (378 times higher than pro)	LGFR*ITGSKNSEAEYGEFPW	923.9	86	1.84×10^{8} (111 times higher than pro)
ProClipA6	IDIR*FSDDNPCVDY	616.2	76	1.4×10^{4}	LGFR*ITGSKNSEAEYGEFPW	923.9	86	3.92×10^{6}
ProClipA6+PAP3	IDIR*FSDDNPCVDY	616.2	76	1.85×10^{8} (1321 times higher than pro)	LGFR*ITGSKNSEAEYGEFPW	923.9	86	1.17×10^{9} (298 times higher than pro)

Table 1. Determination of the *An. gambiae* ClipAs cleavage sites processed by ClipB9Xa and PAP3 using LC-MS/MS analysis. A list of peptides identified as products of PAP3/ClipB9Xa. Chymotrypsin, V8 protease, or LysC were used to process the PAP3/ClipB9Xa cleaved ClipAs or their pro-forms. The table showed the parent ions'm/z, scan time, and peak areas of the peptides. NONE represents no theoretical cleavage sites were predicted; nd means not detected.

Reference

El Moussawi, L., Nakhleh, J., Kamareddine, L., Osta, M.A., 2019. The mosquito melanization response requires hierarchical activation of non-catalytic clip domain serine protease homologs. PLoS Pathog., 15, e1008194.

Gupta, S., Wang, Y., Jiang, H., 2005. Manduca sexta prophenoloxidase (proPO) activation requires proPO-activating proteinase (PAP) and serine proteinase homologs (SPHs) simultaneously. Insect Biochem. Mol. Biol., 35, 241−248.

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.

Cao, X., Gulati, M., Jiang, H., 2017. Serine protease-related proteins in the malaria mosquito, Anopheles gambiae. Insect Biochem. Mol. Biol., 88, 48–62.

Cao, X., Wang, Y., Rogers, J., Hartson, S., Kanost, M.K., Jiang, H., 2020. Changes in composition and levels of hemolymph proteins during metamorphosis of Manduca sexta. Insect Biochem. Mol. Biol., 127, 103489.

Cox, J., Mann, M., 2008. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat. Biotechnol., 26, 1367–1372.

He, Y., Cao, X., Zhang, S., Rogers, J., Hartson, S., Jiang, H., 2016. Changes in the plasma proteome of Manduca sexta larvae in relation to the transcriptome variations after an immune challenge: evidence for high molecular weight immune complex formation. Mol. Cell. Proteomics, 15, 1176–1187.

Jiang, H., Wang, Y., Kanost, M.R., 1998. Pro-phenoloxidase 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., Yu, X.Q., Kanost, M.R., 2003a. Prophenoloxidase-activating proteinase-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 from Manduca sexta hemolymph: a clip-domain serine proteinase regulated by serpin-1J and serine proteinase homologs. Insect Biochem. Mol. Biol., 33, 1049−1060.

Kwon, T.H., Kim, M.S., Choi, H.W., Joo, C.H., Cho, M.Y., Lee, B.L., 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.

Lee, K.Y., Zhang, R., Kim, M.S., Park, J.W., Park, H.Y., Kawabata, S., Lee, B.L., 2002. A zymogen form of masquerade-like serine proteinase homologue is cleaved during pro-phenoloxidase activation by Ca2+ in coleopteran and Tenebrio molitor larvae. Eur. J. Biochem. 269, 4375–4383.

Lee, S.Y., Kwon, T.H., Hyun, J.H., Choi, J.S., Kawabata, S.I., Iwanaga, S., Lee, B.L., 1998. In vitro activation of pro-phenol-oxidase by two kinds of pro-phenol-oxidase-activating factors isolated from hemolymph of coleopteran, Holotrichia diomphalia larvae. Eur. J. Biochem., 254, 50−57.

Lu, Z., Jiang, H., 2008. Expression of Manduca sexta serine proteinase homolog precursors in insect cells and their proteolytic activation. Insect Biochem. Mol. Biol., 38, 89–98.

Wang, Y., Lu, Z., Jiang, H., 2014. Manduca sexta proprophenoloxidase activating proteinase-3 (PAP3) stimulates melanization by activating proPAP3, proSPHs, and proPOs. Insect Biochem. Mol. Biol. 50, 82–91.

Yu, X.Q., Jiang, H., Wang, Y., Kanost, M.R., 2003. Nonproteolytic serine protease homologs are involved in prophenoloxidase activation in the tobacco hornworm, Manduca sexta. Insect Biochem. Mol. Biol., 33, 197−208.

Schnitger AK, Kafatos FC, Osta MA. The melanization reaction is not required for survival of Anopheles gambiae mosquitoes after bacterial infections. J Biol Chem. 2007 Jul 27;282(30):21884-8. doi: 10.1074/jbc.M701635200. Epub 2007 May 30. PMID: 17537726.

Nakhleh J, Christophides GK, Osta MA. The serine protease homolog CLIPA14 modulates the intensity of the immune response in the mosquito Anopheles gambiae. J Biol Chem. 2017 Nov 3;292(44):18217-18226. doi: 10.1074/jbc.M117.797787. Epub 2017 Sep 19. PMID: 28928218; PMCID: PMC5672044.

El Moussawi L, Nakhleh J, Kamareddine L, Osta MA. The mosquito melanization response requires hierarchical activation of non-catalytic clip domain serine protease homologs. PLoS Pathog. 2019 Nov 25;15(11):e1008194. doi: 10.1371/journal.ppat.1008194. PMID: 31765430; PMCID: PMC6901238.

Povelones M, Bhagavatula L, Yassine H, Tan LA, Upton LM, Osta MA, Christophides GK. The CLIP-domain serine protease homolog SPCLIP1 regulates complement recruitment to microbial surfaces in the malaria mosquito Anopheles gambiae. PLoS Pathog. 2013;9(9):e1003623. doi: 10.1371/journal.ppat.1003623. Epub 2013 Sep 5. PMID: 24039584; PMCID: PMC3764210.

Ji Y, Lu T, Zou Z, Wang Y. Aedes aegypti CLIPB9 activates prophenoloxidase-3 in the presence of CLIPA14 after fungal infection. Front Immunol. 2022 Jul 28;13:927322. doi: 10.3389/fimmu.2022.927322. PMID: 35967454; PMCID: PMC9365933.

An C, Budd A, Kanost MR, Michel K. Characterization of a regulatory unit that controls melanization and affects longevity of mosquitoes. Cell Mol Life Sci. 2011 Jun;68(11):1929-39. doi: 10.1007/s00018-010-0543-z. Epub 2010 Oct 17. PMID: 20953892; PMCID: PMC3070200.

Yassine H, Kamareddine L, Chamat S, Christophides GK, Osta MA. A serine protease homolog negatively regulates TEP1 consumption in systemic infections of the malaria vector Anopheles gambiae. J Innate Immun. 2014;6(6):806-18. doi: 10.1159/000363296. Epub 2014 Jul 8. PMID: 25012124; PMCID: PMC4813755.

Zhang X, An C, Sprigg K, Michel K. CLIPB8 is part of the prophenoloxidase activation system in Anopheles gambiae mosquitoes. Insect Biochem Mol Biol. 2016 Apr;71:106-15. doi:

10.1016/j.ibmb.2016.02.008. Epub 2016 Feb 27. PMID: 26926112; PMCID: PMC4828722.

Zhang X, Li M, El Moussawi L, Saab S, Zhang S, Osta MA, Michel K. CLIPB10 is a Terminal Protease in the Regulatory Network That Controls Melanization in the African Malaria Mosquito Anopheles gambiae. Front Cell Infect Microbiol. 2021 Jan 15;10:585986. doi: 10.3389/fcimb.2020.585986. PMID: 33520733; PMCID: PMC7843523.

Sousa GL, Bishnoi R, Baxter RHG, Povelones M. The CLIP-domain serine protease CLIPC9 regulates melanization downstream of SPCLIP1, CLIPA8, and CLIPA28 in the malaria vector Anopheles gambiae. PLoS Pathog. 2020 Oct 12;16(10):e1008985. doi: 10.1371/journal.ppat.1008985. PMID: 33045027; PMCID: PMC7580898.

He X, Cao X, He Y, Bhattarai K, Rogers J, Hartson S, Jiang H. Hemolymph proteins of Anopheles gambiae larvae infected by Escherichia coli. Dev Comp Immunol. 2017 Sep;74:110-124. doi: 10.1016/j.dci.2017.04.009. Epub 2017 Apr 19. PMID: 28431895; PMCID: PMC5531190.

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. Evolutionary dynamics of immune-related genes and pathways in disease-vector mosquitoes. Science. 2007 Jun 22;316(5832):1738-43. doi: 10.1126/science.1139862. PMID: 17588928; PMCID: PMC2042107.

CHAPTER III

SERINE PROTEASE HOMOLOGS CSPH35 AND CSPH242 ACT AS COFACTORS FOR PROPHENOLOXIDASE ACTIVATION IN DROSOPHILA MELANOGASTOR

Qiao Jin[a], Yang Wang[a], Yingxia Hu[a], Steven D. Hartson[b], Haodong Yin, Chao Xiong[a], Tisheng Shan[a], Haobo Jiang[a]

[a] Department of Entomology and Plant Pathology, Oklahoma State University,

Stillwater, OK 74078, USA

[b] Department of Biochemistry and Molecular Biology, Oklahoma State University,

Stillwater, OK, 74078, USA

Key words: clip domain, insect immunity, melanization, hemolymph protein.

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
E-mail: haobo.jiang@okstate.edu

The abbreviations used are: PO and PPO, phenoloxidase and its precursor; PAP, PPO activating protease; SP and SPH, serine protease and its noncatalytic homolog; CLIP, clip-domain SP(H), PGRP, peptidoglycan recognition protein; SDS-PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis.

Abstract

Phenoloxidase (PO), precisely regulated by serine protease (SP) cascade, plays critical roles in catalyzing melanization in insects. Clip domain containing serine protease (cSP) and serine protease homologue (cSPH) majorly consists of the SP cascade pathway. At least in some insects, POs are activated from their precursors PPOs by prophenoloxidase (PPO) activating proteases (PAPs) in the presence of clip-domain SPHs that serve as PAP cofactors. However, it's unclear if PAP and its cofactors are required for PPO activation in *Drosophila*. In this study, we identified cSPH242 and cSPH35 in *D. melanogaster* as cofactors of SP7 (MP2), a PAP can process DmPPO1 into active form. cSPH242 and cSPH35 are the orthologs of *M. sexta* SPH2 and SPH1 separately. Melanization in hemolymph of cSPH242 or cSPH35 single knocking down was inhibited. In the in vitro conditions, DmPPO1 was highly activated in the presence of SP7 and active form of cSPH242 or cSPH35. High level of PO activity suggests that cofactor-assisted PPO activation reaction is indispensable in *Drosophila*.

1. Introduction

Melanization is the arthropod-specific immune response and is considered as one of the most spectacular immune reactions in insects (Nappi and Christensen, 2005). Prophenoloxidase (PPO), which converts monophenols to diphenols, quinones, other reactive intermediates, and then to stable polymers at last, plays critical roles in the black pigment formation (Zhao et al., 2007 and 2011, Cerenius et al., 2008, González-Santoyo et al., 2012, Li et al, 2008, Jan Paul Dudzic et al., 2019). The activation of PPO is precisely regulated by the sequential activation of extracellular serine protease (SP) cascades, a system of serine proteases (SPs), noncatalytic serine protease homologs (SPHs), serpins, and specific recognition proteins. Among the components, SP and SPH, mostly containing one or more amino terminus followed by a linker

and a catalytic serine protease homolog domain (SP) or a noncatalytic serine protease homolog domain (SPH), were non-digestive serine protease and were classified as CLIPDs, CLIPCs, CLIPBs (SPs); and CLIPAs (SPHs). Upon infection, microbial surface molecules stimulate activation of an initiating modular serine protease, which then activates a CLIPC, which in turn activates a CLIPB. The active CLIPB then cleaves and activates an effector molecule (proSpätzle or prophenoloxidase). CLIPAs, lacking proteolytic activity, can regulate other CLIPs activity and forming high M_r weight complex. Specific enzymes, designated PPO activating proteases or PAPs, are responsible for cleavage activation (Kanost and Jiang, 2015, An et al., 2013). At least in some insects, PPO activation occurs in the presence of PAPs and their cofactors (Kwon et al., 2000; Yu et al., 2003, Jin et al., 2022, Ji et al., 2022).

There are 21, 45, 52, 55 and 110 CLIP genes in the genomes of *Apis mellifera, Manduca sexta, Tribolium castaneum, Drosophila melanogaster* and *Anopheles gambiae*, 7, 10, 18, 19 and 55 of which encode cSPHs, respectively (Cao et al., 2015 and 2017; Cao and Jiang, 2018). Some CLIPDs, CLIPCs, and CLIPBs have been identified as serine protease components involved in melanizing and other processes mostly using genetic methods, while biochemical roles of the cSPHs have not yet been defined in nearly all cases. In *D. melanogaster*, the peptidoglycan of gram-positive bacterial will be recognized by PRRs PGRP-SA, and GNBP1, initiating sequential proteolytic cleavage of the SPs ModSP (SP53), Grass (cSP1) Psh (cSP28)/Hayan (cSP33), and SPE (cSP4). This leads to the cleavage of Spz and the activation of Toll signaling in the fat body (middle). On the other hand, this extracellular SP pathway at the position of Hayan and Psh to Sp7 will activate PPO1(An et al., 2013, Yamamoto-Hino and Goto, 2016, Jan Paul Dudzic et al., 2019). Bovine coagulation factor Xa-activated *D. melanogastor* MP2Xa cleaved the purified M. sexta PPOs and yielded low levels of PO activity (An et al., 2013). Nevertheless, it is unclear if

assistantship from cSPH for PPO activation by PAP is essential or not both in the genetic and biochemical ways in *D. melanogastor*.

Although scarface (cSPH142) was identified as a negative regulator for JNK signalling during epithelial morphogenesis (Raphaël Rousset et al., 2010), and masquerade (cSPH79) had effect on nervous and muscle development in Drosophila embryo (Bernadette Murugasu-Oei et al., 1994 and 1996), cSPHs related to immunity have been rarely studied flies. In previous work, we and our colleagues have discovered and built an an SP-SPH-serpin network in *M. sexta* hemolymph for melanization and Toll pathway activation (Kanost and Jiang, 2015; Wang et al., 2020). Also, a platform for predicting functions of SPs and SPHs in holometabolous insects based on the genetic and biochemical data obtained from different species was built (Cao and Jiang, 2018). Recently, five of the ten CLIPAs (SPHI: 1a, 1b, 4 and 101; SPHII: 2) in *M. sexta* were found to form high M_r complexes that assist PPO activation by PAP3 (Jin et al., 2022). By analyzing the phylogenetic relationships along with the transcriptome and proteome data, cSPH242 and cSPH35 were selected as cofactor candidates of PAP in *D. melanogaster*. Single knocking down cSPH242 leads to the failure of melanin deposition on the epithelium of adults. Additionally, hemolymph melanization and PO activity inhibition were observed either silencing cSPH242 or cSPH35. After being activated by *M. sexta* PAP3, purified precursor of cSPH35 from Sf9 cells formed SPHI complexes with decreasing oligomerization states, while cleaved cSPH242, an SPHII, tended to form a lower Mr band, like SPH2 in *M. sexta*. $MP2x_a$ cleave DmPPO1 generated low PO activity. However, in the presence of cSPH242 and cSPH35 pairs, much higher PO activities were yielded in the reactions, indicating that an SPHI-II complex is required for PPO activation by a PAP in *D. melanogaster*. Our work firstly provides key evidence that co-factors of PAP are indispensable for PPO activation in Drosophila using genetic and biochemical methods. Good

applications of direct biochemical investigation were performed here to make up for the shortage of genetic techniques in precisely localizing the positions of certain proteins in a not fully understood pathway.

2. Methods and materials

2.1 Drosophila genetics and single gene (cSPH35 or 242) knock down verification

Crosses were maintained on standard fly food at 25°C (Kapila et al., 2021). The transgenic stocks were obtained or derived from the Bloomington Stock Center and are listed here with corresponding stock numbers (BL#): y[1] sc[*] v[1] sev[21]; P{y[+t7.7] v[+t1.8]=TRiP.HMC03613}attP40 (BL52875), y[1] sc[*] v[1] sev[21]; P{y[+t7.7] v[+t1.8]=TRiP.HMS02730}attP40 (BL44268), y[1] w[*]; P{w[+mC]=Act5C-GAL4}25FO1/CyO, y[+] (BL4414), y[1] w[*]; P{w[+mC]=tubP-GAL4}LL7/TM3, Sb[1] Ser[1](BL5138), y[1] w[*]; P{w[+mC]=Act5C-GAL4}17bFO1/TM6B, Tb[1](BL3954), y[1] w[*]; wg[Sp-1]/CyO; Dr[1]/TM3, Sb[1] (BL59967). Ubiquitous enhancers drive expression of GAL4, which results in transcriptional activation from the GAL4 binding sites (UAS) to drive expression of a gene of interest, is used for cSPH35 and 242 knockdowns by RNAi (GAL4/UAS system). To knock down cSPH35 or 242, the newly merged male adults containing GAL4 driver (BL4414) and driving the RNAi transgene was selected to cross with the virgin female adults of BL52875 or BL44268(ref). F1 offspring from these crosses were collected onto fresh food 1–2 day-post-eclosion for following experiments.

As described in 2.5., cDNA samples from RNA of transgenic lines and controls were produced as templates for knock down verification experiment using qPCR. The primers of cSPH35, 242, or control gene and the details of the qPCR analysis were shown in 2.5.

2.2 Cuticle color change detection in F1 offspring of the crossed transgenic species

F1 offsprings of cSPH35 or 242 knocking down species were collected and pictures from 4–5-day-old adults of F1 offspring of the crossed transgenic species were taken using optical microscope along with the camera (Amscope FMA050). Series of pictures were merged into one using ImageJ (https://imagej.nih.gov/ij/download.html).

2.3 Hemolymph preparation for DmPPO1 analysis, melanizaton observation, and PO activity detection

Sting 1 hole in 0.5ml microtubes and put into 1.5 ml microtubes with removed lid. 20 stinged adults were collected in the 0.5ml microtubes with holes at the bottom, centrifuge the 0.5ml microtubes within the 1.5ml one using 5000 rpm for 1 min at 4°C. Discard the 0.5ml microtubes and around 1 µl of hemolymph at the bottom of 1.5ml tubes was extract using pipette tips.

Hemolymph (1 µl) with PTU (0.1%) from different transgenic flies in a total volume of 12µl was incubated at room temperature for 0 or 30 minutes. The mixtures were subjected to 7.5% SDS-PAGE, transferred onto nitrocellulose membranes, and detected using DmPPO1 antisera as primary antibody and goat-anti-rabbit IgG conjugated to alkaline phosphatase (Bio-Rad) as the secondary antibody, and a BCIP-NBT substrate kit (Bio-Rad) for color development.

To analyze PO activity in *D. melanogaster*, extracted hemolymph (1 µl) was incubated with 150 µl of 2.0 mM dopamine in 50 mM sodium phosphate, pH 6.5 PO immediately, the activity was determined on a microplate reader immediately. The details of PO activity analysis was as described in 2.12.

To observe the melanization (color change) of *D. melanogaster*, 5 µl extracted

hemolymph from different transgenic lines were incubated at 25 °C for 30 minutes by loading on the microplate wells.

2.4 E. faecalis preparation and Infection of RNAi Lines with Knock-Down of cSPH35 and 242

The *E. faecalis* was revived from a glycerol stock maintained at -80 °C by streaking out on an Agar plate and incubating the plate at 37 °C overnight. This plate was then stored at 4°C. The evening before infections were to be performed, a single colony was aseptically picked from the plate and placed in 3 mL of sterile Lysogeny broth at 37 °C with shaking (180 rpm) overnight. The next day, the bacterial suspension was spun down (1 min at 13,000 rpm) and resuspended in LB to a final OD600 of 1.2 (Joanne R Chapman et al., 2020).

4–5-day-old adults, in this case, F1 offspring from the crosses described above (2.1), were infected with *E. faecalis* at a standardized concentration of OD600 = 1.2, by needle (0.1-mm stainless steel needle) pricking in the thorax (Katia Troha et al., 2018), and counted the number of survivors from 24 hours to 10 days post infection.

As described above, for black dots calculation of infected *D. melanogaster*. *E. faecalis* with OD600 = 0.8 was used to infected 4–5-day-old adults to induce the melanization of flies without killing them. The black dots in thoerax cuticle were calculated and observed 5 days post infection. Olympus microscopy (BH2-RFCA) was applied to observe the black dots in thorax and pictures were chaptured using Amscope (FMA050).

2.5. Multiple sequence alignment and analysis of phylogenetic relationships

To identify orthologs of the *D. melanogastor* cSPH35 and 242 in other *Drosophilidae* insects, the cSPH35 and 242 were used as queries in two BLASTP searches of the non-redundant sequence database at GenBank under default conditions. These searches, limited to

Drosophilidae, resulted in a total of 75 hits with similarity higher than 65%. After removal of the sequences lacking a signal peptide or an intact clip domain, 38 sequences were aligned with *M. sexta* SPH4, 1a, 1b, 101 (Jin et al., 2022) to build a phylogenetic tree (Fig. S3), which was used to select the two SPHs and their putative orthologs or paralogs from representative families for a focused study. Multiple sequence alignments of the entire proteins were performed using MUSCLE (Edgar, 2004), one module of MEGA X (Kumar et al., 2018). The aligned sequences were used to construct neighbor-joining trees using MEGA with bootstrap method for the phylogeny test (1000 replications, Poisson model, uniform rates, and complete deletion of gaps or missing data). FigTree 1.4.3 (http://tree.bio.ed.ac.uk/software/figtree) was used to display the phylogenetic trees.

2.6. cDNA cloning and recombinant production of procSPH35, and 242

cSPH242 cDNA was amplified from plasmid FBpp0112536 purchase from flybase (https://flybase.org/) by PCR using primers J1681F (5'-TCATATGGCTCCTCAGCAGAAC) and J1682 (5'-ACTCGAGTGCGGTGTACACAG). The product was cloned into pGEM-T vector (Promega) and, after sequence validation, the *Nde*I-*Xho*I fragments were inserted into the same sites in pMFFMH6. cSPH242/ pMFFMH6 was used to generate a baculovirus to express procSPH242 (GIHDYKDD DDKHMQAPQQN…AVYTAEQKLISEEDLHHHHHH) in insect cells (Sumathipala and Jiang, 2010), and the underlined part is encoded by the cDNA.

A procSPH35 fragment was amplified from plasmid FBpp0079653 purchased from flybase, using primers J1685 (5'-CATATGCAGGACTCTTCCTTG) and J1686 (5'-TTGCTCGAGGGGTGTATAGTGCCT). After cloning and sequence validation, the *Nde*I-*Xho*I fragment was retrieved and inserted into the same sites in pMFFMH6 to generate a baculovirus for producing procSPH35 (GIHDYKDDDDKHMQDSSL…RHYTPEQKLISEEDLHHHHHH)

in insect cells.

proMP2x$_a$ fragment was amplified from pFastBac inserted by proMP2xa (An et al., 2013) using primers J1551 (5’ -CATATGCAAGGAAGTTGTAGG) and J1552 (5’ -CTCGAGGGGACGAATGGTCTC). After TA cloning and sequencing, the *Nde*I-*Xho*I region was inserted into pMFFMH6. SPH101/ pMFFMH6 was used to produce proMP2x$_a$ (GIHDYKDD DDKHMQGSCR … ETIRPEQKLISEEDLHHHHHH) in insect cells, and the underlined part is encoded by SPH101 cDNA.

2.7. Expression and purification of the recombinant proSPHs

*Sf*9 cells at 2.4×10^6 cells/ml in 1.4 L of insect serum-free medium (Invitrogen Life Technologies) were separately infected with the baculovirus stocks at a multiplicity of infection of 10 and grown at 27 °C for 96 h with gentle agitation at 100 rpm. After the cells were removed by centrifugation at 5,000×*g* for 10 min, the pH of the conditioned medium was adjusted to 8.0 using 1.0 M Tris base. Cell debris and fine particles were spun down by centrifugation at 10,000×*g*, and the supernatant was diluted with 20 mM Tris-HCl, pH 8.0 (buffer A) to a final volume of 4.2 L. The solution was applied to a Q-Sepharose FF column (20 ml bed volume) at a flow rate of 5.0 ml/min and, following a washing step with 100 ml buffer A, bound proteins were eluted from the column with a linear gradient of 0−1.0 M NaCl in 240 ml of buffer A at a flow rate of 1.0 ml/min. The proSPH fractions were combined and loaded onto a 10 ml Ni^{2+}-nitrilotriacetic acid agarose column. After washing with 50 ml of 50 mM sodium phosphate, pH 8.0, the bound proteins were eluted with a gradient of 0−0.3 M imidazole in 90 ml of the same buffer. Fractions containing proSPHs were combined, dialyzed against 20 mM Tris-HCl, pH 7.6, and concentrated on Amicon Ultra-30 centrifugal filter devices (Millipore). The protein aliquots were rapidly frozen in liquid nitrogen prior to storage at -80 °C.

2.4. Expression and purification of the recombinant procSPH35 and 242

Sf9 cells at 2.4×10^6 cells/ml in 300 ml of insect serum-free medium (Invitrogen Life Technologies) were separately infected with the baculovirus stocks at a multiplicity of infection of 10 and grown at 27 °C for 96 h with gentle agitation at 100 rpm. After the cells were removed by centrifugation at 5,000×g for 10 min, the pH of the conditioned medium was adjusted to pH 6.4 using 10 N M Hcl. Cell debris and fine particles were spun down by centrifugation at 10,000×g, and 100ml supernatant was diluted with 1mM imidazole in 100ml distilled water. The solution was applied to a DS-Sepharose column (20 ml bed volume) at a flow rate of 1.5 ml/min and, following a washing step with 100 ml DS buffer A (10 mM KPO4, 1mM benzamidine, 0.01% Tween 20, PH=6.4), bound proteins were eluted from the column with a linear gradient of 0−1.0 M NaCl in 240 ml of buffer A at a flow rate of 1.5 ml/min. The procSPH fractions were combined, adjusted to pH 7.5, and loaded onto a 2 ml Ni^{2+}-nitrilotriacetic acid agarose column. After washing with 15 ml of 50 mM sodium phosphate, pH 7.5., the bound proteins were eluted with a gradient of 0−0.3 M imidazole in 20 ml of the same buffer. Fractions containing procSPHs were combined, dialyzed against 20 mM Tris-HCl, pH 7.5, and concentrated on Amicon Ultra-30 centrifugal filter devices (Millipore). The protein aliquots were rapidly frozen in liquid nitrogen prior to storage at -80 °C.

2.8. Profiling of the cSPH35, and 242 mRNA levels in tissues from various stages

To compare the cSPHs' RNA levels in different stages of whole body, cDNA samples (each equivalent to starting with 50 ng total RNA) were incubated with 1× iTag Universal SYBR Green Supermix (Bio-Rad) and specific primers (0.5 mM each) in triplicate in each reaction (10 μl). The ribosomal protein L32 (RPL) primers were J1091 (5'-GTGGTCCATTCCACTTCCGT) and J1092 (5'-TCATCATGTCGCGTGGATCA). Specific qRT-PCR primers were synthesized

for cSPH35 (J1971, 5'-GAAAGTCCATTCACCCTCCA and J1972 5'-TCCTTGCCGAACTTGTTCTT), cSPH242 (J1973, 5'- AAGACGTTGAATCCGACACC and J1974, 5'-CCGCTTCATTTTGCGATACT). Each primer pair was tested to ensure amplification of the target gene only and not close homologs. Efficiency of the amplification was determined individually and confirmed to be 90−110%. Thermal cycling conditions were 95 °C for 2 min and 40 cycles of 95 °C for 10 s and 60 °C for 30 s. After PCR was completed on a CFX Connect Real-Time PCR Detection System (Bio-Rad), melting curves of the products in all reactions were examined to ensure proper shape and Tm values. The mRNA levels were normalized against the internal control of RPL (set at 1.00) using corresponding Ct values for the same cDNA samples and the relative mRNA levels were calculated as 2-ΔCt, where ΔCt = CtcSPH - *CtRPL32*.

2.9. Activation of MP2x$_a$ and PAP3 following by the IEARase amidase detection

To get the active MP2x$_a$, proMP2x$_a$ (0.44 µg/ µl, 1 µl) was activated by Factorxa (0.44 µg/ µl, 1 µl) at 37℃ for 1 hour to the final volume of 4µl. Half of the active mixture (0.22 µg) were incubated with 150 µl 25µM acetyl-Ile-Glu-Ala-Arg-p-nitroanilide (A0180; Sigma) as a chromogenic substrate to test the IEAR amidase. One unit of activity is defined as ΔA405/min=0.001.

To get the active PAP3, proPAP3(0.4 µg/ µl, 1 µl) was incubated with active PAP3 (40 ng/ µl, 1 µl) to the final volume of 12 µl. After incubation at 37℃ for 1 hour, half of the mixture (0.22 µg) was used to test IEAR activity. The method was the same as described above.

2.10. Cleavage of procSPH242, and 35 by PAP3 and electrophoretic mobility changes

M. sexta PAP3 was isolated from pharate pupal hemolymph (Jiang et al., 2003b) and used to activate recombinant proPAP3 (Wang et al., 2014). The *D. melanogastor* procSPH242

and procSP35 were expressed in insect cells, purified from the conditioned media, and used as *M. sexta* PAP3 substrates. The cleavage reactions, controls, and M_r markers were separated by 10% SDS and native PAGE, followed by electrotransfer and immunoblotting using antibody against the hexahistidine tag. To better understand the process of PAP3 cleavage and high M_r complex formation, aliquots of the procSPHs were incubated with different amounts of PAP3 for 1 h at 37 °C. The mixtures and PAP3 control were resolved by 10% SDS and native PAGE, transferred onto nitrocellulose membranes, and detected using *M. sexta* SPH and $(His)_6$ antisera as primary antibody and goat-anti-rabbit/mouse IgG conjugated to alkaline phosphatase (Bio-Rad) as the secondary antibody, and a BCIP-NBT substrate kit (Bio-Rad) for color development.

2.11 Identification and quantification of cleavage sites of procSPH242, and 35 using LC/MS/MS

Around 10 μg the purified *procSPH242, and 35* from insect cells were limited processed using PAP3(1 μg/reaction) for 1h at 37℃, the mixtures or the pro-form of procSPH242, and 35 were separately denatured in urea and digested with chymotrypsin and LysC proteinase (Zhang et al., 2014). Resulting peptides were desalted using C18 affinity media, dried, and redissolved in mobile phase A (0.1% formic acid in H2O). The samples were loaded onto an Acclaim PepMap RSLC C18 column (75 μm × 50 cm, Thermo Fisher) for data-dependent LC-MS/MS analysis as described previously (Cao et al., 2020). Each sample was subjected to the Acclaim column via a gradient of 0–35% mobile phase B (0.1% HCOOH, 80%AcCN, 20% H2O) developed over 120 min as described before (Jin et al., 2022). The survey scans were performed followed by both HCD and CID collisional MS/MS events triggering from parent ions, with scanning of collisional fragment at 15,000 resolutions in the Orbitrap Fusion.

To identify and quantify the limited processed cleavage sites of procSPH242, and 35, a database combined *D. melanogastor* procSPH242, and 35 and background proteins from *M. sexta*,

human, and insect cell were constrcuted, peptide spectrum matches were reviewed in Byonic to check peptides not cut by nonspecific processing enzymes, the details of peak area calculation were same as described previously (Jin et al, 2022).

2.12. ProPO activation and PO activity assay

The active MP2xa was pretreated by incubating proMP2x_a (440 ng/μl, 1 μl) with FactorXa (100 ng/μl, 1 μl) at 37 °C for 1h. The active cSPH35 and 242 were performed by activate procSPH35 (200 ng/μl, 1 μl) and 242(200 ng/μl, 1 μl) using *M. sexta* PAP3 (10 ng/μl, 1 μl) at 25°C for 1h. *M. sexta* serpin3 (100 ng/μl, 1 μl) was added into processed cSPH35 and 242 at 25°C for 30 minutes to inhibit enzyme activity of *M. sexta* PAP3. To activate PPO under different conditions, PPO (1 μl, 270 μg/ml, from *E. coli*) was processed by pretreated MP2x_a (200 ng/μl, 1 μl) with or no active cSPH35 and 242 on ice for 1 hour in a total volume of 25 μl buffer A (0.001% Tween-20, pH 7.5, 20 mM Tris-HCl, 5 mM $CaCl_2$). PO activity was determined on a microplate reader immediately after 150 μl of 2.0 mM dopamine in 50 mM sodium phosphate, pH 6.5, had been added to each sample well (Jiang et al., 2003a).

3. Result

3.1. Evolution of the SPHI and SPHII genes in Drosophilidae insects

In the genome of *D. melanogaster*, 18 genes encoding clip-domain SPHs (66, 101, 79, 142, 35, 242, 166, 121, 58, 125, 128, 231, 69, 156, 94, 93, 64, 9), more than 10 in *M. sexta*, while less than 19 in *Tribolium castaneum*, and 55 in *Anopheles gambiae* (Cao et al., 2015 and 2017; Cao and Jiang, 2018; Miao et al., 2020). The phylogenetical analysis in five insect species (*D. melanogaster*, *M. sexta*, *T. castaneum*, *A. gambiae*, and *Apis mellifera*) indicated *D. melanogaster* cSPH35 and cSPH242 are the unique homologs of *M. sexta* SPHI and SPHII,

separately. To have a comprehensive picture for the evolutionary relationship of cSPHs in *Drosophilidae*, we screened 39 orthologs of DmcSPH35 and cSPH242 from 18 representative species, performed a phylogenetic analysis, and uncovered two main branches in the neighbor-joining tree (Fig. S3), branch I gathered all SPHI genes and branch II gathered all SPHIIs. In the phylogenetic tree, SPHI was present pairing with SPHII in most species, except in *Drosophila suzukii* (Dsu), and *Musca domestica* (Md). There are three cSPHI in *D. suzukii* and two cSPHI in *M. domestica*, suggesting an evolutionary process in similar as SPHIs duplication in Lepidoptera (Jin et al., 2022). Taken together, conservation of SPHI pairing with SPHII in many species in *Drosophilidae* may reveal the functional importance of SPHI-SPHII co-existence and splitting orders of branches I may be related to their putative functional importance as regulators of proPO activation.

3.2. Profiles of D. melanogaster cSPH35 and cSPH242 transcript levels and protein levels

To investigate the abundances of cSPH35 and cSPH242, we compared levels of the cSPH35 and cSPH242 transcripts in multiple tissues at various developmental stages (Fig. S1), similar expression pattern of cSPH35 and cSPH242 at transcript level was observed. cSPH242 was highly expressed except at 4-16 hours at egg stage (*upper*, Fig. S1). While mRNA level of cSPH35 was low during 0-12 hours of the egg stage and high at other stages (*lower*, Fig. S1). To further confirm and extend the expression profiles, we determined the mRNA levels of whole body by qRT-PCR using specific primers of cSPH35 and cSPH242. As shown in Fig 1, the expression patterns of two cSPHs were similar, and compared with larva and pupa stage, mRNA levels relative were low at adult stage. We also found that, after being challenged by bacteria, both cSPH35 and cSPH242 were induced significantly. To know the proteome levels of cSPH35 and cSPH242, we analyzed 17 samples from various developing stages for the whole body of

Drosophila (Casas-Vila et al., 2017, Cao et al., 2018). The protein level of cSPH242 was stable in the whole developing stages, and in terms of cSPH35, protein levels at all developing stages except 4-6 hours at egg stage were abundant and stable. The results above suggested that cSPH35 and cSPH242 were highly expressed both at transcript and protein level.

3.3. Inhibition of hemolymph melanization and culticle melanin deposition in proc*SPH35 and procSPH242 knocking down species*

To investigate the phenotype of single knocking down of cSPH35 or cSPH242, the transgenic flies were constructed as described in 2.1. As shown in Figure 2A, cSPH35 and cSPH242 knocking down species with 60% and 40% efficiency were used to observe the phenotype and extract hemolymph. Interestingly, the cuticle color was less dark in flies with cSPH242 knocking down compared with wild type (Fig. 2D), while there was no difference between cSPH35 RNAi and wild type in term of the cuticle color, the less black of cuticle of cSPH242 RNAi flies indicated the cSPH242 may regulate the formation of cuticle melanin (blackening reaction). On the other hand, hemolymph from wild type, cSPH35, and cSPH242 knocking down species were extracted for DmPPO1 cleavage analysis, PO activity, and blackening detection. After incubation at 25ºC for 10 minutes, nearly 20% of DmPPO1 in the hemolymph of wild type was cleaved into PO (lane1, Fig. 2B), while no active PO1 was recognized from the hemolymph samples of cSPH242 or cSPH35 RNAi knocking down species (lane2-3, Fig. 2B). Furthermore, 27 units of PO activity from 1µl hemolymph of wild type was detected, which is significantly higher than that from the hemolymph samples of cSPH242 or cSPH35 RNAi (Fig. 2C). Consistently, the hemolymph from wild type turned black within 5 minutes at 25ºC, which was much faster than the hemolymph from cSPH242 or cSPH35 RNAi (Fig. 2E), which stayed clear and no change in color within 1 hour. The obvious phenotypes from flies with low expression

cSPH242 or cSPH35 suggested cSPH242 and cSPH35 have functions in regulating hemolymph melanization and culticle melanin deposition of Drosophila.

3.4. Recombinant expression, purification, and characterization of the procSPHs and proMP2x$_a$ from Sf9 cells

To study functions of *cSPHs* in vitro, we amplified the full-length cDNAs of cSPH35 and cSPH242, and subcloned them into pMFFHM6. As described in our previous work (Lu and Jiang, 2008., Jin et al., 2022), upon transposition, the bacmids were isolated for transfecting insect cells and producing high-titer viral stocks through serial infection. Led by the honeybee mellitin signal peptide, the precursor proteins were efficiently secreted into media with the hexahistidine tag fused to their carboxyl-terminus. We adopted the same procedures as proSPHs in *M. sexta* (Lu and Jiang, 2008., Jin et al., 2022) to isolate pro*cSPH35 and procSPH242* from the conditioned media by cationic and nickel affinity chromatography. 0.8 and 1 mg of procSPH35 and procSPH242 were harvested from 100ml of the media, separately. The proteins migrated to 60 kDa (procSPH35) and 63 kDa (procSPH242) on a 10% SDS polyacrylamide gel (Fig.3). The proteins are all homogeneous based on the adjustment of the staining gel (*left panel* of Fig.3) and immunoblot recognized by hexahistidine tag (*right panel* of Fig.3).

We also amplified cDNA of MP2x$_a$ from plasmids and proMP2x$_a$/P6.9FMHaH6 was constructed, and baculovirus for infecting Sf9 cells was made. 100ml media was purified using cationic chromatography, followed by nickel affinity chromatography. Around 0.5 mg pro proMP2x$_a$ was obtained. The recombinant proMP2x$_a$ was migrated as a single band at 50kD on 10% SDS-PAGE (Fig. S4).

3.5. Sequential proteolytic processing of two procSPHs by M. sexta PAP3

In drosophila, only MP2x$_a$ has been identified as the PAP, which can process DmPPO1 to

active form and low PO activity was achieved (An et al., 2013). At least in several insects (mosquito and *M. sexta*), PAP can process both PPO and SPHs at high PO activity can be detected (ref). It's unclear if $MP2x_a$ is an endogenous activating enzyme of cofactor candidate precursors. To investigate this question, we prepared the FactorXa mutated form of $proMP2x_a$, which permits its activation by commercially available bovine $Factorx_a$. SDS-PAGE analysis of purified $proMP2x_a$ indicated that the recombinant protein has a mass of around 50kD (Fig. S4A). Incubation of $Factorx_a$ with $proMP2x_a$ led to the processing of $proMP2x_a$ with a new band appearing around 34kD (Fig.S4A). Activated $proMP2x_a$ resulted in a significant increase in activity (2.7U) as measured by cleavage of the artificial acetyl-Ile-Glu-Ala-Arg-p-nitroanilide (IEARpNA) (Fig.S4B). Pretreated $MP2x_a$ (0.1 µg) were incubated with procSPH35 (0.2 µg) or procSPH242 (0.2 µg) at 37 ºC for 1 hour, separately. The samples were subjected to 10% of SDS-PAGE or native PAGE followed by immunoblotting. Neither procSPH35 nor procSPH242 can be processed by $MP2x_a$ into obviously active forms (Fig. S5A-B), little conformation change was detected in the samples mixing procSPH35 or procSPH242 with $MP2x_a$ on the top the separating gel and stacking gel of native immunoblotting (Fig. S5C), major bands of procSPHs were gathered at the same position as their precursors, suggesting $MP2x_a$ could not activate procSPH35 and procSPH242 effectively. In our previous work of Manduca and *An. gambiae*, MsPAP3 was found to be a strong proteinase that can process SPHs in *M. sexta* (SPH4, 1a, 1b, and 101 and SPH2) and in *An. gambiae* efficiently (Jin et al., 2022), and functional conservation that MsPAP3 was demonstrated at least in these two species. To see the processing of procSPH35 and procSPH242 by MsPAP3, precursors of procSPHs (0.2 µg/ µl, 1 µl) were incubated with MsPAP3 separately. Complete cleavage of procSPH35 yielded two new bands: one band was around 42kDa, while the other band migrated at 30kDa (Fig. 4A, *right*). All procSPH242 was processed by MsPAP3 and

the cleaved form occurred at 42kDa (Fig. 4A, *left*). Native PAGE and immunoblot analysis showed that processing of procSPH35 formed a high Mr complex smear extending from the stacking gel (Fig. 4A, *left*). Similar as the processing of SPHII by MsPAP3 in *M. sexta*, complete digestion of procSPH242 yielded a smear (*orange bar*) that migrated faster than its precursor (*blue bar*).

To examine dynamic processes of the procSPHs cleavage, we incubated aliquots of the procSPH35 and procSPH242 with decreasing amounts of PAP3, separated the reaction mixtures by 10% SDS-PAGE, and examined the products by immunoblot analysis using hexahistidine antibodies (Fig. 5). As the lowest PAP3 level, the 40kDa cSPH35 were detected as major bands and a weak band around 30kDa also occurred (Fig. 5A, left, lane 7 and 8), suggesting that cleavage at the first site occurred first. As the PAP3 concentration increased, the intensity of the 30kDa cSPH35 was increased, and cleavage at 40kDa decreased and then disappeared (Fig. 5A, left, lane 1-3). Precursor of the procSPH35 migrated as one narrow band (Fig. 5C, left, lane 2), cleavage at the first and then second site yielded a smear in the stacking gel and a protein ladder in the separating gel. Ratios of the smear extending to stacking gel increased as more PAP3 were added (Fig. 5C, left). Interestingly, there is a narrow band moving faster than precursor of cSPH242 was weak at the lowest level of PAP3, as PAP3 increased, the narrow band become obvious and cSPH242 smear expanded to the stacking gel (Fig. 5D). Processing of procSPH242 was incomplete and a band at 40kDa was recognized, as PAP3 increased and procSPH242 disappeared (lanes 3−7), the cleaved band intensity increased (Fig. 5B). The results above indicated that MsPAP3 can process procSPH35 and procSPH242 and change their conformation efficiently.

3.6 Identification of the cleavage sites of cSPHs' precursors by M. sexta PAP3 using LC-MS/MS

In *M. sexta*, SPH1 and SPH2 isolated from induced hemolymph of feeding larvae have been first identified to be cleaved after R^{133} and R^{77}(automated Edman degradation) respectively

(Yu et al.,2003), then SPH1 purified from plasma of pharate pupae was also detected being cleaved after R^{82} by peptide mass fingerprint analysis (Wang and Jiang, 2004; Wang et al., 2014). Sequential cleavage from result 3.5 indicated one or multiple PAP specific cleavage sites were present and need to be identified in cSPH35 (SPH1) and cSPH242 (SPH2). To identify the cleavage sites of cSPHs in *D. melanogastor*, various digested enzymes (Chymo, V8, and LysC) were used to process their precursors or active forms of cSPHs separately, then we then used LC-MS/MS to examine the mixtures. A database including cSPH35 (SPH1) and cSPH242 (SPH2) and background proteins was constructed for sequence searching, softwares of PMi-Byonic-Viewer (https://proteinmetrics.com/) and Xcalibur Qual Browser (version 2.2.0.23; TFS, San Jose, CA) were used to identify and quantify targeted peptides. MaxQuant version 1.5.2.28 (Cox and Mann, 2008) and Perseus (https://maxquant.net/perseus/) 1.6.14.0 was used to normalize the detected peptides abundance in total between different treatment. The Second MS spectra were listed in supplement (Fig. S10). The chromatographic peaks were calculated as the relative abundances of detected peptides. In theory, we predicted procSPH242 was cleaved after R^{88}, after being treated by chymotrypsin (cleaved after Y, W, F, M, L), there is no expected peptides detected in procSPH242 (Table 3), while great amount peptide of "TEDGSFDGFGVIDIR" was detected in MsPAP3 cleaved procSPH242 (5.3×10^{8}). We also identified the cleavage sites of procSPH35, two hypothetical cleavage sites (R^{81} and K^{128}) were detected in limited processed samples. The results above suggested that *M. sexta* PAP3 can cleave procSPH242 and procSPH35 as perdiction.

3.7 Effectiveness of the proteolytically processed procSPH242 and procSPH35 as cofactors for proPO activation

MP2x$_a$ was reported as a PAP that can cleave DmPPO1 in Drosophila and yielded low PO activity (An et al., 2013) and our results above (Fig S5) indicated MP2x$_a$ hardly processed procSPH242 and procSPH35 into active forms and cofactors can't be formed. Consequently, although comparing with DmPPO1 mixing with MP2x$_a$, further cleavage of DmPPO1 was obtained by MP2x$_a$ in the presence of procSPH242 and procSPH35 (Fig S7), and a little increase of PO activity was observed (Fig 6B), which is much lower than expected. DmPPO1 and cSPHs can be cleaved by MsPAP3 (Fig 4-6A), while no increase of PO activity (3 units) was detected (Fig 6B). One explanation for the failure of PO activity stimulation by PAP3 and cofactors was that MsPAP3 may be not a good protease to process DmPPO1 into active form. To obtain active cofactor without interfering the cleavage of DmPPO1 by MP2x$_a$, a serpin from *M. sexta* inhibiting MsPAP3 seemed to be necessary to be introduced into DmPPO1 activation system in vitro. Msserpin3 was reported as an inhibitor of MsPAP3 (ref), and it still needs to be proved if MP2x$_a$ can be inhibited by Msserpin3. As it's shown in Fig 6SA, no complex was formed when pre-active MP2x$_a$ was incubated with Msserpin3, the *IEARase* activity of MP2x$_a$ was not decreased in the presence of Msserpin3 compared with control (Fig 6SC), therefore, Msserpin3 was selected as an inhibitor to inhibit PAP3 with no influence on the cleavage of DmPPO1. To optimize PO activity and investigate the function of cSPHs as PAP cofactors, the pretreated cSPHs (incubated with MsPAP3) were mixed with or no MsSerpin3 to inhibit the activity of MsPAP3. The PPOs and the cleaved forms' darkness were measured using Image J 1.53t (https://imagej.nih.gov/ij/download.html), the cleavage efficiency (percentage) was measured using intensity of PO band plus uncleaved PPO band divide PO band intensity, the final number was got using the ratio times 100%. As shown in Fig 6, 4% of DmPPO1 was cleaved by MsPAP3 (lane 7, Fig 6A), with 1.2 units of PO activity (Fig 6B), no further cleavage was

observed if procSPHs were incubated with DmPPO1 together with MsPAP3, yielding low PO activity (1.2 units PO activity without Msserpin3 and 3 units PO activity with Msserpin3), about 40% of DmPPO1 was cleaved by $MP2x_a$ (lane 9, Fig 6A), yielding 13 units (per microgram of DmPPO1) of PO activity (Fig 6B). No further cleavage was performed when DmPPO1 was mixed with $MP2x_a$ and PAP3 (lane 10, Fig 6A), with PO activity about 2 units. However, nearly 60% of DmPPO1 was cleaved into its active form (lane 13, Fig 6A) in the presence of pre-active of cSPHs (with Msserpin3), achiving 128 units of PO activity. The results above indicated cofactor played key role in DmPPO1 activation.

To investigate whether cSPH242 and cSPH35 were necessary for cofactor formation to activate DmPPO1, DmPPO1 was activated by $MP2x_a$ in the presence of pre-active cSPH242, cSPH35 or both by MsPAP3 (with MsPAP3) respectively, after incubation on ice for 1hour, the samples were analyzed using immunoblotting and PO activity kinetics. As its shown in Fig 6A and 6C, when pre-active cSPH242, which are homologues of *M. sexta* SPH2, were incubated with DmPPO1 and pretreated $MP2x_a$ leading to an unobvious enhance process of PPO2 (lane 12, Fig 6A), and the PO activity was 14 unit (Fig 6C), which is significantly lower than 100 units of positive control, suggesting cofactor can't be formed in the absence of cSPH35. While active cSPH35, which are homologues of *M. sexta* SPH1 were mixed with DmPPO1 and pretreated $MP2x_a$, the cleaved band of DmPPO1 was not getting more intense (lane 11, Fig 6A), interestingly, 44 units of PO activity was detected under this condition (Fig 6C), higher than PO activity under the treatment mixing active cSPH242 with DmPPO1 and $MP2x_a$ (14 units), lower than that of positive control (100 units), indicating DmPPO1 can be activated in certain extent in the presence of cSPH35. The results above illustrated that although certain PO activity increase was observed when active cSPH35 was present without pairing with cSPH242 as cofactor, a

significantly higher PO activity and further cleavage in the presence of both cSPH242 and cSPH35 as cofactor, suggesting cSPH242 pairing with cSPH35 as cofactor is indispensable for DmPPO1 activation.

3.8 Comparison of survival rate and melanization after bacterial infection in adults of different transgenic (WT, 35RNAi, and 242RNAi) Drosophila flies

To further investigate and compare the survival rate in cSPHs knocking down and wild type species, 1.2 OD of lethal Gram-positive bacterial infection, *Enterococcus faecalis* (*E. faecalis*), was injured Drosophila as described in *2.2*. The survival analysis was made by calculating the alive flies from 0-8 days post bacteria injury (pbi), and we found that compared with the survival rate of wild-type flies (w1118) septic injury with *E. faecalis* (52%), there is a significantly decrease in flies with cSPH242 (8%) and cSPH35 (12%) knocking down (Fig 7A). We also confirmed there was no statistical difference of the survival rate between flies with cSPH242 and cSPH35 knocking down. The results above suggested that flies with low expression of either cSPH242 or cSPH35 will exhibit high susceptibility compared with wild-type flies. As melanization, regulated by PPO, plays important roles against bacteria in Drosophila (Jan Paul Dudzic et al., 2019), here we also confirmed cSPH242 and cSPH35 was cofactors of DmPAP to activate DmPPO1 in vitro, it's still unclear if cSPH242 and cSPH35 have any regulated function on melanization. To investigate the influence of cofactors on functional melanization response, 0.7 OD of *E. faecalis* was used to prick adult flies, and data blacken of the wound sites was collected after 2 days of pdi. After clean injury, 82% of wild-type flies blacken and 18%was with no black sites (Fig 7C, *left*), whereas low expression of cSPH242 or cSPH35 will led to blacken with the rate of 53% (Fig 7C, *middle*) and 30% (Fig 7C, *right*) respectively, which is much lower than that of wild type. Figure 7B showed the black dots of the

wound sites on thorax of wild type Drosophila and flies with cSPH242 or cSPH35 knocking down. The results above indicated the positive regulation of cSPH242 and cSPH35 on melanization against bacteria in Drosophila adults.

4. Discussion

4.1. Known functions of the clip-domain SPHs in D. melogastor

Like S1A serine proteases, their noncatalytic homologs also play important roles in insect physiological processes, including digestion, development, and defense against pathogen infection (Cao and Jiang, 2018). In M. sexta, the role of serine proteases in digestion, for example, is easy to understand, it is more difficult to rationalize the function of SPHs in gut juice due to their lack of enzymatic activity (Miao et al., 2020). While in Drosophila, cSPHs were rarely studied. Drosophila Masquerade, the orthologous to M. sexta SPH53, a regulator of somatic muscle attachment and axonal guidance (Murugasu-Oei et al., 1996). Both proteins contain five clip domains followed by a protease-like domain. Scarface, the ortholog of *M. sexta* SPH221, one three clip domain containing SPH, was identified as a new transcriptional target of JNK signalling and revealed the first secreted regulator of the JNK pathway acting in a negative-feedback loop during epithelial morphogenesis (Rousset et al., 2010). Here we identified cSPH242 and cSPH35 as two cofactors of PAP to enhance PO activity of DmPPO1. To find proper protease that can activate cSPH242 and cSPH35 endogenously, we tried to use MP2xa to process two cSPHs, however, no cleaved form was observed when pretreated MP2xa was incubated cSPH242 or cSPH35, and the conformation of cSPH242 or cSPH35 was hardly changed to become active forms (Fig. S5). According to the phylogenetic analysis, we screened another PAP candidate, CSP5, to see if it can process DmPPO or cSPHs. Active CSP5 didn’t have IEARase activity. Consequently, we failed to activate DmPPO1 and cSPHs using CSP5

(data not shown). Therefore, introducing an exogenous enzyme into DmPPO activation system, M. sexta PAP3, is critical to identify the function of cSPH242 or cSPH35 as cofactors for PO activivation. While the protease for cSPHs activation still needs to be further explored. In Holotrichia diomphalia, PPAFIII was found to be the upstream protease that can activate cSPHP (PAFII) (S Piao et al., 2005), based on BLAST search result, DmMP1 has 99% sequence identity with HdPPAFIII, suggesting MP1 may have similar function, processing cSPHs in *D. melanogaster*.

4.2. Establishment of cSPH242 pairing with cSPH35 for forming cofactor of PAP in *D. melogastor*

Previous work has indicated only SPHI was essential for PPO acticvation in *H. diomphalia* (S Piao et al., 2005), SPHI and SPHII were essential for PPO activation in Anopheles gambiae and only SPHI was identified as cofactor of PAP in Aedes aegypti. While in lepidoptera, it has been found the formation of cofactor for PAP needs the presence of both SPHI and SPHII (Yu et al., 2003, Wang et al., 2017., Jin et al., 2022). Similarly, here we demonstrated cSPH35 (SPHI) can enhance DmPPO1 activation, however, nearly 2.3 times higher of the PO activities in the presence of both cSPH35 (SPHI) and cSPH242 (SPHII) compared with cSPH35 (SPHI) existence only (Fig. 6C). Our results illustrated the conversation of cofactor composition in Diptera and Lepidoptera to some extent.

4.3. Mechanism for proPO activation in insects

The *D. melanogaster* genome contains three PPO genes, all on the second chromosome (Binggeli O et al., 2014). PPO1 and PPO2 are produced in specialized hemocytes (blood cells), called crystal cells. Both PPO1 and PPO2 but not PPO3 are confirmed as the only source of

hemolymphatic PO contributing to injury-related melanization, while PPO3 was revealed in the encapsulation process, in association with PPO2 (Dudzic et al., 2015). Knocking down of cSPH35 or cSPH242 exhibited failure of injury-related melanization (Fig .2 and Fig .7) in adults of melanization, which is similar with the phenotype from PPO1, and PPO2 knocking out species (Dudzic et al., 2015). As DmPPO1 is involved in the rapid early delivery of PO activity, PPO2 present in the crystals of crystal cells provides a storage form, which can be deployed in a later phase (Binggeli O et al., 2014), both DmPPO1 and PPO2 seem important for bacterial-melanization. Here we demonstrated cSPH35 and cSPH242 are cofactors of PAP for DmPPO1 activation using biochemical way. Because of the lack of knowledge about the protease which can cleave DmPPO2 into active form, we didn't unfold work on the mechanism of PPO2 activation. However, whether cSPH35 and cSPH242 are essential for DmPPO2 activation is still worth further exploration.

The producing of highly active PO in insects seems to be a conserved process more than is currently appreciated. In the silkworm, proPO activating enzyme alone generates active PO by cleaving proPO at Arg51-Phe52 (Satoh et al., 1999), although B. mori SPHI and SPHII both exist (Fig. 1) and may enhance proPO activation. In the beetle H. diomphalia or T. molitor, proPO is activated by HdPPAF1 or TmSPE, a clip-domain SP that cleaved proPO at Arg51-Phe52 but does not yield active PO (Park et al., 2010). In this system, co-presence of HdPPAF2/TmSPH1, an SPHI activated by HdPPAF3/TmSPE, yielded active PO with one subunit cleaved at Arg163-Ala164 in HdPPO (Lee et al., 1998; Lee et al., 2002; Kan et al., 2008). In our previous work of M. sexta, extent of proPO cleavage at Arg51-Phe52 increased after the cofactor (SPH1b-2) was included in the mixture of proPO and PAP1 (Gupta et al., 2005), and PO activity arose disproportionally. Such increases were also observed in H.

diomphalia (Lee et al., 1998; Piao et al., 2005). By integrating the observation of a ternary complex of M. sexta PAP3, cofactor, and proPO (Wang and Jiang, 2004), it was proposed that the high Mr complex of SPH1b and 2 may interact with both PAP and proPO to generate and stabilize highly active PO.

Here we tried to characterize the active form of SPHI, SPHII, and SPHI-SPHII complex to explore the PO activation mechanism in D. melanogastor. The native molecular masses of active cSPH35, cSPH242, or both were determined by electrophoresis in a series of nondenaturing gels at different acrylamide concentrations (4.5%, 7.5%, 10%, 12%, and 15%) and analysed using Ferguson plots of relative mobility (Gallagher, 1995) (Fig. S8). Table1 showed the list for Mr of standard proteins and samples of every lane. As Fig. S9 and Table .2 shown, after calculation, the pro-form of cSPH35 and cSPH242 migrated at 57 kDa and 60 kDa, the cleaved forms of cSPH242 present different aggregated forms including a smaller pattern than pro-form, dimer, trimer, tetramer, and hexamer, the active forms of cSPH35 were aggregated as dimer, trimer, and pentamer. When cSPH35 and cSPH242 were activated together, the form around 55 kDa was still observed, the forms were gathered as dimer, hexamer, octamer, and decamer. It seemed various intermediate complexes were formed and finally a high Mr complex like undecamer or dodecamer. In H. diomphalia, PPAF2 forms a two-hexameric-toroidal structure to bind and activate the 76 kDa PO (cleaved but inactive) (Piao et al., 2005). In the previous model of M. sexta, we thought SPH2 oligomers act as adapters that associate with the SPHI scaffolds, orient PAP and proPO for cleavage under optimal conditions, and perhaps exchange with the leaving PAP to maintain PO in the active conformation. Similarly, in D. melanogaster, the oligomer cSPH242 may act as adapters that anchor the SPHI-SPHII complex,

which provide a platform letting PAP cleave PPO in an ideal condition and extend the active form of PO.

In conclusion, we identified two SPHs function as cofactors for MP2xa to activate DmPPO1. Both cSPH242 and cSPH35 are indispensable for forming cofactor complex. Our work revealed the mechanism of PO activation in the combination of biochemical and genetic approaches. Although *D. melanogaster* is a genetic model, here we demonstrated biochemical methods were more direct to identify the function of certain proteins compared with genetic methods and biochemical methods were effective ways to complement shortage of genetic methods.

Figures

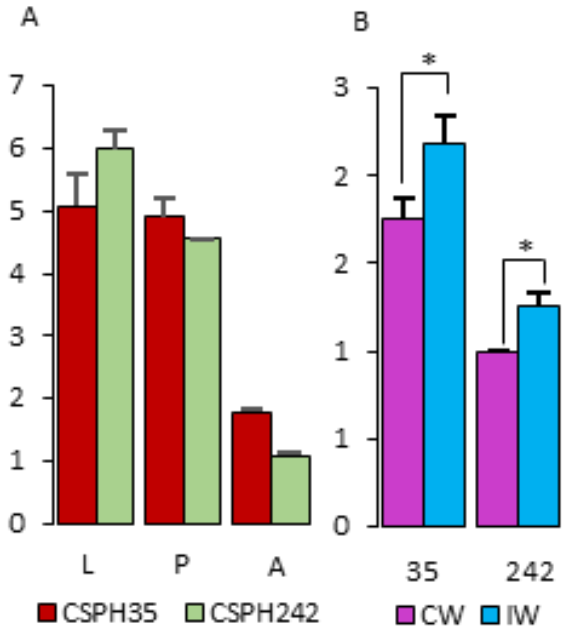


Fig. 1. *D. melanogastor* cSPH35 and cSPH242 mRNA immune inducibility and expression profiles in various developmental stages. The mRNA levels in whole body in three life stages

(A), **c**ontrol and **i**nduced wholebody (CW and IW) **(B)** are calculated based on the C_t values from three biological replicates and plotted as bar graphs (mean ± SEM, n = 3). For panel A, whole body RNA samples from the l**arva** (L) and **pupa** (p) larvae, **adult**-stage (A) prepupae, and **p**upae (P) were prepared for qPCR analysis (*Section*). For panel B, pairwise comparisons were performed between control and induced samples using Student's t-test (*, $p < 0.05$).

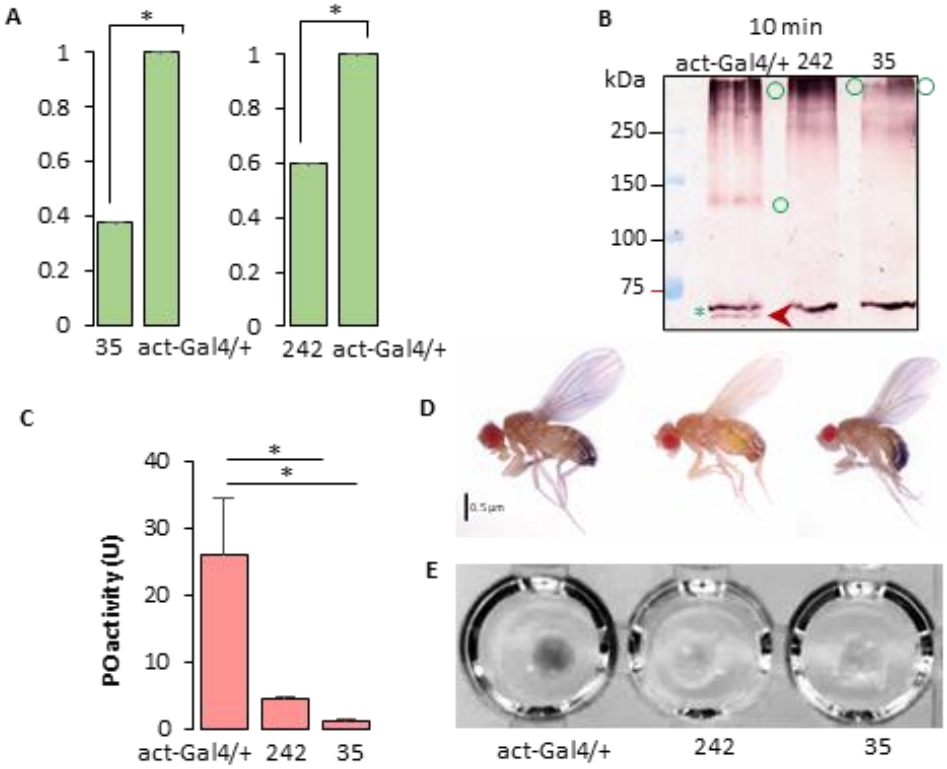


Fig. 2. Knockdown of cSPH35 and cSPH242 mRNA in whole body of *D. melanogastor* in adult stage. (A) Levels of cSPH35 and cSPH242 mRNA, assayed by qPCR, from wholebody of 5d post merged adults in which the act-GAL4 was performed as driver to knock down mRNA level of cSPH35 and cSPH242. Error bars indicate standard deviations. In panel **B**, DmPPO1 cleavage from adults' hemolymph (1 µl from 10 individuals) of control and cSPH35 and cSPH242

knocking down species was indicated. After incubation at room temperature for 5 min (left panel) and 30 min (right panel), hemolymph added in 10 µl buffer A (0.001% Tween-20, pH 7.5, 20 mM Tris-HCl, 5 mM CaCl2) with 0.1% PTU was subjected to 7.5% SDS-PAGE followed by immunoblotting. DmPPO1 and its cleaved form was marked as green astrick and red arrowhead separately. Green circle represents the uncertain bands. For panel **C**, PO activity of *D. melanogastor* hemolymph (1 µl from 10 individuals) in total volume of buffer A was measured immediately after hemolymph extraction using dopamine (2 mM, 150 µl/well) and plotted in the bar graph (mean ± SEM, n = 3). (**D**) The body color comparison of the adults from control and cSPH35 and cSPH242 knocking down species. (**E**) Melanization comparison of the hemolymph (5 µl from 50 individuals for each sample) from act-GAL4 control and cSPH35 and cSPH242 knocking down flies was observed on plate wells separately, after 30 min at room temperature, the picture of the hemolymph color change was captured. 242 and 35 indicated the treatments using cSPH35 and cSPH242 knocking down flies separately.

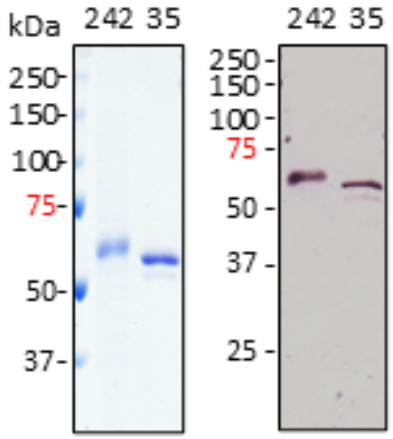


Fig. 3. SDS-PAGE **(A)** and immunoblot analysis **(B)** of the purified *D. melanogastor* procSPH35 and procSPH242. The purified proSPHs were resolved by 10% SDS-PAGE followed by

Coomassie Brilliant Blue (CBB) staining (***A***, 1 μg) or immunoblotting (***B***, 200 ng) using 1:1000 diluted antibody against 6×histidine affinity tag. Positions and sizes (in kDa) of the prestained M_r standards are marked on the *left*, with the 75 kDa marker highlighted *red*.

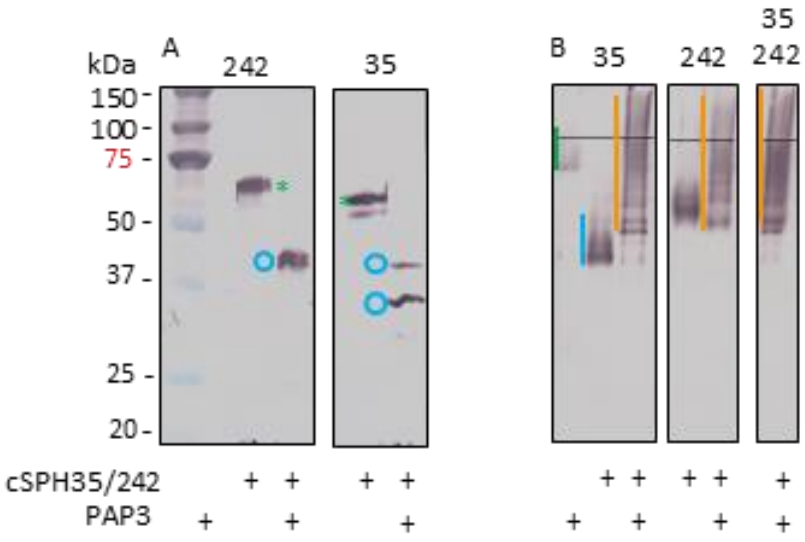


Fig. 4. Immunoblot analysis of the PAP3 proteolytic products of *D. melanogastor* cSPH242 and cSPH35 precursors following 10% SDS **(A)** and native **(B)** polyacrylamide gel electrophoresis (PAGE). The purified proSPHs (200 ng/μl, 1 μl) were separately incubated with PAP3 (40 ng/μl, 1 μl) in 10 μl buffer A (0.001% Tween-20, pH 7.5, 20 mM Tris-HCl, 5 mM $CaCl_2$) at 37 °C for 1 h. The reaction mixtures and controls (40 ng PAP3 and 200 ng proSPH) were treated with 1×SDS sample buffer at 95 °C for 5 min or the same buffer lacking SDS and DTT at 25 °C for 5 min prior to 10% SDS-PAGE and 10% native PAGE. After electrotransfer, immunoblot analysis was performed using 1:1000 diluted anti-specific protein antibody or 1:1000 diluted antibody against the hexahistidine tag (native PAGE). In panel A, procSPH242 (*left* lanes), and major cleavage products (*right* lanes) are marked by *green asterisk*, and *blue circles*, respectively. Positions and sizes of the prestained M_r standards are indicated, with the 75 kDa marker highlighted *red*. In panel B, the dashed line divides the stacking and separating gels. The

smeared bands of PAP3, procSPHs, and cleavage products are marked by *green*, *blue*, and *orange vertical bars*, respectively.

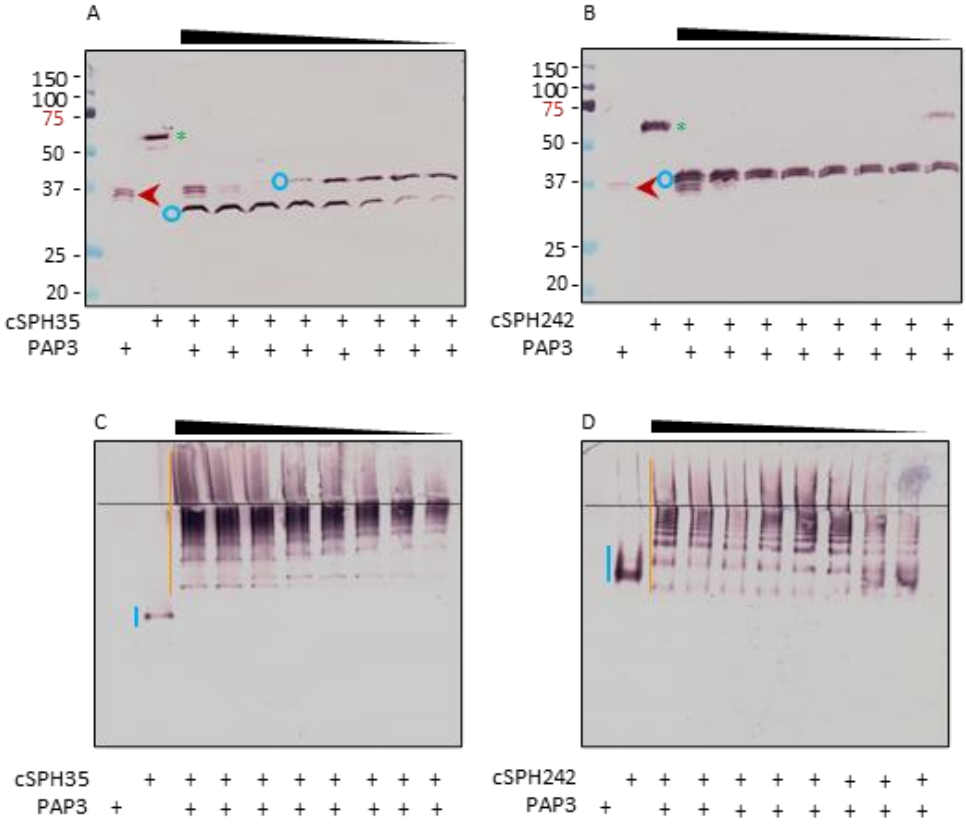


Fig. 5. PAP3 processing of *D. melanogaster* procSPH35, and 242 analyzed by 10% SDS and native PAGE followed by immunoblotting. Aliquots of the purified procSPHs (200 ng) were incubated with 80 (lane 3), 40 (lane 4), 20 (lane 5), 10 (lane 6), 5 (lane 7), 2.5 (lane 8), and 0 (lane 2) ng of PAP3 in 10 μl buffer B at 37 °C for 1 h. The mixtures and PAP3 control (80 ng, lane 1) were subjected to 10% SDS-PAGE (**A** and **B**) and 10% native (**C and D**) PAGE, as described in the legend to Fig. 4. After electrotransfer, immunoblot analysis was performed using antibody against the hexahistidine tag as the first antibody. In panels A and B, PAP3, proSPHs, and major cleavage products are marked by *red arrowheads*, *green asterisks*, and *blue circles*, respectively. Positions and sizes of the prestained M_r standards are indicated, with the 75 kDa

marker highlighted *red*. In panel C and D, the dashed line divides the stacking and separating gels. The smeared bands of procSPHs, and cleavage products are marked by *blue*, and *orange vertical bars*, respectively.

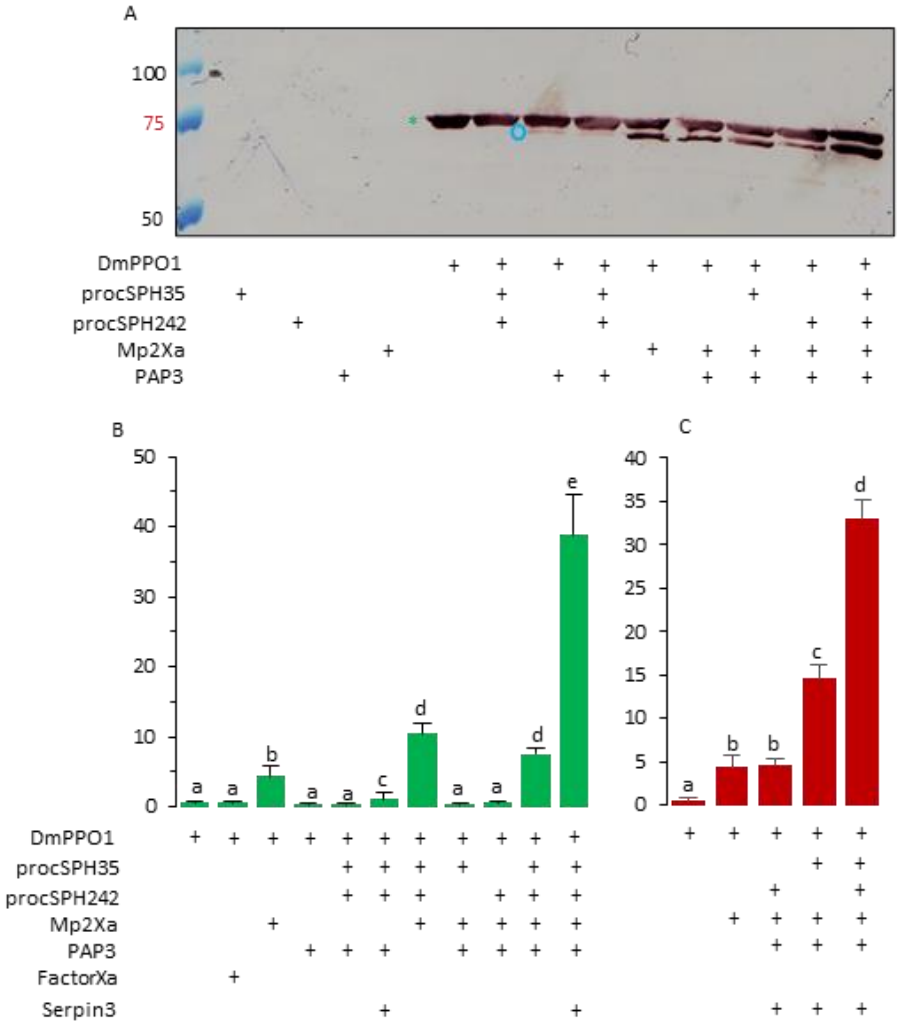


Fig. 6. Effectiveness of the proteolytically processed procSPH242 and procSPH35 as cofactors for DmPPO1 activation (A) proMP2x$_a$ (1 µl, 0.44µg/µl) was activated by FactorXa (1 µl, 0.1µg/µl) at 37 °C for 1 h. And procSPH242 (1µl, 200ng), and cSPH35 (1µl, 200ng) were preprocessed by *M. sexta* PAP3 at 37 °C for 1 h, then PAP3 were inhibited by *M. sexta* serpin3 at room temperature for 30 minutes. Recombinant *D. melanogastor* PPO1 (1 µl, 0.3 µg/ul from

E.coli) was cleaved by MP2x$_a$ (1µl, 200ng/ul) cleaved by FactorXa (1ul, 40ng/ul) with or no active cSPH242 and cSPH35 on ice in a total volume of 20 µl buffer A, after 60 min, half of the mixtures were subjected to 7% SDS-PAGE followed by immunoblotting (**A**). For panel B, the remaining samples were used for PO activity detection by adding 150 µl of 2.0 mM Dopamine (prepared using 50mM sodium phosphate with pH=6.5) in the plate wells, the PO activity was analyzed and plotted as bar graph (**B** and **C**). Statistical analysis was performed using one way ANOVA analysis ($p<0.05$). Three biological replicates were performed and plotted in the bar graphs (mean ± SEM, n = 3).

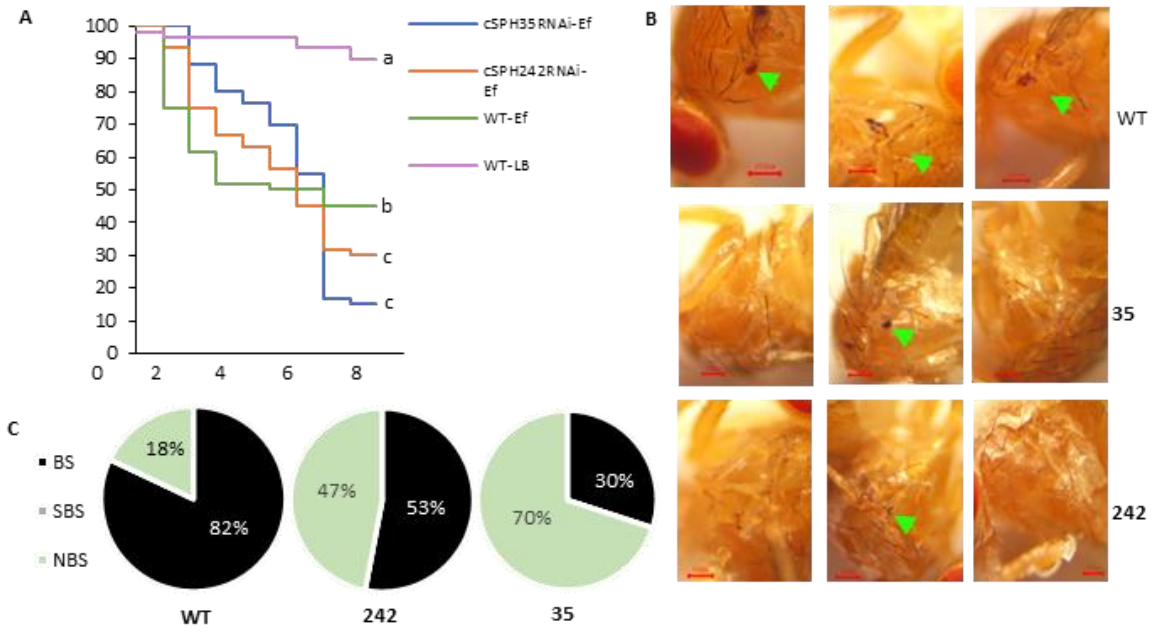


Fig. 7. Comparison of survival rates and melanized dots on thorax cuticle of *D. melanogaster* of wild type (WT), cSPH35, and cSPH242 knocking down flies infected with the indicated *Enterococcus faecalis* (*Ef*) are presented. For panel A, cSPH242 and cSPH35 knocking down and WT flies were challenged at the thorax using 1.2 OD of live *E. faecalis* separately. Infected fliess were incubated at 25 °C. Survival experiments were carried out with 25 flies for each tested genotype. Surviving flies were transferred

daily into fresh vials and counted. Results are expressed as percentage of infected flies at different time points after infection. Each experiment is representative of at least three independent experiments. For panel B and C, 0.8 OD of cultured live *E. faecalis* was used to pierce the thorax of the adults. The black dots formed on the thorax of different genotypes were observed, calculated, and plotted using pie chart (**B** and **C**). The black dots were marked using green arrow heads, horizonal line represented the red scale bar (**B**). BS, SBS, NBS indicated black spot, small black spot, and non-black spot separately (**C**). The experiments were carried out with 25 flies for each tested genotype. Each experiment is representative of at least three independent experiments.

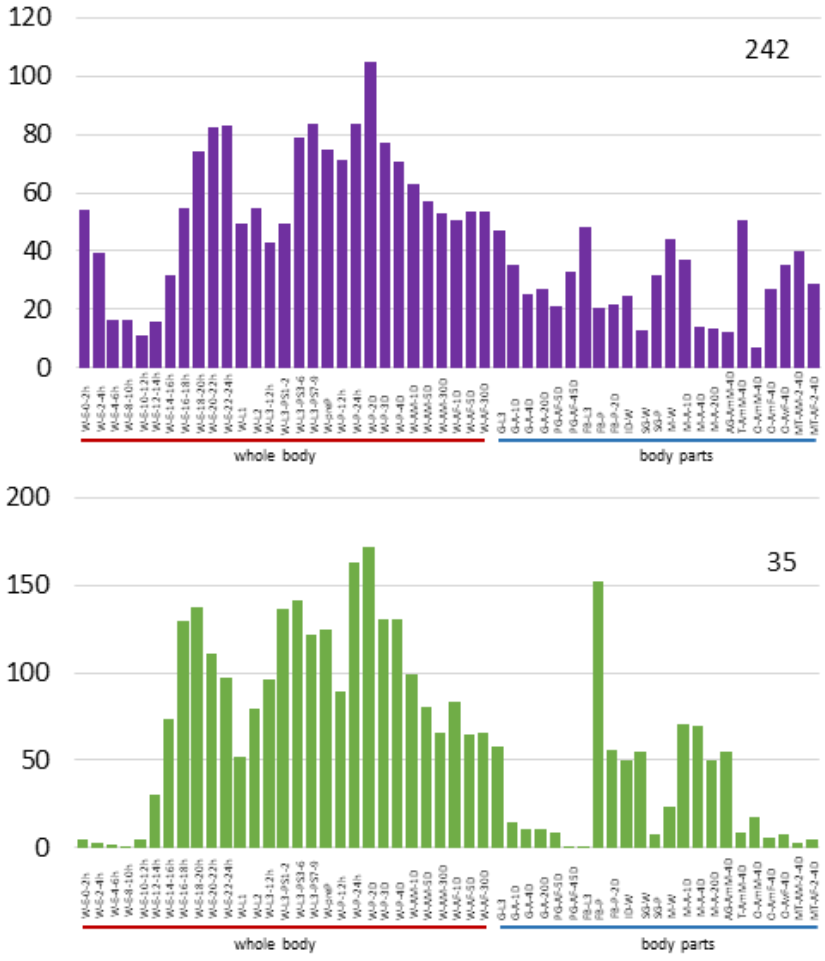


Fig. S1. Transcript profiles of *D. melanogastor* cSPH242, and cSPH35 in different tissue

samples and developing stages. The relative mRNA levels, as represented by FPKM values, are shown in bar graphs. The 54 cDNA libraries were constructed from various tissues and stages.The libraries names are abbreviated using: W, whole insect; E, embryo; L, larval; preP,prepupal; P, pupal; A, adult; M, male; F, female; h, hour; D, day; G, gut; FB, fat body; ID, imaginal discs, SG, salivary glands; C, carcass; AG-AmM, accessary glands of adult mated male; T, testes; O, ovaries of adult mated female; AvF, adult virgin female; MT, Malpighian tubules.

Fig. S2. Abundances of cSPH242, and cSPH35 in *D. melanogaster* at various developmental stages. Relative protein levels in the 16 egg, 16 larval, 20 pupal, 8 female adult, and 8 male adult samples at different time points as represented by $\log_2(\text{LFQ}/5\times10_6 + 1)$ values, are shown in bar

graph. The datasets or library names are abbreviated using E, embryo; L, larval; P, pupal; M, male; F, female; L3c, crawling third instar larva; h, hour; D, day.

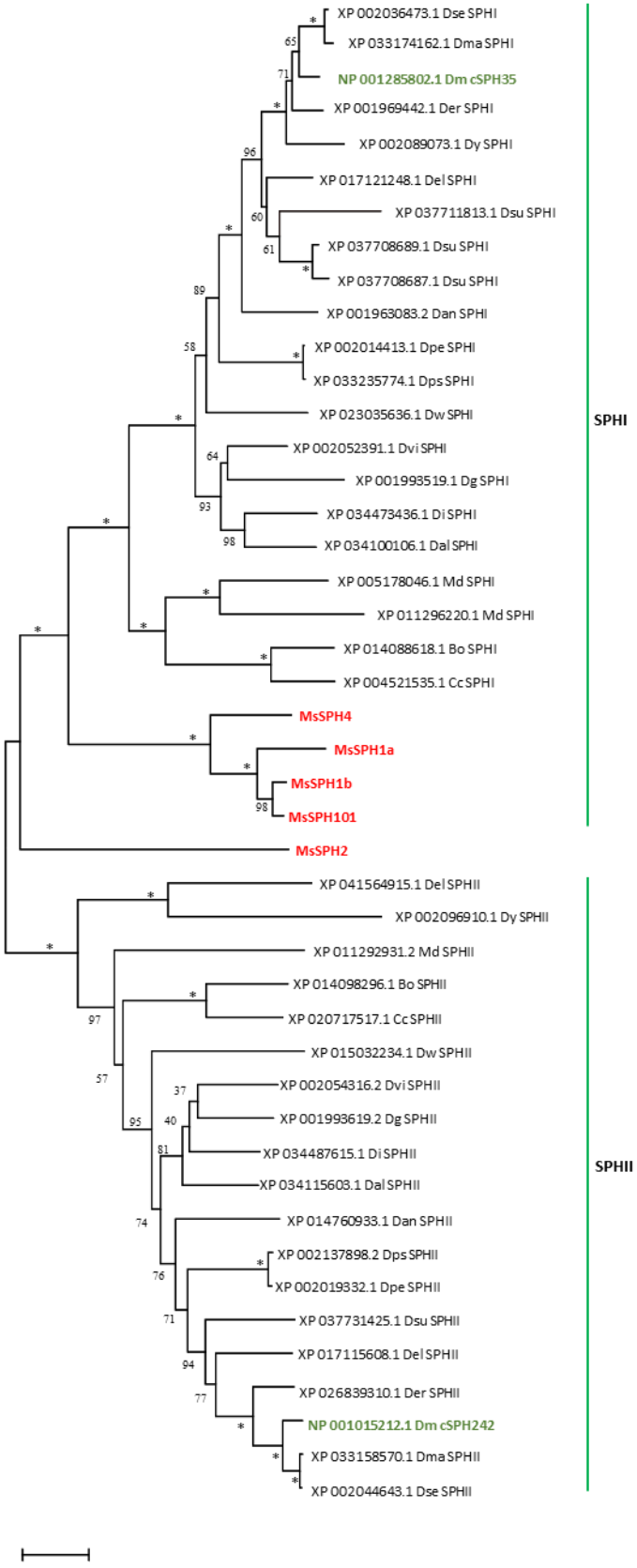

XP 002036473.1 Dse SPHI
XP 033174162.1 Dma SPHI
NP 001285802.1 Dm cSPH35
XP 001969442.1 Der SPHI
XP 002089073.1 Dy SPHI
XP 017121248.1 Del SPHI
XP 037711813.1 Dsu SPHI
XP 037708689.1 Dsu SPHI
XP 037708687.1 Dsu SPHI
XP 001963083.2 Dan SPHI
XP 002014413.1 Dpe SPHI
XP 033235774.1 Dps SPHI
XP 023035636.1 Dw SPHI
XP 002052391.1 Dvi SPHI
XP 001993519.1 Dg SPHI
XP 034473436.1 Di SPHI
XP 034100106.1 Dal SPHI
XP 005178046.1 Md SPHI
XP 011296220.1 Md SPHI
XP 014088618.1 Bo SPHI
XP 004521535.1 Cc SPHI
MsSPH4
MsSPH1a
MsSPH1b
MsSPH101
MsSPH2
XP 041564915.1 Del SPHII
XP 002096910.1 Dy SPHII
XP 011292931.2 Md SPHII
XP 014098296.1 Bo SPHII
XP 020717517.1 Cc SPHII
XP 015032234.1 Dw SPHII
XP 002054316.2 Dvi SPHII
XP 001993619.2 Dg SPHII
XP 034487615.1 Di SPHII
XP 034115603.1 Dal SPHII
XP 014760933.1 Dan SPHII
XP 002137898.2 Dps SPHII
XP 002019332.1 Dpe SPHII
XP 037731425.1 Dsu SPHII
XP 017115608.1 Del SPHII
XP 026839310.1 Der SPHII
NP 001015212.1 Dm cSPH242
XP 033158570.1 Dma SPHII
XP 002044643.1 Dse SPHII
SPHI
SPHII

Fig. S3. Phylogenetic relationships of the SPHI and SPHII subfamily members in *Drosophilidae* insects. Species names and their abbreviations in parenthesis are *Drosophila sechellia* (Dse), *Drosophila mauritiana* (Dma), *Drosophila melanogastor* (Dm), *Drosophila erecta* (Der), *Drosophila elegans* (Del), *Drosophila suzukii* (Dsu), *Drosophila pseudoobscura* (Dps), *Drosophila persimilis* (Dpe), *Drosophila virilis* (Dvi), *Drosophila innubila* (Di), *Drosophila grimshawi* (Dg), *Drosophila albomicans* (Dal), *Drosophila willistoni* (Dw), *Ceratitis capitata* (Cc), *Musca domestica* (Md), *Drosophila yakuba* (Dy). Suggested protein names consist of an abbreviated species name, a group name (SPHI, or SPHII). The *M. sexta* SPHIs *(i.e.*, 1a, 1b, 4, and 101) and SPHII (*i.e.*, 2) are in *red bold* font (Jin et al., 2022). *D. melanogastor* cSPH242 and cSPH35 were in *blue bold* font.

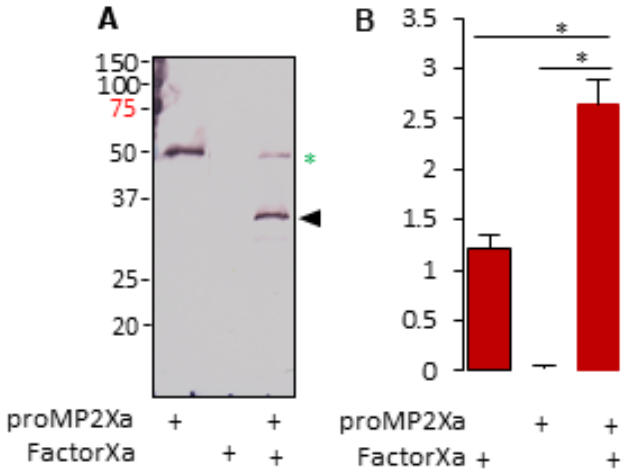


Fig. S4. The activation of MP2x$_a$ and the IEARase activity of MP2x$_a$. The processing of MP2x$_a$ using Factorx$_a$. 0.44 μ g of pro MP2x$_a$ was incubated with 0.1 μ g of FactorXa in the total volume is 10ul, after incubation in 37℃ for 2 hours, the mixture was subjected to the SDS-PAGE followed by immunoblotting (**A**). In panel B, IEARase activity of MP2x$_a$ were detected, 150 μl IEARpNa (25μM) was mixed with samples and the IEARase activity was detected under

405 nm absorption. Three biological replicates were performed and plotted in the bar graphs (mean ± SE, n = 3).

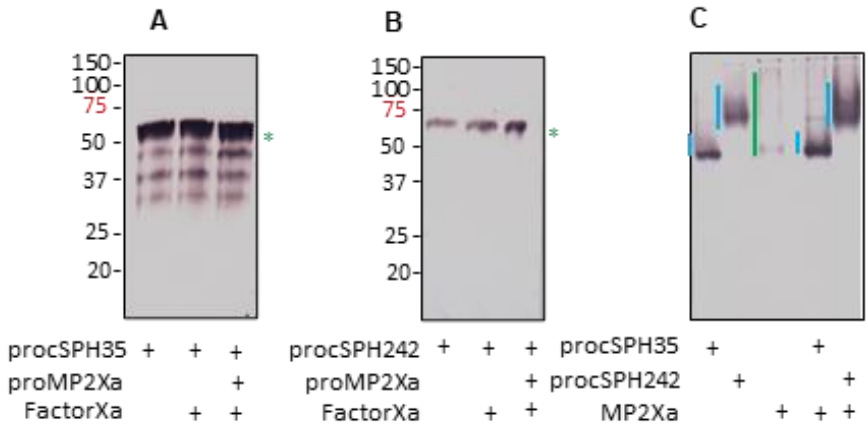


Fig. S5. Processing the precursors of cSPH242 and cSPH35 by pretreated MP2x$_a$. The purified procSPHs (200 ng/μl, 1 μl) were separately incubated with active MP2x$_a$ (40 ng/μl, 1 μl) in 10 μl buffer A at 37 °C for 1 h. The reaction mixtures and controls were treated with 1×SDS sample buffer at 95 °C for 5 min or the same buffer lacking SDS and DTT at 25 °C for 5 min prior to 10% SDS-PAGE and 10% native PAGE. After electrotransfer, immunoblot analysis was performed using 1:1000 diluted anti-specific protein antibody or 1:1000 diluted antibody against the hexahistidine tag (native PAGE). In panel A and B, procSPH35 (*left* lanes) or procSPH242 were marked by *green asterisk*. The positions and sizes of the prestained M_r standards are indicated, with the 75 kDa marker highlighted *red*. In panel B, the dashed line divides the stacking and separating gels. The smeared bands of PAP3, procSPH242 and 35 were marked by *green*, and *blue*, *vertical bars*, respectively.

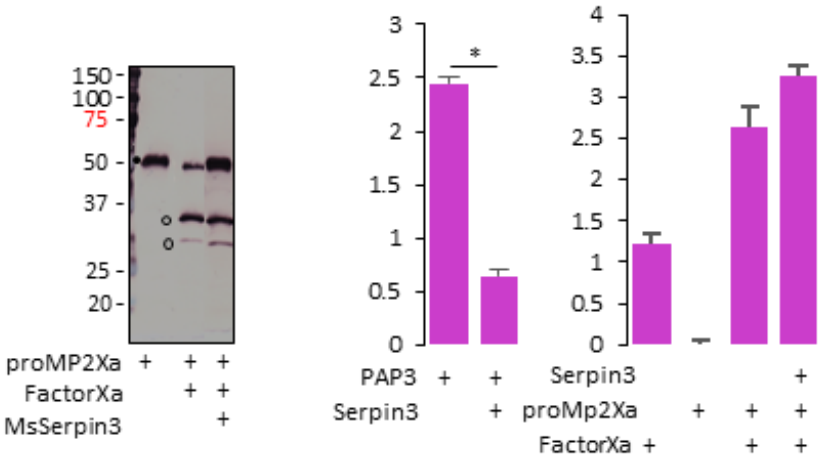


Fig. S6. The inhibition of *M. sexta* PAP3 and *D. melanogastor* MP2X$_a$ using *M. sexta* Serpin3. For panel A, the active MP2xa (1 µl, 0.44µg/µl) was activated as described in figure 6, and *M. sexta* serpin3 (1 µl, 0.3 µg/µl) was added into the mixture of 10 µl in total and incubated at room temperature for 30 minutes. The reaction mixtures and controls were treated with 1×SDS sample buffer at 95 °C for 5 min or the same buffer lacking SDS and DTT at 25 °C for 5 min prior to 10% SDS-PAGE and followed with immunoblotting. For panel B, active 100ng of PAP3 or MP2X$_a$ was incubated with *M. sexta* Serpin 3 (1 µl, 0.3 µg/µl) at room temperature for 30 minutes and IEARase activity was analyzed using 25 µM β IEAR as substrate and plotted as bar graph, pairwise comparisons were performed between control and induced samples using Student's t-test (*, $p < 0.05$). Three biological replicates were performed and plotted in the bar graphs (mean ± SE, n = 3).

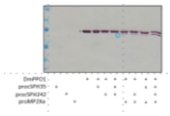


Fig. S7. *D. melanogastor* proPO1 activation using MP2X$_a$, procSPH242, and procSPH35. (A) proMP2xa (1 µl, 0.44µg/µl) was activated by FactorXa (1 µl, 0.1µg/µl) at 37 °C for 1 h. Recombinant *D. melanogastor* PPO1 (1 µl, 0.3 µg/ul from E.coli) was cleaved by MP2x$_a$ (1µl,

200ng/ul) cleaved by FactorXa (1ul, 40ng/ul) with or no procSPH242 and procSPH35 on ice in a total volume of 10 µl buffer A, after 60 min, half of the mixtures were subjected to 7% SDS-PAGE followed by immunoblotting.

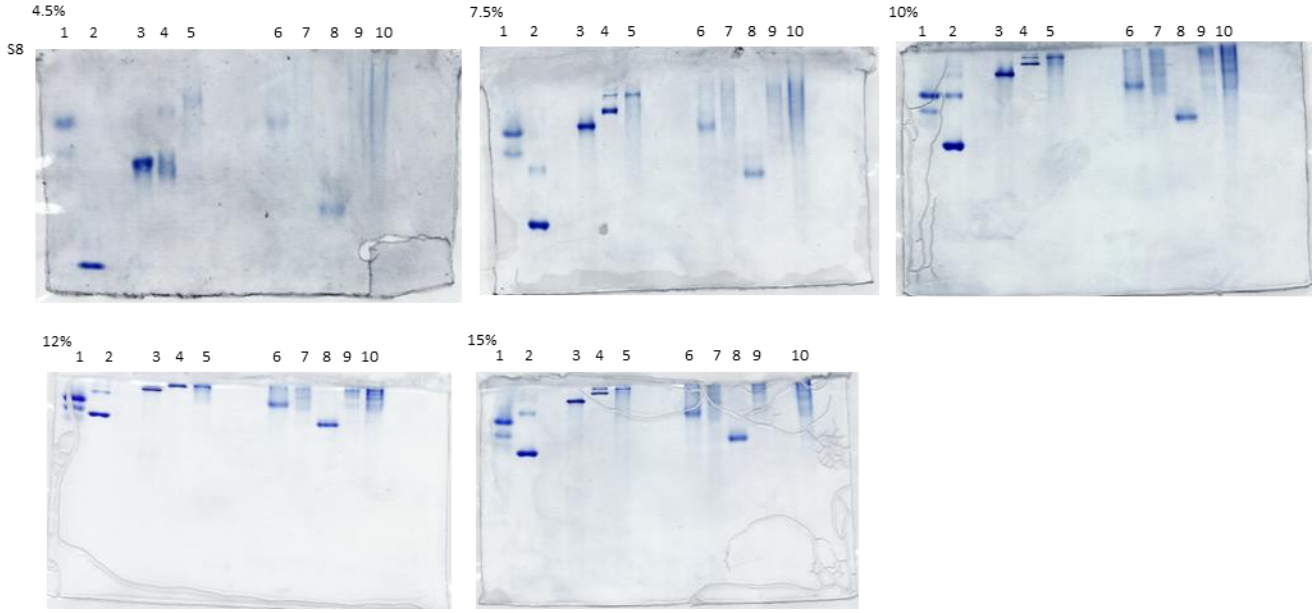


Fig. S8. Analysis of the precursors or cleaved forms of cSPH242 and cSPH35 by *M. sexta* PAP3 by PAGE under non-denaturing conditions.The purified procSPHs (200 ng/µl, 1 µl) were incubated with PAP3 (40 ng/µl, 1 µl) in a total volume of 50 µl bufferA at 37°C for 1 hour. The samples were separated on the native gels with the concentration of 4.5%, 7%, 10%, 12%, and 15%, coomassie blue staining after non-denaturing gel electrophoresis.

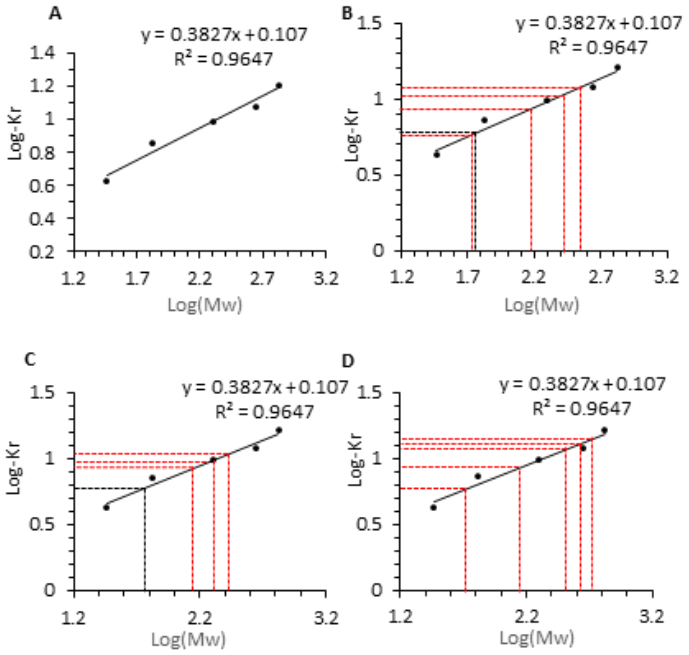


Fig. S9. Ferguson plot (Jiang et al., 1997) were used to estimate the apparent masses of the complexes of active cSPH242 and cSPH35. Molecular mass standards are carbonic anhydrase (29 kDa), BSA monomer and dimer (66 and 132 kDa), β-Amylase (200 kDa), apoferintin (443 kDa), thyroglobulin (669 kDa). procSPH242 (B) and procSPH35 (D) and their complexes after being cleaved by PAP3 were indicated as the *black (pro-form)* and *red (cleaved form) dashed lines*.

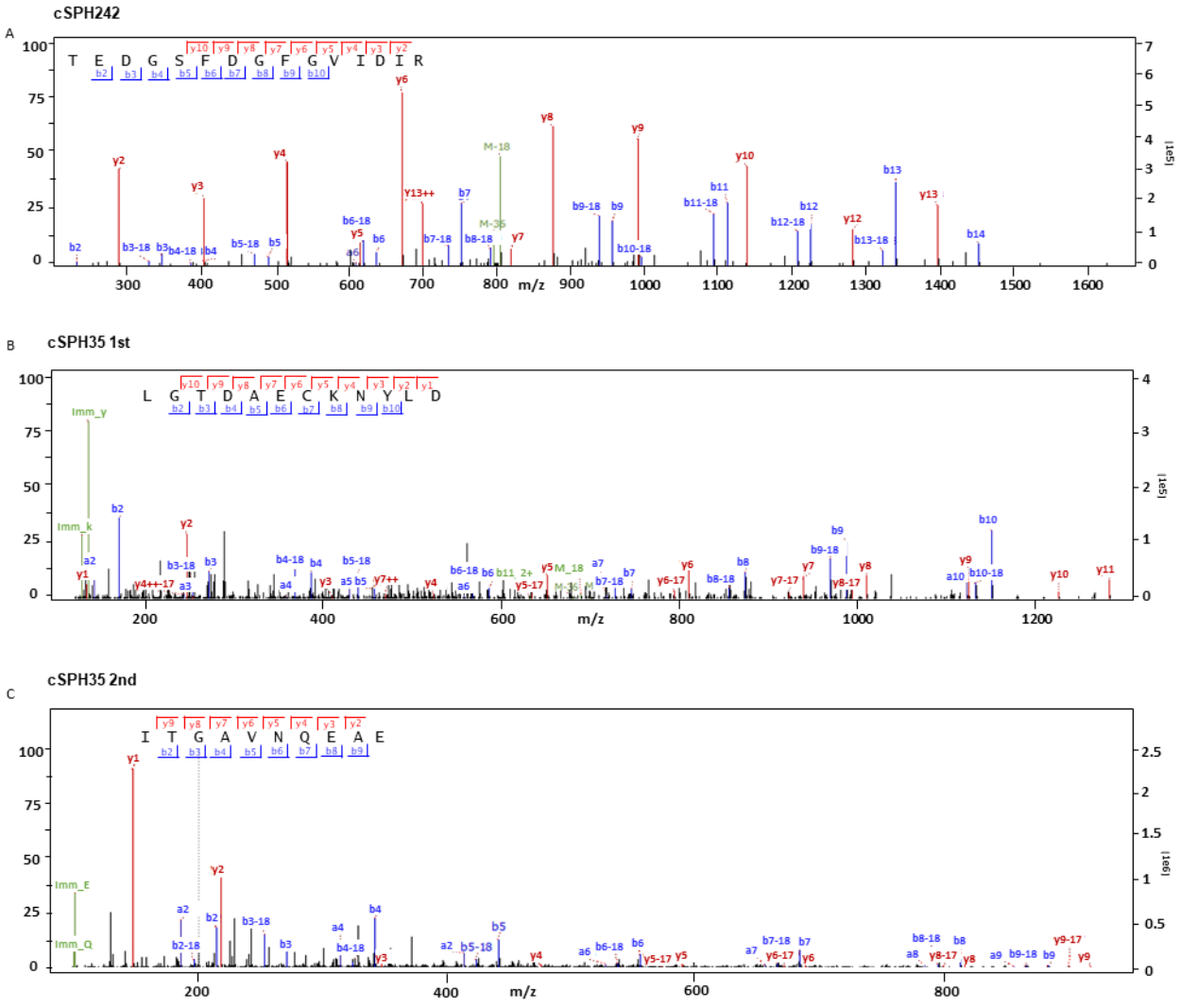


Fig. S10. Second MS spectra of Chymotryspin/V8 protease/ LysC processed peptides from cleavage sites of the purified procSPH242 and procSPH35 or active cSPH242 and cSPH35 by cut by *M. sexta* PAP3. (A) TEDGSFDGFGVIDIR, (B) LGTDAECKNYLD, (C) LGTDAECKNYLD.

Tables

Table1

Lane	Sample name	MW(kDa)
1	Carbon anhydrase	29
2	BSA	66
3	B-Amylase	200
4	Apoferintin	443
5	Thyroglobulin	669
6	procSPH242	
7	procSPH242+PAP3	
8	procSPH35	
9	procSPH35+PAP3	
10	proCSPH35+procSPH242+PAP3	

Table1. A list of samples loaded on the native-PAGE, the lane number, sample name, and molecular weight of standards were listed.

Table2.1

	procSPH242	cSPH242+PAP3	cSPH242+PAP3	cSPH242+PAP3	cSPH242+PAP3
Slope	6.1324	5.9472	8.7459	10.748	12.399
X (lgMW)	1.778496	1.743696	2.181355	2.41528	2.577441
log(-Kr)	0.78763	0.774313	0.941805	1.031328	1.093387
MW(10^x)	60.04767	55.42379	151.829	260.1837	377.9558

Table2.2

	procSPH35	cSPH35+PAP3	cSPH35+PAP3	cSPH35+PAP3
Slope	6.0194	8.0959	9.7258	10.744
X (lgMW)	1.75739	2.093716	2.301869	2.414858
log(-Kr)	0.779553	0.908265	0.987925	1.031166
MW(10^y)	57.19922	124.0841	200.3868	259.9307

Table2.3

	cSPH35+cSPH242+PAP3_1	cSPH35+cSPH242+PAP3_2	cSPH35+cSPH242+PAP3_3	cSPH35+cSPH242+PAP3_4	cSPH35+cSPH242+PAP3_5
Slope	6.0108	8.516	12.21	13.111	14.156
X (lgMW)	1.755768	2.151125	2.56001	2.640804	2.72783
log(-Kr)	0.778932	0.930236	1.086716	1.117636	1.150941
MW(10^y)	56.98593	141.6202	363.0861	437.325	534.3552

Table 2. Parameters of trend lines and predicted molecular weights of complex for precursors and active cSPH242 and cSPH35 were listed. The complexes of PAP3 processed cSPH242 (table 2.1), cSPH35 (table 2.2), or both (table 2.3) were calculated.

Table3

Protein Name	1st cutting site	Intensity of 1st cutting site (Peak area)	2nd cutting site	Intensity of 2nd cutting site (Peak area)
DmCSPH242	TEDGSFDGFGVIDIR*FND	nd	NONE	nd
DmCSPH242+PAP3	TEDGSFDGFGVIDIR*FND	5.3×10^{8}	NONE	nd
DmCSPH35	IDIR*LGTDAECKNYLD	nd	VGFK*ITGAVNQEAE	Not detected
DmCSPH35+PAP3	IDIR*LGTDAECKNYLD	4.61×10^{9}	VGFK*ITGAVNQEAE	2.17×10^{10}

Table3. Determination of the *D. melanogastor* cSPH242 and cSPH35 cleavage sites processed by *M. sexta* PAP3 using LC-MS/MS analysis. A list of peptides identified as processing products of PAP3. Chymotrypsin, V8 protease, or LysC were used to process the PAP3 cleaved cSPH242 and cSPH35, or their pro-forms. The table showed the parent ions' m/z, scan time, and peak areas of the peptides. Peptides labeled as red were the precursors we analyzed and measured before or after the cleaved peptide bonds, which is marked as *. NONE represents no theoretical cleavage sites were predicted; nd means not detected.

Reference

An C, Zhang M, Chu Y, Zhao Z., 2013. Serine protease MP2 activates prophenoloxidase in the melanization immune response of Drosophila melanogaster. PLoS One. 15;8(11):e79533.

Binggeli O, Neyen C, Poidevin M, Lemaitre B, 2014. Prophenoloxidase activation is required for survival to microbial infections in Drosophila. PLoS Pathog. 10(5):e1004067. doi:

Casas-Vila, N., Bluhm, A., Sayols, S., Dinges, N., Dejung, M., Altenhein, T., Kappei, D., Altenhein, B., Roignant, J.Y., Butter, F., 2017. The developmental proteome of *Drosophila melanogaster*. Genome Res. 27, 1273–1285.

Cao, X., Jiang, H., 2015. Integrated modeling of protein-coding genes in the *Manduca sexta*

genome using RNA-Seq data from the biochemical model insect. Insect Biochem. Mol. Biol., 62, 2−10.

Cao, X., Gulati, M., Jiang, H., 2017. Serine protease-related proteins in the malaria mosquito, Anopheles gambiae. Insect Biochem. Mol. Biol., 88, 48–62.

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.

Cao, X., Wang, Y., Rogers, J., Hartson, S., Kanost, M.K., Jiang, H., 2020. Changes in composition and levels of hemolymph proteins during metamorphosis of Manduca sexta. Insect Biochem. Mol. Biol., 127, 103489.

Chapman JR, Dowell MA, Chan R, Unckless RL,2020. The Genetic Basis of Natural Variation in Drosophila melanogaster Immune Defense against Enterococcus faecalis. Genes (Basel). 11(2):234

Cerenius L, Lee BL, Söderhäll K., 2008. The proPO-system: pros and cons for its role in invertebrate immunity. Trends Immunol. 29(6):263-71.

Dudzic JP, Kondo S, Ueda R, Bergman CM, Lemaitre B, 2015. Drosophila innate immunity: regional and functional specialization of prophenoloxidases. BMC Biol. 13:81.

Dudzic JP, Hanson MA, Iatsenko I, Kondo S, Lemaitre B, 2019. More Than Black or White: Melanization and Toll Share Regulatory Serine Proteases in Drosophila. Cell Rep.27(4):1050-1061.

Dudzic JP, Hanson MA, Iatsenko I, Kondo S, Lemaitre B., 2019. More Than Black or White: Melanization and Toll Share Regulatory Serine Proteases in Drosophila. Cell Rep.

27(4):1050-1061.e3.

Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res., 32, 1792–1797.

Gallagher, S. R., eds J. E. Coligan, B. M. Dunn, H. L. Ploegh, D. W. Speicher and P. T. Wingfield, 1995. One-dimensional electrophoresis using nondenaturing conditions. In Current Protocols in Protein Science. 10.3.5–10.3.11.

González-Tokman DM, Munguía-Steyer R, González-Santoyo I, Baena-Díaz FS, Córdoba-Aguilar A., 2012. Support for the immunocompetence handicap hypothesis in the wild: hormonal manipulation decreases survival in sick damselflies. Evolution. 66(10):3294-301.

Gupta, S., Wang, Y., Jiang, H., 2005. Manduca sexta prophenoloxidase (proPO) activation requires proPO-activating proteinase (PAP) and serine proteinase homologs (SPHs) simultaneously. Insect Biochem. Mol. Biol., 35, 241−248.

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

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

Ji Y, Lu T, Zou Z, Wang Y, 2022. Aedes aegypti CLIPB9 activates prophenoloxidase-3 in the

presence of CLIPA14 after fungal infection. Front Immunol. 13:927322.

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

Kwon, T.H., Kim, M.S., Choi, H.W., Joo, C.H., Cho, M.Y., Lee, B.L., 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.

Kumar, S., Stecher, G., Li, M., Knyaz, C., Tamura, K., 2018. MEGA X: molecular evolutionary genetics analysis across computing platforms. Mol. Biol. Evol., 35, 1547−1549.

Lee, K.Y., Zhang, R., Kim, M.S., Park, J.W., Park, H.Y., Kawabata, S., Lee, B.L., 2002. A zymogen form of masquerade-like serine proteinase homologue is cleaved during pro-phenoloxidase activation by Ca2+ in coleopteran and Tenebrio molitor larvae. Eur. J. Biochem. 269, 4375–4383.

Lee, S.Y., Kwon, T.H., Hyun, J.H., Choi, J.S., Kawabata, S.I., Iwanaga, S., Lee, B.L., 1998. In vitro activation of pro-phenol-oxidase by two kinds of pro-phenol-oxidase-activating factors isolated from hemolymph of coleopteran, Holotrichia diomphalia larvae. Eur. J. Biochem., 254, 50−57.

Li H, Tang H, Sivakumar S, Philip J, Harrison RL, Gatehouse JA, Bonning BC., 2008 Insecticidal activity of a basement membrane-degrading protease against Heliothis virescens (Fabricius) and Acyrthosiphon pisum (Harris). J Insect Physiol. 54(5):777-89.

Lu, Z., Jiang, H., 2008. Expression of Manduca sexta serine proteinase homolog precursors in insect cells and their proteolytic activation. Insect Biochem. Mol. Biol., 38, 89–98.

Murugasu-Oei B, Rodrigues V, Yang X, Chia W, 1995. Masquerade: a novel secreted serine protease-like molecule is required for somatic muscle attachment in the Drosophila embryo.

Genes Dev. 9(2):139-54.

Murugasu-Oei B, Balakrishnan R, Yang X, Chia W, Rodrigues V, 1996. Mutations in masquerade, a novel serine-protease-like molecule, affect axonal guidance and taste behavior in Drosophila. Mech Dev. 57(1):91-101.

Miao, Z., Cao, X., Jiang, H., 2020. Digestion-related proteins in the tobacco hornworm, Manduca sexta. Insect Biochem. Mol. Biol., 126, 103457.

Nappi, A.J., Christensen, B.M., 2005. Melanogenesis and associated cytotoxic reactions: applications to insect innate immunity. Insect Biochem. Mol. Biol., 35, 443–459.

Park, J.W., Kim, C.H., Rui, J., Park, K.H., Ryu, K.H., Chai, J.H., Hwang, H.O., Kurokawa, K., Ha, N.C., Söderhäll, I., Söderhäll, K., Lee, B.L., 2010. Beetle immunity. In: Söderhäll, K. (Ed.), Invertebrate Immunity. Adv. Exp. Med. Biol. 708, 163–180.

Piao S, Song YL, Kim JH, Park SY, Park JW, Lee BL, Oh BH, Ha NC, 2005. Crystal structure of a clip-domain serine protease and functional roles of the clip domains. EMBO J. 24(24):4404-14.

Qiao Jin, Yang Wang, Steven D. Hartson, Haobo Jiang, 2022. Cleavage activation and functional comparison of Manduca sexta serine protease homologs SPH1a, SPH1b, SPH4, and SPH101 in conjunction with SPH2. Insect Biochem Mol. Biol. 144,103-762.

Rousset R, Bono-Lauriol S, Gettings M, Suzanne M, Spéder P, Noselli S, 2010. The Drosophila serine protease homologue Scarface regulates JNK signalling in a negative-feedback loop during epithelial morphogenesis. Development. 137(13):2177-86.

Satoh, D., Horii, A., Ochiai, M., Ashida, M., 1999. Prophenoloxidase-activating enzyme of the silkworm, Bombyx mori: purification, characterization, and cDNA cloning. J. Biol. Chem., 274, 7441−7453.

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(6):487-95.

Troha K, Im JH, Revah J, Lazzaro BP, Buchon N, 2018. Comparative transcriptomics reveals CrebA as a novel regulator of infection tolerance in *D. melanogaster*. PLoS Pathog.14(2):e1006847.

Wang, Y., Jiang, H., 2004. Prophenoloxidase (proPO) activation in Manduca sexta: an analysis of molecular interactions among proPO, proPO-activating proteiase-3, and a cofactor. Insect Biochem. Mol. Biol., 34, 731−742.

Wang, Y., Lu, Z., Jiang, H., 2014. *Manduca sexta* proprophenoloxidase activating proteinase-3 (PAP3) stimulates melanization by activating proPAP3, proSPHs, and proPOs. Insect Biochem. Mol. Biol. 50, 82–91.

Wang Y, Jiang H, 2017. Prophenoloxidase activation and antimicrobial peptide expression induced by the recombinant microbe binding protein of *Manduca sexta*. Insect Biochem Mol Biol. 83:35-43.

Wang Y, Yang F, Cao X, Huang R, Paskewitz S, Hartson SD, Kanost MR, Jiang H, 2020. 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. Insect Biochem Mol Biol. 116:103261.

Yu, X.Q., Jiang, H., Wang, Y., Kanost, M.R., 2003. Nonproteolytic serine protease homologs are involved in prophenoloxidase activation in the tobacco hornworm, *Manduca sexta.* Insect Biochem. Mol. Biol., 33, 197−208.

Yamamoto-Hino M, Goto S, 2016. Spätzle-Processing Enzyme-independent Activation of the

Toll Pathway in Drosophila Innate Immunity. Cell Struct Funct. 41(1):55-60.

Zhang, S., Cao, X., He, H., Hartson, S., Jiang, H., 2014. Semi-quantitative analysis of changes in the plasma peptidome of Manduca sexta larvae and their correlation with the transcriptome variations upon immune challenge. Insect Biochem. Mol. Biol., 47, 46−54.

Zhao, P., Li, J., Wang, Y., Jiang, H., 2007. Broad-spectrum antimicrobial activity of the reactive compounds generated *in vitro* by *Manduca sexta* phenoloxidase. Insect Biochem. Mol. Biol., 37, 952−959.

Zhao, P., Lu, Z., Strand, M.R., Jiang, H., 2011. Antiviral, anti-parasitic, and cytotoxic effects of 5, 6-dihydroxyindole (DHI), a reactive compound generated by phenoloxidase during insect immune response. Insect Biochem. Mol. Biol., 41, 645–652.

VITA

QIAO JIN

Candidate for the Degree of
Doctor of Philosophy

Thesis: CHARACTERIZATION OF CO-FACTOR PROTEINS FOR PROPHENOLOXIDASE ACTIVATION IN INSECTS

Major Field: ENTOMOLOGY

Biographical:

Education:

Completed the requirements for the Doctor of Philosophy in Entomology at Oklahoma State University, Stillwater, Oklahoma in Dec, 2022.

Completed the requirements for the Master of Science in Biochemistry and molecular biology at Chinese Academy of Sciences in 2017.

Completed the requirements for the Bachelor of Science in Ecology at Agricultural University of Hebei Province, China in 2013.

Experience:
Doctoral Research at Oklahoma State University from 2017 to 2022.
Master at Chinese Academy of Sciences from 2013 to 2017.
Undergraduate at Agricultural University of Hebei Province from 2009-2013

Professional Memberships:
ESA (Entomological Society of America)