CmHem, a *hemolin*-like gene identified from *Cnaphalocrocis medinalis*, involved in metamorphosis and baculovirus infection

Guangjie Han [Corresp., 1], **Chuanming Li** [1], **Nang Zhang** [1], **Qin Liu** [1], **Lixin Huang** [1], **Yang Xia** [1], **Jian Xu** [1]

[1] Lixiahe District Institute of Agricultural Sciences in Jiangsu, Yangzhou, China

Corresponding Author: Guangjie Han
Email address: hanguangjie177@163.com

Background As a member of the immunoglobulin (Ig) superfamily, Hemolins play a vital role in insect development and defense against pathogens. However, Hemolin has different effects on insect development and innate immune response to baculovirus infection. **Methods and results** In this study, the *hemolin*-like gene from a Crambidae insect, *Cnaphalocrocis medinalis*, *CmHem* was cloned, and i ts role in insect development and baculovirus infeciton was analyzed. A 1528 bp contig as potential *hemolin*-like gene of *C. medinalis* was reassembled from transcriptome. Further, the complete *hemolin* sequence of *C. medinalis* (*CmHem*) was cloned and sequenced. The cDNA of *CmHem* was 1515 bp in length and encoded 408 amino acids. The deduced amino acid of CmHem have relatively low identities (41.9-62.3%) to various insect Hemolins. However, it contains four Ig domains similarity to other insect Hemolins. The expression level of *CmHem* was the highest in eggs, followed by pupae and adults, and maintained a low expression level at larval stage. The synthesized siRNAs were injected into mature larvae, and the RNA interference of *CmHem* was 51.7%. Moreover, the abdominal somites of larvae became straightened, could not pupate normally, then and died. Infection with a baculovirus, *C. medinalis* granulovirus (CnmeGV), the expression levels of *CmHem* in the midgut and fat body of *C. medinalis* were significantly increased at 12 and 24 hours, respectively, and then soon returned to normal levels. **Conclusions** Our results suggested that Hemolin is related to metamorphosis of *C. medinalis*. Exposure to baculovirus induced the phased expression of *hemolin* gene in midgut and fat body of *C. medinalis*, indicated that *hemolin* involved in the immune recognition of Crambidae insects to baculovirus.

CmHem, a *hemolin*-like gene identified from *Cnaphalocrocis medinalis*, involved in metamorphosis and baculovirus infection

Guangjie Han, Chuanming Li, Nan Zhang, Qin Liu, Lixin Huang, Yang Xia, Jian Xu*

Lixiahe District Institute of Agricultural Sciences in Jiangsu/National Agricultural Experimental Station for Agricultural Microbiology in Yangzhou, Yangzhou, 225007, People's Republic of China

*Correspondence: bio-xj@163.com; Tel.: +86-514-8730-2019

ORCID

Guangjie Han: 0000-0002-3920-9413

E-mail:

Guangjie Han: hanguangjie177@163.com

Chuanming Li: liming0595@163.com

Nan Zhang: znfezhangnan@hotmail.com

Qin Liu: bio-lq@126.com

Lixin Huang: 1066803310@qq.com

Yang Xia: xiayeung@163.com

Abstract

Background

As a member of the immunoglobulin (Ig) superfamily, Hemolins play a vital role in insect development and defense against pathogens. However, Hemolin has different effects on insect development and innate immune response to baculovirus infection.

Methods and results

In this study, the *hemolin*-like gene from a Crambidae insect, *Cnaphalocrocis medinalis*, *CmHem* was cloned, and its role in insect development and baculovirus infeciton was analyzed. A 1528 bp contig as potential *hemolin*-like gene of *C. medinalis* was reassembled from transcriptome. Further, the complete *hemolin* sequence of *C. medinalis* (*CmHem*) was cloned and sequenced. The cDNA of *CmHem* was 1515 bp in length and encoded 408 amino acids. The deduced amino acid of CmHem have relatively low identities (41.9-62.3%) to various insect Hemolins. However, it contains four Ig domains similarity to other insect Hemolins. The expression level of *CmHem* was the highest in eggs, followed by pupae and adults, and maintained a low expression level at larval stage. The synthesized siRNAs were injected into mature larvae, and the RNA interference of *CmHem* was 51.7%. Moreover, the abdominal somites of larvae became straightened, could not pupate normally, then and died. Infection with a baculovirus, *C. medinalis* granulovirus (CnmeGV), the expression levels of *CmHem* in the midgut and fat body of *C. medinalis* were significantly increased at 12 and 24 hours, respectively, and then soon returned to normal levels.

Conclusions

Our results suggested that Hemolin is related to metamorphosis of *C. medinalis*. Exposure to baculovirus induced the phased expression of *hemolin* gene in midgut and fat body of *C. medinalis*, indicated that *hemolin* involved in the immune recognition of Crambidae insects to baculovirus.

Keywords: *hemolin*, metamorphosis, immune recognition, infection, *Cnaphalocrocis medinalis*

Introduction

Bacterial recognition in insect is mediated by pattern recognition proteins (PRPs), including peptidoglycan recognition protein (PGRP), gram-negative binding protein (GNBP), β-1,3-glucan recognition protein (βGRP), C-type lectin, apolipophorin-III and hemolin (Eleftherianos et al., 2007; Wang et al., 2018; Wang et al., 2019). Hemolin, previously named P4 protein, belongs to the immunoglobulin (Ig) superfamily, which was described for the first time as cell adhesion molecule which can bind to the bacterial surface against bacteria challenge (Rasmuson and Boman, 1979; Sun et al., 1990). Bacterial exposure has been extensively shown to induce the expression of *hemolin* in insects, including *Manduca sexta* (Eleftherianos et al., 2007), *Antheraea pernyi* (Sun et al., 2015), *Bombyx mori* (Liu et al., 2017), *Plodia interpunctella* (Orozco-Flores et al., 2017), *Actias selene* (Qian et al., 2017), *Spodoptera exigua* (Jung et al., 2019), and the crustacean *Litopenaeus vannamei* (Zuo et al., 2015). However, its role in response to baculovirus infection is still controversial. The mRNA level of *hemolin* was induced to be up-regulated after 24 hours of ApNPV injection in *A. pernyi*. Moreover, the antibacterial activity was not activated, suggesting that *hemolin* was involved in antiviral response as a virus inducible gene (Hirai et al., 2004). *As-HEM* and *Ap-hemolin-like*, homologous genes of *hemolin*, which were cloned from Bombycoidea insects, have also been shown to respond to baculovirus infection by increasing the mRNA expression level and protein production (Qian et al., 2017; Sun et al., 2015). Of note, the *hemolin* expression of fat bodies and hemocytes in *Helicoverpa zea* and *H. virescens* larvae infected with HzNPV were not significantly different from that of the control group (Terenius et al., 2009). So far, the antiviral effect of *hemolin* in Crambidae insects is not known.

In addition, *hemolin* has also been proved to be necessary gene for development (Bettencourt et al., 2002; Liu et al., 2017). Hemolin exists at all developmental stages of *M. sexta.* The content of Hemolin in the hemolymph of naive larvae was the lowest and increased dramatically after pupation (Yu and Kanost, 1999). However, the relative mRNA levels of *S. exigua hemolin* were significantly different among various development stages, with the highest expression level in fourth instar larvae (Jung et al., 2019). Moreover, as a hormone regulating insect development, 20-hydroxyecdysone (20E) can activate the expression of *Hyalophora cecropia hemolin* in the fat body of diapausing pupae, accompanied by ongoing protein synthesis (Roxstrom-Lindquist et al., 2005). Silencing of *hemolin* gene in mated females of *H. cecropia* have resulted in deformed embryos that failed to hatch (Bettencourt et al., 2002). These

results indicated that the expression levels *hemolin* are different from different insects and closely related to development.

Cnaphalocrocis medinalis Guenée (Crambidae, Lepidoptera) has caused rice yield loss by larvae feeding on leaves. The annual occurrence area of *C. medinalis* in China was nearly 14.6 million hectares, forcing farmers to widely use pesticides, which has brought great challenges to the safe production of rice (Fu et al., 2014; Qiu et al., 2018; Zheng et al., 2011). *Cnaphalocrocis medinalis* granulovirus (CnmeGV) is an effective microbial agent against the rice leaffolder. The killing efficacy of CnmeGV with the concentration of 2.5×10^3 Occlusion Bodies (OBs)•g^{-1} reached 71% at 18 days (Xu et al., 2016). The field application showed that it formed persistent infection in the larval population within 48 days and protected arthropods (Xu et al., 2019). However, the progress of CnmeGV infection is slow, which limits its wide use. In this study, the *hemolin* gene of *C. medinalis* (*CmHem*) was cloned and characterized, its role in development and baculovirus infection was confirmed. This basic information is helpful to improve the understanding of the *hemolin* and lay a foundation for further research on improving the infection efficiency of baculovirus by inhibiting *hemolin*.

Materials & Methods

Insect rearing and infection

The larvae of *C. medinalis* were fed with corn leaves in the laboratory, under standard conditions of 14: 10 h (L:D) photoperiod, temperature 28 °C, at relative humidity of 70%. The eggs, different instars larvae, pupae and adults were collected. CnmeGV was purified and prepared for the different concentration as previously described (Han et al., 2016). The early fourth-instar larvae were singled out and starved for 6 hours, then fed with corn leaves soaked in 10^5 OB/ml CnmeGV (Han et al., 2021). Five infected larvae were collected as one replication after 12, 24, 48, 96 hours, respectively. Each treatment was replicated for three times. All larvae were dissected, the midgut and fat bodies were washed in PBS prepared with diethyl pyrocarbonate treated H_2O and collected, respectively. These tissues were immediately frozen in liquid nitrogen, then stored at -80 ℃.

RNA extraction and reverse transcription

All frozen samples were pulverized in liquid nitrogen. Total RNA was extracted using RNA extraction Kit (Code No. 9767) from TaKaRa. The purity and integrity of all RNA samples were assessed using a Nanophotometer® N50

Touch spectrophotometer (IMPLEN, Germany) and confirmed by 1% agarose gel electrophoresis, respectively. The cDNA of all samples were synthesized using the PrimeScript™ RT reagent Kit with gDNA Eraser (Code No. RR047A) from TaKaRa.

cDNA cloning

The *hemolin*-like unigenes were searched in the transcriptom database of *C. medinalis* using local BLAST of TB-tools software (Chen et al., 2020). The identified unigenes were reassembled by multiple alignment as original *hemolin*-like sequence of *C. medinalis.* A pair of primers were designed to amplify the full-length *hemolin* gene of *C. medinalis* (*CmHem*) (CmHemF: TGCCATTTTTGCTGTAGTTTTC; CmHemR: ATGAACCAGAGTTATGGGGATG), using the cDNA of fourth-instar larva as the template. PCR amplification was performed using the 2× Taq Master Mix polymerase (Code No. P112-02) from Vazyme on T100™ Thermal Cycler (Bio-rad, USA) with a condition of 95 °C for 30 s, following by 30 cycles of 5 s at 95 °C, 30 s at 50 °C, 60 s at 72 °C, extending 10 min at 72 °C. The PCR product was purified by 0.8% agarose electrophoresis, then cloned into *pEASY*-T3 vector and sequenced.

Reverse transcription quantitative PCR **(RT-qPCR)**

The relative expression level of *CmHem* (primers for DLHemF: GCCTTCAGAGGTGCTGTTCCG; DLHemR: TCGTCGTCTTTATGCCATTCGTA) was analysed quantitatively using comparative C_T ($2^{-\Delta\Delta CT}$) (Livak and Schmittgen, 2001). The house keeping gene *β-actin* was used as an internal control for normalization (Han et al., 2021). RT-qPCR was performed using TB Green™ Fast qPCR Mix polymerase (Code No. RR430S) from TaKaRa on StepOnePlus Real-Time PCR System (Applied Biosystems).

RNA interference (RNAi) of *CmHem*

Two pairs of specific dsRNA of CmHem, siRNA983(sense: 5′-GGAGUAUAAGUUCAACGUUTT-3′; antisense: 5′-AACGUUGAACUUAUACUCCTT-3′) and siRNA1323 (sense: 5′-GCGAGAUAAUUUGUCGACATT-3′; antisense: 5′-UGUCGACAAAUUAUCUCGCTT-3′) and the non-specific siRNA-c (sense: 5′-UUCUCCGAACGUGUCACGUTT-3′; antisense: 5′-ACGUGACACGYYCGGAGAATT-3′) as negative control were synthesized in GenePharma. All synthesized siRNAs were modified with 2′-Ome and prepared in ddH_2O with concentrations of 1μg/μL, respectively. The siRNA983 and siRNA1323 were mixed by equal volume for treatment. 200 nL of mixed siRNA and siRNA-c were injected into the fat bodies of mature larvae by Nanoject III

(USA), respectively. Each treatment was replicated three times with twelve insect samples.

Sequence analysis

Unigenes of *hemolin*-like were reassembled using Vector NTI. The cDNA and amino acid sequences were analysed by GENEDOC (http://www.psc.edu/biomed/genedoc). The conserved domains were analyzed by conserved domain database (CDD) (www.ncbi.nlm.nih.gov/Structure/cdd) (Lu et al., 2020). The predicted molecular weight and tertiary structure of CmHem protein were analysed by ExPASy (https://www.expasy.org/). The signal peptide was predicted using SignalP 5.0 server (Petersen et al., 2011). Multiple alignment was created by MUSCLE (Edgar, 2004) and the phylogeny of hemolin protein was built by Maximum Likelihood method using MEGA 7.0 (Kumar et al., 2016).

Results

Potential *hemolin*-like gene in the *C. medinalis* transcriptome

A total of 13 unigenes were identified in the transcriptome of *C. medinalis* by searching annotation files (Table 1). Most unigenes were short in length, ranging from 300 to 500 bp. The longest unigene was 676 bp, sharing 50.5% identity to the *hemolin* gene of *Ostrinia furnacalis*. The 13 unigenes were reassembled to obtain a new 1528 bp contig as potential *hemolin*-like gene of *C. medinalis* (Fig. 1). The newly assembled contig had 48 mutation sites. These undetermined bases need to be further confirmed by sanger dideoxy sequencing.

Cloning and characterization of *hemolin* gene from *C. medinalis*

The *hemolin* gene of *C. medinalis* (*CmHem*) was amplified by primers designed with the assembled conting as the reference sequence. The cloned *CmHem* gene was 1515 bp in length and encoded 408 amino acids (GeneBank accession number: MK138364) (Fig. 2). CmHem protein was predicted to contain a signal peptide of 8 amino acids (QAQPVSQA). The predicted isoelectric point and molecular weight were 5.77 and 45.17 kDa, respectively. CmHem has four Ig domains (IG1, IG2, IG3 and IG4) by CDD analysis.

Multiple alignment results showed that the Hemolin protein from *C. medinalis* shared the most similarity with that from *O. furnacalis* (62.3% identity) and the lowest similarity with that from *Plutella xylostella* (41.9% identity). Similar to other Hemolin proteins, CmHem contains 8 cysteine residues forming 4 disulfide bridges with a tryptophan residue packed against the disulfide bond (Fig. 3A). The tertiary structure prediction showed that CmHem had 47.1% identity with the Hemolin of *H. cecropia*, both of which had four Ig domains (Fig. 3B). 22 related Hemolin protein sequences were downloaded from GeneBank for evolutionary analysis. Phylogenetic analysis indicated that the

CmHem had the close genetic distance with the Hemolin of *O. furnacalis*. In addition, different from Lepidoptera, *L. vannamei* belonging to the crustacean, was grouped into a separate cluster (Fig. 3C).

Expression levels of *CmHem* at different developmental stages

To clarify the effect of *hemolin* gene on insect development, the expression levels of *CmHem* in *C. medinalis* at different development stages were analysed by RT-qPCR. The results showed that the expression level of *CmHem* was higher in eggs, pupae and adults, while lower in larvae. In particular, the expression level of *CmHem* in eggs was 310 times higher than that of fifth instar larvae and 3 times of pupae or adults. However, it was no significant difference in the *CmHem* expression among different larval stages (Fig. 4).

The RNAi of *CmHem* in mature larvae of *C. medinalis*

Within 48h of *CmHem* siRNAs injection, the abdominal somites of larvae became straightened and could not pupate normally, then died on day 3 onwards (Fig. 5A). Further, the relative expression of CmHem was analysis after 48 hours of RNAi. Compared to the untreated control, the RNAi efficiency of *CmHem* was 51.7% (Fig. 5B).

Expression analysis of *CmHem* in *C. medinalis* after infection with CnmeGV

The immune response of *CmHem* to CnmeGV challenge in fat body and midgut tissues were analyzed at different stages of infection. The expression level of *CmHem* in midgut tissue of *C. medinalis* was up-regulated at the time after 12 hours of infection, with 22 times higher than that before infection. Subsequently, the expression of *CmHem* decreased without significant difference from that before infection. However, in fat body, the expression level of *CmHem* was up-regulated 4 times higher at the time after 24 hours of infection. After that, the expression level of *CmHem* in fat body also returned to the level before infection.

Discussion

In this study, 13 *hemolin*-like unigenes were identified in the transcriptome of *C. medinalis*. By amplifying the contig and sequencing, a homologous gene of *hemolin* from *C. medinalis* (*CmHem*) was cloned. CmHem is composed of four Ig domains to form a horseshoe structure, which is consistent with other insect Hemolins (Qian et al., 2017; Sun et al., 2015). However, it is different from the LvHemolin of *L. vannamei*, which is composed of seven Ig domains. The last four Ig domains of LvHemolin share high identity with insect Hemolins and form a globular structure, while the role of other three Ig-domains is still unknown (Zuo et al., 2015). As a member of the immunoglobulin superfamily, Down syndrome cell adhesion molecule (Dscam) also contains 10 Ig domains, and is speculated to be a hypervariable

pattern-recognition receptor in insect immunity (Ng and Kurtz, 2020; Watson et al., 2005). In *Drosophila*, the first four Ig domains form a horseshoe structure and Ig5-Ig8 domains show different shapes in the absence of homophilic binding (Hattori et al., 2008). The external facing Ig2 and Ig3 of horseshoe structure play an important role in heterophilic specific binding with pathogens (Li et al., 2018; Ng and Kurtz, 2020). Hence, CmHem with similar structure may also be involved in the identification of pathogens.

Hemolin can bind to lipopolysaccharide (LPS) or lipoteichoic acid on the bacterial surface, and aggregate them to lead hemocytes reaction (Daffre and Faye, 1997; Sun et al., 1990). *Escherichia coli*, *Beauveria bassiana* (Sun et al., 2015), *Micrococcus luteus* (Qian et al., 2017), *Photorhabdus temperata* (Jung et al., 2019) have been widely proved to induce the expression of insect *hemolin*. Interestingly, Hemolin can respond to the viral infection in the superfamily Bombycoidea, but not in the Noctuidae (Qian et al., 2017; Terenius et al., 2009). Our results showed that *CmHem* in the midgut and fat body of *C. medinalis* can responded to CnmeGV infection. However, the expression of *CmHem* soon returned to normal levels after up-regulation, indicating that CmHem may involved in the initial immune recognition. Whether LPS analogs are present on the surface of baculovirus is not clear, or Hemolin can bind to unknown proteins on the surface of the virus. The study found that Hemolin isolated from *H. cecropia*, was confirmed as a specific lectin by the homophilic binding properties analysis and bound to glycosylated surfaces, such as the virion envelope (Bettencourt et al., 1999).

As a PRP, Hemolin is more likely to regulate and stimulate the humoraland cellular immune response through activating a series of signal pathways (Terenius, 2008). In *Drosophila* cell line mbn-2, Hemolin enhanced the protein kinase C (PKC) activity in hemocyte crude extracts and prevented tyrosine phosphorylation of two proteins of 35 and 40 kDa, suggesting that Hemolin is involved in the regulation of the cellular immune responses via a pathway that includes PKC activation and protein tyrosine phosphorylation (Lanz-Mendoza et al., 1996). Hemolin in the pupae of *H. cecropia* has also been shown to be important for the triggering of the prophenoloxidase cascade in the defence against bacterial infections (Terenius et al., 2007). The third intron of *H. cecropia hemolin* contains κB motif, which can be activated by Relish of Toll signaling pathway (Roxström-Lindquist et al., 2002). Inhibition of Toll pathway in *S. exigua* also reduced the expression of *hemolin* (Park and Kim, 2012). These results indicated that the expression of *hemolin* is related to Toll signaling pathway. There have 10 unigenes in Toll/IMD pathway, 2 unigenes in JNK pathway, 8 unigenes in JAK/STAT pathway were identified in *C. medinalis* infection with CnmeGV (Han et al., 2021).

Toll, IMD and JAK-STAT pathways have been shown to be associated with viral infection (Jakubowska et al., 2013; Liu et al., 2015), but the expression of *hemolin* and the interaction of these signaling pathways are unclear in *C. medinalis*.

In addition, the expression of *hemolin* in *H. cecropia* was regulated by 20-hydroxyecdysone (20E) (Roxstrom-Lindquist et al., 2005), and was detected in eggs, larvae, pupa and adults (Jung et al., 2019; Yu and Kanost, 1999), indicating that Hemolin was related to insect development. The synthesization of *M. sexta* Hemolin was very low during the larval feeding perilds, but very high in eggs and pupae (Yu and Kanost, 1999). Hemolin proteins were involved in intercelluar adhesion, including cell proliferation and wound healing (Sato et al., 2016). Meanwhile, specialized structures of the preceding developmental stage have to be completely remodeled to accommodate new structures and behaviors in metamorphosis of insects (Pinet and McLaughlin, 2019). In our study, the expression of *hemolin* in eggs, pupae and adults of *C. medinalis* is significantly higher than that of larvae, and the mature larvae can not pupate normally after *CmHem* interference, suggesting that Hemolin plays a key role in metamorphosis of *C. medinalis*. However, the relative mRNA levels of *hemolin* in eggs of *S. exigua* was found to be the lowest (Jung et al., 2019), showing a low correlation with metamorphosis. 142 and 80 immunoglobulin superfamily proteins were identified in *Drosophila melanogaster* and *Caenorhabditis elegans* , respectively (Vogel et al., 2003). Some proteins that may have a similar structure to Hemolin play a role in the metamorphosis of *S. exigua*. These potential mechanisms deserve further exploration.

Conclusions

As cell adhesion molecule, hemolin has been proved to be able to bind to the bacterial surface against bacteria challenge and is closely related to development of insects. However, the expression pattern of hemolin and its response to baculovirus infection vary among different insects. The expression levels of *hemolin* of *C. medinalis* were higher in eggs, pupae and adults, and inhibiting its expression in mature larvae made it impossible to pupate, suggesting that Hemolin was related to metamorphosis of *C. medinalis*. Infection with baculovirus CnmeGV, the expression levels of *hemolin* in the midgut and fat body of *C. medinalis* were significantly increased at 12 and 24 hours, respectively, and then soon returned to normal levels, showing that exposure to baculovirus induced the phased expression of *hemolin* gene in midgut and fat body of *C. medinalis*. The immune response of insects to viral infections is complex and highly

variable. It is important to study how different insects respond to baculovirus infection in order to better understand insect immunity and develop effective pest control strategies.

Acknowledgements

The authors are grateful to Yurong Lu and Yangsheng Cai for their assistance in laboratory work.

Author contributions

Guangjie Han did most of the experimental work and wrote the manuscript; Chuanming Li analysed the data and performed RNA extraction; Nan Zhang analysed the RNAi efficiency, Qin Liu screened viral sublethal doses and infected moth; Lixin Huang and Yang Xia bred sensitive population; JX designed the experiments and revised the manuscript.

Funding

This work was supported by the Jiangsu Agricultural Science and Technology Innovation Fund (Grant No. ZX(20)1004), Jiangsu Science Project of China (Grant No. BK20181215) and Yangzhou science and technology project (Grant No. YZ2021034).

Competing interests

The authors have no relevant financial or non-financial interests to disclose.

Availability of data and materials

All data generated or analyzed during this study are included in this published article.

References

Bettencourt R, Gunne H, Gastinel L, Steiner H, Faye I. 1999. Implications of hemolin glycosylation and Ca^{2+}-binding on homophilic and cellular interactions. *European Journal of Biochemistry* **266(3)**:964-976. DOI 10.1046/j.1432-1327.1999.00934.x.

Bettencourt R, Terenius O, Faye I. 2002. Hemolin gene silencing by ds-RNA injected into Cecropia pupae is lethal

to next generation embryos. *Insect Molecular Biology* **11(2)**:267-271. DOI 10.1046/j.1365-2583.2002.00334.x.

Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, He Y, Xia R. 2020. TBtools: an integrative toolkit developed for interactive analyses of big biological data. *Molecular Plant* **13(8)**:1194-1202. DOI 10.1016/j.molp.2020.06.009.

Daffre S, Faye I. 1997. Lipopolysaccharide interaction with hemolin, an insect member of the Ig-superfamily. *FEBS Letters* **408(2)**:127-130. DOI 10.1016/s0014-5793(97)00397-9.

Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* **32(5)**:1792-1797. DOI 10.1093/nar/gkh340.

Eleftherianos I, Gokcen F, Felfoldi G, Millichap PJ, Trenczek TE, ffrench-Constant RH, Reynolds SE. 2007. The immunoglobulin family protein Hemolin mediates cellular immune responses to bacteria in the insect *Manduca sexta*. *Cellular Microbiology* **9(5):**1137-1147. DOI 10.1111/j.1462-5822.2006.00855.x.

Fu XW, Li C, Feng HQ, Liu ZF, Chapman JW, Reynolds DR, Wu KM. 2014. Seasonal migration of *Cnaphalocrocis medinalis* (Lepidoptera: Crambidae) over the Bohai Sea in northern China. *Bulletin of Entomological Research* **104(5)**:601-609. DOI 10.1017/S0007485314000376.

Han G, Liu Q, Li C, Xu B, Xu J. 2021. Transcriptome sequencing reveals *Cnaphalocrocis medinalis* against baculovirus infection by oxidative stress. *Molecular Immunology* **129**:63-69. DOI 10.1016/j.molimm.2020.10.020.

Han GJ, Xu J, Liu Q, Li CM, Xu HX, Lu ZX. 2016. Genome of *Cnaphalocrocis medinalis* granulovirus, the first Crambidae-infecting betabaculovirus isolated from rice leaffolder to sequenced. *PLoS One*, **11(2)**:e0147882. DOI 10.1371/journal.pone.0147882.

Hattori D, Millard SS, Wojtowicz WM, Zipursky SL. 2008. Dscam-mediated cell recognition regulates neural circuit formation. *Annual Review of Cell and Developmental Biology* **24**:597-620. DOI 10.1146/annurev.cellbio.24.110707.175250.

Hirai M, Terenius O, Li W, Faye I. 2004. Baculovirus and dsRNA induce Hemolin, but no antibacterial activity, in *Antheraea pernyi*. *Insect Molecular Biology* **13(4)**:399-405. DOI 10.1111/j.0962-1075.2004.00497.x.

Jakubowska AK, Vogel H, Herrero S. 2013. Increase in gut microbiota after immune suppression in baculovirus-infected larvae. *PLoS Pathogens* **9(5)**:e1003379. DOI 10.1371/journal.ppat.1003379.

Jung J, Sajjadian SM, Kim Y. 2019. Hemolin, an immunoglobulin-like peptide, opsonizes nonself targets for phagocytosis and encapsulation in *Spodoptera exigua*, a lepidopteran insect. *Journal of Asia-Pacific Entomology*

22(3):947-956. DOI 10.1016/j.aspen.2019.08.002.

Kumar S, Stecher G, Tamura K. 2016. MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution* **33(7)**:1870-1874. DOI 10.1093/molbev/msw054.

Lanz-Mendoza H, Bettencourt R, Fabbri M, Faye I. 1996. Regulation of the insect immune response: the effect of hemolin on cellular immune mechanisms. *Cellular Immunology* **169(1)**:47-54. DOI 10.1006/cimm.1996.0089.

Li XJ, Yang L, Li D, Zhu YT, Wang Q, Li WW. 2018. Pathogen-specific binding soluble down syndrome cell adhesion molecule (Dscam) regulates phagocytosis via membrane-bound Dscam in crab. *Frontiers in Immunology* **9**:801. DOI 10.3389/fimmu.2018.00801.

Liu W, Liu J, Lu Y, Gong Y, Zhu M, Chen F, Liang Z, Zhu L, Kuang S, Hu X. 2015. Immune signaling pathways activated in response to different pathogenic micro-organisms in *Bombyx mori*. *Molecular Immunology* **65(2)**:391-397. DOI 10.1016/j.molimm.2015.02.018.

Liu X, Li J, Wang Q, Xia H, Chen K. 2017. Functional analysis of hemolin gene from silkworm, *Bombyx mori*-immune and development. *ISJ-Invertebrate Survival Journal* **14**:330-339. DOI 10.25431/1824-307X/isj.v14i1.330-339.

Livak KJ, Schmittgen TD. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. *Methods* **25(4)**:402-408. DOI 10.1006/meth.2001.1262.

Lu S, Wang J, Chitsaz F, Derbyshire MK, Geer RC, Gonzales NR, Gwadz M, Hurwitz DI, Marchler GH, Song JS. 2020. CDD/SPARCLE: the conserved domain database in 2020. *Nucleic Acids Research* **48(D1)**:D265-D268. DOI 10.1093/nar/gkz991.

Ng TH, Kurtz J. 2020. Dscam in immunity: a question of diversity in insects and crustaceans. *Developmental and Comparative Immunology* **105**:103539. DOI 10.1016/j.dci.2019.103539.

Orozco-Flores AA, Valadez-Lira JA, Oppert B, Gomez-Flores R, Tamez-Guerra R, Rodríguez-Padilla C, Tamez-Guerra P. 2017. Regulation by gut bacteria of immune response, *Bacillus thuringiensis* susceptibility and hemolin expression in *Plodia interpunctella*. *Journal of Insect Physiology* **98**:275-283. DOI 10.1016/j.jinsphys.2017.01.020.

Park J-A, Kim Y. 2012. Eicosanoid biosynthesis is activated via Toll, but not Imd signal pathway in response to fungal infection. *Journal of Invertebrate Pathology* **110(3)**:382-388. DOI 10.1016/j.jip.2012.04.015.

312 **Petersen TN, Brunak S, Heijne GV, Nielsen H. 2011.** SignalP 4.0: discriminating signal peptides from
313 transmembrane regions. *Nature Methods* **8(10)**:785-786. DOI 10.1038/nmeth.1701.
314 **Pinet K, McLaughlin KA. 2019.** Mechanisms of physiological tissue remodeling in animals: Manipulating tissue,
315 organ, and organism morphology. *Developmental Biology* **451(2)**:134-145. DOI 10.1016/j.ydbio.2019.04.001.
316 **Qian C, Wang F, Zhu B, Wang L, Wei G, Sun Y, Li S, Liu C. 2017.** Identification of a hemolin protein from *Actias*
317 *selene* mediates immune response to pathogens. *International Immunopharmacology* **42**:74-80. DOI
318 10.1016/j.intimp.2016.11.020.
319 **Qiu L, Liu Q, Lu X, Zhan Z. 2018.** Susceptibility of *Cnaphalocrocis medinalis* (Güenée) in Fujian to six insecticides.
320 *Journal of Southern Agriculture* **49(1)**:62-67. DOI 10.3969/j.issn.2095-1191.2018.01.10.
321 **Rasmuson T, Boman HG. 1979.** Insect immunity—V. Purification and some properties of immune protein P4 from
322 haemolymph of *Hyalophora cecropia* pupae. *Insect Biochemistry* **9(3)**:259-264. DOI 10.1016/0020-1790(79)90003-
323 9.
324 **Roxstrom-Lindquist K, Assefaw-Redda Y, Rosinska K, Faye I. 2005.** 20-Hydroxyecdysone indirectly regulates
325 Hemolin gene expression in *Hyalophora cecropia*. *Insect Molecular Biology* **14(6)**:645-652. DOI 10.1111/j.1365-
326 2583.2005.00593.x.
327 **Roxström-Lindquist K, Lindström-Dinnetz I, Olesen J, Engström Y, Faye I. 2002.** An intron enhancer activates
328 the immunoglobulin-related Hemolin gene in *Hyalophora cecropia*. *Insect Molecular Biology* **11(5)**:505-515. DOI
329 10.1046/j.1365-2583.2002.00359.x.
330 **Sato AC, Bosch RV, Will SE, Alvarez-Flores MP, Goldfeder MB, Pasqualoto KF, da Silva BA, de Andrade SA,**
331 **Chudzinski-Tavassi AM. 2016.** Exploring the in vivo wound healing effects of a recombinant hemolin from the
332 caterpillar *Lonomia obliqua*. *Journal of Venomous Animals and Toxins Including Tropical Diseases* **22**:36. DOI
333 10.1186/s40409-016-0093-4.
334 **Sun SC, Lindström I, Boman HG, Faye I, Schmidt O. 1990.** Hemolin: an insect-immune protein belonging to the
335 immunoglobulin superfamily. *Science* **250(4988)**:1729-1732. DOI 10.1126/science.2270488.
336 **Sun Y, Wang L, Qian C, Dai L, Li J, Zhang C, Zhu BJ, Liu CL. 2015.** Molecular cloning and expression analysis
337 of a hemolin-like molecule from *Antheraea pernyi*. *International Immunopharmacology* **26(1)**:65-71. DOI
338 10.1016/j.intimp.2015.03.010.

Terenius O. 2008. Hemolin-A lepidopteran anti-viral defense factor? *Developmental and Comparative Immunology* **32(4)**:311-316. DOI 10.1016/j.dci.2007.09.006.

Terenius O, Bettencourt R, Lee SY, Li W, Söderhäll K, Faye I. 2007. RNA interference of Hemolin causes depletion of phenoloxidase activity in *Hyalophora cecropia*. *Developmental and Comparative Immunology* **31(6)**:571-575. DOI 0.1016/j.dci.2006.09.006.

Terenius O, Popham H, Shelby K. 2009. Bacterial, but not baculoviral infections stimulate Hemolin expression in noctuid moths. *Developmental and Comparative Immunology*, **33(11)**:1176-1185. DOI 10.1016/j.dci.2009.06.009.

Vogel C, Teichmann SA, Chothia C. 2003. The immunoglobulin superfamily in *Drosophila melanogaster* and *Caenorhabditis elegans* and the evolution of complexity. *Development* **130(25)**:6317-6328. DOI 10.1242/dev.00848.

Wang X, Luo H, Zhang R. 2018. Innate immune responses in the Chinese oak silkworm, *Antheraea pernyi*. *Developmental and Comparative Immunology* **83**:22-33. DOI 10.1016/j.dci.2017.12.010.

Wang X, Zhang Y, Zhang R, Zhang J. 2019. The diversity of pattern recognition receptors (PRRs) involved with insect defense against pathogens. *Current Opinion in Insect Science* **33**:105-110. DOI 10.1016/j.cois.2019.05.004.

Watson FL, Püttmann-Holgado R, Thomas F, Lamar DL, Hughes M, Kondo M, Rebel VI, Schmucker D. 2005. Extensive diversity of Ig-superfamily proteins in the immune system of insects. *Science* **309(5742)**:1874-1878. DOI 0.1126/science.1116887.

Xu J, Li CM, Han GJ, Liu Q. 2016. *Cnaphalocrocis medinalis* (Guenée) granulovirus, a potential microbial agent against the rice leaf folder *Cnaphalocrocis medinalis* Güenée (Lepidoptera: Pyralidae). *Egyptian Journal of Biological Pest Control* **26(4)**:675-681.

Xu J, Liu Q, Li C, Han G. 2019. Field effect of *Cnaphalocrocis medinalis* granulovirus (CnmeGV) on the pest of rice leaffolder. *Journal of Integrative Agriculture* **18(9)**:2115-2122. DOI 10.1016/S2095-3119(18)62097-0.

Yu XQ, Kanost MR. 1999. Developmental expression of *Manduca sexta* Hemolin. *Archives of Insect Biochemistry and Physiology* **42**:198-212. DOI 10.1002/(SICI)1520-6327(199911)42:3<198::AID-ARCH4>3.0.CO;2-G.

Zheng X, Ren X, Su J. 2011. Insecticide susceptibility of *Cnaphalocrocis medinalis* (Lepidoptera: Pyralidae) in China. *Journal of Economic Entomology* **104(2)**:653-658. DOI 10.1603/ec10419.

Zuo H, Li H, Wei E, Su Z, Zheng J, Li C, Chen Y, Weng S, He J, Xu X. 2015. Identification and functional analysis of a Hemolin like protein from *Litopenaeus vannamei*. *Fish & Shellfish Immunology* **43(1)**:51-59. DOI

10.1016/j.fsi.2014.12.004.

Figure legend

Fig.1 The reassembly of *hemolin*-like gene in *C. medinalis* transcriptome.

Fig. 2 Full-length cDNA sequence of *CmHem* gene and amino acid sequence from *C. medinalis*.

The 1515 bp sequence of *CmHem* has 1227 bp of 408 amino acid protein-coding region, 53 bp putative 5′ untranslated region and 235 bp 3′ untranslated region. The signal sequence is underlined. Four groups of the immunoglubin domain (IG1, 2, 3, 4) are boxed.

Fig. 3. The characteristic of CmHem protein from *C. medinalis*. (A) Multiple alignment of Hemolin proteins between *C. medinalis* and other insects. Pentagrams and triangles point to cysteine residues and tryptophan residues, respectively. (B) Predicted tertiary structure of CmHem protein by ExPASy. The CmHem protein contains four immunoglobulin domains (IG1, IG2, IG3 and IG4). (C) Phylogenetic analysis of Hemolin proteins from *C. medinalis* and other species. The phylogenetic tree was constructed by MEGA 7.0 using Maximum Likelihood method.

Fig. 4. The relative expression levels of *CmHem* at different developmental stages of *C. medinalis*.

1^{st}, 2^{nd}, 3^{rd}, 4^{th} and 5^{th} represent different instar larvae. The values are represented as mean ± SD. Different small letters above histograms indicated significant differences among different treatments at the 0.05 level (ANOVA).

Fig. 5. The phenotype of mature larvae after *CmHem* interference at 72 hours (A) and the relative expression levels of *CmHem* after 48 hours of interference (B).

Fig. 6. The relative expression levels of *CmHem* in midguts (A) and fat bodies (B) of *C. medinalis* after CnmeGV infection.

Samples were collected after 12, 24, 48 and 96 hours of infection for differential expression analysis of *CmHem* (ANOVA, $p<0.0001$).

Figure 1

Fig.1 The reassembly of *hemolin*-like gene in *C. medinalis* transcriptome.

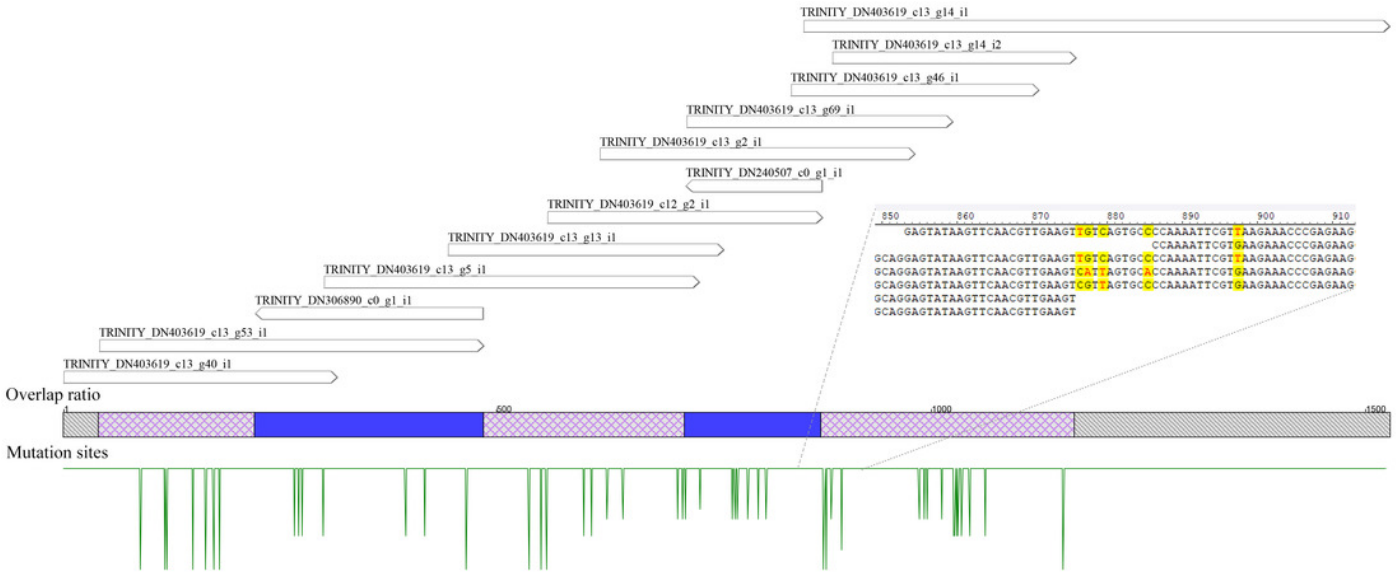

Figure 2

Fig. 2 Full-length cDNA sequence of *CmHem* gene and amino acid sequence from *C. medinalis*.

The 1515 bp sequence of *CmHem* has 1227 bp of 408 amino acid protein-coding region, 53 bp putative 5′ untranslated region and 235 bp 3′ untranslated region. The signal sequence is underlined. Four groups of the immunoglubin domain (IG1, 2, 3, 4) are boxed.

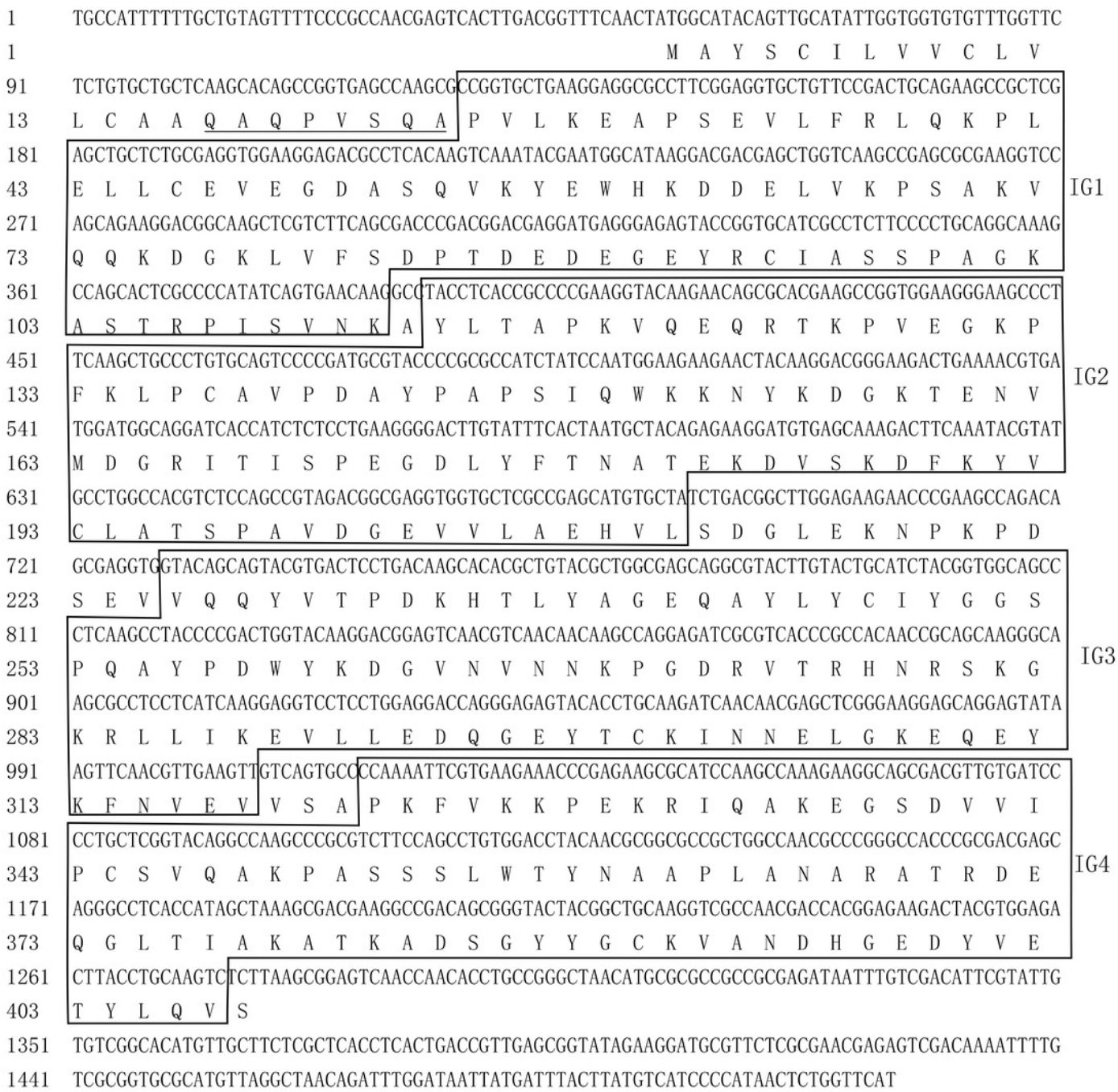
1 TGCCATTTTTTGCTGTAGTTTTCCCGCCAACGAGTCACTTGACGGTTTCAACTATGGCATACAGTTGCATATTGGTGGTGTGTTTGGTTC
1 M A Y S C I L V V C L V
91 TCTGTGCTGCTCAAGCACAGCCGGTGAGCCAAGCGCCGGTGCTGAAGGAGGCGCCTTCGGAGGTGCTGTTCCGACTGCAGAAGCCGCTCG
13 L C A A Q A Q P V S Q A P V L K E A P S E V L F R L Q K P L
181 AGCTGCTCTGCGAGGTGGAAGGAGACGCCTCACAAGTCAAATACGAATGGCATAAGGACGACGAGCTGGTCAAGCCGAGCGCGAAGGTCC
43 E L L C E V E G D A S Q V K Y E W H K D D E L V K P S A K V
271 AGCAGAAGGACGGCAAGCTCGTCTTCAGCGACCCGACGGACGAGGATGAGGGAGAGTACCGGTGCATCGCCTCTTCCCCTGCAGGCAAAG
73 Q Q K D G K L V F S D P T D E D E G E Y R C I A S S P A G K
361 CCAGCACTCGCCCCATATCAGTGAACAAGGCCTACCTCACCGCCCCGAAGGTACAAGAACAGCGCACGAAGCCGGTGGAAGGGAAGCCCT
103 A S T R P I S V N K A Y L T A P K V Q E Q R T K P V E G K P
451 TCAAGCTGCCCTGTGCAGTCCCCGATGCGTACCCCGCGCCATCTATCCAATGGAAGAAGAACTACAAGGACGGGAAGACTGAAAACGTGA
133 F K L P C A V P D A Y P A P S I Q W K K N Y K D G K T E N V
541 TGGATGGCAGGATCACCATCTCTCCTGAAGGGGACTTGTATTTCACTAATGCTACAGAGAAGGATGTGAGCAAAGACTTCAAATACGTAT
163 M D G R I T I S P E G D L Y F T N A T E K D V S K D F K Y V
631 GCCTGGCCACGTCTCCAGCCGTAGACGGCGAGGTGGTGCTCGCCGAGCATGTGCTATCTGACGGCTTGGAGAAGAACCCGAAGCCAGACA
193 C L A T S P A V D G E V V L A E H V L S D G L E K N P K P D
721 GCGAGGTGGTACAGCAGTACGTGACTCCTGACAAGCACACGCTGTACGCTGGCGAGCAGGCGTACTTGTACTGCATCTACGGTGGCAGCC
223 S E V V Q Q Y V T P D K H T L Y A G E Q A Y L Y C I Y G G S
811 CTCAAGCCTACCCCGACTGGTACAAGGACGGAGTCAACGTCAACAACAAGCCAGGAGATCGCGTCACCCGCCACAACCGCAGCAAGGGCA
253 P Q A Y P D W Y K D G V N V N N K P G D R V T R H N R S K G
901 AGCGCCTCCTCATCAAGGAGGTCCTCCTGGAGGACCAGGGAGAGTACACCTGCAAGATCAACAACGAGCTCGGGAAGGAGCAGGAGTATA
283 K R L L I K E V L L E D Q G E Y T C K I N N E L G K E Q E Y
991 AGTTCAACGTTGAAGTTGTCAGTGCCCCAAAATTCGTGAAGAAACCCGAGAAGCGCATCCAAGCCAAAGAAGGCAGCGACGTTGTGATCC
313 K F N V E V V S A P K F V K K P E K R I Q A K E G S D V V I
1081 CCTGCTCGGTACAGGCCAAGCCCGCGTCTTCCAGCCTGTGGACCTACAACGCGGCGCCGCTGGCCAACGCCCGGGCCACCCGCGACGAGC
343 P C S V Q A K P A S S S L W T Y N A A P L A N A R A T R D E
1171 AGGGCCTCACCATAGCTAAAGCGACGAAGGCCGACAGCGGGTACTACGGCTGCAAGGTCGCCAACGACCACGGAGAAGACTACGTGGAGA
373 Q G L T I A K A T K A D S G Y Y G C K V A N D H G E D Y V E
1261 CTTACCTGCAAGTCTCTTAAGCGGAGTCAACCAACACCTGCCGGGCTAACATGCGCGCCGCCGCGAGATAATTTGTCGACATTCGTATTG
403 T Y L Q V S
1351 TGTCGGCACATGTTGCTTCTCGCTCACCTCACTGACCGTTGAGCGGTATAGAAGGATGCGTTCTCGCGAACGAGAGTCGACAAAATTTTG
1441 TCGCGGTGCGCATGTTAGGCTAACAGATTTGGATAATTATGATTTACTTATGTCATCCCCATAACTCTGGTTCAT
IG1
IG2
IG3
IG4

Figure 3

Fig. 3. The characteristic of CmHem protein from *C. medinalis*.

(A) Multiple alignment of Hemolin proteins between *C. medinalis* and other insects. Pentagrams and triangles point to cysteine residues and tryptophan residues, respectively. (B) Predicted tertiary structure of CmHem protein by ExPASy. The CmHem protein contains four immunoglobulin domains (IG1, IG2, IG3 and IG4). (C) Phylogenetic analysis of Hemolin proteins from *C. medinalis* and other species. The phylogenetic tree was constructed by MEGA 7.0 using Maximum Likelihood method.

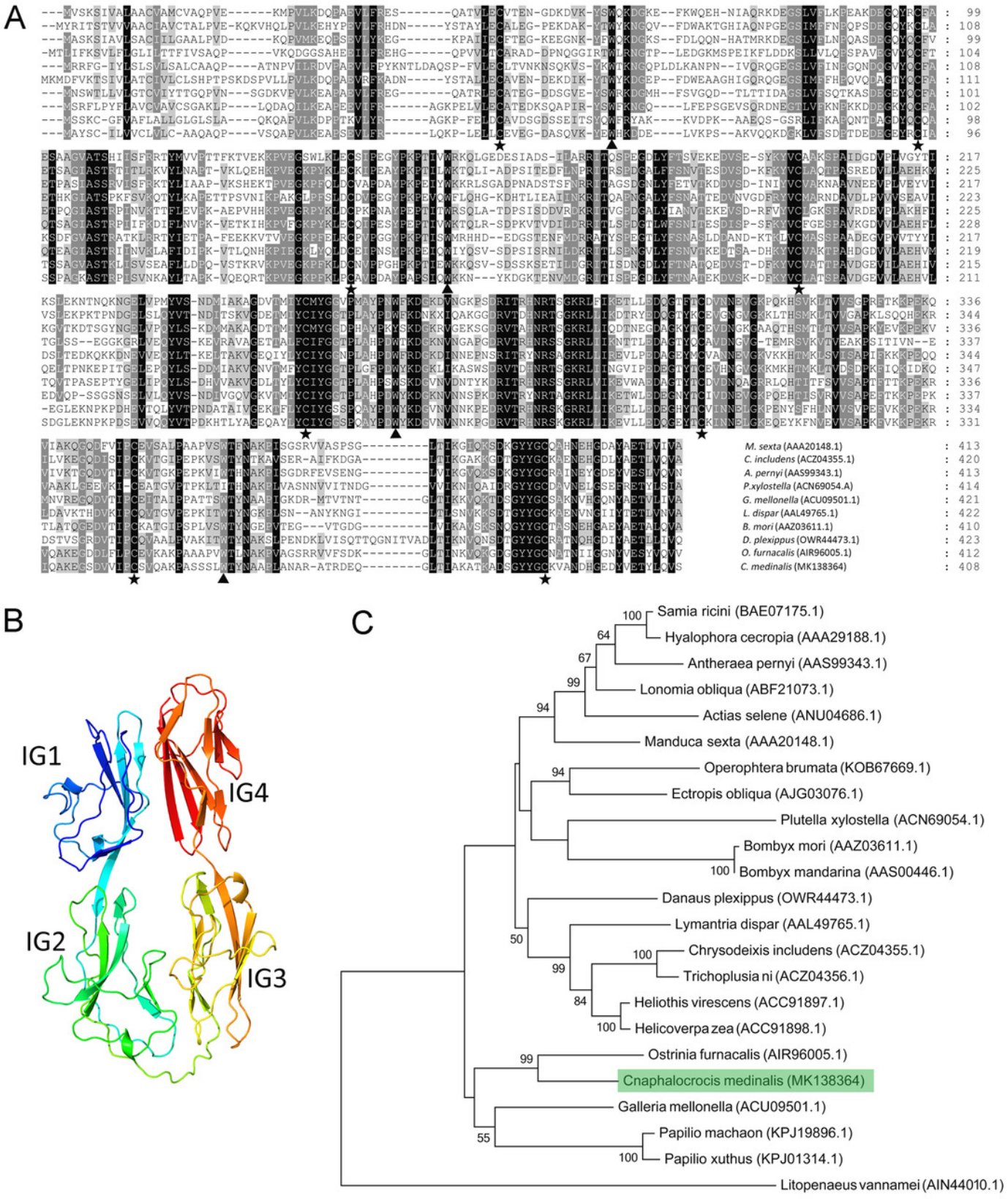
A
M. sexta (AAA20148.1)
C. includens (ACZ04355.1)
A. pernyi (AAS99343.1)
P.xylostella (ACN69054.A)
G. mellonella (ACU09501.1)
L. dispar (AAL49765.1)
B. mori (AAZ03611.1)
D. plexippus (OWR44473.1)
O. furnacalis (AIR96005.1)
C. medinalis (MK138364)
B
IG1
IG2
IG3
IG4
C
Samia ricini (BAE07175.1)
Hyalophora cecropia (AAA29188.1)
Antheraea pernyi (AAS99343.1)
Lonomia obliqua (ABF21073.1)
Actias selene (ANU04686.1)
Manduca sexta (AAA20148.1)
Operophtera brumata (KOB67669.1)
Ectropis obliqua (AJG03076.1)
Plutella xylostella (ACN69054.1)
Bombyx mori (AAZ03611.1)
Bombyx mandarina (AAS00446.1)
Danaus plexippus (OWR44473.1)
Lymantria dispar (AAL49765.1)
Chrysodeixis includens (ACZ04355.1)
Trichoplusia ni (ACZ04356.1)
Heliothis virescens (ACC91897.1)
Helicoverpa zea (ACC91898.1)
Ostrinia furnacalis (AIR96005.1)
Cnaphalocrocis medinalis (MK138364)
Galleria mellonella (ACU09501.1)
Papilio machaon (KPJ19896.1)
Papilio xuthus (KPJ01314.1)
Litopenaeus vannamei (AIN44010.1)

Figure 4

Fig. 4. The relative expression levels of *CmHem* at different developmental stages of *C. medinalis*.

1st, 2nd, 3rd, 4th and 5th represent different instar larvae. The values are represented as mean ± SD. Different small letters above histograms indicated significant differences among different treatments at the 0.05 level (ANOVA).

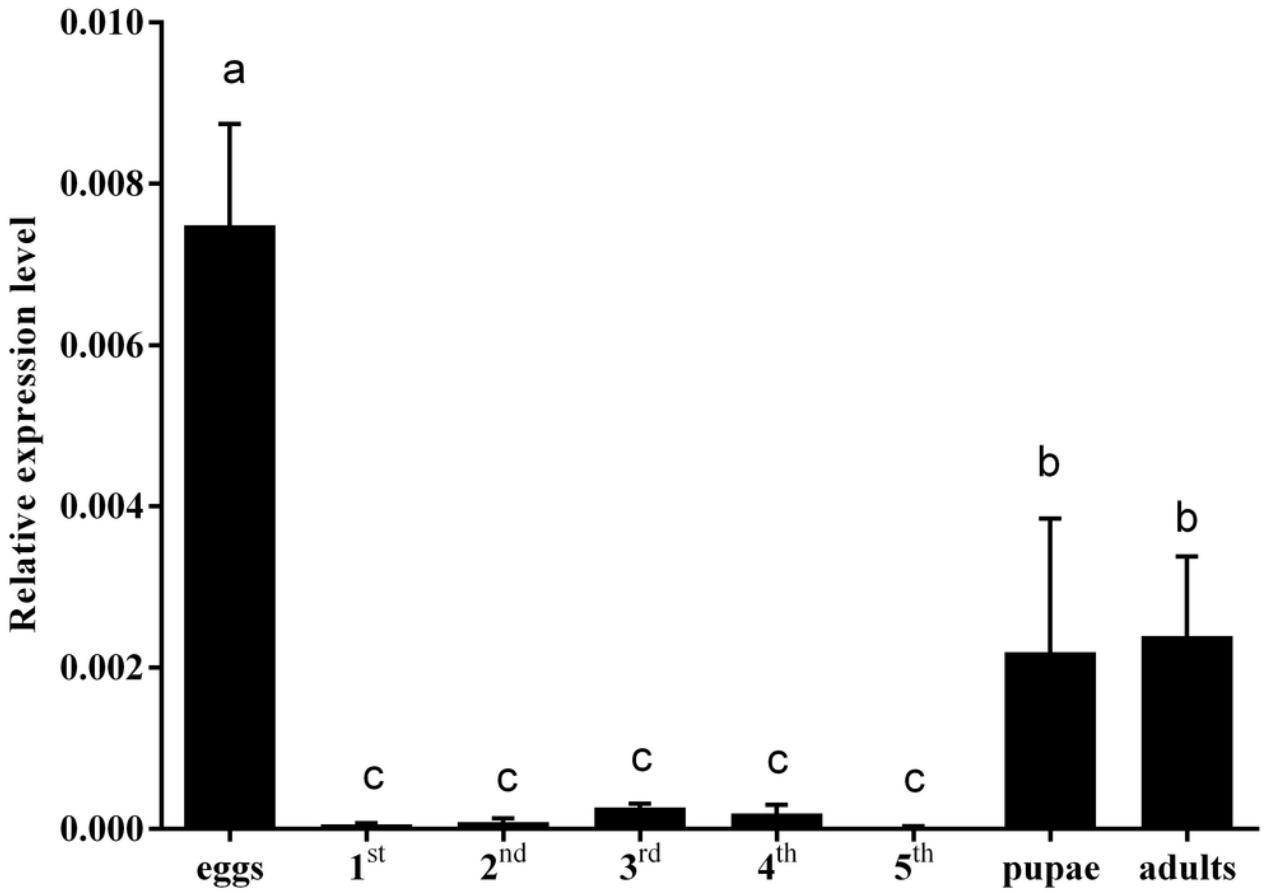

Figure 5

Fig. 5. The phenotype of mature larvae after *CmHem* interference at 72 hours (A) and the relative expression levels of *CmHem* after 48 hours of interference (B).

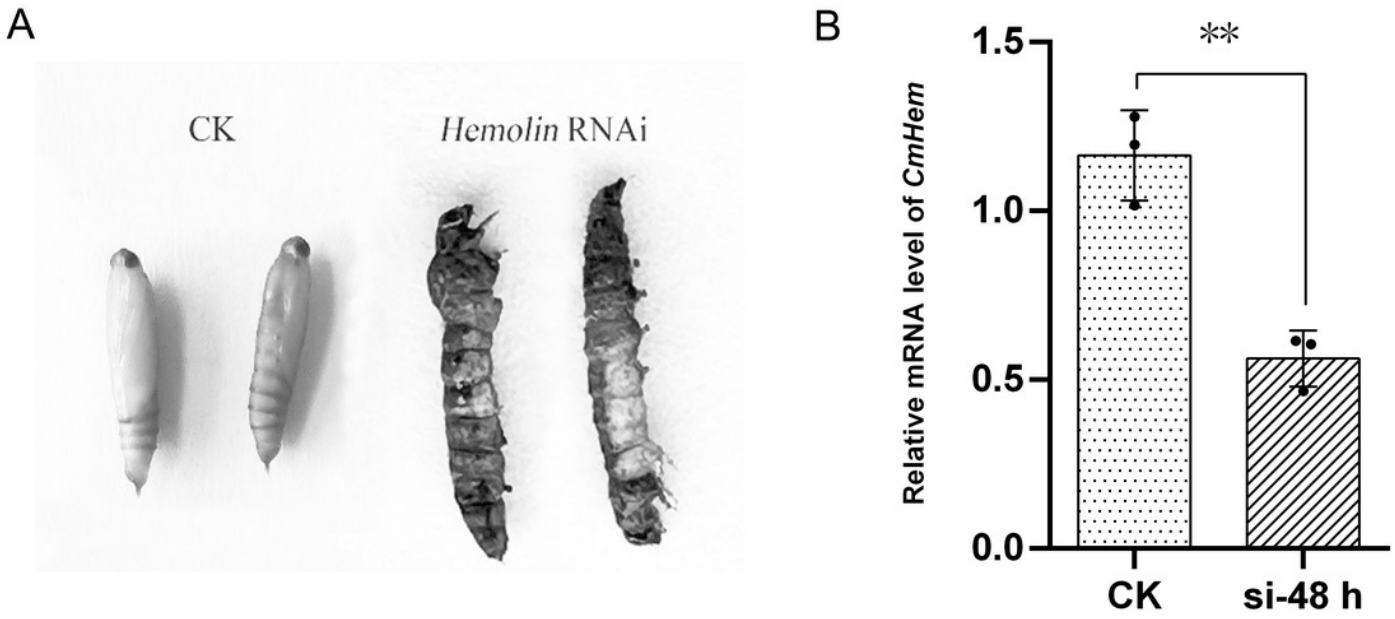

Figure 6

Fig. 6. The relative expression levels of *CmHem* in midguts (A) and fat bodies (B) of *C. medinalis* after CnmeGV infection.

Samples were collected after 12, 24, 48 and 96 hours of infection for differential expression analysis of *CmHem* (ANOVA, $p<0.0001$).

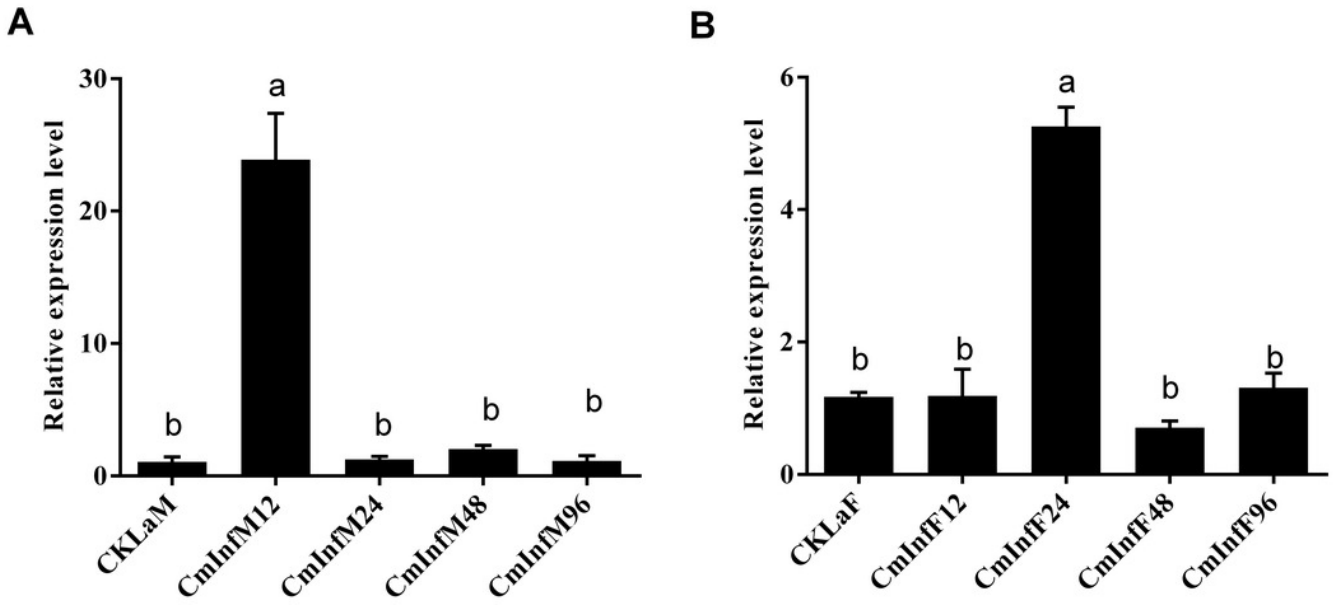

Table 1(on next page)

Table 1 The *hemolin*-like unigenes in the transcriptome of *C. medinalis*

1 **Table 1 The *hemolin*-like unigenes in the transcriptome of *C. medinalis***

geneID	Length (bp)	Identity %	Evalue	Score	NR_Description
TRINITY_DN306890_c0_g1_i1	293	57.8	1.10E-20	107.8	hemolin [Ostrinia furnacalis]
TRINITY_DN403619_c12_g2_i1	324	73.6	4.80E-41	175.6	hemolin [Ostrinia furnacalis]
TRINITY_DN403619_c13_g2_i2	363	72	7.30E-46	191.8	hemolin [Ostrinia furnacalis]
TRINITY_DN403619_c13_g5_i1	440	69.4	3.00E-54	219.9	hemolin [Ostrinia furnacalis]
TRINITY_DN403619_c13_g13_i1	318	77.1	9.50E-42	177.9	hemolin [Ostrinia furnacalis]
TRINITY_DN403619_c13_g14_i1	676	50.5	6.40E-19	103.2	hemolin [Ostrinia furnacalis]
TRINITY_DN403619_c13_g14_i2	486	52.9	1.70E-18	101.3	hemolin [Ostrinia furnacalis]
TRINITY_DN403619_c13_g40_i1	316	46.8	8.10E-17	95.1	hemolin [Danaus plexippus]
TRINITY_DN403619_c13_g46_i1	286	51.6	8.30E-21	108.2	hemolin [Ostrinia furnacalis]
TRINITY_DN403619_c13_g53_i1	447	50	5.10E-33	149.4	hemolin [Ostrinia furnacalis]
TRINITY_DN403619_c13_g69_i1	310	65.3	3.50E-33	149.4	hemolin [Ostrinia furnacalis]
TRINITY_DN487348_c0_g1_i1	302	56.4	5.70E-04	52.4	hemolin [Samia ricini]
TRINITY_DN240507_c0_g1_i1	216	72.1	7.20E-17	94.7	hemolin [Ostrinia furnacalis]

2
3
4