Gene 617 (2017) 1–7

Contents lists available at ScienceDirect

Gene

journal homepage: www.elsevier.com/locate/gene

Research paper

Functional characterization of the aspartic proteinase cathepsin D in the beet armyworm (*Spodoptera exigua*)

Tinghao Kang [a], Ruoheng Jin [a], Yunhua Zhang [a], Hu Wan [a], Kwang Sik Lee [b], Byung Rae Jin [a,b,*], Jianhong Li [a,*]

[a] Hubei Insect Resources Utilization and Sustainable Pest Management Key Laboratory, College of Plant Science and Technology, Huazhong Agricultural University, Wuhan 430070, PR China
[b] College of Natural Resources and Life Science, Dong-A University, Busan 604-714, Republic of Korea

ARTICLE INFO

Article history:
Received 12 February 2017
Received in revised form 21 March 2017
Accepted 24 March 2017
Available online 25 March 2017

Keywords:
Aspartic proteinase
Spodoptera exigua
Beet armyworm
Metamorphosis
Cathepsin D
RNAi

ABSTRACT

In insects, proteolytic enzymes are involved in food digestion and the metamorphosis process. In the present study, the full-length cDNA of an aspartic proteinase, *Spodoptera exigua* cathepsin D (SeCatD), was cloned, and its functions in metamorphosis were characterized. SeCatD contains an open reading frame of 1152 nucleotides, encoding a 384-amino acid polypeptide including a signal peptide and two functional domains (family A1 propeptide of amino acids (19–45) and a cathepsin D-like domain of 327 amino acids (55–381)). Three-dimensional structure analysis indicated that Asp66 and Asp251 may play important role in hydrolysis. Recombinant SeCatD was expressed in Sf9 insect cells and verified via SDS-PAGE and Western blot, the molecular mass of the expressed SeCatD was approximately 42 kDa. The enzyme had an optimal pH value of 3 for activity. In addition, the tissue expression profile of SeCatD during metamorphosis was obtained, and the data demonstrated that SeCatD was expressed increasingly in the fat body and midgut, but not in the epidermis. Finally, injection of dsRNA-SeCatD into the fifth-instar larvae significantly reduced SeCatD expression and larvae survival rate compared to a dsRNA-GFP treatment. These data imply that SeCatD may function during metamorphosis and may represent a target for insect control.

© 2017 Elsevier B.V. All rights reserved.

1. Introduction

Endoproteinases are normally characterized according to their origin, such as papain, pepsin or trypsin, and categorized based on the amino acid residues or metal ions involved in the catalytic sites, including serine (EC 3.4.21), cysteine (EC 3.4.22) (or thiol), aspartyl (EC 3.4.23) (or carboxyl) and metalloproteinases (EC 3.4.24) (Blanco-Labra et al., 1996; Wilhite et al., 2000). In the aspartic peptidases, a pair of aspartic residues bind to the target protein and act together to activate the catalytic water molecule. In some aspartic peptidases, the second aspartic residue can be replaced by other amino acids, such as histidine, glutamic acid or asparagine (Barrett et al., 2004). Cathepsin D is a member of the A1 family of aspartic peptidases, and is identified as a lysosomal enzyme in majority insect species, but only classified as a digestive enzyme in Cyclorrhapha Diptera (Padilha et al., 2009).

The first investigation focused on the low pH level proteolytic enzymes revealed that aspartic enzymes play essential roles in digestion in *Musca domestica* midgut (Greenberg and Paretsky, 1955). Cathepsin D (CatD, EC 3.4.23.5) was firstly identified in *M. domestica* (Lemos and Terra, 1991). Further study indicates that cathepsins are related to a number of serious pathologies, such as neurodegenerative diseases, rheumatoid arthritis, atherosclerosis and cancer, that result from cathepsin deficiency and lead to lysosome function disorders (Kinser and Dolph, 2012; Kuronen et al., 2009; Myllykangas et al., 2005). In the beginning of insect life, CatD participates in yolk mobilization, in which it is responsible for proteolytic cleavage of yolk proteins that are working in synchrony during yolk granule acidification to supply the nutrients for embryogenesis (Abreu et al., 2004). In the growth stage, insects face challenges from plants, which generate a series of proteinase inhibitors (PIs) that interfere with the digestive process of exogenous proteins to protect themselves (Wilhite et al., 2000). Unfortunately, the resistance appears to be co-evolved, which implies that insects have a sophisticated system to digest and assimilate exogenous substances. A previous study showed that PIs trigger CatD overexpression as a defense against PIs (Ahn and Zhu-Salzman, 2009). During the developmental stages, cathepsins have an essential function in metamorphosis. Investigations in *Bombyx mori* reveal that the transcription and translation levels of BmCatB and BmCatD are obviously high in the developmental stages, which implies that cathepsins might be involved in organ degeneration (Gui et al., 2006; Lee et al., 2009).

Abbreviations: CatD, cathepsin D; SeCatD, *Spodoptera exigua* cathepsin D; PIs, proteinase inhibitors; AcNPV, *Autographa californica* nucleopolyhedrovirus; ORF, open reading frame.

* Corresponding authors.
E-mail addresses: brjin@dau.ac.kr (B.R. Jin), jianhl@mail.hzau.edu.cn (J. Li).

http://dx.doi.org/10.1016/j.gene.2017.03.035
0378-1119/© 2017 Elsevier B.V. All rights reserved.

Here, we focused on SeCatD, which might play essential roles during metamorphosis. Firstly, we cloned the SeCatD gene from the cDNA library of *S. exigua* and performed bioinformatics analysis according to the cDNA sequence as well as the deduced amino acid sequence. We expressed and purified recombinant SeCatD from the engineered baculovirus infected Sf9 insect cells. Furthermore, the hydrolyzation activity of recombinant SeCatD was determined with hemoglobin as the substrate. Finally, the expression profile of the SeCatD was measured via Northern blot, and the dsRNA-SeCatD was injected into the 5th instar larvae to gain further insight into SeCatD functions during the metamorphosis.

2. Materials and methods

2.1. Animals

Beet armyworm (*S. exigua*) was reared at 25 ± 1 °C with a relative humidity of 60 ± 5% under a 16-h light/8-h dark photoperiod via an artificial diet, as previously described (Wan et al., 2012).

2.2. cDNA library screening and bioinformatic analysis

Total RNA was isolated from the whole bodies of *S. exigua* larvae using TRIzol reagent (Invitrogen, Carlsbad, CA, USA), and a cDNA library was constructed based on this RNA. The SeCatD was obtained via sequencing the cDNA library using an ABI 310 automated DNA sequencer (Perkin-Elmer Applied Biosystems, Foster City, CA, USA) and identified by BLAST software against the expressed sequence tags database from NCBI (http://www.ncbi.nlm.nih.gov/BLAST). The plasmid containing this gene was isolated using the Wizard Mini-Prep Kit (Promega, Madison, WI, USA) and confirmed via sequencing. Subsequently, the amino acid sequence was deduced and compared using the BLAST software from NCBI for searching similar sequences and identifying the functional domains using the auxiliary module Conserved Domains tool (http://www.ncbi.nlm.nih.gov/cdd/). Meanwhile, the signal peptide of the protein was identified with the SignalP 4.1 Server (http://www.cbs.dtu.dk/services/SignalP/), and relevant parameters (molecular mass and isoelectric point) were predicted using the Compute pI/Mw tool (http://web.expasy.org/compute_pi/).

Based on the BLAST results, the amino acid sequences with higher similarity and query cover were selected and aligned with the SeCatD amino acid sequence using MacVector (ver. 6.5, Oxford Molecular Ltd.). The phylogenetic analysis was performed via the MEGA 7.0 using the Maximum Likelihood algorithm by Bootstrap analysis (cut off 50) based on 1000 replicates (Kumar et al., 2016). For further insight into the catalytic mechanism of the SeCatD, the 3-D structure was constructed via SWISS-MODEL (http://swissmodel.expasy.org/) (Arnold et al., 2006).

2.3. Recombinant SeCatD expression and purification

The Bac-to-Bac® Baculovirus Expression System (Invitrogen) including the *Autographa californica* nucleopolyhedrovirus (AcNPV) and the *Spodoptera frugiperda* (Sf9) insect cell line were used to produce a recombinant virus expressing recombinant SeCatD. The PCR amplification was performed to obtain the SeCatD sequence from *pBluescript-SeGSTo* using the genetically engineered forward primer 5′-GAATTCATGGGGAAATTACTGTTA-3′ and the reverse primer 5′-AAGCTTTTAATGATGATGATGATGATGTCAGACGGCGCGCGCGAA-3′. To facilitate the purification of the recombinant protein, the reverse primer was designed with an additional His-tag. The PCR cycling conditions were as follows: 94 °C for 3 min, 30 cycles of amplification (94 °C for 30 s, 58 °C for 60 s, and 72 °C for 90 s), and 72 °C for 5 min. The PCR product was purified and inserted into the pGEM-T vector (Promega), before being inserted into the expression vector pFastBac1. Subsequently, the engineered plasmids were isolated via the Wizard mini kit (Promega) and confirmed using the ABI 310 automated DNA Sequencer. Following the Bac-to-Bac Baculovirus Expression System (Invitrogen) protocol, the recombinant bacmid DNA with the SeCatD sequence was obtained via homologous recombination in MAX efficiency DH10Bac cells (containing bacmid and helper). The Cellfectin II reagent (Invitrogen) was employed as a transfection enhancer to increase the efficiency of recombinant bacmid DNA transfected into the Sf9 insect cells. TC100 medium (Gibco BRL) supplemented with 10% fetal bovine serum (FBS, Gibco BRL) was used to culture the transfected cells at 27 °C. After 7 days, the Sf9 insect cells were further propagated in TC100 medium at 27 °C. The recombinant proteins were purified using the MagneHis ™ Protein Purification System (Promega). The protein concentration and molecular weight were determined using the Bio-Rad Protein Assay Kit and 12% SDS-polyacrylamide gel electrophoresis (SDS-PAGE), respectively.

2.4. In vitro relative enzymatic activity assays

The proteolytic activity of SeCatD was estimated based on its ability to hydrolyze hemoglobin via a protocol described previously, with slight modifications (Ahn and Zhu-Salzman, 2009). In brief, the hemoglobin served as a substrate that was denatured in 100 mM formic acid (pH 3.5) at a concentration of 2.5%. The recombinant SeCatD 20 μg in 100 μL was added to 50 μL of denatured hemoglobin. Subsequently, 850 μL of reaction buffer (Britton-Robinson Buffer Solution pH 2, 3, 4, 5, 6, 7 and 8) was mixed with the reaction system, respectively, and incubated in a water bath at 37 °C for 30 min. Next, 500 μL of 10% trichloroacetic acid was added to the reaction system and incubated at 37 °C for 10 min to precipitate the residual substrate. The sedimentation was removed by centrifugation at 15,000 × *g* for 10 min. The supernatant was transferred to a cuvette and absorbance was measured using an ultraviolet spectrophotometer (SHIMADZU UV-1800) at 280 nm. The relative enzyme activity was calculated by comparison each enzyme activity to the highest enzyme activity.

2.5. Collection of tissues, RNA isolation and qRT-PCR

During the developmental stages (from the second day of 4th instar larva to prepupa), different tissues (epidermis, fat body and midgut) were collected from healthy *S. exigua* by dissection under a stereomicroscope (Zeiss, Jena, Germany) with a cooled platform. The tissues were washed with phosphate-buffered saline (PBS; 140 mM NaCl, 27 mM KCl, 8 mM Na_2HPO_4, and 1.5 mM KH_2PO_4, pH 7.4) and fast frozen in liquid nitrogen before storage at −80 °C for further RNA extraction. TRIzol reagent was (Invitrogen, Carlsbad, CA, USA) used to purified total RNA. According to the manufacturer's protocol, no >0.2 g tissue was homogenized with 1 mL TRIzol for obtaining high quality RNA from different tissues (3 biological replicates). The RNA (2 μg) was employed to synthesize cDNA using the RevertAid First Strand cDNA Synthesis Kit (ThermoFisher, Waltham, MA, USA). After that, qRT-PCR was conducted with the first-strand cDNA, and PCR cycling conditions were as follows: initial denaturation at 95 °C for 30 s, followed by 40 cycles of 95 °C for 5 s and 60 °C for 10 s. After all reactions, a melting curve analysis was conducted from 55 to 95 °C. Each amplification reaction was performed in a 10-μL volume with 5 μL of SoFast™ EvaGreen® Supermix (Bio-Rad, Hercules, CA, USA) and 100 nM of each primer in an iQ2 Optical System (Bio-Rad). The *SeRPL*10 (Zhu et al., 2014) (forward primer: 5′-GGCTACGGTCGACGACTTCCC-3′, reverse primer: 5′-GCAGCCTCAT GCGGATGTGGAAC-3′) was employed as the reference to normalize the expression level of SeCatD (forward primer: 5′-TGAATGCCA GTGTATCAGTA-3′, reverse primer: 5′-GAGACAGCAGTGACCTAG-3′) in different treatments, RNAi efficiency was calculated based on the $2^{-\Delta\Delta CT}$ method (Livak and Schmittgen, 2001).

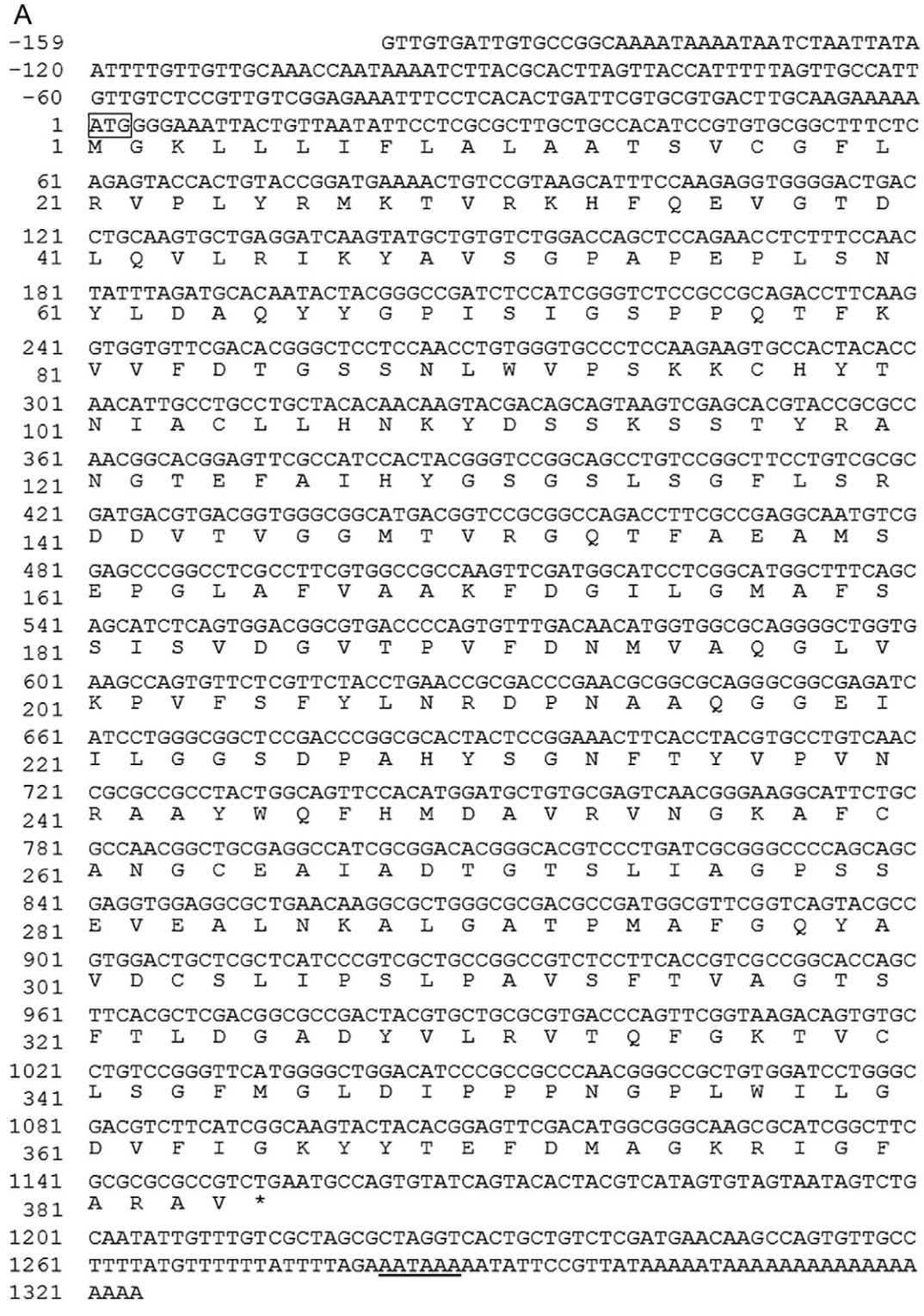

Fig. 1. Cloning of SeCatD cDNA. (A) The nucleotide and predicted amino acid sequences of the SeCatD cDNA and SeCatD protein, respectively. The start codon (ATG) is indicated with a box, and the termination codon is indicated with an asterisk. (B) Amino acid sequence alignment of SeCatD with other known CatDs. Identical domains are shown in solid boxes; the cleaved site of the signal peptide is indicated via an arrowhead; the cleaved site of the family A1 propeptide is indicated via an inverted triangle; the identical cysteines are indicated via a circle and other identical residues are identical via a rhombus. The proline loop is indicated by a black box. The sources for the sequences used in the alignment are as follows: SeCatD (this study, GenBank accession no: KX827245), *Bombyx mori* cathepsin D precursor (BmCatD, GenBank accession no: NP_001037351), *Callosobruchus maculatus* putative gut cathepsin D-like aspartic protease (CmCatD, GenBank accession no: ACO56332), *Aedes aegypti* aspartic protease (mLAP, GenBank accession no: AAA29350), *Sitophilus zeamais* lysosomal aspartic protease-like precursor (ppCAD1, GenBank accession no: BAH24176), *Musca domestica* aspartic proteinase (SAP1, GenBank accession no: NP_001273807). (C) The phylogenetic tree of known lysosomal and digestive cathepsin Ds via MEGA 7.0 using the Maximum Likelihood algorithm by Bootstrap analysis (cut off 50) based on 1000 replicates. The branch of SeCatD is marketed with a green oval. The SeCatD is classified into the lysosomal group. (D) The predicted tertiary structure of SeCatD. The structure was built using homology modeling in SWISS-MODEL. Asp66 and Asp251 are predicted catalytic active site residues.

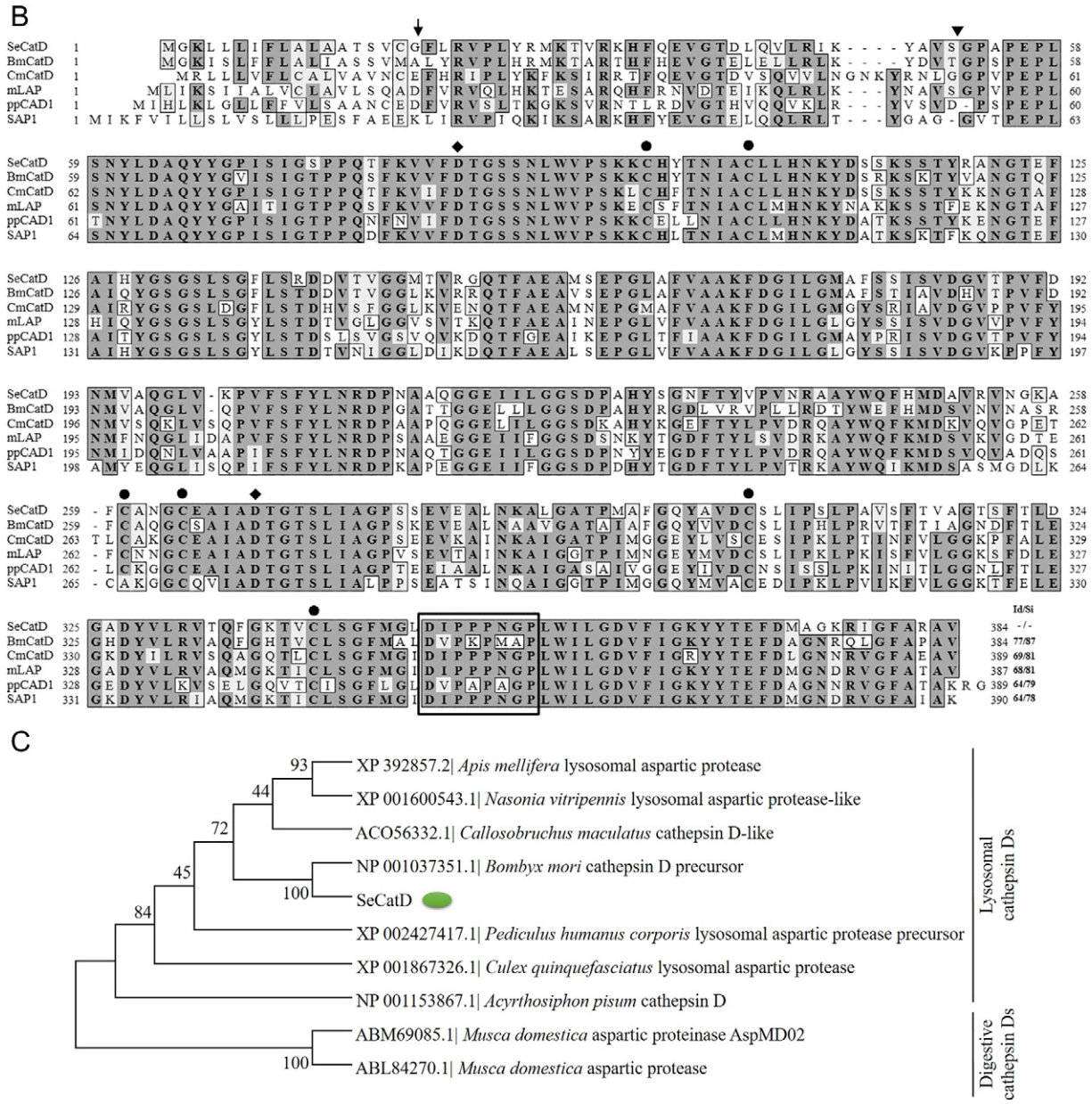

Fig. 1 (*continued*).

2.6. Effects of RNAi

The dsRNA of SeCatD and GFP were synthesized and purified via the T7 RiboMAX™ Express RNAi System (Promega). Corresponding to nucleotides of the SeCatD, primers were designed with a T7 promoter site, forward primer: 5′-GGATCCTAATACGACTCACTATAGGGCCTACTGGCAGTTCCACAT-3′, reverse primer: 5′-GGATCCTAATACGACTCACTATAGGGAGACAGCAGTGACCTAGCG-3′. GFP primers were also designed with a T7 promoter site, forward primer: 5′-GGATCCTAATACGACTCACTATAGGCAGTGCTTCAGCCGCTAC-3′, reverse primer: 5′-GGATCCTAATACGACTCACTATAGGGTTCACCTTGATGCCGTTC-3′. The concentrations of dsRNA were determined by an ultraviolet spectrophotometer (SHIMADZU UV-1800) at 260 nm and then diluted to 1 μg/μL with PBS. Each larva was injected with 5 μL diluted dsRNA of SeCatD or GFP at the first day of 5th instar. The whole bodies of larva were collected at 24, 48 and 72 h post-injection to calculate the RNAi efficiency. Total RNA was isolated and converted to cDNA. After that, the RNAi efficiency was measured via qRT-PCR as described above.

To investigate the functions of SeCatD in the metamorphosis process, each group (100 larvae at the first day of 5th instar) was injected with 5 μg of dsRNA-SeCatD or dsRNA-GFP with three replicates, respectively. Next, survival rates following different treatments were calculated, and images were obtained via a digital camera.

2.7. Statistical analysis

The data were expressed as the mean ± SE values from triplicate experiments and were analyzed for statistical significance using a one-way analysis of variance (ANOVA) followed by Duncan's multiple range test using the SPSS statistical software package (version 11.5; SPSS, Inc.). Differences with P-values <0.05 were considered to be statistically significant for all treatments.

Fig. 1 (*continued*).

3. Results

3.1. Cloning and bioinformatic analysis of SeCatD

In this study, the full-length sequence of SeCatD was obtained from the cDNA library, and the sequencing data indicated that SeCatD contains an open reading frame (ORF) of 1152 nucleotides encoding a 384-amino acid polypeptide (Fig. 1A). The protein parameters of the theoretical molecular mass (41,085 Da) and isoelectric point (6.99) were calculated using the Compute pI/Mw tool. Based on the results of signal peptide analysis, SeCatD contains a signal peptide and the cleavage site is between the Gly18 and Phe19 positions. Moreover, a functional domains module identified a family A1 propeptide of 27 amino acids (19–45) and a cathepsin D-like domain of 327 amino acids (55–381) (Fig. 1B). The nucleotide sequence of SeCatD has been submitted to GenBank (GenBank accession number: KX827245).

The alignment of multiple amino acid sequences shows that the SeCatD sequence is much more closely related to the cathepsin D of *Bombyx mori* BmCatD (87% protein sequence similarity) than the cathepsin D of *Callosobruchus maculatus* in the midgut CmCatD (81% protein sequence similarity). The proline loop was identified and indicated by a black box, which confirmed that SeCatD is a lysosomal enzyme (Fig. 1B). The phylogenetic analysis exhibits that SeCatD closes to the cathepsin D precursor of *Bombyx mori* and belongs to the lysosomal cathepsin D group (Fig. 1C). The tertiary structure of SeCatD was constructed based on the template of an aspartic proteinase (PDB: 3PSG) from *Sus scrofa domestica* by the SWISS-MODEL program (Hartsuck et al., 1992). The structure of SeCatD contains two hemispheres that form a valley to catalyze protein hydrolyzation and Asp66 and Asp251 may play an important role in this process based on the structural site (Fig. 1D).

3.2. Expression, purification and enzyme activity of recombinant SeCatD

To generate the recombinant SeCatD, the SeCatD ORF with a His-tag was inserted into the baculovirus transfer vector pFastBac1 under the control of the AcNPV polyhedrin promoter. The recombinant baculovirus was generated in Sf9 insect cells by transfection with bacmid DNA. Recombinant SeCatD was purified from the recombinant baculovirus-infected insect cells. The SDS-PAGE and Western blot analysis indicate that the molecular mass of recombinant SeCatD is approximately 42 kDa (Fig. 2).

To demonstrate the enzymatic characteristics of recombinant SeCatD, the optimal reaction pH value was measured, and the data show that the catalytic ability peaks at pH 3.0 and then dramatically falls to almost 0% at pH 5.0 (Fig. 3).

3.3. Expression profile of SeCatD and function analysis during metamorphosis

To illustrate the functions of SeCatD during the metamorphosis stages, the transcriptional expression profile of SeCatD was monitored via qRT-PCR. The SeCatD shows a higher expression level in the fat body when compared with the midgut. The highest signal was detected

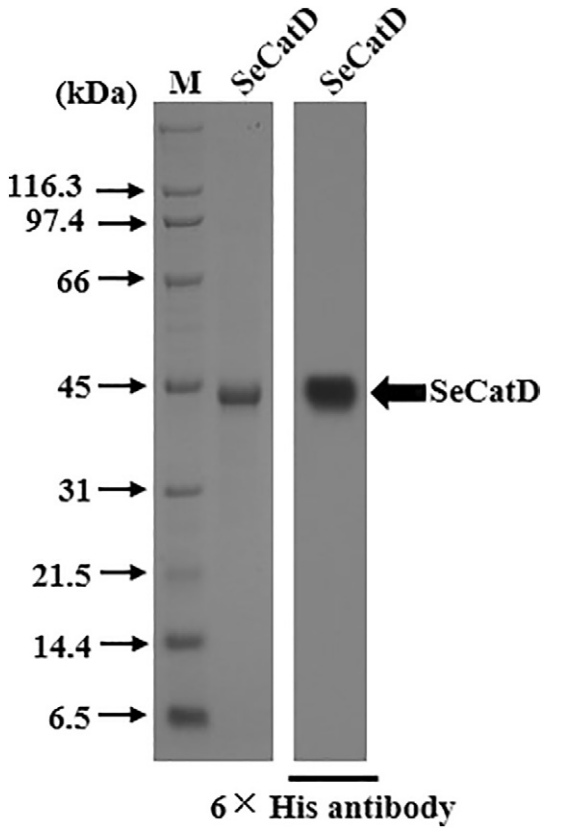

Fig. 2. Expression and purification of recombinant SeCatD. Recombinant SeCatD was expressed in baculovirus-infected Sf9 insect cells and purified using the MagneHis™ Protein Purification System. The samples were subjected to 12% SDS-PAGE; molecular weight standards were used as size markers.

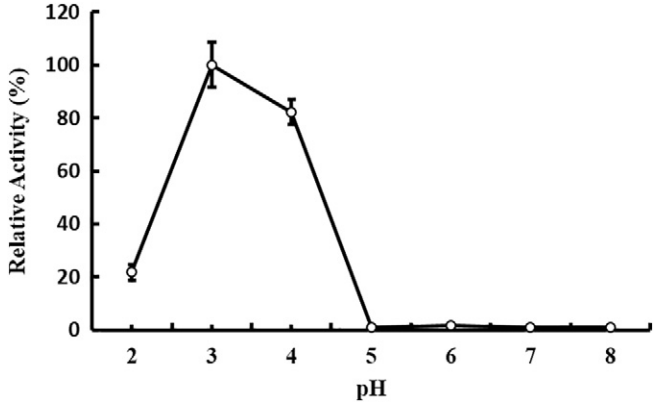

Fig. 3. Relative enzyme activity of recombinant SeCatD at different pH values. Recombinant SeCatD expressed in baculovirus-infected insect sf9 cells. The final absorbance of hydrolyzed hemoglobin was measured spectrometrically and calculated at different pH values. The error bars represent the mean ± SE values (n = 3).

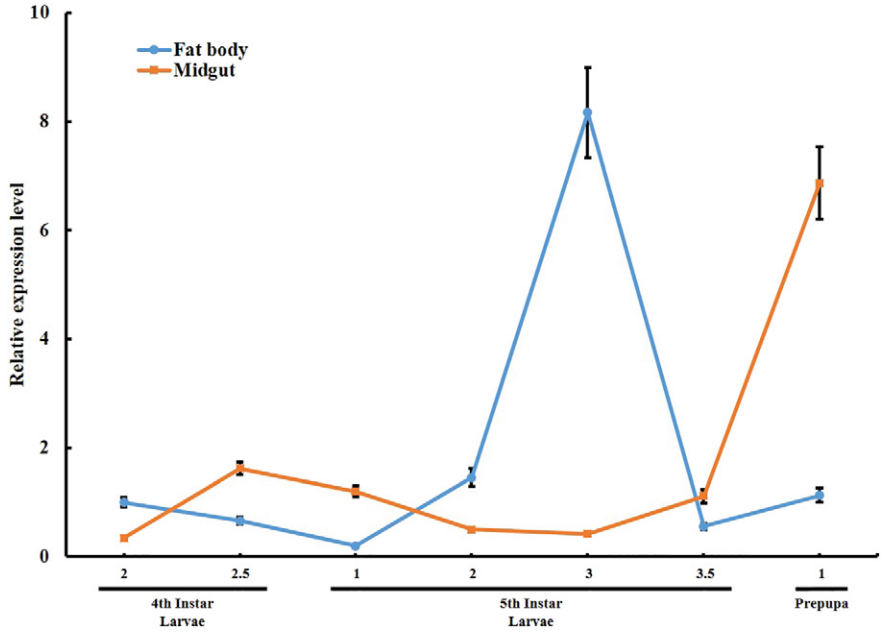

Fig. 4. SeCatD expression in *S. exigua* tissues. Total RNA was isolated from the epidermis, fat body, and midgut. The expression profile of SeCatD during *S. exigua* metamorphosis stages in the epidermis, fat body, and midgut was analyzed by qRT-PCR. The signal of the epidermis cannot be detected by qRT-PCR.

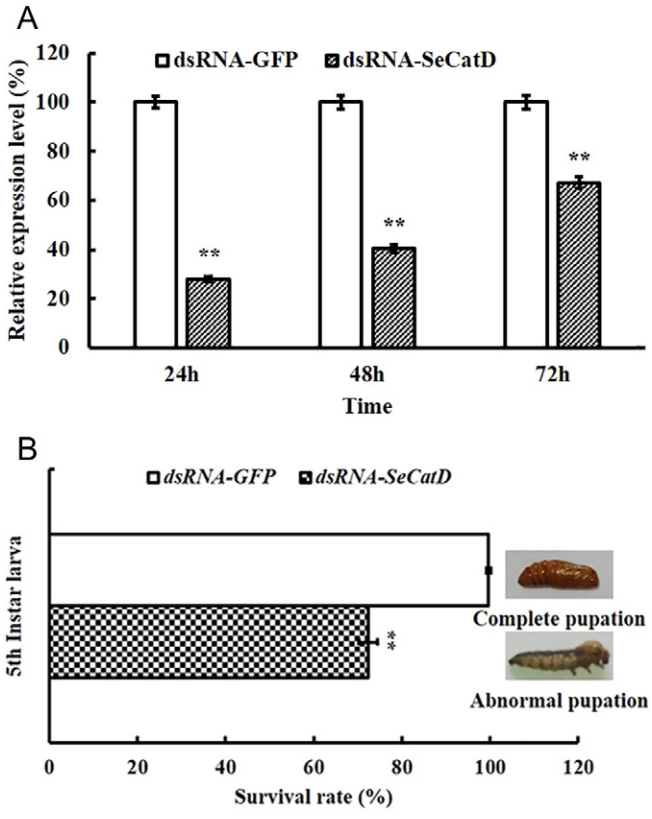

Fig. 5. Effects of SeCatD RNAi on *S. exigua* metamorphosis stages. (A) the SeCatD expression profile post dsRNA-SeCatD or dsRNA-GFP mediated treatment at 24, 48 and 72 h. The dsRNA-SeCatD was injected into l first of 5th instar larvae. The control was injected with the dsRNA-*GFP*. The total RNA was isolated from the fat body at each day post treatment. The expression profile of SeCatD was determined via qRT-PCR, and the ribosomal protein L10 gene was used as the internal control. (B) The survival rate in the dsRNA-SeCatD or dsRNA-GFP mediated treatment. The survival rates are the means of three assays. The error bars represent the mean ± SE values (n = 3). Significant differences are indicated with one asterisk at $P < 0.05$ and with two asterisks at $P < 0.01$.

at the third day of 5th instar larva in the fat body (Fig. 4). The signal could not be detected in the epidermis at any time point (data not show).

To further unwrap the functions of SeCatD, dsRNA-SeCatD was synthesized and injected into the fat body of larvae and the silent efficiency was determined. As shown in Fig. 5A, the expression level dropped substantially at 24 h post-injection (approximately 30%) and increased to almost 60% at 72 h post injection (Fig. 5A). Measuring the survival rate after dsRNA-SeCatD injection was conducted to exhibit the essential functions of SeCatD during the metamorphosis stages. The results show that the survival rate of 5th instar decreased to 75% after injection (Fig. 5B).

4. Discussion

In the current investigation, a novel gene encoding SeCatD was firstly identified and reported from the beet armyworm *S. exigua*. The ORF of SeCatD contains 1152 nucleotides encoding a 384-amino acid polypeptide. Furthermore, the amino acid sequence alignment analysis indicates that SeCatD has high similarity to the cathepsin D of *Bombyx mori* BmCatD (87% protein sequence similarity) and both of the two CatDs have been predicted to contain signal peptides and N-glycosylation sites (Gui et al., 2006). Moreover, two conserved domains were identified including a family A1 propeptide of 27 amino acids (19–45) and a cathepsin D-like domain of 327 amino acids (55–381). Furthermore, the phylogenetic analysis indicated that SeCatD is a lysosomal cathepsin D. The 3-D structure analysis illustrated that the two amino acids Asp66 and Asp251, which are located at the either side of the binding area, might play an essential role in the catalytic reaction. From the above, our primary study of SeCatD (based on the cDNA and amino acids sequence) strongly suggests that SeCatD is a member of the insect CatD family. These characteristics indicate that SeCatD is structurally and functionally similar to the other lysosomal CatDs.

To further investigate the protease characteristics of SeCatD in vitro, recombinant SeCatD was expressed via the infected sf9 cell line and purified through Ni particles. Subsequently, the molecular mass was determined by SDS-PAGE and Western blot, and the results demonstrate that

the recombinant SeCatD possesses a similar molecular mass to the existent CatDs, approximately 45 kDa (Balczun et al., 2012).

The cathepsin D-like aspartic proteinase is considered an important protease that is involved in digesting foreign proteins and is not sensitive to the cysteine proteinase inhibitors from plants, which indicates that it is an essential alternative proteinase that counters the plant self-protection strategy (Matsumoto et al., 2009). To measure the catalytic activity of recombinant SeCatD, hemoglobin served as the substrate. In this study, the optimal reactive pH value of recombinant SeCatD was pH 3.0, which is consistent with the results of other studies in insects including *Musca domestica*, *Triatoma infestans*, and *Tribolium castaneum* (Padilha et al., 2009; Balczun et al., 2012; Blanco-Labra et al., 1996).

The expression levels of CatDs have been well investigated in various types of insects, and the recent results show that three CatDs from *Musca domestica* are steadily expressed in the midgut (Padilha et al., 2009). Other research in *Bombyx mori* indicates that expression can be detected in the fat body during the metamorphosis stages (Lee et al., 2009). Unfortunately, the evidence for the expression of CatDs in the epidermis has not yet been found. Our results exhibit similar tendencies as the expression profile, which suggested that SeCatD is expressed the highest in the fat body at the third day of 5th instar larva and in midgut during the pre-pupal period. Consistent with previous studies, the expression has not been detected in the epidermis.

For further investigation on the functions of CatDs during the metamorphosis stages, previous studies down-regulated the expression level of CatDs and monitored the pupation rate of larvae. The results show that the non-pupation rate increased with treatment duration (Gui et al., 2006). The SeCatD dsRNA was designed to decrease the expression level of the gene and evaluate the survival rate of 5th instar larvae in our study. The qRT-PCR results confirmed that our dsRNA-SeCatD was able to significantly reduce the expression level of the target gene, and the survival rate was substantially lower than the dsRNA-GFP treatment. These data imply that SeCatD might possess essential functions during metamorphosis stages. The down-regulation of SeCatD transcription might affect organ degeneration and regeneration, which are essential processes during the pre-pupal and pupal stages.

In conclusion, the results of our study demonstrate that SeCatD is a member of the CatD family, and this is the first time a CatD has been identified in beet armyworm. The bioinformatics analysis provides a molecular basis for understanding the catalytic mechanisms of aspartic proteinases. The enzyme activity experiment proved that SeCatD is a functional proteinase and additional data implied that SeCatD plays important roles during metamorphosis stages, which might become the potential target site of pesticide design in the future.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (NSFC No. 31272069).

References

Abreu, L.A., Valle, D., Manso, P.P., Façanha, A.R., Pelajo-Machado, M., Masuda, H., Masuda, A., Vaz Jr., I., Lenzi, H., Oliveira, P.L., Logullo, C., 2004. Proteolytic activity of *Boophilus microplus* yolk pro-cathepsin D (BYC) is coincident with cortical acidification during embryogenesis. Insect Biochem. Mol. Biol. 34, 443–449.

Ahn, J.E., Zhu-Salzman, K., 2009. CmCatD, a cathepsin D-like protease has a potential role in insect defense against a phytocystatin. J. Insect Physiol. 55, 678–685.

Arnold, K., Bordoli, L., Kopp, J., Schwede, T., 2006. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. Bioinformatics 22, 195–201.

Balczun, C., Siemanowski, J., Pausch, J.K., Helling, S., Marcus, K., Stephan, C., Meyer, H.E., Schneider, T., Cizmowski, C., Oldenburg, M., Höhn, S., Meiser, C.K., Schuhmann, W., Schaub, G.A., 2012. Intestinal aspartate proteases TiCatD and TiCatD2 of the haematophagous bug *Triatoma infestans* (Reduviidae): sequence characterisation, expression pattern and characterisation of proteolytic activity. Insect Biochem. Mol. Biol. 42, 240–250.

Barrett, A., Rawlings, N., Woessner, J., 2004. Handbook of Proteolytic Enzymes. second ed. Elsevier Academic Press, Amsterdam.

Blanco-Labra, A., Martínez-Gallardo, N.A., Sandoval-Cardoso, L., Delano-Frier, J., 1996. Purification and characterization of a digestive cathepsin D proteinase isolated from *Tribolium castaneum* larvae (Herbst). Insect Biochem. Mol. Biol. 26, 95–100.

Greenberg, B., Paretsky, D., 1955. Proteolytic enzymes in the house fly, *Musca domestica* (L.). Ann. Entomol. Soc. Am. 48, 46–50.

Gui, Z.Z., Lee, K.S., Kim, B.Y., Choi, Y.S., Wei, Y.D., Choo, Y.M., Kang, P.D., Yoon, H.J., Kim, I., Je, Y.H., Seo, S.J., Lee, S.M., Guo, X., Sohn, H.D., Jin, B.R., 2006. Functional role of aspartic proteinase cathepsin D in insect metamorphosis. BMC Dev. Biol. http://dx.doi.org/10.1186/1471-213X-6-49.

Hartsuck, J.A., Koelsch, G., Remington, S.J., 1992. The high-resolution crystal structure of porcine pepsinogen. Proteins 13, 1–25.

Kinser, R.D., Dolph, P.J., 2012. Cathepsin proteases mediate photoreceptor cell degeneration in *Drosophila*. Neurobiol. Dis. 46, 655–662.

Kumar, S., Stecher, G., Tamura, K., 2016. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol. Biol. Evol. 33, 1870–1874.

Kuronen, M., Talvitie, M., Lehesjoki, A.E., Myllykangas, L., 2009. Genetic modifiers of degeneration in the cathepsin D deficient *Drosophila* model for neuronal ceroid lipofuscinosis. Neurobiol. Dis. 36, 488–493.

Lee, K.S., Kim, B.Y., Choo, Y.M., Yoon, H.J., Kang, P.D., Woo, S.D., Sohn, H.D., Roh, J.Y., Gui, Z.Z., Je, Y.H., Jin, B.R., 2009. Expression profile of cathepsin B in the fat body of *Bombyx mori* during metamorphosis. Comp. Biochem. Physiol. B Biochem. Mol. Biol. 154, 188–194.

Lemos, F.J.A., Terra, W.R., 1991. Properties and intracellular distribution of a cathepsin D-like proteinase active at the acid region of *Musca dometica* midgut. Insect Biochem. 21, 457–465.

Livak, K.J., Schmittgen, T.D., 2001. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. Methods 25, 402–408.

Matsumoto, I., Asakura, T., Ohmori, T., Tamura, T., Abe, K., 2009. Cathepsin D-like aspartic proteinase occurring in a maize weevil, *Sitophilus zeamais*, as a candidate digestive enzyme. Biosci. Biotechnol. Biochem. 73, 2338–2340.

Myllykangas, L., Tyynelä, J., Page-McCaw, A., Rubin, G.M., Haltia, M.J., Feany, M.B., 2005. Cathepsin D-deficient *Drosophila* recapitulate the key features of neuronal ceroid lipofuscinoses. Neurobiol. Dis. 19, 194–199.

Padilha, M.H., Pimentel, A.C., Ribeiro, A.F., Terra, W.R., 2009. Sequence and function of lysosomal and digestive cathepsin D-like proteinases of *Musca domestica* midgut. Insect Biochem. Mol. Biol. 39, 782–791.

Wan, H., Yuan, M., You, H., Li, J., Jin, B.R., 2012. Molecular cloning and characterization of a diapause-specific peptide in the beet armyworm, *Spodoptera exigua*. J. Asia Pac. Entomol. 15, 369–373.

Wilhite, S.E., Elden, T.C., Brzin, J., Smigocki, A.C., 2000. Inhibition of cysteine and aspartyl proteinases in the alfalfa weevil midgut with biochemical and plant-derived proteinase inhibitors. Insect Biochem. Mol. Biol. 30, 1181–1188.

Zhu, X., Yuan, M., Shakeel, M., Zhang, Y., Wang, S., Wang, X., Zhan, S., Kang, T., Li, J., 2014. Selection and evaluation of reference genes for expression analysis using qRT-PCR in the beet armyworm *Spodoptera exigua* (Hübner) (Lepidoptera: Noctuidae). PLoS One 9, e84730.