Genetics and Molecular Research

Characterization and functional analysis of a chitin synthase gene (*HcCS1*) identified from the freshwater pearlmussel *Hyriopsis cumingii*

H.F. Zheng[1,2], Z.Y. Bai[1], J.Y. Lin[1], G.L. Wang[1], J.L. Li[1,3]

[1]Key Laboratory of Freshwater Aquatic Genetic Resources,
Shanghai Ocean University, Ministry of Agriculture, Shanghai, China
[2]East China Sea Fisheries Research Institute, Chinese Academy of Fishery Sciences,
Key Laboratory of East China Sea and Oceanic Fishery Resources Exploitation,
Ministry of Agriculture, Shanghai, China
[3]Aquaculture Division, E-Institute of Shanghai Universities, Shanghai Ocean University,
Shanghai, China

Corresponding author: J.L. Li
E-mail: jlli@shou.edu.cn

Genet. Mol. Res. 14 (4): 19264-19274 (2015)
Received August 5, 2015
Accepted October 9, 2015
Published December 29, 2015
DOI http://dx.doi.org/10.4238/2015.December.29.36

ABSTRACT. The triangle sail mussel, *Hyriopsis cumingii*, is the most important freshwater pearl mussel in China. However, the mechanisms underlying its chitin-mediated shell and nacre formation remain largely unknown. Here, we characterized a chitin synthase (CS) gene (*HcCS1*) in *H. cumingii*, and analyzed its possible physiological function. The complete ORF sequence of *HcCS1* contained 6903 bp, encoding a 2300-amino acid protein (theoretical molecular mass = 264 kDa; isoelectric point = 6.22), and no putative signal peptide was predicted. A myosin motor head domain, a CS domain, and 12 transmembrane domains were found. The predicted spatial structures of the myosin head and CS domains were similar to the electron microscopic structure of the heavy meromyosin subfragment

of chicken smooth muscle myosin and the crystal structure of bacterial cellulose synthase, respectively. This structural similarity indicates that the functions of these two domains might be conserved. Quantitative reverse transcription PCR results showed that *HcCS1* was present in all detected tissues, with the highest expression levels detected in the mantle. The *HcCS1* transcripts in the mantle were upregulated following shell damage from 12 to 24 h post-damage, and they peaked (approximately 1.5-fold increase) at 12 h after shell damage. These findings suggest that *HcCS1* was involved in shell regeneration, and that it might participate in shell and nacre formation in this species via chitin synthesis. *HcCS1* might also dynamically regulate chitin deposition during the process of shell and nacre formation with the help of its conserved myosin head domain.

Key words: *HcCS1*; Triangle sail mussel; Sequence and structure analysis; Expression profiles

INTRODUCTION

Chitin, a polymer of *N*-acetyl-β-d-glucosamine, is the most widespread amino polysaccharide in nature, and its estimated annual production is nearly comparable to that of cellulose (Andersen et al., 1995). Chitin mainly exists in arthropod exoskeletons, fungal cell walls, and nematode eggshells (Merzendorfer and Zimoch, 2003). In insects, chitin is a major component of the arthropod cuticle, and it is an integral part of peritrophic matrices (Merzendorfer, 2006). Depending on the insect species, the chitin content constitutes over 40% of the exuvial dry mass, and it varies substantially with different cuticle types, even in a single organism (Kramer et al., 1995). Thus, chitin synthesis is essential for insect development.

The biosynthetic chitin pathway in insects was first established by Jaworski et al. (1963), and a subsequent study reported that chitin synthase was the last key enzyme in this pathway (Merzendorfer, 2006). In contrast to fungi that possess multiple genes encoding chitin synthase (CS) isoforms (Munro and Gow, 2001), insects only contain two different CS genes that encode two enzyme classes, CS class A (CSA) and CS class B (CSB), which originated from a common ancestor via different splicing events. CSAs are specifically expressed in the epidermis, and are related ectodermal cells; however, the expression of CSBs is restricted to gut epithelial cells that produce peritrophic matrices (Merzendorfer and Zimoch, 2003).

Chitin is an essential component of the mollusk shell and nacre, which are the hard tissues consisting of calcium carbonate crystals and an organic matrix that is composed of chitin, silk fibroin protein, and acidic macromolecules (Furuhashi et al., 2009). Covalent modification of chitin with silk-derivatives can create an amphiphilic interface that acts as a self-organizing template to form the periodic matrix texture during nacre biomineralization (Weiss et al., 2002, 2009). Suzuki et al. (2009) showed that chitin could recruit an acidic matrix protein, Pif, which regulates nacre formation in the pearl oyster *Pinctada fucata*.

The triangle sail mussel (*Hyriopsis cumingii*) is the most important mussel species used for commercial freshwater pearl production in China. In contrast to the seawater pearl oyster, the triangle sail mussel has a different evolutionary position, and its shell and nacre contain aragonite prisms instead of the calcite prisms found in seawater oysters (Bai et al., 2013). To date, the mechanisms of shell and nacre formation, especially chitin-mediated shell and nacre formation, in the

triangle sail mussel remain largely unknown. In this study, we cloned a complete ORF sequence of a CS gene in the triangle sail mussel (*HcCS1*), and then analyzed the characteristics of its amino acid sequence and protein structure. We then investigated its potential physiological functions in an attempt to determine whether *HcCS1* was involved in shell regeneration. These results will aid in the elucidation of the mechanisms underlying shell and nacre formation in this freshwater pearl mussel.

MATERIAL AND METHODS

Ethics statement

The handling of mussels was conducted in accordance with the care and use of animals for scientific purpose guidelines set by the Institutional Animal Care and Use Committee (IACUC) of Shanghai Ocean University (Shanghai, China).

Total RNA isolation and first-strand cDNA synthesis

Total RNA was extracted from the selected mussel tissues using the RNAiso Plus method (TaKaRa, Dalian, China), and it was then digested with DNase I (TaKaRa) following the manufacturer protocol. First-strand cDNA was synthesized with 5 µg DNA-free total RNA as the PCR template using a First-Strand cDNA Synthesis Kit (Clontech, USA), according to the manufacturer instructions. The harvested cDNA was stored for further experiments.

Cloning the *HcCS1* open-reading frame (ORF)

A partial cDNA sequence was harvested via high-throughput transcriptome sequencing using the mRNA extracted from pooled mantle tissue (Bai et al., 2013). To amplify the 5'-end of the *HcCS1* ORF sequence, one pair of primers (RaceF: 5'-TAC GGC TGC GAG AAG ACG ACA GAA G-3' and RaceR: CAC TGG ATG CCC TCA AAC TCA TAG TC) was subsequently designed and synthesized. A rapid amplification of cDNA ends (RACE) cDNA library was constructed with a Clontech Smart cDNA Amplification kit using RNA from the mantle, according to the manufacturer instructions. The PCR was conducted under the following parameters: 95°C for 5 min; 28 cycles at 94°C for 30 s, 53°C for 30 s, and 72°C for 2 min; and 72°C for 8 min. The amplified fragment was then sub-cloned into a pMD-19T vector prior to sequencing by a commercial company (Sangon, China). The complete *HcCS1* ORF sequence was obtained by overlapping these two fragments.

Bioinformatic analysis

Translation of the amino acid sequence was conducted using online software (http://web.expasy.org/translate/). The similarities of *HcCS1* and other CSs were analyzed using the online Basic Local Alignment Search Tool Program (BLASTp) (http://blast.ncbi.nlm.nih.gov/Blast.cgi). Multiple amino acid sequence alignments were generated with the GENEDOC software and the ClustalX 2.0 program (http://www.ebi.ac.uk/tools/clustalw2). Prediction of the domain topology was performed via http://smart.embl-heidelberg.de/. The theoretical molecular weight and isoelectric point were calculated using an online software source (http://web.expasy.org/compute_pi/). MEGA 4.0 was used to construct phylogenetic trees, and 1000 bootstraps were conducted to assess the reliability of the neighbor-joining tree. The 3-D structure of the CS domain was predicted through

the Phyre server (Kelley and Sternberg, 2009), using the crystal structure of bacterial cellulose synthase (PDB ID: 4hg6, chain A) as a template. The 3-D structure of the myosin head domain was modeled by homology, and it was based on the spatial structure template of the heavy meromyosin subfragment of chicken smooth muscle myosin (PDB ID: 3j04, chain A) using the SWISS-MODEL workspace (http://swissmodel.expasy.org/workspace/). Discovery Studio Visualizer 4.0 was used to generate the predicted 3-D structures of these two domains.

Animal and tissue collection

Live triangle sail mussel individuals (shells of approximately 12 cm in length) were harvested from WeiminAquaculture Farm of Wuyi (Zhejiang Province, China), and were temporarily cultured in flowing, aerated freshwater tanks for 1 week before processing. Healthy individuals were selected to investigate *HcCS1* expression profiles. To analyze temporal expression profiles after shell damage, V-shaped damage was inflicted on the shell near the adductor muscle in each of the 25 mussels in the experimental group. An additional five healthy mussels composed the control group. The mantle was harvested at 0, 3, 6, 9, 12, and 24 h post-damage by dissection, and it was then used for total RNA extraction. To investigate *HcCS1* tissue distribution, hemolymph was collected from the heart of each mussel with a syringe, and was centrifuged at 1500 rpm for 15 min at 4°C to isolate hemocytes. The supernatant was discarded, and the hemocyte pellet was collected for RNA extraction. Subsequently, the mantle, gill, liver, stomach, intestine, kidney, and foot from healthy mussels were also collected, washed with sterile PBS, and pooled from at least three mussels for subsequent total RNA extraction. Additionally, two batches of RNA samples previously extracted at different times and stored in liquid nitrogen were used to eliminate the batch difference.

Real-time PCR

Quantitative reverse transcription PCR (qRT-PCR) was conducted in a real-time thermal cycler (ABI, USA) using synthesized cDNA as the template for detecting the relative *HcCS1* expression levels, based on a previous protocol (Li et al., 2013). One primer pair (*Hc*CS1RF: TGG GTA CGA TCA TAA GCA TT and *Hc*CS1RR: CAA ACC ATA GAT TAG GCA GA) was designed to produce a 134-bp amplicon. Another primer pair (actinRF: ACG GAT AAC ACA AGG AAA GGA and actinRR: ATG GAT GGA AAC ACG GCT CT) was used to amplify the corresponding DNA fragment as a reference. The 20-μL reaction mixture consisted of 10 μL 2X Premix Ex Taq, 2 μL cDNA, and 4 μL of each primer. qRT-PCR was programmed as follows: 95°C for 5 min; 40 cycles at 95°C for 10 s and 60°C for 50 s; and a melt from 60° to 95°C. All tests were conducted three times using individual templates. The $2^{-\Delta Ct}$ method was used to calculate the relative expression level of *HcCS1* in different tissues, and the$2^{-\Delta\Delta Ct}$ algorithm was applied in expression profile analyses. Significant differences were determined by one-way analysis of variance and the Duncan test for multiple-range comparison using SPSS 13.0, with significant levels accepted at $P < 0.05$.

RESULTS

Cloning the *HcCS1* ORF

RACE-PCR analyses were used to obtain the *HcCS1* ORF sequence from the triangle

sail mussel. Based on the partial *HcCS1* cDNA sequence, the 5'-end of the ORF sequence was obtained using RACE technology with one pair of primers (RaceF and RaceR). The full-length ORF sequence was determined using the overlapping regions of these two fragments. The ORF sequence contained 6903 bp, encoding a 2300-amino acid protein (GenBank accession No. KR149328). The deduced amino acid sequence is shown in Figure 1. The inferred HcCS1 protein had a theoretical molecular mass of 264 kDa and an isoelectric point of 6.22, but no putative signal peptide was predicted. A myosin motor head domain (residues 1-733) and a CS domain (residues 1537-1802) were found in the deduced protein. Using the TMHMM program, 12 transmembrane domains were identified. A preliminary model of the HcCS1 protein was constructed based on the highlighted sequence marked in Figure 2. As shown in this model, both the myosin head domain and the CS domain were located intracellularly.

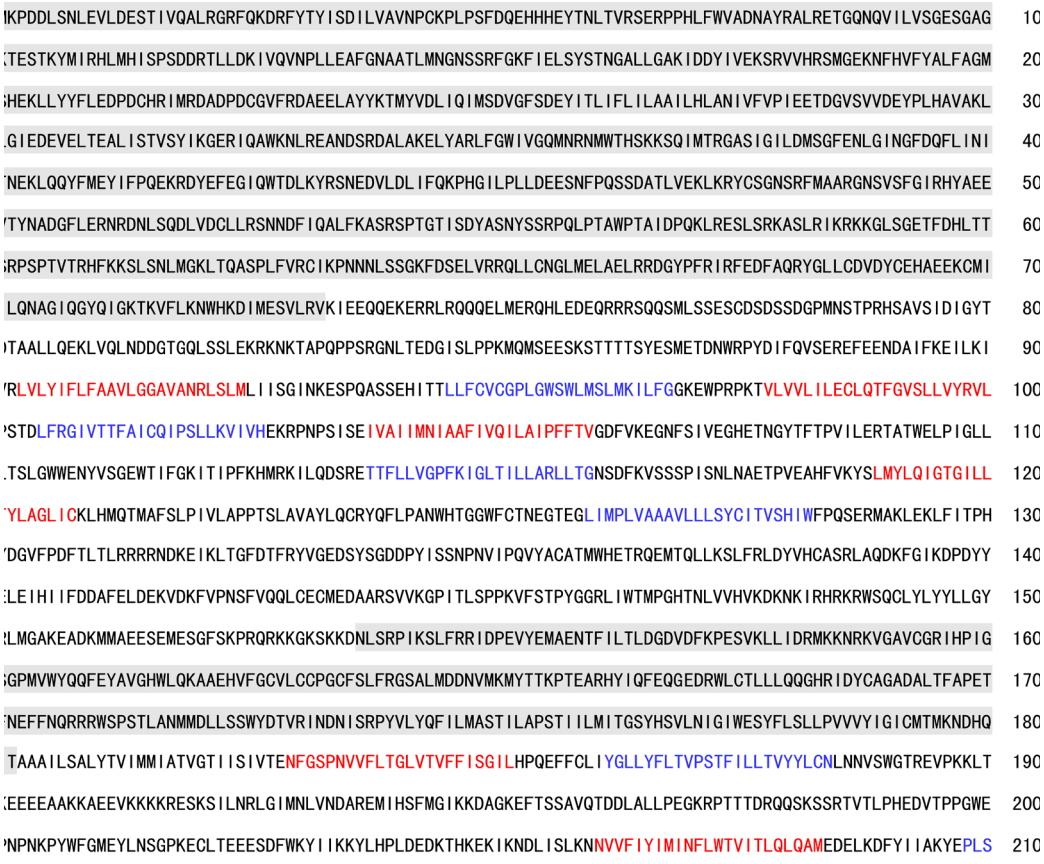

Figure 1. Translated amino acid sequences of the *HcCS1* opening read frame (ORF). Both the myosin motor domain (MYSc) at the N-terminal and the CS domain are shadowed. The stop codon is indicated by an asterisk, and the motifs of the active center (EDR and NQRRW) in the CS domain are surrounded by red frames. The transmembrane domain sequences are shown in red (in-out direction) or blue (out-in direction) letters.

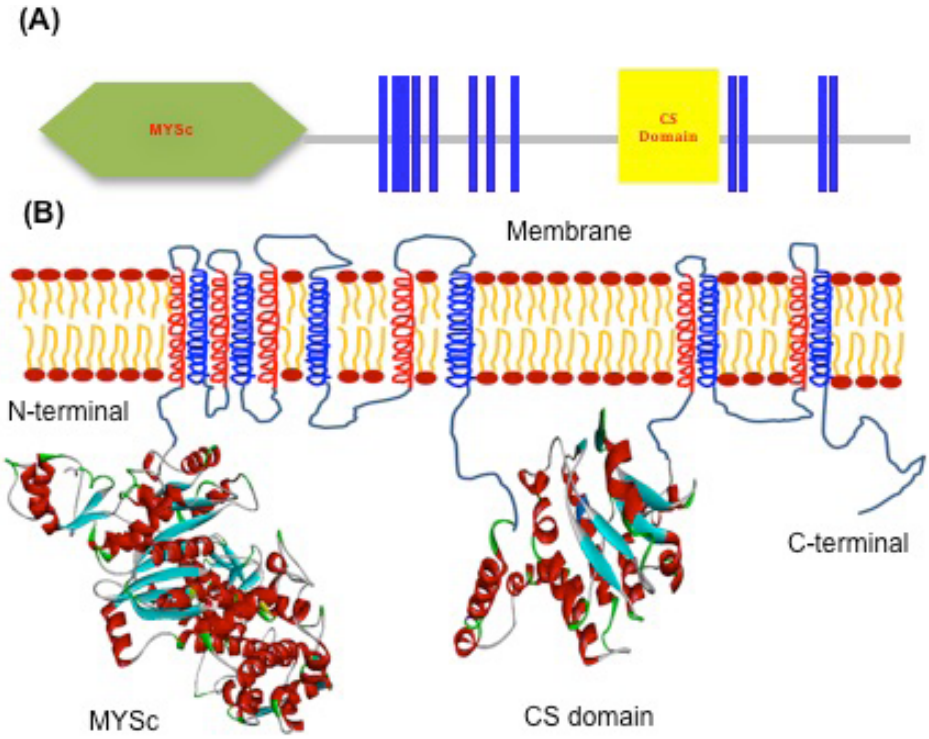

Figure 2. Diagrams of *HcCS1* structural elements. Schematic representation of the domain topology of *HcCS1* was predicted with the online SMART software. It contains a myosin head (MYSc) domain, a chitin synthase (CS) domain, and 12 transmembrane (TM) domains. **A.** The predicted 3-D structure of the MYSc domain, the predicted 3-D structure of the CS domain, and the 12 TM domains (six in-out TM and six out-in TM) in the membrane were modeled and then arranged based on the topology of the above domains (**B**).

Similarity and phylogenetic analyses

The BLASTp search analyses demonstrated that *Hc*CS1 shared the highest identity (67%) with both *P. fucata Pf*CHS1 (BAF73720) and *Atrina rigida Ar*CS1 (AAY86556), and it shared 65% identity with *Mytilus galloprovincialis Mg*CS1 (ABQ08059) and *Crassostrea gigas* CS3 (EKC25899). It shared lower identity with insect CSs, such as 36, 35, and 35% identity with *Acromyrmex echinatio*r CS (EGI66236), *Tribolium castaneum* CS1 (NP_001034491), and *Aphis glycines* CS1 (AFJ00066), respectively. Alignments of four mollusk CS homologs showed that *HcCS1* contained two conserved motifs present in three other CSs (Figure 3), suggesting that *HcCS1* might possess the same physiological function as other mollusk CS homologs.

Based on the BLASTP results and the reported CS homologs in Mollusca, arthropods, and parasites, a phylogenetic tree was constructed to show their evolutionary relationships (Figure 4). In this tree, CS homologs were grouped into two clusters. *HcCS1* together with *ArCS1*, *MgCS1*, and *CgCS3* formed a small cluster, and the other CS homologs from insects and parasites belonged to the larger cluster, indicating that *HcCS1* had a closer evolutionary relationship with other mollusk CSs.

Figure 3. Multiple alignments of chitin synthases from *Hyriopsis cumingii*, *Pinctada fucata*, *Atrina rigida*, and *Mytilus galloprovincialis*. Identical amino acids are in white letters with a black background. The amino acid sequences between two arrows are chitin synthase domains, and blue frames surround the predicted active site residues of the chitin synthase domains.

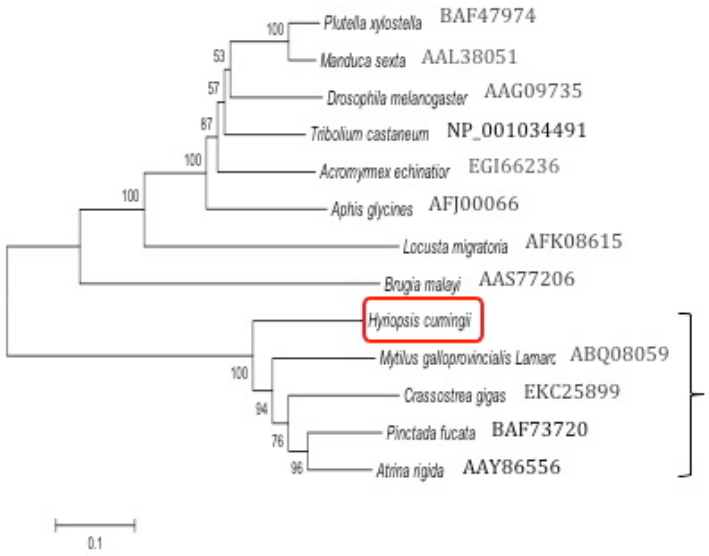

Figure 4. Phylogenetic analysis of chitin synthases from shellfish and insects based on the BLASTp results. A neighbor-joining tree was constructed using the MEGA 4.0 software, and was adjusted by 1000 boot strap replicates. The corresponding GenBank accession No. are shown. HcCS1 from *Hyriopsis cumingii* is marked with the red box, andthe mollusk group is indicated with a brace.

Spatial structures of the CS domain and myosin head domain

Knowing that the mature *HcCS1* peptide shares no more than 30% identity with CSs from the RCSB protein data bank, the Phyre server was used to predict the 3-D model. Figure 2 shows that the predicted 3-D model of *HcCS1* is comprised of eight α-helices and seven β-sheets, which is similar to the crystal structure of bacterial cellulose synthase. The active center was located at the center of the fourth α-helix. In addition, since the myosin head domain of *HcCS1* shared the highest identity (36%) with the heavy meromyosin subfragment of chicken smooth muscle myosin stored in the RCSB protein data bank, the spatial structure was predicted with the SWISS-MODEL workspace based on the 3-D structure of this protein. Figure 2 shows that the predicted 3-D model of the *HcCS1* myosin head domain, consisting of 29 α-helices and 18 β-sheets, was similar to the electron microscopic structure of the heavy meromyosin subfragment of chicken smooth muscle myosin, which suggested that the *HcCS1* myosin head domain may function as a myosin head and act as a force transducer.

Tissue distribution of *HcCS1*

qRT-PCR was conducted to analyze the tissue distribution of *HcCS1*. Figure 5A shows that *HcCS1* was detected in the hemocyte, mantle, gill, liver, stomach, intestine, kidney, and foot. The highest *HcCS1* expression levels were detected in the mantle, and the lowest expression levels were detected in the intestine. In addition, the expression levels of *HcCS1* transcripts in the liver and gill were higher than that observed in the hemocyte, stomach, kidney, or foot.

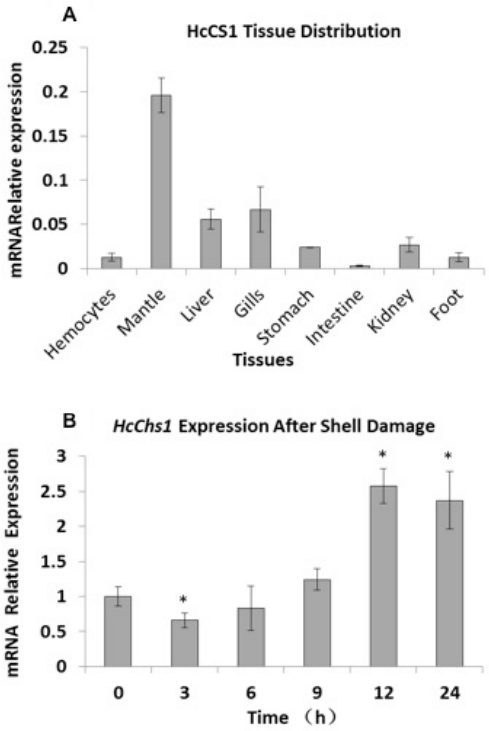

Figure 5. Relative expression level of *HcCS1* in different tissues. The transcripts from the hemocyte, mantle, liver, gill, stomach, intestine, kidney, and foot were detected by real-time PCR. β-actin was amplified as an internal control. **A.** The temporal expression profile of *HcCS1* in the mantle after shell damage was revealed by qRT-PCR. The asterisks indicate significant differences ($P < 0.05$) compared to the normal sample.

Expression profiling after challenge

The temporal expression profile of *HcCS1* after shell damage was examined by qRT-PCR. The *HcCS1* transcripts in the mantle were slightly decreased at 3 h after shell damage, but were upregulated from 12 to 24 h post-damage, peaking (approximately 1.5-fold) at 12 h post-damage (Figure 5B). Moreover, the amount of *HcCS1* was elevated (approximately 1.0-fold) at 24 h (lower than 12 h), showing that *HcCS1* expression began to recover gradually. These results suggest that *HcCS1* was involved in the process of shell regeneration.

DISCUSSION

The triangle sail mussel is the most important and widely used freshwater pearl mussel in China. The nacre, located at both the inner shell layer and outer coating of pearls, has long been appreciated for its beauty, but the mechanisms of nacre formation remain unclear. Chitin is an important organic component of the insect cuticle and mollusk nacre (Merzendorfer, 2006; Weiss et al., 2013). In this study, we identified a CS gene from the triangle sail mussel, and we found that it shared similar characteristics with CS homologs from seawater mussels and insects in both sequence and structure. Additional experiments confirmed that it was involved in shell regeneration.

Insect CSs are large transmembrane proteins with slightly acidic isoelectric points ranging from 6.10 to 6.70, and their theoretical molecular mass is between 160 and 180 kDa. Each has an *N*-terminal domain, a catalytic domain, and a *C*-terminal domain. Similar to insect CSs, *HcCS1* is also slightly acidic with a theoretical isoelectric point of 6.22, but its theoretical molecular mass (~264 kDa) is much larger than that seen in insects. This is mainly because *Hc*CS1 contains a special myosin head domain at the *N*-terminal of the protein. An early report showed that both *ArCS1* from *A. rigida* and *MgCS1* from *M. galloprovincialis* also had the myosin head domain (Weiss et al., 2006), and analyses of *PfCS1* (identified from *P. fucata*) indicated that a myosin head domain was also present (Suzuki et al., 2007). These results suggest that the unique myosin head domain may be a signature structure of mollusk CSs that is absent in insects and parasites. Furthermore, phylogenetic analyses demonstrated that CSs from mollusks were more closely related than those from insects and parasites, suggesting that *HcCS1* may have a special physiological function that differs from insect and parasite CSs.

In addition to the conserved myosin head domain at the intracellular *N*-terminal, *HcCS1* does not differ significantly from the transmembrane architecture and conserved motifs found in insect CSs. The catalytic domains of *HcCS1* and other mollusk CS catalytic domains are highly conserved, and they contain two unique motifs, EDR and QRRRW, which are present in insect CSs (Merzendorfer, 2006). Furthermore, 12 transmembrane domains were predicted based on the protein sequence of *HcCS1*, and this was less than the 13, 15, and 16-18 transmembrane domains found in *MyCS1*, *ArCS1* and *PfCS1*, and insect CSs, respectively (Weiss et al., 2006; Suzuki et al., 2007). Although the number of transmembrane domains is smaller than that in reported mollusk and insect CSs, we speculated that *HcCS1* still possesses the major physiological functions observed in insects and other mollusks because its catalytic center is highly conserved.

Previous studies reported that two different kinds of CS genes existed that encoded CSA and CSB in insects. Insect CSA is specifically expressed in normal epidermal cells, while CSB is restricted to gut epithelial cells (Merzendorfer and Zimoch, 2003). In this study, we identified a homologous CS gene in the triangle sail mussel, and found that it was mainly distributed in the mantle. We concluded that this CS gene was more similar to insect CSA in terms of the classification standard of insect CSs, and we therefore designated it as *HcCS1*. Knowing that the mollusk shell and nacre originate from mantle secretions, the mantle tissue might be similar to insect epidermis, which produces the cuticle. Our study also showed that *HcCS1* was upregulated upon shell damage, and that it was involved in shell regeneration. These findings further confirmed that *HcCS1* was more like insect CSA, suggesting that *Hc*CS1 might participate in shell and nacre formation via chitin synthesis. In addition, the myosin head domain of *ArCS1* is not only capable of producing chitin, but it acts as a force transducer, which affects the mineralization process by creating force fields that regulate chitin deposition in a hypothetical scenario (Weiss et al., 2006). Since *HcCS1* contains a myosin head domain, it might also regulate chitin deposition during the process of shell and nacre formation in a similar dynamic way.

In conclusion, in this study we identified a homologous CS gene, *HcCS1*, from the triangle sail mussel, which is an important fresh water pearl mussel in China. *HcCS1* functionally resembled insect CSA in that it participates in shell and nacre formation by producing chitin. Moreover, *HcCS1* may also affect the mineralization process of the shell and nacre via its myosin head domain, which performs a regulatory function in chitin deposition. These findings may shed light on the mechanisms of shell and nacre formation in freshwater pearl mussels.

Conflicts of interest

The authors declare no conflict of interest.

ACKNOWLEDGMENTS

Research supported by the National Science and Technology Support Program (#2012BAD26B04), the National Natural Science Foundation of China (#31001110), and the Shanghai Collaborative Innovation Center for Aquatic Animal Genetics and Breeding (#ZF1206).

REFERENCES

Andersen SO, Hojrup P and Roepstorff P (1995). Insect cuticular proteins. *Insect Biochem. Mol. Biol*. 25: 153-176.

Bai Z, Zheng H, Lin J, Wang G, et al. (2013). Comparative analysis of the transcriptome in tissues secreting purple and white nacre in the pearl mussel *Hyriopsis cumingii*. *PLoS One* 8: e53617.

Furuhashi T, Schwarzinger C, Miksik I, Smrz M, et al. (2009). Molluscan shell evolution with review of shell calcification hypothesis. *Comp. Biochem. Physiol. B Biochem. Mol. Biol*. 154: 351-371.

Jaworski E, Wang L and Margo G (1963). Synthesis of chitin in cell-free extracts of *Prodenia eridania*. *Nature* 198: 790.

Kelley LA and Sternberg MJ (2009). Protein structure prediction on the Web: a case study using the Phyre server. *Nat. Protoc*. 4: 363-371.

Kramer KJ, Hopkins T and Schaefer J (1995). Applications of solids NMR to the analysis of insect sclerotized structures. *Insect Biochem. Mol. Biol.* 25: 1067-1080.

Li XC, Zhu L, Li LG, Ren Q, et al. (2013). A novel myeloid differentiation factor 88 homolog, SpMyD88, exhibiting Sp Toll-binding activity in the mud crab *Scylla paramamosain*. *Dev. Comp. Immunol*. 39: 313-322.

Merzendorfer H (2006). Insect chitin synthases: a review. *J. Comp. Physiol. B* 176: 1-15.

Merzendorfer H and Zimoch L (2003). Chitin metabolism in insects: structure, function and regulation of chitin synthases and chitinases. *J. Exp. Biol.* 206: 4393-4412.

Munro CA and Gow NAR (2001). Chitin synthesis in human pathogenic fungi. *Med. Myc*. 39: 41-53.

Suzuki M, Sakuda S and Nagasawa H (2007). Identification of chitin in the prismatic layer of the shell and a chitin synthase gene from the Japanese pearl oyster, *Pinctada fucata*. *Biosci. Biotechnol. Biochem*. 71: 1735-1744.

Suzuki M, Saruwatari K, Kogure T, Yamamoto Y, et al. (2009). An acidic matrix protein, Pif, is a key macromolecule for nacre formation. *Science* 325: 1388-1390.

Weiss IM, Renner C, Strigl MG and Fritz M (2002). A simple and reliable method for the determination and localization of chitin in abalone nacre. *Chem. Mater.* 14: 3252-3259.

Weiss IM, Schonitzer V, Eichner N and Sumper M (2006). The chitin synthase involved in marine bivalve mollusk shell formation contains a myosin domain. *FEBS Lett*. 580: 1846-1852.

Weiss IM, Kaufmann S, Heiland B and Tanaka M (2009). Covalent modification of chitin with silk-derivatives acts as an amphiphilic self-organizing template in nacre biomineralisation. *J. Struct. Biol*. 167: 68-75.

Weiss IM, Luke F, Eichner N, Guth C, et al. (2013). On the function of chitin synthase extracellular domains in biomineralization. *J. Struct. Biol*. 183: 216-225.