Fish & Shellfish Immunology 35 (2013) 581–588

Contents lists available at SciVerse ScienceDirect

Fish & Shellfish Immunology

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

Molecular characterization and expression analysis of four cathepsin L genes in the razor clam, *Sinonovacula constricta*

Donghong Niu [a,b,1], Kai Jin [a,1], Lie Wang [a], Bingbing Feng [c], Jiale Li [a,b,*]

[a] Key Laboratory of Exploration and Utilization of Aquatic Genetic Resources and College of Fisheries and Life Science, Shanghai Ocean University, 999 Hucheng Ring Road, Shanghai 201306, China
[b] Shanghai Engineering Research Center of Aquaculture, Shanghai 201306, China
[c] Jiangsu Station for Fisheries Technology Extension, Nanjing 210036, Jiangsun, China

ARTICLE INFO

Article history:
Received 4 February 2013
Received in revised form
29 May 2013
Accepted 2 June 2013
Available online 10 June 2013

Keywords:
Cathepsin L
CTSL
Cysteine protease
Immune response
Sinonovacula constricta

ABSTRACT

Cathepsin L (CTSL) is a lysosomal cysteine protease involved in immune responses in vertebrates. However, few studies exist regarding the role of cathepsin L in bivalves. In this study, we isolated and characterized four cathepsin L genes from the razor clam *Sinonovacula constricta*, referred to as CTSL1, CTSL2, CTSL3 and CTSL4. These four genes contained typical papain-like cysteine protease structure and enzyme activity sites with ERWNIN-like and GNFD-like motifs in the proregion domain and an oxyanion hole (Gln) and a catalytic triad (Cys, His and Asn) in the mature domain. Expression analysis of the four transcripts revealed a tissue-specific pattern with high expression of CTSL1 and CTSL3 in liver and gonad tissues and high expression of CTSL2 and CTSL4 in liver and gill tissues. During the developmental stages, the four transcripts showed the highest expression in the juvenile stage; however, CTSL3 had a much higher expression level than the other three transcripts during embryogenesis. The four transcripts showed significant changes in expression as early as 4 h or 8 h after infection with *Vibrio anguillarum*. The fact that bacterial infection can induce expression of the four CTSL transcripts suggests that these transcripts are important components of the innate immunity system of the clam.

1. Introduction

Lysosomes are membrane-bound cytoplasmic organelles that contain hydrolytic enzymes [1]. Endogenous and exogenous macromolecules are delivered to lysosomes to maintain cellular homeostasis [2]. The proteolytic enzymes in lysosomes assist in the breakdown of foreign materials through membrane destabilization [3].

Cathepsins are a family of lysosomal proteases that are classified into three groups based on the amino acid residues in their active sites: cysteine protease (cathepsin B, C, F, H, K, L, O, S,W, and Z), serine protease (cathepsin A and G) and aspartic protease (cathepsin D and E) [4,5]. In particular, cathepsin L is an important member of the cysteine proteases, consisting of a signal peptide, a pro-peptide and a mature enzyme with four active residues: Gln, Cys, His and Asn. Similar to many other endoproteases, cathepsin L is synthesized as an inactive proenzyme and activated by autocatalysis or other proteases [6]. Signal peptides are responsible for translocation into the endoplasmic reticulum during ribosomal protein expression. Pro-peptides are responsible for protein folding of the catalytic domain and preventing the premature activation of the catalytic domain as a high affinity reversible inhibitor [7].

Cathepsin L functions in physiological and pathological processes, participates in intracellular digestion within the vacuoles of cells and is secreted extracellularly into the lumen of the hepatopancreas and stomach in shrimp [8]. Cathepsin L is also involved in yolk processing during oocyte maturation and embryogenesis. For example, cathepsin La was implicated in yolk processing during oogenesis and embryogenesis in zebrafish by analyzing gene expression during development and in adult tissues [9]. In rainbow trout, cathepsin L, together with cathepsin D, have activities that are associated with yolk processing during vitellogenesis [10,11]. Additionally, cathepsin L is in a class of parasite proteases that has been implicated in critical parasitic functions. In the Chinese liver fluke, *Clonorchis sinensis*, cathepsin L facilitated the invasion of the cercariae into the intermediate fish host by enzyme secretion [12]. The salmon louse, *Lepeophtheirus salmonis*, is a parasitic copepod

* Corresponding author. Key Laboratory of Exploration and Utilization of Aquatic Genetic Resources and College of Fisheries and Life Science, Shanghai Ocean University, 999 Hucheng Ring Road, Shanghai 201306, China. Tel.: +86 021 61900401; fax: +86 021 61900405.
E-mail address: jlli2009@126.com (J. Li).
[1] These authors contributed equally to the manuscript.

1050-4648/$ – see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.fsi.2013.06.001

that feeds on the mucus, skin and blood of salmonids. Cathepsin L activity was observed in the secretory/excretory products, possibly indicating a role in immune evasion and parasite establishment on the host [13]. Moreover, in the immune responses of vertebrates against pathogen invasion, cathepsin L might be involved in Toll-like receptor 9 signaling, activation of interleukin-8, regulation of defense in activity and antigen presentation [14]. Cathepsin L has been commonly associated with malignancy in several cancer types [15]. Due to their emerging roles in immune responses, cathepsin L and its inhibitors have become attractive therapeutic targets [14]. However, in the innate immune system, cathepsin L serves as a chemical barrier against microbial invasion [16]. Cathepsin L has been found in rock bream [17] and orange-spotted grouper [18] and functions as a barrier against invading pathogens.

Despite the well-documented role of cathepsin L in the immune response among vertebrates, its functionality within the innate immune system of invertebrates is unknown. Reports have suggested that cathepsin L is responsible for lysing pathogenic bacteria of the Chinese mitten crab [1] and Chinese white shrimp [19], indicating its potential role in the immune defense system. Few studies exist on the role cathepsin L in bivalves, and these studies have only evaluated the role of cathepsin L the pearl oyster (*Pinctada fucata*) [20], bay scallop (*Argopecten irradians*) [21], and pacific oyster (*Crassostrea gigas*) [22].

The razor clam *Sinonovacula constricta*, a commercially important species of bivalve, is widely distributed in intertidal zones and estuarine water in China, Japan, and Korea. The razor clam is one of the four major aquacultural clams in China. This clam has been suffering from juvenile and summer mortalities because of degradation of germplasm resources and pollution from high-density aquaculture and pathogens [23]. To enhance the yields and quality of this clam, research on its innate immune mechanisms is needed. This work aimed to study cathepsin L gene characterization, expression regulation and its immune responses to *Vibrio anguillarum* in the razor clam *S. constricta*.

2. Materials and methods

2.1. Identification and sequencing of CTSL cDNAs

The pearl oyster cathepsin L (ACQ90252) was used as a query to search cDNAs encoding cathepsin L from the transcriptome sequences (SRA062228) of *S. constricta* using *tblastn*. Contigs were assembled using Vector NTI 10.0 (Invitrogen, Carlsbad, CA) to identify sequences that potentially contain full open reading frames (ORFs). These sequences were re-sequenced in both directions with forward and reverse primers (Table 1) using Sanger technology on the ABI3730 platform (Applied Biosystems). Assembly and comparison of existing transcripts with de novo sequences allowed high certainty regarding transcript identity and sequence accuracy.

2.2. Sequence analysis

The cathepsin L protein sequences were identified by simple keyword searches. Protein sequences retrieved from NCBI were used for ORF and domain searches, alignment, and phylogenetic reconstruction. ORFs were predicted using the Open Reading Frame Finder (http://www.ncbi.nlm.nih.gov/gorf/gorf.html), and signal peptides and mature domains were identified by the Simple Modular Architecture Research Tool (SMART) (http://smart.embl-heidelberg.de).

2.3. Phylogenetic analysis

Cathepsin L sequences of human, mouse, frog, chicken, zebrafish and oyster retrieved from databases were aligned using the ClustalW2 program (http://www.ebi.ac.uk/ Tools/clustalw2/). Phylogenetic tree was constructed using the neighbor-joining (NJ) method based on the deduced full-length amino acid sequences with 10,000 bootstrapping replications within the Molecular Evolutionary Genetics Analysis (MEGA 4.0) package. Data were analyzed using Poisson correction, and gaps were removed by complete deletion.

2.4. Sample preparation and bacterial challenge

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

All experiment clams were obtained from the Yuejingyang Farm, Ninghai City, Zhejiang Province, China. Embryos and larvae were reared in hatching tanks at 25–27 °C. Embryos and larvae from different developmental stages were collected and stored in RNA storage liquid. These stages included post-fertilization embryos, trochophore, veliger, umbo larvae, creeping larvae and juvenile clams.

For the bacterial challenge, clams with an average body weight of 9.2 g and an average body length of 5.5 cm were kept at 25–27 °C in 30‰ saline. The clams were acclimated for one week before processing. The four control and four challenge groups (50 clams each group) were designed. The treatments were as follows: (1) control group (phosphate-buffered saline, 50 μL PBS injected) and (2) *V. anguillarum*-challenged group (injection). The clams were

Table 1
Primers used for the study of the four CTSL genes.

Gene	GenBank no.	Isotig no.	Amplicon sizes (bp)	Use for primer	Primer sequences (5′–3′)
CTSL-1	KC874989	isotig09527	221	For qRT-PCR	CGAGGATGAAACTCTGAGACG
					TCACCTTGATGTTGCTTGGA
			1478	For complete cDNA	AAGAGACGAAAAACCAAGATGTTC
					TTAGAATGCGAGTTTGATTAGGTG
CTSL-2	KC874990	isotig21070	239	For qRT-PCR	GATGGACCAGGCTTTCACCTAC
					CCTGTACAACTGGAAGCTGGTG
			1096	For complete cDNA	AATCGTCATCAGAGTTTGTTTTGA
					CACGGAACTTCAATTAAGCAAGA
CTSL-3	KC874991	isotig07968	241	For qRT-PCR	AACATGTCTTCGCCACCTTC
					CTGGCCTTAGTTTCCACATT
			1331	For complete cDNA	CATGTGATATTCATGAACCATGTG
					AAGCTGTATGACTTGATGGCAATA
CTSL-4	KC874992	isotig18217	248	For qRT-PCR	CCGACTCTTCCAGGTCTTCA
					CCACCTGCGACATGTTGAT
			1225	For complete cDNA	TTGTAAGATGTCAAAACACCTCGT
					GGAGACGCTTTAAAGGATACAGTC