Fish & Shellfish Immunology 62 (2017) 332–340

Contents lists available at ScienceDirect

Fish & Shellfish Immunology

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

Full length article

Identification of cvSI-3 and evidence for the wide distribution and active evolution of the I84 family of protease inhibitors in mollusks

Qinggang Xue [a, b, c, *], Jean-Phillipe Beguel [c], Julie Gauthier [d], Jerome La Peyre [c]

[a] Zhejiang Key Laboratory of Aquatic Germplasm Resources and College of Biological and Environmental Sciences, Zhejiang Wanli University, Ningbo, Zhejiang 315100, China
[b] Laboratory for Marine Fisheries Science and Food Production Processes, Qingdao National Laboratory for Marine Science and Technology, Qingdao, 266071, China
[c] School of Animal Sciences, Louisiana State University Agricultural Center, Baton Rouge, LA 70803, USA
[d] Loyola University, Department of Biological Sciences, New Orleans, LA 70118, USA

ARTICLE INFO

Article history:
Received 25 September 2016
Received in revised form
25 January 2017
Accepted 28 January 2017
Available online 31 January 2017

Keywords:
Serine protease inhibitors
Bivalve
Oyster immunity
Disease resistance
Molecular evolution

ABSTRACT

Protease inhibitors are an extremely diverse group of proteins that control the proteolytic activities of proteases and play a crucial role in biological processes including host defenses. The I84 family of protease inhibitors in the MEROPS database currently consists of cvSI-1 and cvSI-2, two novel serine protease inhibitors purified and characterized from the eastern oyster *Crassostrea virginica* plasma and believed to play a role in host defense and disease resistance. In the present study, a third member of I84 family, named cvSI-3, was identified from *C. virginica* by cDNA cloning and sequencing. The full cvSI-3 cDNA was composed of 342 bp including a 255 bp open reading frame (ORF) that encodes an 84-amino acid peptide. The mature cvSI-3 molecule was predicted to have 68 amino acid residues after removal of a 16-amino acid signal peptide, with a calculated molecular mass of 7724.5 Da and a theoretical isoelectric point (pI) of 6.28. CvSI-3 amino acid sequence shared 41% identity with cvSI-2 and 37% identity with cvSI-1, which included 12 conserved cysteines. Quantitative real-time PCR determined that cvSI-3 gene expressed primarily in oyster digestive glands. Real-time PCR also detected that cvSI-1, cvSI-2 and cvSI-3 expression levels in digestive glands varied significantly, with cvSI-2 showing the highest expression level and cvSI-3 the lowest. Additionally, a significant correlation was detected between cvSI-2 and cvSI-3 mRNAs levels. Searches into sequence databases using cvSI-1, cvSI-2 and cvSI-3 as queries retrieved ESTs suggesting the possible existence of at least 9 more I84 family members in eastern oysters and of I84 family protease inhibitors in various bivalve and gastropod species. Moreover, orthologs of all *C. virginica* I84 family members or potential member genes were found to be present in the *C. gigas* genome, and their distributions among species provided important information about the evolution of the I84 family of protease inhibitors. It appears that the I84 family of protease inhibitors is widely distributed and actively evolving in the Phylum Mollusca.

1. Introduction

Protease inhibitors are an extremely diverse group of proteins that control the proteolytic activities of proteases and play a crucial role in biological processes including host defenses. They have traditionally been classified on the basis of their specific inhibitory activity towards one of four mechanistic classes of proteases, and thus have been named either serine, cysteine, aspartic or metalloprotease inhibitors [1]. Recently, a sequence similarity based classification system has been proposed to classify all protease inhibitors into protein families and a database (i.e., MEROPS; http://merops.sanger.ac.uk/) has been developed [2]. The latest MEROPS database update (i.e., Release 10.0) includes 81 families of protease inhibitors [3].

Two serine protease inhibitors, cvSI-1 and cvSI-2 were purified and characterized from the plasma of the eastern oyster, *Crassostrea virginica* [4–6], and as a result the MEROPS database (i.e., Release 8.5) created a new protease inhibitor family, I84, in August 2009 to

* Corresponding author. Zhejiang Key Laboratory of Aquatic Germplasm Resources and College of Biological and Environmental Sciences, Zhejiang Wanli University, 8 South Qianhu Road, Ningbo, Zhejiang 315100, China.
E-mail address: qxue@zwu.edu.cn (Q. Xue).

http://dx.doi.org/10.1016/j.fsi.2017.01.040
1050-4648/

host these two novel protease inhibitors. Purified cvSI-1 is a 7609.6 Da protein consisting of 71 amino acids whereas cvSI-2 has a molecular weight of 7202.96 Da consisting of 67 amino acids. Enzyme kinetic characterization determined that both proteins inhibit, in a slow tight-binding manner, subtilisin, trypsin and perkinsin, a major subtilisin-like protease secreted by *Perkinsus marinus,* a protozoan parasite of eastern oysters. *In situ* hybridization detected that cvSI-1 and cvSI-2 genes are expressed mainly by the basophil cells of digestive tubules in eastern oyster digestive glands. The amino acid sequences of the two proteins determined by mass spectrometry sequencing of purified proteins and predicted from cDNA, shared high similarity but showed no significant homology to any known sequence in protein databases.

While the biological roles of these novel protease inhibitors are not fully known, results of several studies suggest they are involved in host defenses and disease resistance. It was shown that purified cvSI-1 inhibited *P. marinus* propagation in vitro in a dose-dependent manner and oysters selected for increased resistance to *P. marinus* consistently expressed more cvSI-1 than unselected oysters that were highly susceptible to the parasite [6]. CvSI-1 gene expression level in eastern oysters 8–21 days after challenge with *P. marinus* was also negatively correlated with parasite infection intensity [7]. CvSI-1 genetic variations in single nucleotide polymorphisms (SNPs) in coding region and indel in promoter region were associated with eastern oyster resistance to *P. marinus* [8,9]. Additional evidence for cvSI-1 and cvSI-2 involvement in resistance to *P. marinus* in Cortez oysters (*Crassostrea corteziensis)* was also revealed as these oysters which appear less affected by *P. marinus*, expressed significantly more cvSI-1 and cvSI-2 mRNA than eastern oysters after challenge with the parasite [10]. The protease inhibitors are also likely involved in host defenses against other pathogens as oysters differing in resistance to Roseovarius oyster disease (ROD) showed different patterns of cvSI-1 and cvSI-2 gene expression following *P. marinus* injection [11]. Apparently, maintaining high gene expression and activity level is a key factor for the functionality of cvSI-1 and cvSI-2 in host defense.

The objectives of the present research were to (1) clone and sequence the full length cDNA of cvSI-3, a third member of the I84 family of protease inhibitors in *C. virginica*, (2) determine in which tissue cvSI-3 gene was expressed, (3) compare cvSI-1, cvSI-2 and cvSI-3 gene expression levels and determine their contribution to plasma serine protease inhibitory activity, and (4) analyze available EST sequences and Pacific oyster genome sequences to gain some insights into the evolution of the I84 family of protease inhibitors. Such information should help us better understand the complexity and potential functions of this novel family of protease inhibitors.

2. Materials and methods

2.1. Oysters and sampling

Eastern oysters, about 8–12 cm in shell length, and grown at the Louisiana Sea Grant oyster hatchery in Grand Isle, Louisiana were collected in March 2009 to clone cvSI-3 cDNA, in December 2009 (average water temperature 15.4 ± 1.5 °C; average salinity 17.3 ± 2.2 psμ) to measure cvSI-3 mRNA levels in various tissues and in April 2014 (average water temperature 21.8 ± 2.0 °C; average salinity 17.5 ± 5.9 psμ) to compare cvSI-1, cvSI-2 and cvSI-3 mRNA levels in digestive glands and plasma subtilisin inhibitory activity. Oysters were maintained in a 1000 L recirculating seawater system at salinity of 15 at 20 °C for 1 week prior to hemolymph and tissue collection. Oyster hemolymph was collected from the adductor muscle sinus using a 3-ml syringe equipped with a 25-gauge needle through a notch on the dorsal side of the shell. After centrifugation of the hemolymph for 15 min at 300 g, 4 °C, the supernatant was collected and centrifuged again for 15 min at 2500 g, 4 °C. The resulting supernatant was used as plasma and stored at −20 °C before analysis; the hemocyte pellets were kept at −80 °C prior to RNA extraction. After hemolymph withdrawal, oysters were shucked and tissue pieces of ~50 mg were excised from the digestive gland, gills, labial palps, mantle, and adductor muscle, immersed separately in 500 μL of RNALater (Qiagen) at 4 °C overnight and then stored at −80 °C until used for RNA extraction.

2.2. cvSI-3 cDNA cloning and sequencing

The consensus EST sequence CVEST03 identified in previous research [5] was chosen as a template to design gene specific primers (Table 1) to clone the complete cvSI-3 cDNA using 3'- and 5'- rapid amplification cDNA ends (RACE). Primers were designed using Primer3 [12]. Total RNA was extracted from the digestive gland tissues of an oyster collected in March 2009 using RNeasy Min Kit (Qiagen). RNA concentration and purity was determined by A260/A280 using a NanoDrop 2000c Spectrophotometer (Thermo Scientific) and the integrity was verified by agarose electrophoresis. An aliquot of 3 ng RNA was processed to include a 5′ linker sequence on capped (full length) transcripts, and cDNAs were synthesized using the GeneRACER™ Oligo dT primer and Superscript III (Invitrogen). RACE reactions were performed using the GeneRACER™ Kit (Invitrogen) and the RACE products were purified by agarose gel electrophoresis and cloned using the TOPO TA Cloning® Kit for Sequencing (Invitrogen). The recombinant plasmids were sequenced with CEQ 8800 Genetic Analysis System (Beckman Coulter, Fullerton, CA). Two clones were sequenced to confirm 3′ end, and three clones were sequenced to confirm 5′ end. As the overlapping region sequences were identical in all 5 sequenced clones (i.e., 2 clones for 3′ end confirmation and 3 clones for 5′ end confirmation) whole gene re-sequence was not performed after 3′ and 5′ end sequence determination. The complete cDNA sequence was compared with sequences in the National Center for Biological Information (NCBI) GenBank databases using the nucleotide-nucleotide BLAST search [13]. An amino acid sequence was deduced from the cDNA sequence data using the Translate tool at the ExPASy server, and the deduced protein sequence was analyzed using SignalP 3.0 server to predict a signal peptide by both neural networks and hidden Markov models [14].

2.3. Determination of cvSI-1, cvSI-2 and cvSI-3 mRNA expression

Total RNAs were extracted from tissue samples using RNeasy® Mini Kit and treated with DNase using RNase-Free DNase Set (QIAGEN) to prevent DNA contamination according to manufacturer's instructions. RNA concentrations were measured using the A260/A280 ratio with a Nanodrop spectrometer (Thermo Scientific). RNA quality was analyzed using RNA nanochips and Agilent RNA 6000 nanoreagents (Agilent Technologies) according to manufacturer's instructions and Integrity Numbers (RIN) of the RNAs used in the research were greater than 9.40. Total cDNAs were synthesized with 650 ng of total RNAs in a final volume of 20 μL using Omniscript RT Kit with Oligo-dT Primers and RNase Inhibitor (QIAGEN) according to manufacturer's instructions.

Real-time PCR assays were performed in a Bio-Rad CFX96™ Real-Time C1000 Thermal Cycler using the synthesized total cDNA as template and primers specific for cvSI-1, cvSI-2, or cvSI-3 designed using Primer3 [12] (Table 1). The reactions were carried out in a total volume of 10 μL containing 4 μL cDNA at 1.25 ng/μl, 0.5 μL of each primer at 10 μM and 5 μL of SsoFast™ EvaGreen® Supermix (Bio-Rad). The cycling conditions consisted of Taq polymerase activation for 30 s at 95 °C, followed by 40 cycles of denaturation at 95 °C for 15 s and annealing/elongation at 60 °C for 45 s.