Molecular characterization of two kazal-type serine proteinase inhibitor genes in the surf clam Mesodesma donacium exposed to Vibrio anguillarum

Waleska Maldonado-Aguayo, Gustavo Núñez-Acuña, Valentina Valenzuela-Muñoz, Jacqueline Chávez-Mardones, Cristian Gallardo-Escárate*

Laboratory of Biotechnology and Aquatic Genomics, Interdisciplinary Center for Sustainable Aquaculture Research (INCAR), University of Concepción, PO. Box 160-C Concepción, Chile

A R T I C L E   I N F O

Article history:
Received 29 November 2012
Received in revised form 20 February 2013
Accepted 11 March 2013
Available online 22 March 2013

Keywords:
Serine protease inhibitor
SPI
Mesodesma donacium
Vibrio anguillarum
Gene expression

A B S T R A C T

This study reports two kazal-type serine protease inhibitors (KPI) identified in a cDNA library from the surf clam Mesodesma donacium, and characterized through Rapid Amplification of cDNA Ends (RACE). The KPIs, denoted as MdSPI-1 and MdSPI-2, presented full sequences of 1139 bp and 781 bp respectively. MdSPI-1 had a 5’ untranslated region (UTR) of 175 bp, a 3’UTR of 283 bp and an open reading frame (ORF) of 681 pb that encodes for 227 amino acids. MdSPI-2 showed a 5’UTR of 70 bp, a 3’UTR of 279 bp and an ORF of 432 bp that encodes for 144 amino acids. Both sequences presented two kazal-type tandem domains. Phylogenetic analysis of MdSPI-1 and MdSPI-2 shows a main clade composed by other bivalve species and closely related crustaceans. Real time PCR analysis showed that MdSPI-1 is mainly up-regulated in mantle, foot, gills and muscle tissues, while MdSPI-2 is expressed principally in foot tissue. Moreover, to evaluate the immune response of MdSPI-1 and MdSPI-2, infections with Vibrio anguillarum were performed. Herein, MdSPI-1 and MdSPI-2 transcription expression were significantly up-regulated at 2 and 8 h post-challenge. Our results suggest that MdSPI-1 and MdSPI-2 are important humoral factors of innate immunity in M. donacium.

© 2013 Elsevier Ltd. All rights reserved.

1. Introduction

Serine protease inhibitors (SPI) are a super-family of proteins with physiological functions, such as blood coagulation [1], inflammation [2], metamorphosis [3], complement system [4] and innate immune response. In all these process, SPIs are mainly responsible for maintaining cellular homeostasis, inhibiting undesirable proteolytic cascades [5] or inhibiting exogenous protease secreted by pathogenic microorganisms that use the protease to penetrate the host and invade new tissues [6]. Among the families of serine protease inhibitors there is the I1, which is composed by proteins containing the kazal-type domain and therefore are called kazal-type SPIs (KPI) [7]. These have a single or multi kazal domain, which can inhibit different proteases [6]. The latter are characterized by a well conserved amino acid sequence, with six cysteines responsible of the formation of disulfide bridges (C1x (3) C2x (7) PVC1X (3) GX2YXNC4X (6) CSX (12) C) [8]. The specificity of the inhibitor is determined by the P1 residue located in the second residue after the second cysteine conserved from the kazal domain [9].

Serine protease inhibitors have been identified in mollusks as humoral factors present in haemolymph and related to components of innate immune response. For instance, in species like the Pacific oyster Crassostrea gigas [10,11], soft-shell clam Mya arenaria [12], bay scallop Argopecten irradians [13] and Zhikong scallop Chlamys farreri [14], have been characterized as kazal-type inhibitors (KPI) capable of blocking the action of some proteases such as subtilisin A and trypsin. Furthermore, KPI immune response has been studied in the shrimp Penaeus monodon and the river crab Procambarus clarkii, evidencing the presence of kazal-type domains with inhibitory activity against subtilisin, trypsin and chymotrypsin [5,15]. Four serine protease inhibitors in Fenneropenaeus chinensis have the capacity to inhibit subtilisin and proteinase K and have been associated with the innate antibacterial and antifungal immune response system [6].

The surf clam Mesodesma donacium, the model species for our study, has socio-economic importance all along the Chilean coast. There have been few studies of this species at the molecular level, which has limited the development of DNA markers to assess this over-exploited population at genetic level [16]. The objective of this study was to characterize two EST sequences homologs to kazal-type SPI identified by 454 pyrosequencing from M. donacium transcriptome, and determine their transcription
patterns in different tissues using qPCR and their role in antibacterial immune response.

2. Materials and methods

2.1. Experimental design

Forty specimens with the same stage of gonadal maturation were obtained from Coquimbo, Chile (31°55'S, 71°30'W). The specimens were used to characterize the SPIs mRNA, and also in challenge trials against *Vibrio anguillarum*. Tissues from the foot, mantle, gills and muscle were fixed in RNALater RNA Stabilization Reagent (Ambion, USA) and stored at −80 °C for subsequent RNA extraction. The remaining specimens were first acclimatized for a week in filtered seawater at 17 °C with constant airing. The specimens were then divided into a control group (N = 18) and a challenge group (N = 18). For the challenge trials, 100 μL of *V. anguillarum* (3.2 × 10^6^ cell/ml) in Luria Bertani (LB) culture medium was injected into the adductor muscle of *M. donacium* specimens, while the control group was injected with 100 μL of LB medium. Both groups were sampled at different exposure times: 2, 4, 6, 8, 16 and 32 h post-injection. Around 300 mg of foot, mantle, gills and muscle samples were dissected and fixed in RNALater RNA Stabilization Reagent (Ambion, USA), and stored at −80 °C for subsequent RNA extraction.

2.2. Extraction of total ARN

Total RNA was extracted from 100 mg of the previously described tissues using TRIzol Reagent (Invitrogen®, Life Technologies, USA) according to the manufacturer’s instructions. The concentration and purity of the extracted RNA was determined by spectrophotometry in a Nanodrop ND-1000 (NanoDrop® Technologies, Inc) and its integrity was visualized by electrophoresis in denaturing agarose gels at 1.2% stained with ethidium bromide at 0.001%. Subsequently, from 200 ng of RNA of each sample, cDNA was synthesized using the RevertAid™ H Minus First Strand cDNA Synthesis kit (Thermo Scientific, Maryland, USA), in accordance with the supplier’s instructions.

2.3. KPI identification and characterization

We identified two annotated contigs for KPIs from a cDNA library for *M. donacium* obtained by 454 pyrosequencing of gill and mussel tissue from adult individuals of *M. donacium* that are available for download at the Dryad Digital 1600 depository (http://datadryad.org/) under the access http://dx.doi.org/10.5061/dryad.8jd18. For both sequences annotated as KPIs, specific primers were designed for Primer 3 in the software Geneious 5.1.3 (Biomatters, New Zealand) [17] (Table 1). The partial sequence of each contig was then amplified by PCR in a final volume of 12.5 μl of reaction. The amplification conditions were: 94 °C for 2 min and 30 s (holding stage), 35 cycles of 94 °C for 30 s (denaturing), 60 °C (T° annealing) for 30 s, 72 °C for 45 s and 72 °C for 5 min. The PCR product was visualized through electrophoresis in agarose gels at 1.2% and, then sent to Macrogen Inc. (Korea) for sequencing in an ABI 3730xl capillary sequencer (Applied Biosystems). Subsequently, the sequences were analyzed with Geneious 5.1.3 software, resulting in two partial sequences termed MdSPI-1 and MdSPI-2. Based on these sequences, specific new primers were designed (Table 1) for amplification of the 3° and 5°UTR ends, using the FirstChoice® RLM RACE kit (Ambion®, Life Technologies, USA) in accordance with the manufacturer’s instructions. The fragments obtained for the 5° and 3°UTR ends were ligated in the cloning vector TOPO TA Cloning kit (Invitrogen™, Life Technologies, Carlsbad, CA, USA) and transformed into electrocompetent bacteria of

<table>
<thead>
<tr>
<th>Primers</th>
<th>Sequence (5‘–3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MD_8873F(MdSPI1)</td>
<td>ATCCCAAGAAGCTAAAGCA</td>
</tr>
<tr>
<td>MD_8873R(MdSPI1)</td>
<td>TCTGCGTCTCCGACAGAT</td>
</tr>
<tr>
<td>MD_3594F(MdSPI2)</td>
<td>GACGCTGGTCTTCATATTGA</td>
</tr>
<tr>
<td>MD_3594R(MdSPI2)</td>
<td>CTCTGGTTGAGTCTCCAGAT</td>
</tr>
<tr>
<td>QMD_8873F</td>
<td>CGAGTCGCAACAGAAGAA</td>
</tr>
<tr>
<td>QMD_8873R</td>
<td>CAGAGGAAACAAATGCGAGA</td>
</tr>
<tr>
<td>QMD_3594F</td>
<td>TGGTGTTGAGATCGACATTC</td>
</tr>
<tr>
<td>QMD_3594R</td>
<td>CAGTCGCAACAGAAGAA</td>
</tr>
<tr>
<td>MD_atub_1F</td>
<td>ACCGACTTTCGCGAAGCA</td>
</tr>
<tr>
<td>3594_5'R inner</td>
<td>AAGCTCTTGTATCCCGGAGCA</td>
</tr>
<tr>
<td>3594_5'R outer</td>
<td>TATGGTCTGACCTGCTGCTG</td>
</tr>
<tr>
<td>3594_3'F outer</td>
<td>GGTGCTGACCTGCTGCTG</td>
</tr>
<tr>
<td>Md_atub_1F inner</td>
<td>GGTGCTGACCTGCTGCTG</td>
</tr>
<tr>
<td>Adapter 5</td>
<td>CTAATACGACTCACTATAGT</td>
</tr>
<tr>
<td>3593_3'inner</td>
<td>TTAATACGACTCACTATAGT</td>
</tr>
<tr>
<td>Adapter 3′</td>
<td>TTAATACGACTCACTATAGT</td>
</tr>
</tbody>
</table>

| Table 1 | Sequence of the oligonucleotide primer used in the study. |

Escherichia coli JM109 in LB/amp/IPTG/Xgal plates overnight at 37 °C. The positive clones were selected and purified to obtain plasmids using the E.Z.N.A® Plasmid DNA Mini Kit II (Omega Bio-tek, Doraville, GA, USA). The plasmids obtained were sequenced in both directions and their sequences were then assembled using the software Geneious 5.1.3.

Blastn and Blastx searches were conducted against the non-redundant Genbank database (http://www.ncbi.nlm.nih.gov/) to identify the conserved sites of the kazal domain. Subsequently, a multiple alignment was made of the kazal domains of *M. donacium* and the kazal-type domains publicly available for the species: *P. clarkii* (hpCPSI), *P. monodon* (SPIPm2 and SPIPm), *F. chenensis* (hpFCPSI1), *Litopenaeus vannamei* (hcvLSPI and hplVSPI), *C. farrelli* (CRKZSPI), *C. gigas* (CvSPI) and *A. irritans* (AISPI). Finally, to determine clusters among the amino acid sequences of the kazal domains, a phylogenetic analysis was conducted using the Neighbor-Joining method adjusted to 1000 iterations, and including *Hidra magnipapillata* as outgroup.

2.4. Expression profiles of *MdSPI* in *M. donacium*

Sustainable primers were designed for qPCR reactions for the sequences obtained from *MdSPI-1* (QMD_8873F, QMD_8873R) and *MdSPI-2* (QMD_3594F, QMD_3594R) (Table 1). As well, the genes α- tubulin, β-tubulin and GAPDH were identified and assessed as endogenous controls for the relative expression qPCR assays. After an assessment of efficiency, α-tubulin was selected as an endogenous control. All the primers were designed to amplify ~100 bp of PCR products.

The qPCR runs were performed with StepOnePlus™ (Applied Biosystems, Life Technologies, USA) using the comparative ΔCt method [18]. This method requires efficiency values close to 100%, therefore we analyzed the dynamic range to calculate the reaction efficiency of every gene. To perform this, we conducted five cDNA dilution series, beginning with 80 ng of template and a dilution factor of 1:5. Each reaction was conducted with a volume of 10 μl using Maxima® SYBR Green/ROX qPCR Master Mix (Thermo Scientific, USA). The amplification conditions were: 95 °C for 10 min (holding stage), 40 cycles of 95 °C for 30 s (denaturation), T° (annealing) 30 s, 72 °C for 30 s (extension). To determine the presence of single amplicons, non-specific products, gDNA contamination and primer dimers, a melting curve was carried out following the amplification,
from 60 °C to 95 °C and taking fluorescence data every 0.03 °C. Once the reaction efficiency of all the genes (including the endogenous controls) was between 90 and 110%, we proceeded to evaluate the expression profile of MdSPIs in specimens after being challenged with *V. anguillarum*. The data obtained were analyzed with the Kruskal–Wallis test with Statistica software (Version 7.0, StatSoft, Inc.), given the non-parametric distribution of the data. Statistically significant differences were accepted with a *p* value < 0.05.

### 3. Results

#### 3.1. Characterization of MdSPIs mRNA

The full length obtained for *MdSPI-1* cDNA was 1181 bp, which is composed by 217 bp of 5′-end untranslated region (UTR), 283 bp of 3′UTR and 681 bp of ORF that encodes for a sequence of 227 amino acids. Equally, a full length of 781 bp was obtained for *MdSPI-2*, containing a 5′UTR of 70 bp, a 3′UTR of 279 bp and an ORF of 432 bp that encoding for 144 amino acids. Both sequences presented the conserved polyadenylation signal of AATAA and a poly a tail (A). As well, both sequences presented the characteristic Kazal domain, with six conserved cysteines responsible for forming disulfide bridges (C1x (3) C2X (7) PVCX (3) CXZXNXCX (6) CSX (12)). The translated sequence of *MdSPI-1* contains two kazal-type domains separated by fifteen amino acid residues and the P1 amino acids, phenylalanine (Phe) and alanin (Ala) for each respective domain (Fig. 1). Similarly, *MdSPI-2* presented two kazal-type domains separated by 37 amino acid residues, and leucine (Leu) and lysine (Lys) were observed in the P1 residue of each domain (Fig. 2).

Nucleotide sequences of *MdSPI-1* and *MdSPI-2* were analyzed by Blastx algorithm against the EST database from GenBank. The recorded e-values were 3E-9 and 5E-19 for *MdSPI-1* and *MdSPI-2* respectively, with a high level of homology to sequences described as kazal-type serine protease inhibitors (PROSITE motif PS00282). Both sequences were homologous to the KPI of marine invertebrates such as *C. gigas* (EKC34738), *Azuampecten farreri* (ACD88987) and *A. irradians* (AAX38588). The amino acid sequences deduced from *MdSPI-1* and *MdSPI-2* were compared to the GenBank database using Blastp, resulting in 38% identity and an e-value of 7E-11 for *MdSPI-1* and 50% identity and an e-value of 5E-12 for *MdSPI-2* with respect to the KPI of *C. gigas* (EKC34738). The alignment of the amino acid sequences of both transcripts for *M. donacium* with those of other marine invertebrates evidenced the conserved sites for the cysteines characteristic of the kazal-type domain and the diversity of P1 residues among the different domains (Fig. 3). The phylogenetic analysis showed a consensus topology with well-sustained nodes by high bootstrap values. It can be noted that the *MdSPI* sequences form a separate clade supported by a bootstrap value of 68%, which is closely related to the clade of the bivalves (*C. gigas, A. irradians* and *Pinctada fucata*) and in a second order with a clade composed of crustacean species (*P. clarkii, F. chinensis* and *P. monodon*) (Fig. 4). It should be noted that the bootstrap values in the basal nodes are higher than those of the distal nodes.

#### 3.2. MdSPIs expression patterns in *M. donacium*

The relative expression of *MdSPIs* in the different tissues of *M. donacium* not challenged was evaluated by RT-qPCR. We found higher levels of expression of *MdSPI-1* in the mantle than in the foot, gills and muscle, with significant differences supported by a *p* value < 0.005. *MdSPI-1* was expressed 5 times more in tissue from the mantle than from the foot, 3 times more than in tissue from the gills and 4 times than muscle tissue (Fig. 5). The highest levels of expression of *MdSPI-2* were found in the foot, being 10 times higher than in the gills, 9 times more than in the muscles and 5 times more than in the mantle than in the gills, 9 times more than in the muscles and 5 times more than in the foot, gills and muscle, with significant differences from the control (Fig. 6). The expression analysis after exposure to *V. anguillarum* were carried using a cDNA of mantle tissue. The relative expression of both *MdSPIs* gradually increased compared to the control groups (Fig. 6). *MdSPI-1* expression peaked at 2 hpi (Fig. 6A), while *MdSPI-2* expression peaked at 8 hpi, with significant differences from the control (*p* < 0.05) at all times evaluated.

![Nucleotide and deduced amino acid sequence of *MdSPI-1* gene from *Mesodesma donacium* (GenBank accession no. KC160531). The two Kazal-type domain are shadowed in black boxes. The reading frame is shadowed in gray boxes. The classical polyadenylation signal is underlined. The poly-A tail is labeled with a black rectangle. The asterisk (*) indicates the stop codon.](image1.png)
Specifically, *MdSPI-1* increased its mRNA two times in challenged group compared to control group at 2 hpi, and it reached up to four times in respect to controls at 8 hpi. Meanwhile, *MdSPI-2* increased its mRNA levels in challenged group up to nine times more than control groups at 2 hpi, and then its levels decreased.

### 4. Discussion

#### 4.1. Serine protease inhibitors and their role in the innate immunity of marine invertebrates

Serine protease inhibitors play an important role in the innate immunity of marine invertebrates, acting as modulators of proteases from endogenous and pathogenic microorganisms [8]. In this study we identified two mRNA encoding for the serine protease inhibitor in *M. donacium*, termed *MdSPI-1* and *MdSPI-2*. The complete mRNA was characterized and the amino acid sequences of both inhibitors were deduced. Two domains were identified in both mRNA sequences that are homologous to those described for kazal-type serine protease inhibitors. A typical or canonical Kazal domain is composed of 40–60 amino acid residues including some spacer amino acids [8]. This was found in the deduced amino acid sequences *MdSPI-1* where the first and second domain had 42 and 41 amino acid. *MdSPI-2* also showing amino acid 43 and 49 for the first and the second domain respectively. Kazal domain similar lengths have been found in other marine invertebrates such as in the swimming crab *Portunus trituberculatus* [19], *L. vannamei* [14] and *P. clarkii* [13]. The inhibitory specificity of a kazal domain arises mainly from the P1 residue, which varies widely, with more one domain being possible in the same sequence [5]. Each domain acts independently, inhibiting different protease. Generally, inhibitors with the amino acids Arg or Lys in the P1 residue trend to inhibit Trypsin, while those with the amino acids Ala or Ser inhibit elastase, and those with Tyr, Pro, Met, Leu or Phe inhibit chymotrypsin [20]. For example, in *Bombix mori*, it was established by recombinant enzymatic characterization of a kazal-type serine protease inhibitor, that there is inhibitory action against subtilisin featured by Ala is present as a P1 residue [21]. Nine kazal-domains have been identified in *F. chinensis*, corresponding to five Asp, Thr, two Gl and Ala, the latter showing inhibitory activity against elastase [6]. Consequently, the role of *MdSPIs* could be inferred according to the amino acid in the P1 position of the kazal domain. The P1 residues of the *MdSPI-1* sequence in its first domain presents an Ala residue, the latter suggesting inhibitory activity against subtilisin, acting in this way as a component of the defense against the protease of pathogenic bacteria as well as those secreted by of *V. anguillarum* [21,22]. A Phe was found in the P1 residue of the second *MdSp1* domain, suggesting inhibitory activity against trypsin and chymotrypsin enzymes [23]. In turn, *MdSPI-2* presents the Leu residue in its first domain, providing specificity of the inhibitor for chymotrypsin or subtilisin, according to what is described in *Pacifastacus leniusculus* [24] and *P. clarkii* [25]. The second *MdSPI-2* domain presents a Lys residue with inhibitory activity against trypsin according to what has been observed in *P. monodon* [26,27], *H. magnipapillata* [28] and *C. farreri* [29]. Overall, the SPIs characterized in this work suggest that both identified enzymes are important humoral factors of innate immunity in *M. donacium*.

A clade was observed in the phylogenetic analysis formed by the kazal-type domains of *MdSPI-1* and *MdSPI-2*. The two sequences of

---

**Fig. 2.** Nucleotide and deduced amino acid sequence of *MdSPI-2* gen from *Mesodesma donacium* (GenBank accession no. KC160532). The two Kazal-type domain are shadowed in black boxes. The reading frame are shadowed in gray boxes. The classical polyadenylation signal is underlined. The poly-A tail is labeled with a black rectangle. The asterisk (*) indicates the stop codon.
M. donacium have a closer phylogenetic relationship to each other than either has with other bivalve species, and in turn this bivalve clade has a relatively distant phylogenetic relationship with other kazal-type inhibitors of crustacean species. It is important to keep in mind that this analysis is not a taxonomic inference. Nevertheless, it is possible to suggest that these inhibitors have specific immune system functions, probably coded by different genes for each species. Our results are similar to those obtained with P. clarkii, where it was found that the crustacean clade was evolutively distant from serine protease inhibitors identified in other bivalve mollusks like C. farreri [15].

4.2. Patterns of KPIs expression in marine invertebrates

The expression of serine protease inhibitors in marine invertebrates is generally associated with hemocytes, as have been shown in studies with different invertebrates like P. monodon, where SPIm2 was identified from a cDNA library of hemocytes, with a high level of gene expression [30]. Similar patterns have been observed with P. clarkii [25], C. farreri [29], and L. vannamei [14], while in F. chinensis SPI expression has only been observed in the hepatopancreas [6]. The relative expression levels of MdSPI-1 showed higher levels of expression in the mantle than in other tissue, while MdSPI-2 was expressed mainly in the foot. It is possible that high level of MdSPI-1 expression in the mantle is due to the direct contact of this tissue with the external medium and consequently with pathogenic microorganisms, thus requiring the expression of genes that act as barriers to contagious pathogens. On the other hand, our results suggest that the higher level of expression of MdSPI-2 in the foot than in the mantle, gills and muscle could be related to the residue P1s found in this sequence, which correspond to Lys and Leu, where it can be inferred that their objective protease are trypsin and chymotrypsin, both proteins involved in digestion [31].

The relative expression of the mRNA of MdSPIs after the bacterial challenge had a time-dependent response, with statistically
significant differences from the control at all the exposure times. Both MdSPI-1 and MdSPI-2 transcripts were overexpressed after the injection with *V. anguillarum*. Similar results have been reported for *P. clarkii*, where *hcPcSI1* showed positive regulation in response to a bacterial challenge and negative regulation in response to a viral challenge. The authors suggested that the SPI response is associated with an innate defense system that is only antibacterial [32]. With respect to bivalve mollusks, the levels of mRNA of *CfKZSPI* in *C. farreri* are over-expressed 3 h after infection with *V. anguillarum*, decaying by 6 h and increasing again at 24 h [29]. In the case of *A. irradians*, it was determined that expression peaked at 16 h post-infection with *V. anguillarum* [13]. Consequently, the differences in expression of *MdSPI-1* and *MdSPI-2* at the analyzed time points could indicate that different transcription factors or activation cascades are regulate by the two transcripts. As well, the qPCR results suggest that in the *V. anguillarum* challenge, there is positive regulation of *MdSPIs* expression, suggesting its participation in the regulation of innate antibacterial system in *M. donacium*. In conclusion, the present study contributes relevant information at the molecular level about two genes associated with innate immunity in *M. donacium* derived from NGS transcriptome sequencing, and their relation with presence of marine pathogens such as *V. anguillarum*.

**Acknowledgments**

This work has been supported by FONDECYT (1120397) and FONDAP (15110027) from CONICYT – Chile.

**References**


[30] Ponprateep S, Tassanakajon A, Rimphaniatchayakit V. A Kazal type serine proteinase SPIPm2 from the black tiger shrimp Penaeus monodon is capable of neutralization and protection of hemocytes from the white spot syndrome virus. Fish & Shellfish Immunology 2011;31:1179–85.
