Comparative Biochemistry and Physiology, Part A 150 (2008) 320–325

Contents lists available at ScienceDirect

Comparative Biochemistry and Physiology, Part A

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

Molecular cloning and mRNA expression of cathepsin C gene in black tiger shrimp (*Penaeus monodon*)

Lihua Qiu, Shigui Jiang *, Jianhua Huang, Weifang Wang, Dianchang Zhang, Qiaer Wu, Keng Yang

The South China Sea Fisheries Research Institute, Chinese Academy of Fishery Sciences, Guangzhou, 510300, PR China

ARTICLE INFO

Article history:
Received 2 November 2007
Received in revised form 26 March 2008
Accepted 1 April 2008
Available online 7 April 2008

Keywords:
Molecular cloning
Cathepsin C
mRNA expression
Penaeus monodon

ABSTRACT

Cathepsin C (dipeptidyl-peptidase I, DPPI) is a lysosomal cysteine proteinase belonging to the papain superfamily, which is capable of removing dipeptides sequentially from the amino terminus of peptide and protein substrates. In the present study, the cDNA of a cathepsin C was cloned from black tiger shrimp *Penaeus monodon* (designated PmcathepsinC) by homology cloning and rapid amplification of cDNA ends (RACE) approaches. The full-length cDNA of PmcathepsinC consisted of 2051 nucleotides with a canonical polyadenylation signal sequence AATAAA and a poly(A) tail, and an open reading frame (ORF) of 1350 bp encoding a polypeptide of 449 amino acid residues with a predicted molecular weight of 50.0 kDa and theoretical isoelectric point of 5.65. The high identity of PmcathepsinC with Cathepsin C in other organisms indicated that PmcathepsinC should be a new member of the Cathepsin C family. By fluorescent quantitative real-time PCR, mRNA transcript of PmcathepsinC was detectable in all the examined tissues with higher level in ovary and heart. The temporal expression of PmcathepsinC mRNA in the hepatopancreas was up-regulated by lipopolysaccharide (LPS) stimulation and reached the maximum level at 4 h post-stimulation, and then dropped back to the original level gradually. These results indicated that PmcathepsinC was a constitutive and inducible acute-phase protein that perhaps involved in the immune defense of *P. monodon*.

© 2008 Elsevier Inc. All rights reserved.

1. Introduction

Cysteine proteinases are ubiquitously present in almost all life forms. These enzymes catalyze the hydrolysis of many proteins with different specificities and are considered to play an important role in intracellular protein degradation and turnover (McDonald et al., 1969; Coffey and de Duve, 1968; Katunuma and Kominami, 1989). Cathepsin C (dipeptidyl-peptidase I, DPPI) is a lysosomal cysteine proteinase that belongs to the papain superfamily (Turk et al., 1997). It is capable of removing dipeptides sequentially from the amino terminus of peptide and protein substrates (McDonald et al., 1969; Metrione and MacGeorge, 1975). This enzyme has been postulated to involved in cell growth (Doughty and Gruenstein, 1987), alimentary tract (Ishidoh et al., 1991), neuraminidase activation (D'Agrosa and Callahan, 1988), and platelet factor XIII activation (Lynch and Pfueller, 1988). It is known that cysteine proteinases of the papain family except cathepsin C are monomers consisting of R- and L-domains (Turk et al., 1997), whereas cathepsin C has a unique structure, consisting of four identical subunits, each composed of three polypeptide chains: the pro-region, the heavy and the light chain (Dolenc et al., 1995; Cigić et al., 1998; Cigić et al., 2000). It is controversial whether these chains are joined by a disulfide bond (Dolenc et al., 1995; Cigić et al., 1998); however, it has been postulated that the pro-region contains an intramolecular disulfide bond(s) and that it is glycosylated (Cigić et al., 1998; Cigić et al., 2000). For cathepsin C to exert its proteinase activity, it is crucial that the pro-cathepsin C is to be cleaved by a certain proteinase such as cathepsin L or S (Dahl et al., 2001), followed by forming such an oligomeric structure.

Recent studies have demonstrated the existence of cathepsin C in many animals, such as: *Xenopus laevis* (GenBank accession No. BC056109), *Danio rerio* (GenBank accession No. NM-214722), *Marsupenaeus japonicus* (GenBank accession No. AB104735). But there are very few reports on the function of the cathepsin C in crustaceans. In kuruma prawn (*M. japonicus*), it was demonstrated that cathepsin C was involved in the final stage of oocyte maturation (Qiu et al., 2005). There are no reports on the gene function in crustacean immunology so far.

Shrimp have no acquired adaptive immune system. Their defense is believed to depend entirely on an innate, non adaptive mechanism to resist pathogen invasion (Gross et al., 2001) and inoculation against viruses has no effect. Understanding the interaction between host and pathogen will be helpful in controlling infectious diseases in shrimp. Although an increasing number of immune function related genes in shrimp have been reported, the genes involved in immune response against infection still remain unclear. Both cathepsin C and L are lysosomal cysteine proteases, but cathepsins L was highly expressed

* Corresponding author. Biotechnology and Aquaculture Laboratory, The South China Sea Fisheries Research Institute, Chinese Academy of Fishery Sciences, 231 Xingangxi Road, Guangzhou 510300, PR China. Tel.: +86 20 89108 338 or 316; fax: +86 20 84451442.

E-mail address: jiangsg@21cn.com (S. Jiang).

1095-6433/$ – see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.cbpa.2008.04.001

during the WSSV infection in pacific white shrimp (*Litopenaeus vannamei*), suggesting that it is involved in the defense response against the virus (Zhao et al., 2007). So we suppose cathepsin C might have other functions in crustacean we do not know.

The main objectives of this study are (1) to clone the full-length cDNA of cathepsin C from black tiger shrimp, (2) to investigate the expression pattern of cathepsin C gene in the tissues, (3) to provide information about if LPS could induce the expression of cathepsin C in black tiger shrimp.

2. Materials and methods

2.1. Animals and immune challenge

The healthy black tiger shrimp (*P. monodon*), weighing about 60–300 g, were purchased from Sanya, Hainan province, PR China. Sixty shrimp were maintained in aerated seawater (salinity 30 ppt) at 24–25 °C for three days before processing.

For gene cloning, three shrimp weighing about 300 g were employed and kept in the tank. 200 μL of LPS solution (100 ng mL^{-1}, resuspended in water) was injected into the muscle of each shrimp. 6 h later, the ovary from the three shrimp was collected, mixed, and subjected to total RNA extraction.

For the challenge experiment, forty shrimp were employed. 50 μL LPS (100 ng mL^{-1}) was injected into the muscle of each shrimp and they were used as the stimulated group. The untreated shrimp and shrimp injected with 50 μL water were used as the blank and the control group, respectively. The injected shrimp were returned to seawater tanks, and 3 individuals from the blank, control and stimulated group respectively were randomly collected at 2, 4, 6, 8, 10, and 16 h post-injection. At each time point, the hepatopancreas from the three individuals was collected and mixed. They were subjected to total RNA extraction.

2.2. RNA isolation and cDNA synthesis

Total RNA was isolated from hepatopancreas, ovary, heart, muscle, blood and brain (weight about 50 mg) of three shrimp by using Trizol reagent (Invitrogen, USA). The first-strand cDNA was synthesized from 2 μg of DNase-treated (Promega RQ1 DNase I) total RNA by M-MLV reverse transcriptase (Promega, USA) at 42 °C for 50 min with oligo-dT-adaptor primer (Table 1). The cDNA was used as the template for PCR reactions in gene cloning and expression analysis.

2.3. Gene cloning and sequence analysis

A pair of degenerated primers, dXCF (nucleotide position 266–287 bp) and dXCR (nucleotide position 765–787 bp), was designed based on the conserved regions of cathepsin C gene sequences from *X. laevis* (GenBank accession No. BC056109), *D. rerio* (GenBank accession No. NM-214722), *Homo sapiens* (GenBank accession No. X87212), *M. japonicus* (GenBank accession No. AB104735). The initial PCR was performed with dXCF and dXCR in a 25 μL reaction volume containing 2.5 μL of 10×PCR buffer, 1.5 μL of $MgCl_2$ (25 mmol L^{-1}), 2.0 μL of dNTP (2.5 mmol L^{-1}), 1.0 μL of each primer (10 μmol L^{-1}), 15.8 μL of PCR-grade water, 0.2 μL of Taq polymerase (Takara) (5 U μL^{-1}) and 1 μL of cDNA mix. The PCR temperature profile was 94 °C for 5 min followed by 34 cycles of 94 °C for 40 s, 57 °C for 40 s, 72 °C for 1 min and the final extension step at 72 °C for 10 min. The obtained PCR products (about 500 bp) were separated by 1.2% agarose gel, and then purified by an agarose gel purification kit. The purified PCR product was ligated with the PMD20-T vector (Takara, Japan), and transformed into the competent *Escherichia coli* cells. The recombinants were identified through blue-white color selection and screened with M13 forward and reverse primers. Three of the positive clones were sequenced on an ABI3730 Automated Sequencer (Applied Biosystem, USA). Sequences generated were analyzed for similarity with other known sequences using the BLAST programs (http://www.ncbi.nim.nih.gov/).

With the obtained partial sequence, three specific primers, sense primer F1 and reverse primers R1 and R2 (Table 1), were designed to clone the full length of PmcathepsinC by rapid amplification of cDNA ends (RACE) methods. In 3′ RACE-PCR, PCR reaction was performed with primer F1 (nucleotide position 311–330 bp) and adaptor primer (Table 1). The PCR profile was as follows: 94 °C, 5 min; 94 °C, 45 s; 60 °C, 30 s; 72 °C, 45 s; 35 cycles; 72 °C, 10 min. For 5′ RACE-PCR, the first-strand cDNA was tailed with poly (C) at the 5′ end using terminal deoxynucleotidyl transferase (Takara, Japan). PCR was performed initially with primer R1 (nucleotide position 364–384 bp) and Oligo-dG, followed by semi-nested PCR with R2 (nucleotide position 308–327 bp) and Oligo-dG. The PCR products were gel-purified and subjected to sequencing. The resulting sequences were verified and subjected to cluster analysis.

Multiple sequence alignments were performed using the CLUSTAL W program at the European Bioinformatics Institute (http://www.ebi.ac.uk). Analyses of the deduced amino acid sequences were conducted using the programs PSORT (Kenta Nakai, National Insitute Basic Biology), Scan Prosite (EXPASy Molecular Biology Server) and Predict Protein (EMBL-Heidelberg). The phylogenetic tree was constructed by the neighbor-joining (NJ) method using the programs of CLUSTAL X1.83 (Thompson et al., 1997) and MEGA3.1 (Kumar et al., 2004).

2.4. Quantification of PmcathepsinC expression by quantitative real-time PCR

Real-time quantitative PCR was performed with the SYBR Green 2×Supermix (Applied Biosystems, USA) on an ABI 7300 Real-Time Detection System (Applied Biosystems, USA) to investigate the expression of PmcathepsinC. Two specific primers, rccF and rccR (Table 1) were used to amplify a PCR product of 240 bp. β-actin (GenBank accession No. EF087977) was chosen as the reference gene for internal standardization (Zhao et al., 2007). Two β-actin primers ractinF and ractinR (Table 1) were used to amplify a β-actin gene fragment of 110 bp as the internal control for qRT-PCR. The qRT-PCR amplifications were carried out in triplicates in a total volume of 20 μL containing 10 μL of 2×Supermix (Applied Biosystems, USA), 5 μL of the 1:5 diluted cDNA, 1 μL each of forward and reverse primer and 3 μL PCR-grade water, The qRT-PCR program was 50 °C for 2 min, 95 °C for 10 min, followed by 40 cycles of 94 °C for 15 s, 61 °C for 30 s, 72 °C 30 s. Melting curve analysis of amplification products was performed at the end of each PCR reaction to confirm that only one PCR product was amplified and detected. After the PCR program, qRT-PCR data from three replicate samples were analyzed with a 7300 System SDS Software v1.3.0 (Applied Biosystems, USA) to estimate transcript copy numbers for each sample. To maintain

Table 1
Oligonucleotide primers used

Primer name	(5′→3′) nucleotide sequence
dXCF	T(C/A/G)A TCT A(T/C)A ACCAGG G(T/C)T T(C/T)G A
dXCR	CCAT(A/G)GA(A/G)GCAAA(G/A)G(C/A)GTA(G/A)CA
F1	ACTTCGCTTTCTCGCTCTGG
R1	CAGCCCGTCAGTGTCTTGTC
R2	GAGCGAGAAAGCGAAGTAGG
Oligo-dT adaptor	GGCCACGCGACTAGTAC$(T)_{16}$
Adaptor	GGCCACGCGACTAGTAC
Oligo-dG	GGGGGGGGGGGGGG
Ractin F	GCCCTTGCTCCTTCCACTATC
Ractin R	CCGGACTCTTCGTACTCATCCT
rccF	GATTCTGACCAGCAACCACCA
rccR	TACAGGCTCCATAGTAACCTCCAA

consistency, the baseline was set automatically by the software. The comparative C_T method was used to analysis the expression level of black tiger shrimp cathepsin C. The C_T for the target amplification of cathepsin C and the C_T for the internal control β-actin were determined for each sample. Differences between the C_T for the target and the internal control, called $^{\Delta}C_T$, were calculated to normalize the differences in the amount of total nucleic acid added to each reaction and the efficiency of the RT-PCR. The blank group was used as the reference sample, called the calibrator. The $^{\Delta}C_T$ for each sample was subtracted from the $^{\Delta}C_T$ of the calibrator; the difference was called $^{\Delta\Delta}C_T$ value. The expression level of black tiger shrimp cathepsin C could be calculated by $2^{-\Delta\Delta C_T}$, and the value stood for an *n*-fold difference relative to the calibrator. The average cycle threshold (C_T) measurement for the three determinations were used in calculations of relative expression using β-actin as the internal control. The data obtained from RT-PCR analysis were subjected to one-way analysis of variance (one-way ANOVA) followed by an unpaired, two-tailed *t*-test. Differences were considered significant at $P<0.05$.

Table 2
Homology of PmcathepsinC with other cathepsin C proteins

	Identities (%)				GenBank accession number
	Full-length black tiger shrimp cathepsin C	Propeptide	Heavy chain	Light chain	
P. monodon	100	100	100	100	EU026137
M. japonicus	90	89	92	95	AB104735
P. troglodytes	52	38	63	79	XM-508684
H. sapiens	52	39	62	79	X87212
B. Taurus	51	38	60	78	AB060542
G. gallus	51	36	62	77	XM-417207
D. rerio	51	36	60	79	NM-214722
X. laevis	52	35	61	80	BC056109

```
                                     CCCAGAGTCCCAGAGTCAGCTGATTCA    27
ATGCTCCTGTGGGTCCTAGCAGCACTGTCCCTCGCGGTACGAGTCACGTGGGCTGACACG    87
M  L  L  W  V  L  A  A  L  S  L  A  V  R  V  T  W  A  D  T      20
CCCGCGAACTGCACTTACGACGACATCAGAGGAGACTGGTTGCTGTATGAGACCGAGAGA    147
P  A  N  C  T  Y  D  D  I  R  G  D  W  L  L  Y  E  T  E  R      40
ACTGGTGATGCTGGAATTGACTGTGAAGACATGGGCCCGAATATCCACAGGACGGAAGTG    207
T  G  D  A  G  I  D  C  E  D  M  G  P  N  I  H  R  T  E  V      60
AGCTTAATGTATCCCAACATAGCAGTCGATGAGTATGGAAATCGAGGGACTTGGACCATG    267
S  L  M  Y  P  N  I  A  V  D  E  Y  G  N  R  G  T  W  T  M      80
ATCTATAACCAGGGTTTCCAAAACAAGGTCGCAGGCAGATCGTACTTCGCTTTCTCGCTC    327
I  Y  N  Q  G  F  Q  N  K  V  A  G  R  S  Y  F  A  F  S  L      100
TGGGAAAAGACAGGCGACACCGTGCTGTCCATCTGCGACAAGACACTGACGGGCTGGTCG    387
W  E  K  T  G  D  T  V  L  S  I  C  D  K  T  L  T  G  W  S      120
AGGGACATCACCGTCAGGAACTGGGCCTGCTACAGGGCGCAGAAGAAGGCGGCCGCGCCT    447
R  D  I  T  V  R  N  W  A  C  Y  R  A  Q  K  K  A  A  A  P      140
CACAAGACGCACGTCCTCCAGCGAATGGCTCAGGAGAACTCCCTTTACAAGTACAACAAG    507
H  K  T  H  V  L  Q  R  M  A  Q  E  N  S  L  Y  K  Y  N  K      160
GACCTGATCAGCGCCATCAACTCGCAACAAAGCAGTTGGGTCGCTGACGTGTACCCGCAG    567
D  L  I  S  A  I  N  S  Q  Q  S  S  W  V  A  D  V  Y  P  Q      180
TATGAGACTATGACCCTTATGGACCACCTCAGGCGCAACGGGGGAAGAGCGTCCGCCGTC    627
Y  E  T  M  T  L  M  D  H  L  R  R  N  G  G  R  A  S  A  V      200
GCCAGTCGAGCTCGCTCAGGACCTGTAAGCTTCATGACCCGCCTGCGTGACCTGTCTCTG    687
A  S  R  A  R  S  G  P  V  S  F  M  T  R  L  R  D  L  S  L      220
CCCGAGGCGTGGGACTGGCGCAACGTGAGCGGGGTCAACTACGTGTCTCCGGTCAGGAAC    747
P  E  A  W  D  W  R  N  V  S  G  V  N  Y  V  S  P  V  R  N      240
CAGGGCAACTGTGGATCCTGCTATGCCTTTGCCTCCATGGGTGGCCTAGAATCTCGAGTT    807
Q  G  N  C  G  S  C  Y  A  F  A  S  M  G  G  L  E  S  R  V      260
AGGATTCTGACCAGCAACCACCAGAAGCCAGTTTTTGCCCCGCAGGATATCGTAGGATGT    867
R  I  L  T  S  N  H  Q  K  P  V  F  A  P  Q  D  I  V  G  C      280
TCCAAGCTCTCTCAAGGCTGTGAGGGTGGTTTCCCGTTCCTTATCGCTGGCAGGTATGCA    927
S  K  L  S  Q  G  C  E  G  G  F  P  F  L  I  A  G  R  Y  A      300
CAAGATGTTGGTGTTGTGCTTGAGGACTGTTACCCCTATGAAGGCAAAGATGACACATGC    987
Q  D  V  G  V  V  L  E  D  C  Y  P  Y  E  G  K  D  D  T  C      320
GTAAGGACTACCTGCACCAAACATTACACAGCTTATTACCGTTATGTTGGAGGTTACTAT    1047
V  R  T  T  C  T  K  H  Y  T  A  Y  Y  R  Y  V  G  G  Y  Y      340
GGAGCCTGTAATGAAGAGGAGATGAAACTAGCTCTAATCAAGGGTGGACTTCTCATCGTG    1107
G  A  C  N  E  E  E  M  K  L  A  L  I  K  G  G  L  L  I  V      360
GGTTTGGAAGTCTACGATGATTTCCTCCACTACAAAGGCGGCATTTACCATCACACAGGT    1167
G  L  E  V  Y  D  D  F  L  H  Y  K  G  G  I  Y  H  H  T  G      380
CTCCAGGACAGATTTAATCCTCTTGAGCTGACGAATCACGCAGTATTACTCGTCGGATAC    1227
L  Q  D  R  F  N  P  L  E  L  T  N  H  A  V  L  L  V  G  Y      400
GGGGAAGACGAGGCAACTGGAGAAAAGTACTGGAGCGTTAAGAATTCCTGGGGTGAAGAC    1287
G  E  D  E  A  T  G  E  K  Y  W  S  V  K  N  S  W  G  E  D      420
TGGGGAGAGGATGGTTACTTTAGGATTCGTCGTGGCGTTGACGAGTGTGCCATCGAGTCC    1347
W  G  E  D  G  Y  F  R  I  R  R  G  V  D  E  C  A  I  E  S      440
ATGGCTGTGGAAGCTGTGCCCATCCCATAA                                  1377
M  A  V  E  A  V  P  I  P  *                                    449
ATTTGTAGAGCAAAAAATAGATAATGAGAATAATTATAAGATGTGTTATGAAAGATTTAA    1437
ATGTTGCATTTGGCTTTTTTGGCAAGTTGTGACAGACTTACATTGGTAATGTTGCCTGAC    1497
AGAATTACTTGTTTTCTTAATTAGATAAATATAAAACTTTTTAAAAATGTCTTCTCAAAT    1557
TTAAAAAAATATATATTCAGAGCAATTTTTTTCATGTAGCAGTAATCAAAATATTCTCAT    1617
TATATTGGATCCTCACACAAATAGTTACTAATTAAAGAAGACAATAAATAAAGCATAAGT    1677
TAAGAATAGATTTCGAGACAAATTGAACTGCAATAATTTTGTGTTTTTATGATTGTAATT    1737
CCTTATGGGAAATTGCCCTTTTAACAACAACAAAAAGAATTTGGGATCCATAGGCTCATC    1797
TTGCCACTGCATATATTATGATGTATATATCTTTTCCAGCATATTGCTTCTTATTATTTT    1857
CTTTTTAATGTTTAGAGAGGTACATATTATCAACAATATTTATTGTTTATATACTGTCA    1917
TTGTATAACATCTTAAGATACATTTTCTACAATTCTTTTTATATAATAGTCAACAATAAT    1977
TATCATCCATGTAAAAGATAAAAAAATATATGCATTCAAAATAAATATTTCATAATACAA    2037
AAAAAAAAAAAAAA                                                  2051
```

Fig. 1. Nucleotide and deduced amino acid sequences of PmcathepsinC. The polyadenylation signal sequence was highlighted and underlined; the start code was in the bold and the asterisk showed the stop code.

3. Results

3.1. cDNA cloning and sequence analysis of PmcathepsinC gene

A 2051 bp nucleotide sequence representing the complete cDNA sequence of PmcathepsinC was obtained by overlapping the three fragments. The sequence was deposited in GenBank under accession No. EU026137 and the deduced amino acid sequence is shown in Fig. 1. The complete sequence of PmcathepsinC cDNA contained a 5′ untranslated region (UTR) of 27 bp, a 3′ UTR of 674 bp with a canonical polyadenylation signal sequence AATAAA and a polyA tail, and an open reading frame (ORF) of 1350 bp encoding a polypeptide of 449 amino acids with predicted molecular mass of 50.0 kDa and theoretical isoelectric point of 5.65.

3.2. Homology and phylogenetic analysis of PmcathepsinC

The deduced amino acid sequence of PmcathepsinC showed very high homology with the sequences of *M. japonicus* (90% Identity), *D. rerio* (51% Identity), *X. laevis* (52% Identity) and even the mammalians *H. sapiens* (52% Identity), *B. taurus* (51% Identity) (Table 2). The software analysis indicated that there existed a putative signal peptide of 18 amino acids (position 1–18aa), a long propeptide of 201 amino acids (position 19–219aa) and a putative mature peptide region of 230 (position 220–449aa) in PmcathepsinC. The mature protein, comprised of a heavy chain (position 220–382aa) and a light chain (position 383–449aa), contained three catalytic active sites (Cys^{247}, His^{393} and Asn^{415}) that were highly conserved in all papain family members (Fig. 2). The N-terminus of the propeptide and the heavy and light chains of the mature peptide were all well conserved as compared with those of the other animals, and a high identity was detected in the heavy and light chains of the mature peptide region (Table 2, Fig. 2).

Based on the amino acid sequence of cathepsin C, a phylogenetic tree was constructed by using the programs of CLUSTAL X1.83 and MEGA2.1 (Fig. 3). All the vertebrate cathepsin C genes and invertebrate cathepsin C genes were clustered together. In the phylogenetic tree, the black tiger shrimp shows the closest relationship with the *M. japonicus*, the result is similar with the result of the BLAST. The relationships revealed in the phylogenetic tree were in agreement with the concept of traditional taxonomy.

3.3. Tissue distribution of the PmcathepsinC transcripts

Real-time quantitative PCR was employed to quantify the PmcathepsinC expression in the tissues of hepatopancreas, ovary, heart, muscle, blood and brain. The amplification specificity for PmcathepsinC and β-actin was determined by analyzing the dissociation curves. Only one peak presented in the dissociation curves for both the

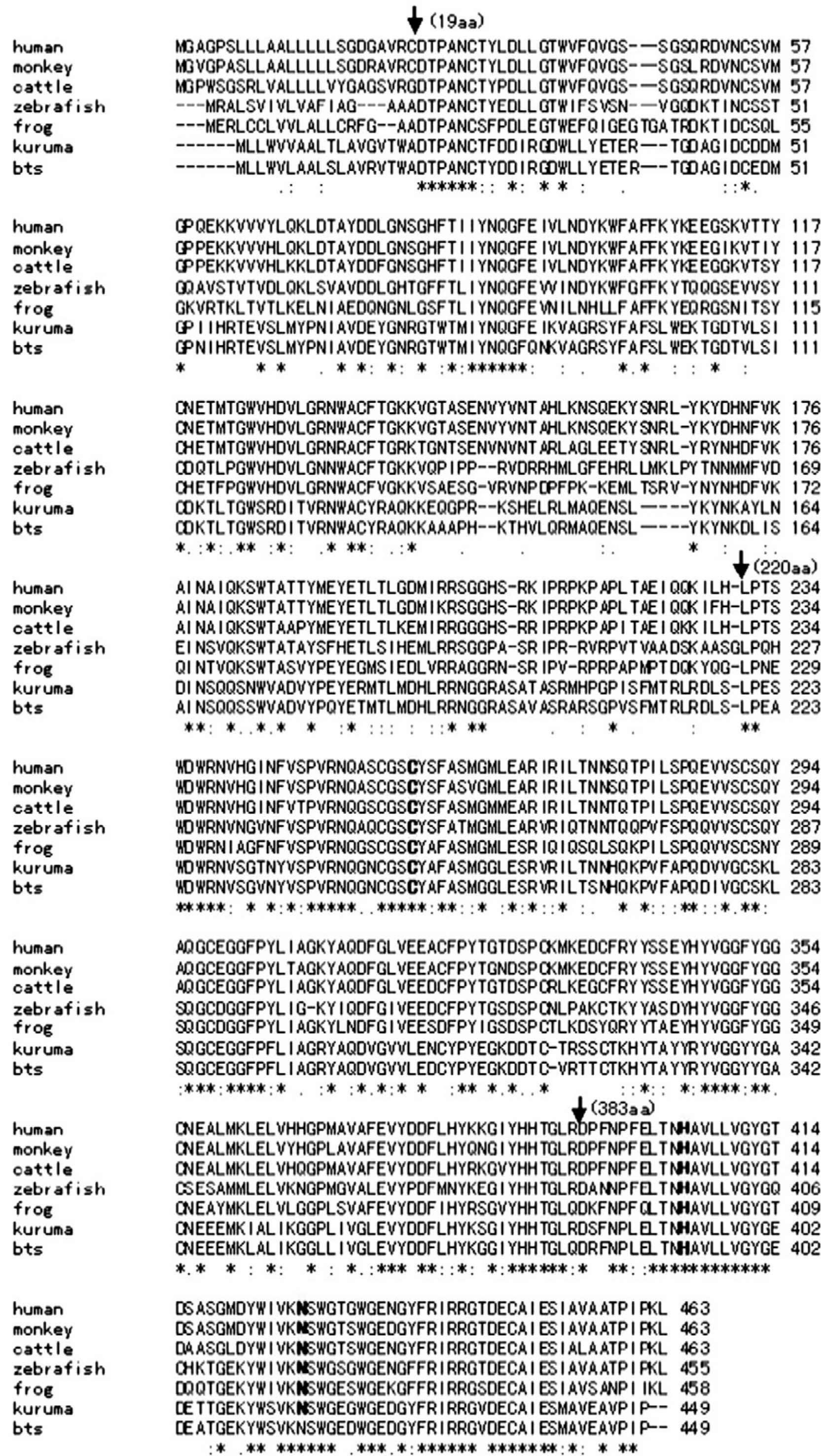

Fig. 2. Multiple alignments of PmcathepsinC with other known cathepsin C proteins. Identical and similar sites are shown with asterisk (*) and dots (. or :), '*' indicates identical residues, ':' indicates residues with more similar properties, '.' indicates residues with some similar properties, the blank indicates residues with opposite properties; the arrows indicate the cleavage sites of signal peptide, propeptide, the heavy and light chains of the mature peptide; the conserved catalytic residues, cysteine, histidine and asparagines, are highlighted and bolded. bts, black tiger shrimp; kuruma, kuruma prawn.

PmcathepsinC and β-actin gene (data not shown), indicating that the amplifications were specific.

PmcathepsinC mRNA was found to be constitutively expressed in all the examined tissues with significant variation of expression level. There was a high-level expression of PmcathepsinC in ovary, heart and hepatopancreas, while a low-level expression in blood, muscle and brain. The highest level of cathepsin C expression was detected in ovary (Fig. 4).

3.4. Quantification of PmcathepsinC mRNA expression after LPS stimulation

The temporal expression of the PmcathepsinC transcript in hepatopancreas of shrimps after LPS stimulation was shown in Fig. 5. During the first 2 h after LPS stimulation, the cathepsin C mRNA remained at a low level. At 4 h after stimulation, the expression of the PmcathepsinC was up-regulated and there was a significant increase in

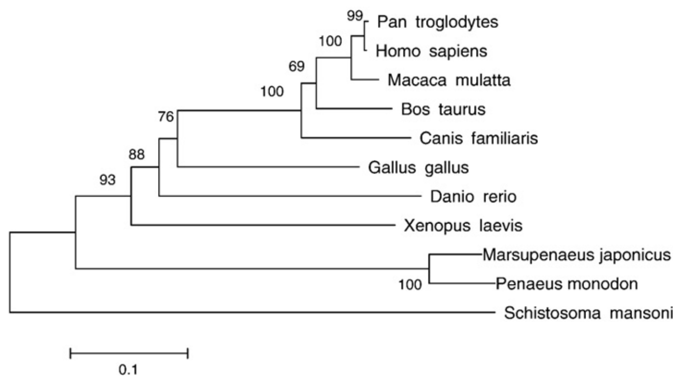

Fig. 3. Phylogenetic tree constructed by the neighbor-joining method based on the cathepsin C sequences. Bootstrap support values for the NJ tree are shown at the nodes (out of 1000 replicates).

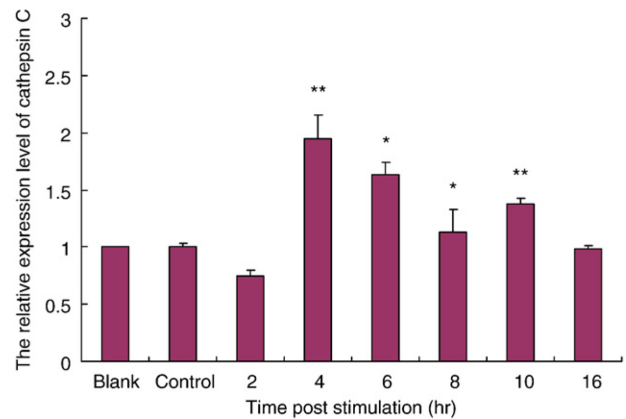

Fig. 5. The temporal expression of PmcathepsinC transcript in hepatopancreas after LPS challenge. The mRNA expression of PmcathepsinC and β-actin was measured at 2, 4, 6, 8, 10, 16 h post-injection and blank groups. Vertical bars represented the mean±S.E. ($N=3$). Significant differences across control were indicated with an asterisk at $P<0.05$, and with two asterisks at $P<0.01$.

the relative abundance of PmcathepsinC mRNA. At 4 and 6 h post-LPS stimulation, the cathepsin C gene expression level was 2.1 and 1.9 fold higher than that observed in the control group, respectively. As time progressed, the expression of PmcathepsinC mRNA decreased and recovered to the original level at 16 h post-stimulation. An unpaired, two-tailed *t*-test with blank and challenged groups showed statistically significant difference in PmcathepsinC gene expression at 2 h and 4 h ($P<0.05$) post-stimulation. However, no significant difference was observed in other time point in challenge group (Fig. 5).

4. Discussion

In recent years, several papers about cathespins L in the shrimp had been reported (Hu and Leung, 2004, 2007; Zhao et al., 2007). In Pacific white shrimp (*L. vannamei*), cathespins L was highly expressed during the WSSV infection, suggesting that it is involved in the defense response against the virus (Zhao et al., 2007). In *Metapenaeus ensis*, cathepsin L was cloned and was also expressed in a wide array of tissues, especially in the hepatopancreas (Hu and Leung, 2004, 2007). Cathepsins L and C are all belong to the papain superfamily. But only one cathespins C gene from *M. japonicus* (GenBank accession No. AB140735) was deposited in the GenBank (Qiu et al., 2005).

In the present study, the full-length cDNA encoding cathepsin C (named as PmcathepsinC) was cloned from the black tiger shrimp *Penaeus monodon*. Alignment analysis indicated that the deduced amino acid sequence of black tiger shrimp cathepsin C was homologous to other known cathepsin C. As in mammalian species, the black tiger shrimp pro-cathepsin C protein predicted from the cDNA sequence consists of three functional domains: a signal peptide, a uniquely long propeptide and a mature peptide (including a heavy and light chain). The three amino acid residues known as putative catalytic active sites in all papain family members were detected in the deduced amino acids of black tiger shrimp cathepsin C and also conserved across the invertebrate and vertebrate species. In the deduced amino acids, no potential glycosylation site for post-translational modification was found. However, in kuruma prawn cathepsin C two glycosylation sites were detected (Qiu et al., 2005), in mammalian cathepsin C three to four sites were described (Dahl et al., 2001; Pham et al., 1997) and five in the flatworm (Brindley et al., 1997).

It had been reported that cathepsin C in mammals was ubiquitously expressed in tissues, particularly high in lung, kidney, placenta, liver, spleen, and intestines (Qiu et al., 2005; Kominami et al., 1992; Rao et al., 1997). In the present study, the mRNA expression of PmcathepsinC could be detected in all studied tissues with the higher level in ovary and heart. The expression pattern of black tiger shrimp cathepsin C in tissues was somewhat different from that of cathepsin Cs in other animals. In mouse, cathepsin C is abundant in the liver, and is little detectable in the heart and brain (Pham et al., 1997). Human cathepsin C was expressed at a high level in lung and kidney, and at low level in heart and liver, while not expressed in the brain (Pham et al., 1997; Rao et al., 1997; Smyth et al., 1995). So the result indicated that although cathepsin C gene could be expressed constitutively in different tissues, the expression level might be different with the species of the animals. The physiological roles of cathepsin C have been found to be not only confined to intracellular protein degradation, but also involved in cell growth (Metrione and MacGeorge, 1975) and the activation of several granule-associated serine proteases, such as granzymes A and B, cathepsin G, neutrophil elastase and chymases (Wolters et al., 2001; Pham and Ley, 1999; McGuire et al., 1993; Nauland and Rijken, 1994). Cathepsin C is thus thought to be one of the major processing enzymes known so far (Turk et al., 2001). In kuruma prawn, cathepsin C gene was found to be an up-regulated gene involved in the final stages of oocyte maturation, and displayed a significantly high expression level at the early cortical rod stage (Qiu et al., 2005). In our results, the expression level of PmcathepsinC was highest in the ovary, not the immune tissues. So we deduced that the PmcathepsinC perhaps possessed the same function as the kuruma prawn cathepsin C in regulating the ovary growth.

The innate immune system is the first line of defense to protect the host in the first hours to days of infection (Lee and Soderhall, 2002). The first distinct phase of the immune response in shrimp is approximately in the first 12 h after challenge (Bachere et al., 2004). In order to better understand the response of black tiger shrimp

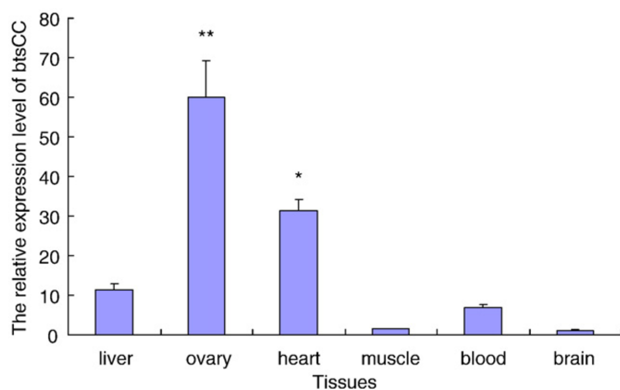

Fig. 4. The expression level of PmcathepsinC transcript in different tissues. Vertical bars represented the mean±S.E. ($N=3$). hep, hepatopancreas. Significant differences across calibrator (brain) were indicated with an asterisk at $P<0.05$, and with two asterisks at $P<0.01$.

cathepsin C exposed to the potential pathogens, its expression in hepatopancreas after stimulation with LPS at different time point was investigated by real-time RT-PCR in the present study. LPS acts as a powerful stimulator of innate immunity in diverse eukaryotic species (Ulevitch and Tobias, 1995; Lemaitre et al., 1996). Some marine invertebrates such as the Atlantic horseshoe crab and the giant African snail were also highly sensitive to LPS (Iwanaga, 2002; Biswas and Mandal, 1999). In the present study, after LPS treatment, the expression level of black tiger shrimp cathepsin C was not changed significantly during the first 2 h after LPS stimulation, and then up-regulated and increased significantly at 4 h after LPS stimulation. The expression level of black tiger shrimp cathepsin C at 4 h post-LPS stimulation was the highest, and from the 6 h post-LPS stimulation the expression level became to decrease. The result indicted that black tiger shrimp cathepsin C was a constitutive and inducible acute-phase protein. In invertebrates, cysteine proteases represent a major component of the lysosomal proteolytic system, and are responsible for intracellular protein degradation (Knop et al., 1993). In 2007, there were reports about that cathepsin L was responsive to virus challenge (Robalino et al., 2007; Zhao et al., 2007). Although there are no studies reporting that the cathepsin C gene could take part in the immune response, in our study we found that LPS could induce the expression of the cathepsin C gene in the black tiger shrimp, suggesting that it was involved in the defense response against bacterial infections. However, the functioning of the gene in disease resistance should be explored further in a more comprehensive study.

Acknowledgments

This study was supported by the National Science Foundation of China (No. 30571447), National “863” Project of China (No. 2006AA10A406), Agriculture Department Project of China (06-05-01B) and National Project of China (2006BAD01A13).

References

Bachere, E., Gueguen, Y., Gonzalez, M., de Lorgeril, J., Garnier, J., Romestand, B., 2004. Insights into the anti-microbial defense of marine invertebrates: the penaeid shrimps and the oyster Crassostrea gigas. Immunol. Rev. 198, 149–168.

Biswas, C., Mandal, C., 1999. The role of amoebocytes in endotoxin-mediated coagulation in the innate immunity of *Achatina fulica* snails. Scand. J. Immunol. 49, 131–138.

Brindley, P.J., Kalinna, B.H., Dalton, J.P., Day, S.R., Wong, J.Y., Smythe, M.L., cManus, D.P., 1997. Proteolytic degradation of host hemoglobin by schistosomes. Mol. Biochem. Parasitol. 89, 1–9.

Cigic´, B., Kriz˘aj, I., Kralj, B., Turk, V., Pain, R.H., 1998. Stoichiometry and eterogeneity of the pro-region chain in tetrameric human cathepsin C. Biochim. Biophys. Acta 1382, 143–150.

Cigic´, B., Dahl, S.W., Pain, R.H., 2000. The residual pro-part of cathepsin C fulfills the criteria required for an intramolecular chaperone in folding and stabilizing the human proenzyme. Biochemistry 39, 12382–12390.

Coffey, J.W., de Duve, C., 1968. Digestive activity of lysosomes. I. The digestion of proteins by extracts of rat hepatopancreas lysosomes. J. Biol. Chem. 243, 3255–3263.

D'Agrosa, R.M., Callahan, J.W., 1988. In vitro activation of neuraminidase in the beta-galactosidase–neuraminidase–protective protein complex by cathepsin C. Biochem. Biophys. Res. Commun. 157, 770–775.

Dahl, S.W., Halkier, T., Lauritzen, C., Dolenc, I., Pedersen, J., Turk, V., Turk, B., 2001. Human recombinant pro-dipeptidyl peptidase I (cathepsin C) can be activated by cathepsins L and S but not by autocatalytic processing. Biochemistry 40, 1671–1678.

Dolenc, I., Turk, B., Pungercic, G., Ritonja, A., Turk, V., 1995. Oligomeric structure and substrate induced inhibition of human cathepsin C. J. Biol. Chem. 270, 21626–21631.

Doughty, M.J., Gruenstein, E.I., 1987. Cell growth and substrate effects on characteristics of a lysosomal enzyme (cathepsin C) in Duchenne muscular dystrophy fibroblasts. Biochem. Cell. Biol. 65, 617–625.

Gross, P.S., Bartlett, T.C., Browdy, C.L., Chapman, R.W., Warr, G.W., 2001. Immune gene discovery by expressed sequence tag analysis of hemocytes and hepatopancreas in the Pacific white shrimp, *Litopenaeus vannamei*, and the Atlantic white shrimp, *L. setiferus*. Dev. Comp. Immunol. 25, 565–577.

Hu, K.J., Leung, P.C., 2004. Shrimp cathepsin L encoded by an intronless gene has predominant expression in hepatopancreas, and occurs in the nucleus of oocyte. Comp. Biochem. Physiol., B 137, 21–33.

Hu, K.J., Leung, P.C., 2007. Food digestion by cathepsin L and digestion-related rapid cell differentiation in shrimp hepatopancreas. Comp. Biochem. Physiol., B 146, 69–80.

Ishidoh, K., Muno, D., Sato, N., Kominami, E., 1991. Molecular cloning of cDNA for rat cathepsin C. Cathepsin C, a cysteine proteinase with an extremely long propeptide. J. Biol. Chem. 266, 1632–16317.

Iwanaga, S., 2002. The molecular basis of innate immunity in the horseshoe crab. Curr. Opin. Immunol. 14, 87–95.

Katunuma, N., Kominami, E. (Eds.), 1989. Intracellular Proteolysis: Mechanisms and Regulations. Japan Scientific Press, Tokyo, Japan, pp. 3–23.

Kominami, E., Ishidoh, K., Muno, D., Sato, N., 1992. The primary structure and tissue distribution of cathepsin C. Biol. Chem. Hoppe-Seyler 373, 367–373.

Knop, M., Schiffer, H.H., Rupp, S., Wolf, D.H., 1993. Vacuolar/lysosomal proteolysis: proteases, substrates mechanisms. Curr. Opin. Cell Biol. 5, 990–996.

Kumar, S., Tamura, K., Nei, M., 2004. MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment. Brief. Bioinform. 5, 150–163.

Lee, S.Y., Soderhall, K., 2002. Early events in crustacean innate immunity. Fish Shellfish Immunol. 12, 421–437.

Lemaitre, B., Nicolas, E., Michaut, L., Reichhart, J.M., Hoffmann, J.A., 1996. The dorsoventral regulatory gene cassette spatzle/Toll/cactus controls the potent antifungal response in *Drosophila* adults. Cell 86, 973–983.

Lynch, G.W., Pfueller, S.L., 1988. Thrombin-independent activation of platelet factor XIII by endogenous platelet acid protease. Thromb. Haemost. 59, 372–377.

McDonald, J.K., Callahan, P.X., Zeitman, B.B., Ellis, S., 1969. Inactivation and degradation of glucagon by dipeptidyl aminopeptidase I (cathepsin C) of rat hepatopancreas. J. Biol. Chem. 244, 6199–6208.

McGuire, M.J., Lipsky, P.E., Thiele, D.L., 1993. Generation of active myeloid and lymphoid granule serine proteases requires processing by the granule thiol protease dipeptidyl peptidase I.J. Biol. Chem. 268, 2458–2467.

Metrione, R.M., MacGeorge, N.L., 1975. The mechanism of action of dipeptidyl aminopeptidase. Inhibition by amino acid derivatives and amines; activation by aromatic compounds. Biochemistry 14, 5249–5252.

Nauland, U., Rijken, D.C., 1994. Activation of thrombin-inactivated singlechain urokinase-type plasminogen activator by dipeptidyl peptidase I (cathepsin C). Eur. J. Biochem. 223, 497–501.

Pham, C.T.N., Ley, T.J., 1999. Dipeptidyl peptidase I is required for the processing and activation of granzymes A and B in vivo. Proc. Natl. Acad. Sci. U. S. A. 96, 8627–8632.

Pham, C.T.N., Armstrong, R.J., Zimonjic, D.B., Popescu, N.C., Payan, D.G., Ley, T.J., 1997. Molecular cloning, chromosomal localization, and expression of murine dipeptidyl peptidase I. J. Biol. Chem. 272, 10695–10703.

Qiu, G.F., Yamanoa, K., Unuma, T., 2005. Cathepsin C transcripts are differentially expressed in the final stages of oocyte maturation in kuruma prawn *Marsupenaeus japonicus*. Comp. Biochem. Physiol., B 140, 171–181.

Rao, N.V., Rao, G.V., Hoidal, J.R., 1997. Human dipeptidyl-peptidase I. Gene characterization, localization, and expression. J. Biol. Chem. 72, 10260–10265.

Robalino, J., Almeida, J.S., McKillen, D., Colglazier, J., Trent, H.F., Chen, Y.A., Peck, M.E.T., Browdy, C.L., Chapman, R.W., Warr, G.W., Gross, P.S., 2007. Insights into the immune transcriptome of the shrimp *Litopenaeus vannamei*: tissue-specific expression profiles and transcriptomic responses to immune challenge. Physiol. Genomics 29, 44–56.

Smyth, M.J., McGuire, M.J., Thia, K.Y., 1995. Expression of recombinant human granzyme B. A processing and activation role for dipeptidyl peptidase I. J. Immunol. 154, 6299–62305.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., Higgins, D.G., 1997. The CLUSTALX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 25, 4876–4882.

Turk, B., Turk, V., Turk, D., 1997. Structural and functional aspects of papain-like cysteine proteinases and their protein inhibitors. Biol. Chem. 378, 141–150.

Turk, V., Turk, B., Turk, D., 2001. Lysosomal cysteine protease: facts and opportunities. EMBO J. 20, 4629–4633.

Ulevitch, R.J., Tobias, P.S., 1995. Receptor-dependent mechanisms of cell stimulation by bacterial endotoxin. Annu. Rev. Immunol. 13, 437–457.

Wolters, P.J., Pham, C.T., Muilenburg, D.J., Ley, T.J., Caughey, G.H., 2001. Dipeptidyl peptidase I is essential for activation of mast cell chymases, but not tryptases, in mice. J. Biol. Chem. 276, 18551–18556.

Zhao, Zhi-Ying, Yin, Zhi-Xin, Weng, Shao-Ping, Guan, Hao-Ji, Li, Se-Dong, Xing, Ke, Chan, Siu-Ming, He, Jian-Guo, 2007. Profiling of differentially expressed genes in hepatopancreas of white spot syndrome virus-resistant shrimp (*Litopenaeus vannamei*) by suppression subtractive hybridization. Fish Shellfish Immunol. 22, 520–534.