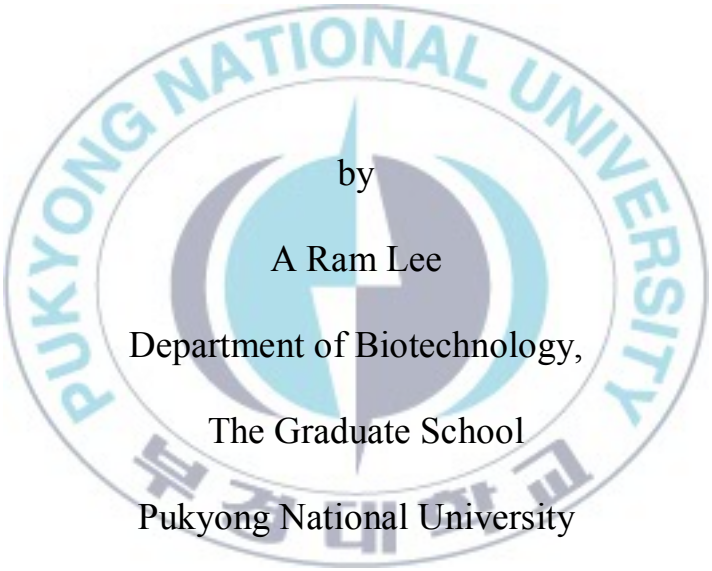
Thesis for the Degree of Master of Engineering

Molecular cloning, expression, characterization and enzymatic analysis of cathepsin L from starfish (*Asterina pectinifera*)



by

A Ram Lee

Department of Biotechnology,

The Graduate School

Pukyong National University

February 2011

Molecular cloning, expression, characterization and enzymatic analysis of cathepsin L from starfish (*Asterina pectinifera*)

별불가사리로부터의 카텝신 L 의 분자생물학적 클로닝, 발현, 그리고 특성화와 효소적 분석

Advisor : Prof. Hyung Ho Lee

by

A Ram Lee

A thesis submitted in partial fulfillment of the requirements
for the degree of

Master of Engineering

in Department of Biotechnology, The Graduate School,
Pukyong National University

February 2011

Molecular cloning, expression, characterization and enzymatic analysis of cathepsin L from starfish (*Asterina pectinifera*)

A dissertation

by

A Ram Lee

Approved by :

_____
(Chairman) Joon Ki Chung

_____ _____
(Member) Nam Gyu Park **(Member)** Hyung Ho Lee

February 23, 2011

CONTENTS

LIST OF TABLES

LIST OF FIGURES

별불가사리로부터의 카텝신 L의 분자생물학적 클로닝, 발현,

그리고 특성화와 효소적 분석

이 아 람

부 경 대 학 교 대 학 원 생 물 공 학 과

요 약

Cathepsin L은 단백질 분해의 초기과정에 관여하는 cysteine protease들 중 하나이다.

본 연구는 기존의 알려진 유사 유전자 서열 정보를 바탕으로 하여 RACE를 수행, 별 불가사리 cathepsin L(*ApCtL*)의 cDNA를 클로닝하였다. 별 불가사리로부터 동정된 cathepsin L 유전자는 327개의 아미노산을 암호화하는 984 bp의 open reading frame을 가진다. Cathepsin L의 propeptide region 내에 ERF/WNIN motif가 존재함으로써 이것이 명백하게 cathepsin L group이라는 것을 보여주며, 계통 유전학적 분석 결과 다른 종의 cathepsin L에 비해 초창기에 분화되어 나온 것으로 사료된다.

RT-PCR을 수행한 결과, 별 불가사리의 조사된 모든 조직에서 cathepsin L이 발현되는 것을 볼 수 있었다.

Pro-mature enzyme인 proApCtL은 fusion protein인 glutathione S-transferase를 포함하는 pGEX-4T-1 vector에 삽입하여 *E.coli* 균주인 DH5α 내에 과발현시켰다. 재조합 단백질인 proApCtL은 61KDa의 분자량을 가진다. proApCtL의 활성은 gelatin zymography를 통한 방법과 fluorogenic 펩타이드 기질인 Z-Phe-Arg-AMC을 이용한 방법으로 측정되었다. 적정 pH는 pH 8.0 이다.

1. Introduction

Proteolytic enzymes are classified into four groups ; serine proteases, cysteine proteases, aspartic proteases, and metallo-proteases (Turk et al., 2000). The cysteine proteases, including the papain superfamily, which called lysosomal cathepsins, are function in intracellular protein degradation, and a universal group of proteolytic enzyme that catalyze the hydrolysis of many different proteins. Cysteine proteases are found in viruses, bacteria, plants, invertebrates, and vertebrates (Lecaille et al., 2002, Berti et al., 1995, Barrett and Storer, 2001).

Cathepsin L is a member of cysteine protease, which is important protease in the initiation of protein degradation (Turk et al., 2000) and able to degrade extracellular matrix proteins such as fibronectin, myosin, actin, elastin, cytosolic protein, and collagen (Karrer et al., 1993, Kominami et al., 1991). According to sequence analysis, cathepsin L has cysteine, histidine, and asparagine as catalytic triad residues. Cysteine proteases including cathepsin L contain ERF/WNIN motif in the propeptide region. Cathepsin L performs many biological function related tissue degeneration, bone resorbtion, parasitic infection and immune system disease (Lecaille et al., 2002, Lindeman et al., 2004, Li et al., 2010).

Cathepsin L has been reported from lots of marine organisms such as mud

loach (*Misgurnus mizolepis*; Ahn et al., 2010), pearl oyster (*Pinctada fucata*; Ma et al., 2010), norway lobster (*Nephrops norvegicus*; Le Boulay et al., 1995), mitten crab (*Eriocheir sinensism*; Li et al., 2010), sea cucumber (*Stichopus japonicas*; Zhu et al., 2008), sea urchin (*Tetrapygus niger*; Morin et al., 2008). However, starfish cathepsin L has not been completely characterized with regard to its enzymatic properties and physiological functions. In this study, we describe the cDNA cloning, tissue-typic expression, and enzymatic characterization of the recombinant cathepsin L protein in the starfish.

2. Materials and methods

2.1. cDNA synthesis from starfish and rapid amplification of cDNA ends (RACE)

The total RNA was isolated from starfish using the TRIzol® (Invitrogen) according to the manufacturer's instructions and cDNA was synthesized from this isolated mRNA using the Transcriptor First Strand cDNA Synthesis Kit (Roche) and then used as the template for amplification. In an effort to identify the cathepsin L of the starfish, degenerate oligonucleotides, designed to target regions around the highly conserved region of cathepsin L (sense primer, CtL-F, Table 1; antisense primer, CtL-R, Table 1) were utilized in the amplification of cDNAs from a starfish cDNA mixture. Rapid amplification of cDNA ends (RACE) was used to clon cDNA containing 5'- and 3'-end of cathepsin L from a cDNA mixture which was prepared from total RNA of starfish using a SMART™ RACE cDNA amplification kit (Clontech). The 5'-end of the cathepsin L was obtained by 5' RACE-PCR using the specific primer ApCtL-GSP-R1 with Universal Primer Mix (UPM) for first round-PCR and the second specific primer ApCtL-GSP-R2 with Nested Universal Primer (NUP) for second nested PCR, respectively. The two primer sets used for 3' RACE-PCR were ApCtL-3'F2 with UPM and ApCtL-3'F3 with NUP (see Table 1) for obtaining 3'-end region. After

amplification, the RACE products were subcloned into a pGEM T-Easy vector (Promega) and then transformed into *E. coli* DH5α competent cells according to the manufacturer's instructions. The *E. coli* clones containing the recombinants were overlaid with 100 μg/ml of ampicillin, 0.4 mM isopropyl-β-thiogalactopyranoside (IPTG) and 40 μg/ml 5-bromo-4-chloro-3-indoly-β-D-galactoranoside (X-Gal) in the Luria-Bertani (LB) agar plate. White colonies were randomly chosen, cultivated and used for extraction of plasmid DNA. Plasmid DNA was prepared from *E. coli* using a LaboPass™ Plasmid Mini Purification Kit (COSMO GENETECH). DNA sequencing was conducted using a T7 promoter/ SP6 primers in the COSMO co, Ltd. (Seoul, Korea).

2.2. Sequence and phylogenetic analysis

Nucleotide and predicted amino acid sequences were analyzed using DNAsis for Windows version 2.5 (Hitachi software engineering), BioEdit Sequence Alignment Editor (Hall, 1999) and BLAST programs in the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/BLAST/). The signal sequence and putative cleavage site of *ApCtL* was identified using the SignalP 3.0 (http://www.cbs.dtu.dk/services/SignalP). Predictions of the pro-region cleavage sites and active sites were

based on alignment of the cathepsin L protein sequences with the vertebrate orthologues. Multiple sequence alignments were constructed using CLUSTAL W version 1.8 (Thompson et al., 1994) and adjusted with the BioEdit Sequence Alignment Editor. The phylogenetic tree was constructed using the Neighbor-Joining Method and plotted with MEGA version 3.0 (Kumar et al., 2004).

2.3. Expression studies by RT-PCR

In order to analyze the tissue expression of the *ApCtL* mRNA, RT-PCR was conducted using muscle, gonad, tubefeet, liver, stomach and body-rind tissues from healthy *A. pectinifera* specimens. Total RNA was isolated using TRIzol® (Invitrogen) in accordance with the manufacturer's instructions, and purified RNA was quantified by optical density at 260 nm using a UV spectrophotometer (Ultrospec 6300 pro, Amersham Biosciences). Two micrograms of total RNA from the *A. pectinifera* tissues were reverse-transcribed with Transcriptor First Strand cDNA Synthesis Kit (Roche), in accordance with the manufacturer's instructions. The specific primers for starfish cathepsin L were ApCtL-RT-For and ApCtL-RT-Rev (Table 1). Starfish GAPDH and β-actin were utilized as the internal controls (Ap-GAPDH-real-F and Ap-GAPDH-real-R, Ap-bactin-F and Ap-bactin-R,

for the starfish, Table 1). All of the PCR was run as follows: 94 ℃ for 5 min, 25 cycles of 94 ℃ for 30 s, 55 ℃ for 30 s and 72 ℃ 30 s, and a final 7 min of elongation at 72 ℃. The resultant PCR products were separated on 1.5% agarose/TAE gels containing ethidium bromide and visualized with a Gel Doc image analysis system (Bio-Rad). The PCR products were purified via agarose gel extraction (QIAquick® Gel Extraction kit) and sequenced (COSMO co, Ltd., DNA Sequencing Service, Seoul, Korea).

2.4. Expression and purification of recombinant proApCtL in E. coli

To prepare an expression vector suitable for production of recombinant starfish cathepsin L in *E. coli*, a 939 bp DNA fragment containing the coding sequence for the A. *pectinifera* procathepsin L (*proApCtL*) was generated by PCR amplification. The primers (EcoRI-proApCtL-F, 5'-CCG GAATTCGCCAGCCCTGACCTGGACCAAGAA-3'; XhoI-proApCtL-R, 5'-GCCGCTCGAGTTAGACGAGCGGGTAGCTGGCGTT-3') harbor E*coR*I/*Xho*I restriction sites (underlined), allowing for the cloning of the amplified DNA in a predicted orientation into pGEX-4T-1 (Amersham Pharmacia Biotech). Recombinant plasmid (*proApCtL/pGEX*) was transformed into *E. coli* strain DH5α. Transformed cells were grown in LB broth (100 ml) containing 100 mg/ml ampicillin at 37 ℃ for approximately

16 hr, diluted 1/100 with the same medium, and grown to an A_{600} of 0.6. Next, isopropyl-β-D-thiogalactopyranoside (IPTG) was added to a final concentration of 0.4 mM, and the incubation was continued for 3 hr. Cells were collected by centrifugation, washed, and resuspended in 0.2 volumes of phosphate buffered saline (PBS), lysed by using a sonicator (Vibra cell, Sonics & materials Inc, USA) at a setting of 40%, and centrifuged at 20,000 × *g* for 20 min at 4 °C. The soluble supernatant was subjected to glutathione-Sepharose 4B column (Pharmacia Biotech Co., USA) that had been equilibrated with PBS. After washing the column with equilibration buffer, the protein was eluted in elution buffer with 50 mM Tris/pH 8.0, 10 mM reduced glutathione (Sigma). The fractions containing sufficient amounts of active enzymes were pooled, and then dialyzed and concentrated using centricon 10 concentrators (Amicon). Purified proApCtL protein was used for SDS-PAGE, western blotting and enzyme activity assay.

2.5. SDS-PAGE, western blotting and zymography

Purified proApCtL enzyme was analyzed by 10% SDS-PAGE. All samples were denatured in a buffer containing 60 mM Tris/pH 6.8, 25% glycerol, 2% SDS, 14.4 mM 2-mercaptoethanol and 0.1% bromophenol blue, boiled for 5 min, and separated by 10% SDS-PAGE (Bio-Rad). Stained

molecular weight markers (GE, USA) were run as standards on each gel. After electrophoresis, the gel was stained with Coomassie brilliant blue R-250. Western blotting was performed using mouse monoclonal anti-GST antibody (1: 2000, Santa Cruz Biotechnology). The substrate zymography was performed by a modified procedure using gels with gelatin (Sigma), azocasein (Sigma), fibrinogen (Sigma), and bovine serum albumin (BSA) (Sigma) (0.1% w/v) as described earlier (Heussen and Dowdle, 1980) with slight modifications. To prepare the zymography, 30 ul of a protease sample was mixed with 10 ul of 4x SDS-sample buffer (0.5 M Tris/pH 6.8, 10% SDS, 20% glycerol and 0.02% bromophenol blue) without reducing the agent and without boiling. The sample was then applied to the gel and electrophoresed using a Bio-Rad Mini-Protean system (Bio-Rad) with a constant current of 12 mA per gel at 4 °C. After electrophoresis, the gels were immersed in 100 ml of 2.5% (v/v) Triton X-100 for 1 h to remove SDS and were washed once with incubation buffer (0.1 M Tris/ pH 8.0, containing 1 mM DTT). Next, the gels were immersed in the incubation buffer for 18 hr at 37 °C. Subsequently, the gels were washed with water and stained in 5% methanol/10% acetic acid/water containing 0.1% Coommassie Brilliant Blue R-250. Protease bands appeared as clear zones on a blue background.

2.6. Enzyme activity assays

The cathepsin L activity was assayed according to the modified method of Barret and Kirschke (1981). The optimum pH for enzymatic activity was determined using a sodium acetate buffer in pH ranges of 3-10 with Z-Phe-Arg-7-amido-4-methylcoumarin hydrochloride (Z-FR-AMC; Sigma) as substrates. Briefly, 10 ul of recombinant proApCtL enzyme in 85 ul of 0.1 M Tris/ pH 8.0, containing 1 mM DTT were preincubated at 37°C for 2 hr, and the enzyme reaction was initiated by adding 5 ul of 1 mM Z-FR-AMC at 37 °C for 10 min. The 7-amido-4-methylcoumarin (AMC) was measured using a Microplate Fluorometer (Packard Co. USA) at an excitation wavelength of 380 nm and an emission wavelength of 460 nm.

Substrate specificities were investigated using Z-Gly-Pro-Arg-AMC (Sigma), Z-Gly-Gly-Arg-AMC (Sigma), Z-Arg-Arg-AMC (Sigma), Ala-Ala-Phe-AMC (Sigma), Suc-Leu-Tyr-AMC, Z-Leu-Leu-Glu-AMC (Sigma), Suc-Ile-Ala-AMC, Suc-Leu-Leu-Val-Tyr-AMC and Z-Val-Val-Arg-AMC with 0.1 M Tris/ pH 8.0, containing 1 mM DTT, respectively. Substrates were added to a final concentration of 100 μM.

2.7. Effect of enzyme inhibitors, metal ions and detergents

The effects of enzyme inhibitors on protease activity were studied using

Z-Phe-Arg-AMC as the fluorogenic substrate. The following known proteinase inhibitors were tested: *trans*-Epoxysuccinyl-L-leucyl-amido (4-guanidino) butane (E-64; Sigma), Antipain (Sigma), and Leupeptin (USB Co., USA) for cysteine protease inhibitor; *N*-ethylmaleimide (NEM; Sigma), Phenylmethylsulphonyl fluoride (PMSF; Sigma), Chymostatin (Sigma), and Aprotinin (Sigma) for serine protease inhibitor; Ethylene diamine tetraacetic acid (EDTA; Sigma), Ethylene glycol-bis(beta-aminoethyl ether)-N,N,N',N'-tetraacetic acid (EGTA; Sigma) and 1,10- Phenanthroline (Sigma) for metalloproteinase inhibitor; Pepstatin A (Sigma) for aspartic protease inhibitor.

The effect of various metal ions (1 and 5 mM) on enzyme activity was investigated using $ZnSO_4$, $CuSO_4$, $CoCl_2$, KCl, $MgSO_4$, $CaCl_2$ and $HgCl_2$. The effects of some surfactants (Brij 35, Triton X-100, Tween 20 and SDS) on enzyme stability were also examined. The effects of enzyme inhibitors, metal ions and detergents (0.01 and 0.05%) on the activity of proApCtL protease were studied at pH 8 and 37 °C. All the experiments were replicated three times. The values shown in tables and graphical data represent the mean of three assays (±standard deviation). All the analyses were performed using Microsoft Excel and SigmaPlot software package.

3. Results & Discussion

3.1 Cloning and sequence analysis of A. pectinifera cathepsin L cDNA

The complete sequence of starfish cathepsin L (*ApCtL*) cDNA was obtained by combining DNA sequence of the initial cDNA clone, and 5' and 3'-RACE cDNA amplification PCR products. The full-length cDNA yielded a 1,693 bp sequence comprised a 98-bp 5'-untranslated region (5'-UTR) and a 984-bp coding region including a stop codon (TAA), followed by a 613-bp 3'-UTR containing one potential poly-adenylation signal, AATAAA (underlined in the Fig. 1). The nucleotide sequence of *ApCtL* was predicted to encode for a preproprotein of 327 amino acids, which contained a 15-residue putative signal peptide analyzed with signalIP (Martoglio and Dobberstein, 1998), a 94-residue propeptide and the 218-residue mature enzyme (Fig. 1). The propeptide of mammalian cathepsins has been shown to be crucial for proper enzyme folding, for the stabilization of its structure upon exposure to pH changes, and for microsomal and lysosomal targeting (Turk et al., 2000). Many members of the papain family harbor a proline (P) residue at position 2 in the mature enzyme. This is also observed in *ApCtL*. The proline may serve to prevent unwanted N-terminal proteolysis (Rawlings and Barret, 1994). All cysteine proteases harbor a conserved active site consisting of cysteine, histidine, and asparagine residues. The

cysteine residue (Cys25 based on mature *ApCtL* numbering) was embedded within a highly conserved peptide sequence, CGSCWAFS. The histidine residue (His166; *ApCtL* numbering) was adjacent to small amino acid residue, glycine. Asparagine (Asn186; *ApCtL* numbering) was a component of the Asn-Ser-Trp (NSW) motif.

Fig. 2 shows an alignment of *ApCtL* with the sequences of other cathepsins. The amino acid sequence of *ApCtL* harbors the ERF/WNIN motif (ERWNIN; *ApCtL*), which is highly conserved in the cathepsin L family. In addition, *ApCtL* also contains the GNFD motif in its proregion, which is generally conserved in most of the cysteine proteases of the papain superfamily (Vernet et al., 1995; Turk et al.,2000). The presence of the ERWNIN/GNFD motif in the pro-region of *ApCtL* clearly indicates that this protein is related to the cathepsin L group and is apart from the cathepsin B subfamily. A comparison of the amino acid sequence of *ApCtL* with different species of cathepsin L showed 46-54% identities and 44-51% identities with *Homo sapiens* cathepsin L group (L, V, S and K). However, *ApCtL* did not exhibit a high degree of identities with cathepsin B-like (B, C and X, 16-18%), cathepein A (5%) and cathepisn D (8%). In order to determine the evolutionary relationship of *ApCtL* with other cathepsin families, a phylogenetic tree was constructed. Phylogenetic analysis was

conducted with the amino acid sequences of other cathepsinLs and human cathepsins obtained from GenBank using neighbor-joining methods (Fig. 3). On the basis of a comprehensive phylogenetic analysis, the enzymes of the Family C1 peptidases (i.e. the papain superfamily of cysteine proteases) could be divided into two primary evolutionary branches, branches A and B. Branch A includes cathepsins B, C and W. Branch B includes the cathepsin L-like enzymes, a group including papain, cathepsin L, cathepsin V, cathepsin S, cathepsin K, cathepsin H, cathepsin W and cathepsin F (Santamaría et al.,1999; Tingaud-Sequeira and Cerdà, 2007; Dacks et al., 2008). *ApCtL* was more closely related to the cathepsin L subfamily (L, V, S, K, H, W and F) than to the cathepsin B, X, A, C, and D subfamilies.

3.2. Tissue-typic expression of ApCtL

Distribution of *ApCtL* mRNA transcripts in different organs was examined by RT-PCR using *ApCtL*-specific primers (Table 1). Relative gene expression levels were normalized using the GAPDH and β-actin genes. As shown in Fig. 4 expression of *ApCtL* was observed in all of the tissues. The expression pattern of *ApCtL* was revealed in high levels in the muscle, liver, stomach. Existence of cathepsin L was involved in parasite-host interaction. Additional many expression of cathespsin L was related with immune

system to protect from outer molecules (Rao et al., 2007). Owing to direct water exposure, cathepsin L in stomach of starfish was expressed high level. Whereas cathepsin K was highly expressed kidney in goldfish and osteoclasts in human (Inaoka et al., 1995; Harikrishnan et al., 2010). Expression of cathepsin S was higher spleen and lung in human (Shi et al., 1994).

3.3. Enzymatic characterization of recombinant proApCtL

In order to assess the functional and enzymatic characteristics of *ApCtL*, the cDNA encoding for *proApCtL* was expressed in *E. coli* as a fusion protein with glutathione S-transferase (GST). The recombinant *proApCtL/pGEX* was overexpressed in *E. coli* DH5α as a fusion protein. The overproduced soluble GST-fusion protein (proApCtL) was then applied to glutathione-Sepharose 4B column chromatography. The proApCtL fusion protein band had a high purity and the correct size was 61 kDa by SDS-PAGE and Western blot analysis (Fig. 5).

The purified proApCtL activity was quantified by measuring the cleavage of a synthetic fluorogenic peptide substrate, Z-FR-AMC. In previous reports, Z-FR-AMC derivatives were shown to be efficiently hydrolyzed by both cathepsin B and L, whereas Z-RR-AMC derivatives were found to be

efficiently hydrolyzed by only cathepsin B (Barrett and Kirschke, 1981) although to a lower degree compared with Z-FR-AMC hydrolysis. Chemical properties of protease S_2 pocket is regarded to prefer like Z-RR-AMC substrate. Cathepsin L has alanine at P_2 position which cannot donate arginine binding, whereas cathepsin B proteases has acidic group at this position and stabilizes hydrophilic guanadino group of arginine (Sajid and Mckerrow, 2002).

As can be seen in Fig. 6, the proApCtL protein was demonstrated to evidence a high level of activity at pH 8. Also proteolytic activity of proApCtL protein was determined by gelatin zymography. The purified proApCtL was capable of hydrolyzing 0.1% gelatin at pH 8 (Fig. 7) but was not able to hydrolyze protein substrates such as azocasein, fibrinogen, and BSA (data not shown). Interestingly, using the synthetic substrate, Z-FR-AMC, the proApCtL protein displayed activity over a wide range of pH (pH 6.5 to 10) with optimal activity occurring at pH 8 (Fig. 6), but the activity was profoundly reduced at pH values of 6 (lysosomal pH) and below. Mammalian lysosomal cysteine proteases are unstable at neutral pH, but the proApCtL protein was stable for neutral and alkaline pH.

We also compared proApCtL activity on various substrates conjugated with aminomethylcoumarin as the fluorescent chromophore (Table 2). The

highest levels of AMC release activity were seen from Z-FR-AMC, Z-GGR-AMC, Z-GPR-AMC, and Z-VVR-AMC. The Z-FR-AMC substrate was hydrolyzed 4-fold more efficiently than the Z-RR-AMC substrate.

To identify of proApCtL as a cysteine protease was performed via inhibitor assay. Inhibitors for serine, aspartic and metallo-proteases show no influence for proApCtL activity (Table 3). The enzymatic activity of proApCtL was reduced significantly or blocked completely by the two tested cysteine protease inhibitors, antipain and leupeptin. Additionally, many metal ions and detergents have been shown no influence the activity of proApCtL (Table 4). The proApCtL enzyme was inactivated by $ZnSO_4$, $CuSO_4$, $HgCl_2$ and $CoCl_2$, and SDS.

.

Table 1

Oligonucleotide primers used for *ApCtL* amplification and expression studies.

Primer name	5'-3' sequence	Information
CtL-F	CABTGGVANCWVTGGAAGAA	Primers used to obtain initial fragments
CtL-R	CCRCAGTGRTTGTKYCKGTCYTTKGCCAT-GT	
ApCtL-GSP-R1	TGCGGCCCTGAGTAGACAGGATGTAC	Specific primers for 5' and 3' RACE
ApCtL-GSP-R2	TGGAAGGACTCGTGACTGGCGTCGAT	
ApCtL-3'F2	TTCCAGTACGTCCACGACAACATG	
ApCtL-3'F3	ATGGACGAAAAGGCACTCCA	
Universal UP-Long Primer A Mix (UPM) UP-Short Nested Universal Primer A (NUP)	CTAATACGACTCACTATAGGGCAAGCAGT-GGTATCAACGCAGAGT	Universal primers for 5' and 3' RACE
	CTAATACGACTCACTATAGGGC	
	AAGCAGTGGTATCAACGCAGAGT	
Ap-GAPDH-real-F	TGGGATAACCGAGGGACTGATGA	Primers for expression studies
Ap-GAPDH-real-R	CCAGTGGATGCTGGGATGATGT	
Ap-bactin-real-F	GGTCATCACCATCGGCAACG	
Ap-bactin-real-R	ACGGATCTCCACGTCGCACT	
ApCtL-RT-For	AGCACAACCAGCGATTCG	
ApCtL-RT-Rev	TCTTGACACCAGTGACCCAT C	

Table 2

Substrate specificity of proApCtL.

Substrates	Concentration (μM)	Activity (%)
Z-Phe-Arg-AMC (FR)	50	100.00
Z-Gly-Pro-Arg-AMC (GPR)	50	93.04 ± 8.10
Z-Gly-Gly-Arg-AMC (GGR)	50	97.02 ± 10.50
Z-Arg-Arg-AMC (RR)	50	27.95 ± 1.89
Ala-Ala-Phe-AMC (AAF)	50	1.67 ± 0.05
Suc-Leu-Tyr-AMC (LY)	50	0.17 ± 0.08
Z-Leu-Leu-Glu-AMC (LLE)	50	4.69 ± 0.51
Suc-Ile-Ala-AMC (IA)	50	0.18 ± 0.12
Suc-Leu-Leu-Val-Tyr-AMC (LLVY)	50	5.72 ± 0.19
Z-Val-Val-Arg-AMC (VVR)	50	83.26 ± 4.77

Table 3

Effect of various protease inhibitors on the enzymatic activity of the proApCtL. The purified proApCtL was pre-incubated with the indicated inhibitors and assayed for residual activity using Z-Phe-Arg-AMC as the fluorogenic substrate.

Inhibitiors	Specificity	Concentration (mM)	Inhibition (%)
Control	-	-	0
E-64	All cysteine proteinases	0.1	33.58 ± 7.82
		0.2	38.46 ± 9.28
		0.3	42.09 ± 0.91
		0.5	51.98 ± 1.27
Antipain	Serine/cysteine proteinases	0.1	88.44 ± 0.56
Chymostatin	Serine/cysteine proteinases	0.1	76.98 ± 3.09
Leupeptin	Cysteine/trypsin-like Serine proteinases	0.1	89.32 ± 1.08
NEM	Serine proteinases	0.1	2.87 ± 3.91
PMSF	Serine proteinases	0.1	0 ± 6.15
Aprotinin	Serine proteinases	0.1	70.50 ± 2.41
EDTA	Metallo proteinases	0.1	3.48 ± 6.78
EGTA	Metallo proteinases	0.1	3.40 ± 4.52
1,10-phenanthroline	Metallo proteinases	0.1	10.30 ± 5.73
Pepstatin A	Aspartic proteinases	0.1	0

Table 4

Effect of metal ions and various detergents on proApCtL activity.

Agent	Relative activity (%)	
	1mM	**5mM**
Control	100.00	100.00
$ZnSO_4$	83.59 ± 2.16	79.36 ± 0.70
$CuSO_4$	89.83 ± 1.21	81.71 ± 0.07
$MgSO_4$	101.77 ± 1.67	114.45 ± 0.51
$HgCl_2$	89.75 ± 0.78	19.61 ± 0.99
$CaCl_2$	102.35 ± 0.67	102.71 ± 0.40
KCl	100.73 ± 0.76	102.61 ± 0.48
$CoCl_2$	89.89 ± 1.62	58.88 ± 0.11
	0.01%	**0.05%**
Triton X-100	101.79 ± 1.31	95.96 ± 0.47
Tween 20	99.65 ± 0.18	81.63 ± 1.99
SDS	21.28 ± 0.32	6.53 ± 0.35
Brij 35	99.56 ± 0.86	97.03 ± 1.26

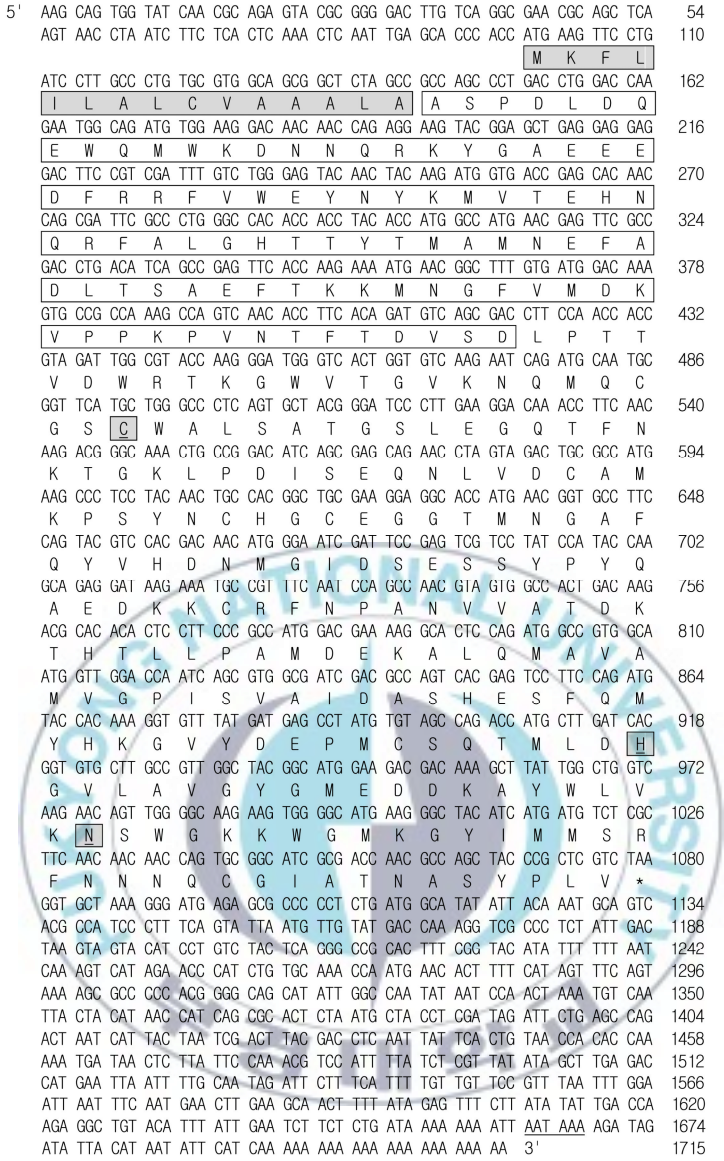


Fig. 1. Nucleotide and deduced amino acid sequence of starfish cathepsin L cDNA (*ApCtL*). The shaded box and the open box in the amino acid sequences indicate the putative signal peptide (pre) and the propeptides of

ApCtL, respectively. The active site triad residues Cys25, His166, and Asn186 are indicated in the *thickly underlined* and shaded boxes. The *asterisk* (*) at the end of the amino acid sequences shows the stop codon. The polyadenylation consensus sequences (AATAAA) are *double underlined*.

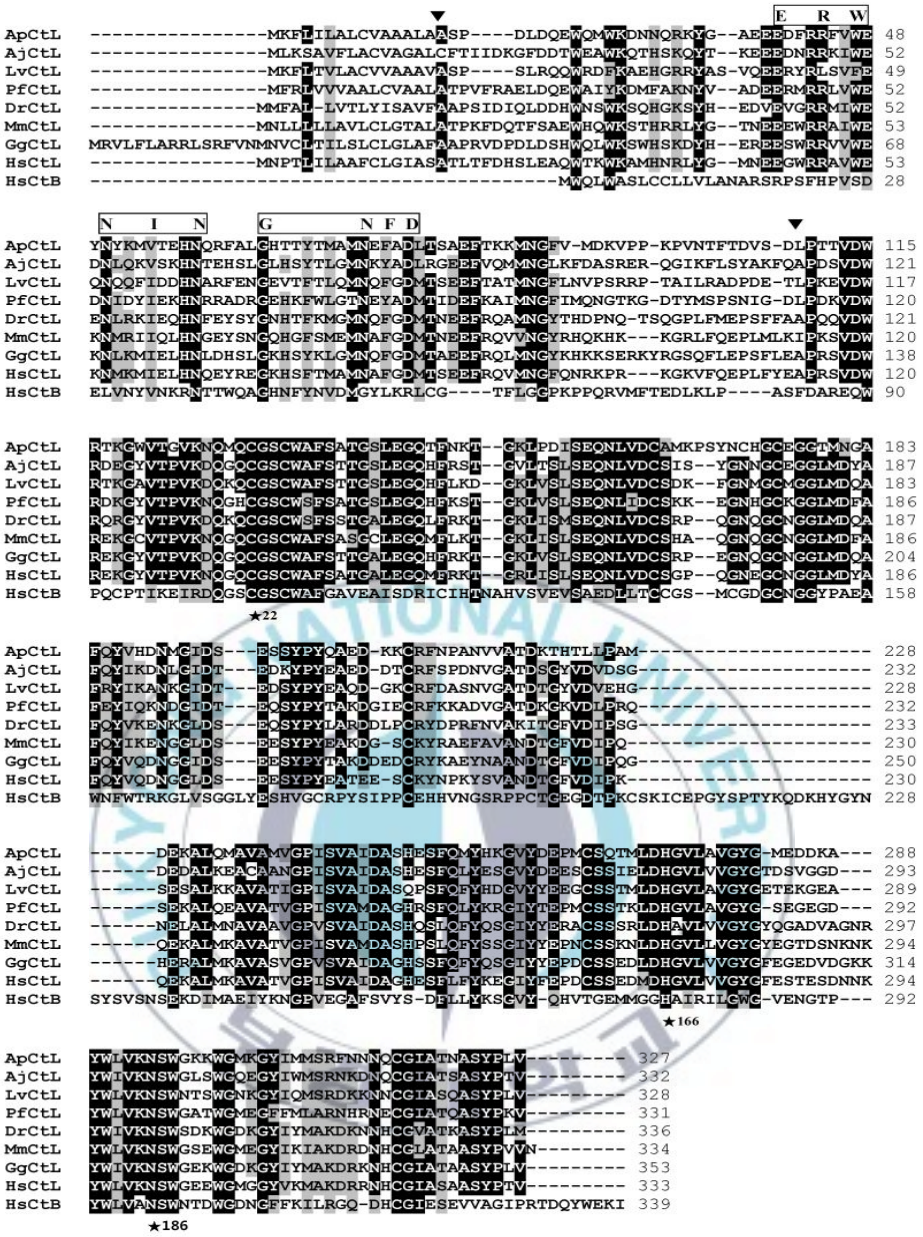


Fig. 2. Multiple amino acid sequence alignment of starfish (*Asterina pectinifera*) cathepsin L between different species of cathepsin L and human cathepsin L. Identical amino acid residues are *darkly shaded*, similar amino

acids are *lightly shaded*, unrelated residue have a *white background*, and amino acid numbers are shown on the right. The cysteine proteinase catalytic triad residues (C, H, and N) are shown in the stars (★). Conserved inter-spaced motifs in the pro-region, ERFNIN and GNFD are indicated in box. GenBank accession numbers are as follows: *AjCtL*, *Apostichopus japonicas* cathepsin L (ABW98676); *PfCtL1*, *Pinctada fucata* cathepsin L (ADC52430); *PfCtL2*, *Pinctada fucata* cathepsin L (ADC52431); *DrCtL*, *Danio rerio* cathepsin L (CAA69623); *LvCtL*, *Litopenaeus vannamei* cathepsin L (CAA68066); *GgCtL*, *Gallus gallus* cathepsin L (NP_001161481); *MmCtL*, *Mus musculus* cathepsin L (AAD32137); *HsCtL*, *Homo sapiens* cathepsin L (CAA75029); *HsCtB*, *Homo sapiens* cathepsin B (AAH10240).

.

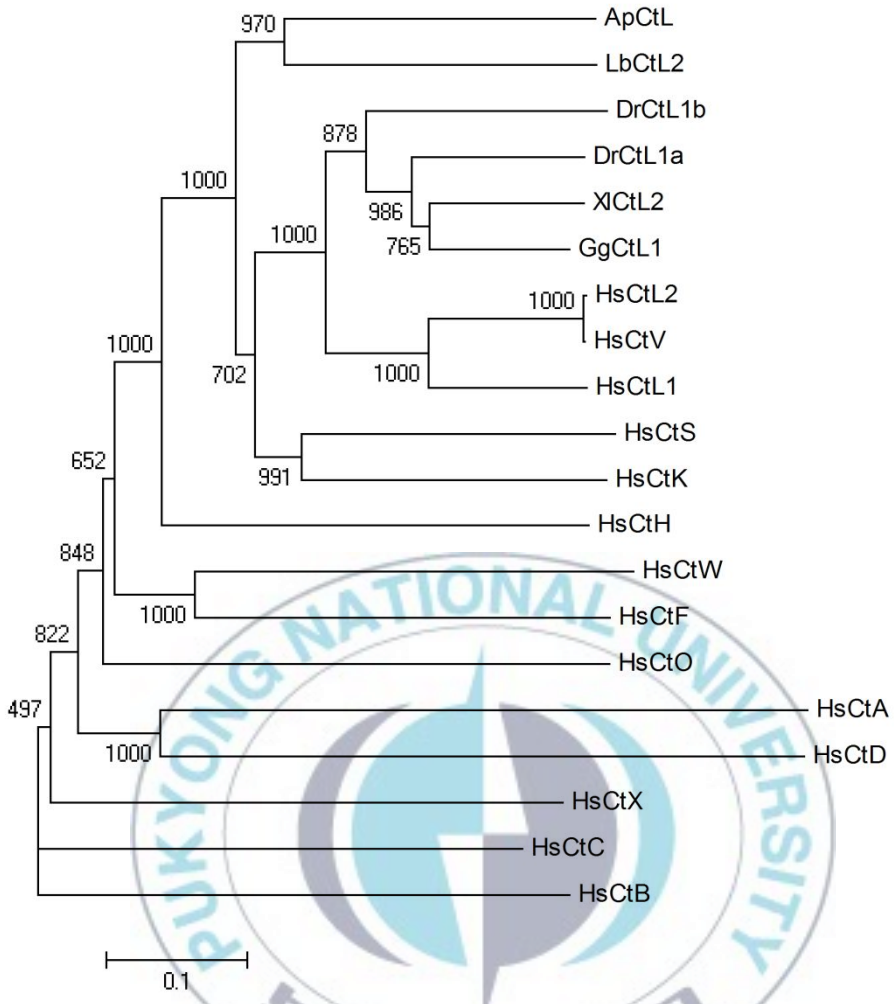


Fig. 3. Phylogenetic relationships of *ApCtL* among representative mammalian groups based on the cathepsin genes. In this neighbor-joining phylogram, all individuals are represented and the branches are based on the number of inferred substitutions as indicated by the bar. GenBank accession numbers are as follows: *LbCtL2*, Lubomirskia baicalensis cathepsin L2 (CAI91575); DrCtL1b, Danio rerio cathepsin L1b (NP_571273); DrCtL1a,

Danio rerio cathepsin L1a (NP_997749); XlCtL2, Xenopus laevis (NP_001087489); GgCtL1, Gallus gallus (NP_001161481); HsCtL2, Homo sapiens cathepsin L2 (CAA75029); HsCtV, Homo sapiens cathepsin V (BAA25909); HsCtL1, Homo sapiens cathepsin L1 (AB463468.1); HsCtS, Homo sapiens cathepsin S (NP_004070); HsCtK, Homo sapiens cathepsin K (NP_000387); HsCtH, Homo sapiens cathepsin H (NP_004070); HsCtW, Homo sapiens cathepsin W (NP_001326); HsCtF, Homo sapiens cathepsin F (NP_003784); HsCtO, Homo sapiens cathepsin O (NP_001325); HsCtA, Homo sapiens cathepsin A (P10619); HsCtD, Homo sapiens cathepsin D (AAP35556); HsCtX, Homo sapiens cathepsin X (AAA95998); HsCtC, Homo sapiens cathepsin C (AAQ08887); HsCtB, Homo sapiens cathepsin B (AAH10240).

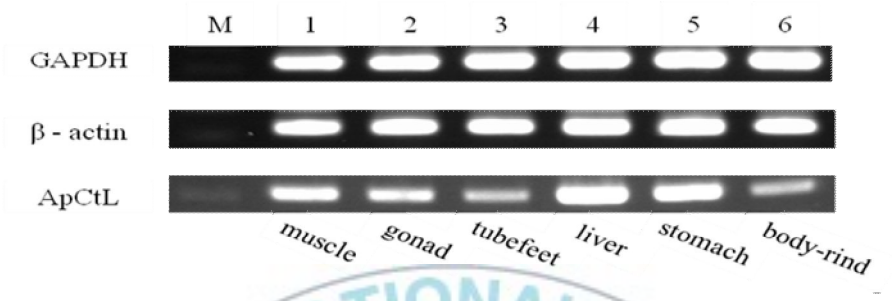
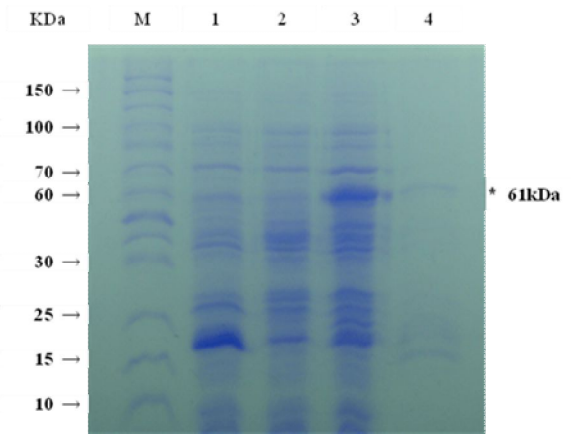


Fig. 4. Tissue-typic expression of the *ApCtL* mRNA. Total RNA was isolated from various *Asterina pectinifera* tissues, and 0.5 μg was subjected to RT-PCR analysis using the amplimers for the *A. pectinifera* cathepsin L,GAPDH and β-actin.

A



B

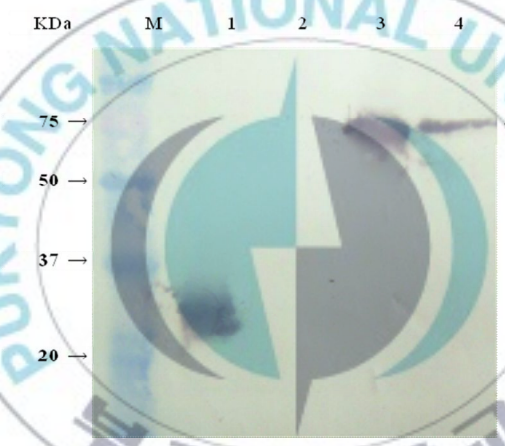


Fig. 5. Expression of recombinant proApCtL. Purification of recombinant proApCtL cysteine protease fused with glutathione S-transferase (GST) from overexpression in E. coli.

(A) Coomassie blue staining after SDS-PAGE. The asterisk (*) indicates GST-fused proApCtL. The lanes were labeled as follows: M, standard size marker; 1, overexpressed GST protein; 2, non-induced GST-fused

proApCtL; 3, overexpressed GST-fused proApCtL (37 °C); 4, glutathione-Sepharose 4B affinity column purified proApCtL.

(B) Western blot analysis. M: prestained protein size marker, lane 1: expressed GST protein (37 °C) reacted with monoclonal anti-GST antibody (positive control), lane 2: non-inducedproApCtL protein (negative control), lane 3, overexpressed GST-fusedApCtL (37 °C); lane 4, glutathione-Sepharose 4B affinity column purified proApCtL.

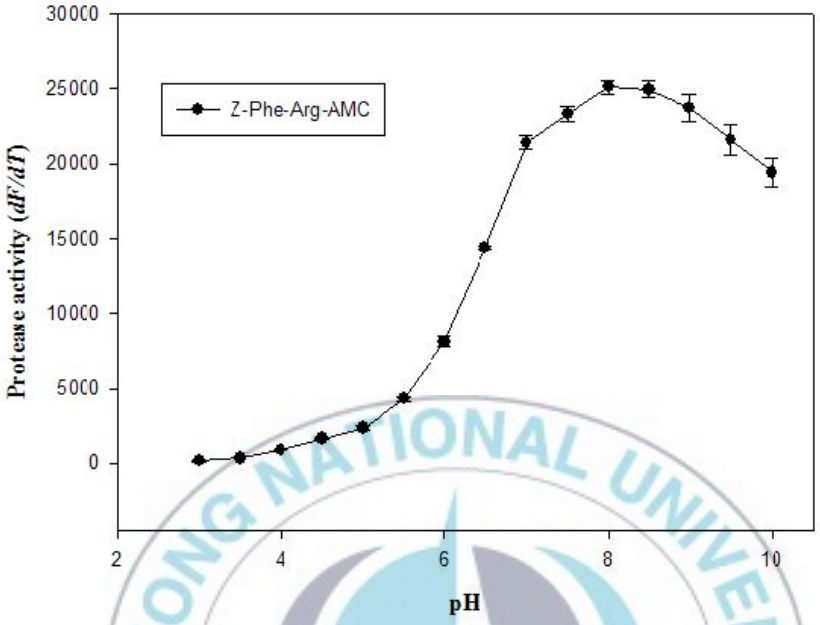


Fig. 6. pH dependency of proApCtL. Proteolytic activity was assayed against Z-Phe-Arg-AMC and was expressed in terms of arbitrary fluorescence units per time (*dF/dt*). The points and bars show the average values and S.D. of 3 independent experiments, respectively.

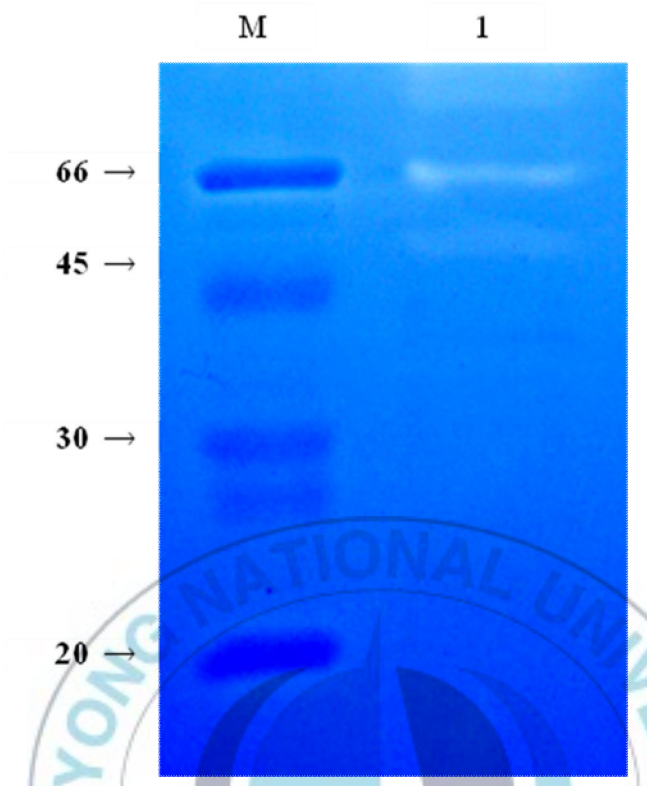


Fig. 7. Gelatin zymography of purified proApCtL. The purified enzyme was subjected to electrophoresis on a 10% polyacrylamide gel containing 0.1% gelatin. The gel was incubated overnight in 0.1 M Tris/ pH 8.0, containing 1 mM DTT at 37 °C . After staining with 0.1% Coommassie Brilliant Blue R-250, the gel was destained with destaining solution (40% methanol and 7% acetic acid). Areas of proteolysis appear as clear regions within the gel. Lane 1, purified proApCtL protein.

References

Ahn S.J., Sung J.H., Kim N.Y., Lee A.R., Jeon S.J., Lee J.S., Kim J.K., Chung J.K., Lee H. H., Molecular Cloning, Expression, and Characterization of Cathepsin L from Mud Loach. (*Misgurnus mizolepis*) Appl. Biochem. Biotechnol. 10 (2010), 8964-8966

Barrett A.J., Kirschke H., Cathepsin B, Cathepsin H, and cathepsin L. Methods Enzymol. 80 (1981), 535-561.

Barrett A.J., Rawlings N.D., Evolutionary lines of cysteine peptidases. J. Biol. Chem. 382 (2001), 727-733.

Berti P.J., Storer A.C., Alignment/phylogeny of the papain superfamily of cysteine proteases. J. Mol. Biol. 246 (1995), 273-283.

Dacks, J.B., Kuru, T., Liapounova, N.A., Gedamu, L., Phylogenetic and primary sequence characterization of cathepsin B cysteine proteases from the oxymonad flagellate monocercomonoides. J. Eukaryot. Microbiol. 55 (2008), 9-17.

Hall T.A., BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp. Ser. 41(1995), 95-98.

Harikrishnan R., Kim M.C., Kim J.S., Han Y.J., Jang I.S., Balasundaram C., Heo M.S., Immune response and expression analysis of cathepsin K in goldfish during Aeromonas hydrophila infection. Fish & Shellfish Immunology 28 (2010), 511-516.

Heussen C., Dowdle E.B., Electrophoretic analysis of plasminogen activators in polyacrylamide gels containing sodium dodecyl sulfate and copolymerized substrates. Anal. Biochem. 102 (1980), 196-202.

Inaoka T, Bilbe G, Ishibashi O, Tezuka K, Kumegawa M, Kokubo T., Molecular cloning of human cDNA for cathepsin K: novel cysteine proteinase predominantly expressed in bone. Biochem. Biophys. Res. Commun. 206 (1995), 89-96.

Ishidoh K., Kominami E., Procathepsin L degrades extra cellular matrix

proteins in the presence of glycosaminoglycans in vitro. Biochem. Biophys. Res. Commun. 217(1995), 624-631

Karrer K.M., Peiffer S.L., DiTomas M.E., Two distinct gene subfamilies within the family of cysteine protease genes. Proc. Natl. Acad. Sci. U. S. A. 90 (1993), 3063-3067.

Kominami E., Ueno T., Muno D., Katunuma N., The selective role of cathepsins B and D in the lysosomal degradation of endogenous and exogenous proteins. FEBS Letters 287 (1991), 189-192.

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

Le Boulay C., Van Wormhoudt A., Sellos D., Molecular cloning and sequencing of two cDNAs encoding cathepsin L-related cysteine proteinases in the nervous system and in the stomach of the Norway lobster (*Nephrops norvegicus*). Comp. Biochem. Physiol. 111B(1995), 353-359.

Lecaille F., Kaleta J., Brömme D., Human and parasitic papain-like cysteine proteases: their role in physiology and pathology and recent developments in inhibitor design. Chem. Rev. 102 (2002), 4459-4488.

Lindeman J.H., Hanemaaijer R., Mulder A., Dijkstra P.D., Szuhai K., Bromme D., et al. Cathepsin K is the principal protease in giant cell tumor of bone. AmJ. Pathol. 165 (2004), 593-600.

Li W.W., Jin X.K., He L., Jiang H., Gong Y.N., Xie Y.N., Wang Q., Molecular cloning, characterization, expression and activity analysis of cathepsin L in hinese mitten crab, *Eriocheir sinensis*. Fish & shellfish Immunology 29 (2010), 1010-1018

Ma J.J., Zhang D.C., Jiang J.J., Cui S.G., Pu H.L., Jiang S.G., Molecular characterization and expression analysis of cathepsin L1 cysteine protease from pearl oyster *Pinctada fucata*. Fish & Shellfish Immunology 29 (2010), 501-507

Martoglio B., Dobberstein B., Signal sequences: more than just greasy

peptides. Trends Cell Biol. 8 (1998), 410-415.

Morin V., Sanchez A., Quinones K., Huidobro J.G., Iribarren C., Bustos P., Puchi M., Genebie're A.M., Imschenetzky M., Cathepsin L Inhibitor I Blocks Mitotic Chromosomes Decondensation During Cleavage Cell Cycles of Sea Urchin Embryos. J. Cell. Physiol. 216 (2008), 790-795

Rao Y.Z., Yang T.B., cDNA cloning, mRNA expression and recombinant expression of a cathepsin L-like cysteine protease from Neobenedenia melleni (Monogenea: Capsalidae). Aquaculture 269 (2007), 41-53.

Rawlings N.D., Barret A.J., Families of cysteine peptidases. Methods Enzymol. 244 (1994), 461-486.

Saitou, N., Nei, M., The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4 (1987), 406-25.

Sajid, M., McKerrow, J.H., Cysteine protease of parasitic organisms. Mol. Biochem. Parasitol. 120 (2002), 1– 21.

Santamaría, I., Velasco, G., Pendás, A.M., Paz, A., López-Otín, C., Molecular cloning and structural and functional characterization of human cathepsin F, a new cysteine proteinase of the papain family with a long propeptide domain. J. Biol. Chem. 274 (1999), 13800-13809.

Shi G.P., Webb A.C., Foster K.E., Knoll J.H., Lemere C.A., Munger J.S., Chapman H.A., Human cathepsin S: chromosomal localization, gene structure, and tissue distribution. J. Biol. Chem. 269 (1994), 11530-11536.

Thompson J.D., Higgins D.G., Gibson T.J., Clustal W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. Nucleic Acids Res. 22 (1994), 4673-4680.

Tingaud-Sequeira, A., Cerdà, J., Phylogenetic relationships and gene expression pattern of three different cathepsin L (Ctsl) isoforms in zebrafish: Ctsla is the putative yolk processing enzyme. Gene 386(2007), 98-106.

Turk B., Turk D., Turk V., Lysosomal cysteine proteases: more than scavengers. Biochim. Biophys. Acta. 1477 (2000), 98-111.

Vernet T., Berti P.J., Montigny C., Musil R., Tessier D.C., Ménard R., Magny M.C., Storre A.C., Thomas D.Y., Processing of the papain precursor. The ionization state of a conserved amino acid motif within the Pro region participates in the regulation of intramolecular processing. Journal of Biological Chemistry 270 (1995), 10838-10846.

ABSTRACT

Molecular cloning, expression, characterization and enzymatic analysis of cathepsin L from starfish (*Asterina pectinifera*)

A Ram Lee

Department of Biotechnology, The Graduate School,

Pukyong National University

Abstract

Cathepsin L is a member of cysteine protease, which is important protease in the initiation of protein degradation. In this study, the cDNA of starfish (*Asterina pectinifera*) cathepsin L (*ApCtL*) was cloned by the combination of homology molecular cloning and rapid amplification of cDNA ends (RACE). The full-length of *ApCtL* gene contains an open reading frame of 984bp encoding 327 amino acids. The presence of ERF/WNIN motif which is conserved in the propeptide region of cathepsin L clearly reveals that *ApCtL* is cathepsin L group. Phylogenetic comparisons of *ApCtL* to other cathepsin groups have shown ApCtL in the early lineage of the cathepsin L group. The results of RT-PCR analysis show expression pattern of *ApCtL* observed in all of the tissues. The pro-mature enzyme of

ApCtL (proApCtL) was overexpressed in *E.coli* DH5α as a fusion protein with glutathione S-transferase in pGEX-4T-1vector. The molecular weight of recombinant proApCtL was a 61 kDa. The activity of proApCtL was detected by gelatin zymography and cleaving synthetic fluorogenic peptide substrates, Z-Phe-Arg-AMC. The optimal pH for the protease activity was 8.0.

ACKNOWLEDGEMENT

부족한 저에게 학문의 길을 열어주시고 본 논문이 완성되기까지 정성껏 도와주신 저의 지도교수님이신 이 형호 교수님께 머리 숙여 감사드립니다. 교수님 가르침의 은혜 절대 잊지 않겠습니다. 또한 생물공학과 홍 용기 교수님, 공 인수 교수님, 김 중균 교수님, 박 남규 교수님, 김 성구 교수님 그리고 정 귀택 교수님께도 감사의 말씀 드립니다. 또한 석사과정 내내 저에게 많은 조언을 해주신 정 준기 교수님께도 감사의 말씀드립니다.

처음 실험실에 들어와서 아무것도 모르는 저에게 하나하나 꼼꼼하게 가르쳐주시고 졸업하는 지금까지 챙겨주시는 저에게 있어서 정말 아빠 같은 분인 김 무상 박사님. 박사님이 가르쳐 주신 것들 정말 평생 잊지 않겠습니다. 부족한 저를 이끌어 주셔서 감사합니다. 지금은 미국에 가 계시지만 항상 모르는 것이 있을 때마다 꼼꼼하게 하나하나 메모체크까지 해주시면서 가르쳐주신 상중선배님, 석사 과정동안 함께 보내며 저에게 정말 많은 것을 가르쳐주신 나영이 언니에게도 감사의 말씀 드립니다. 정말 좋은 결과 있기를 바랍니다.

우리 분생 아가들이자 제 선배인 진욱이 오빠, 혜진이 너무너무 고맙고, 사랑한다는 말 전하고 싶습니다. 약리학실험실의 상환선배님께도 감사드리고, 희성이 오빠, 재웅이 오빠, 희영이, 혜인이 다들 너무 고맙다는 말 전하고 싶습니다. 희성이 오빠, 진욱이 오빠, 혜진이, 희영이 모두 좋은 결과 얻어서 석사 졸업하기를 바랍니다. 정기오빠, 진영이, 기륜이, 해리는 아직 실험실에 들어온지 얼마 안되었지만 앞으로 친해지길 바라며 차근차근 실험 잘 배워나가길 바랍니다.

항상 옆에서 도와주시고 응원해주시는 소울메이트 혜진이 언니, 힘들 때 정말로 큰 힘이 되어주셔서 감사드립니다. 또 제가 이렇게 졸업을 할 수 있도록 도와주신 샤방샤방 상현이 언니, 언니의 그 밝은 모습을 본받아 저도 매사에 긍정적인 사람이 되었음 합니다. 지금은 졸업을 하고 서울에 있음에도 불구하고 저에게 많은 조언을 해 주신 주은이 언니에게도 감사드리고, 실험을 하는데 있어서 의문이 생겼을 때 도와주신 똑똑한 서 정길 박사님께도 감사의 말씀 드립니다.

제가 힘이 들 때 아무 말 없이 한 잔 해준 영림이 오빠와 지웅이 오빠 너무너무 고맙고, 오빠들이 힘들 때 이제는 제가 힘이 되었음 합니다.

과 사무실에 있는 지영이 언니와 수경이 언니 두 분께도 부족한 저를 챙겨주

신 것 감사드리고, 함께 졸업하는 우리 지숙이 언니, 발 다쳐서까지 열심히 일하는 언니를 더 챙겨주지 못해서 미안하고, 힘든 것조차도 함께 할 수 있어서 좋았습니다. 함께 석사과정을 지내왔고, 함께 졸업하는 석주오빠, 승준오빠, 혜림이, 경은이, 유리 모두 멋진 사회생활을 하길 바랍니다. 힘이 들 때 커피 한 잔 타 주시던 수근선배님, 다른 실험실임에도 불구하고 저를 잘 따랐던 유경이 혜진이 모두 좋은 결과를 얻기 바랍니다.

제가 힘이 들 때, 항상 옆에서 위로가 되주었던 영원한 내 친구 희영이, 성은이 너무너무 고맙고, 영원한 내 자기들 혜림이, 은정이, 정희, 희영이와 핸드폰 때문에 친구가 되어 지금은 너무나도 소중한 내 보물 나방, 개그맨 못지 않게 재미있는 경태, 태우, 어엿한 사회인이 되어 제 몫을 하고 있는 양기언니, 민희, 부경대 입학홍보대사 우리 이쁜이들 보미나, 윤진이, 혜진이, 이제는 정말로 내 친구가 되어버린 재목이, 명근이, 수영이, 정민이, 종성이. 말로 표현하지 못할 정도로 너무너무 고마운 내친구들이 있어서 지금의 제가 있을 수 있었습니다.

학위과정 동안 시도 때도 없이 변하는 언니의 못된 성격을 다 받아주면서 힘들 때 언니마냥 달래주고, 보듬어주었던 속 깊은 내 동생 아영이, 언제까지나 나의 보살핌을 받아야만 할 것 같아도 벌써 의젓하게 커서 누나의 옆자리를 지켜주는 듬직한 내 막내 동생 재선이, 두 동생들이 있어서 공부를 하는데 많은 힘을 얻었습니다.

끝으로 세 남매를 착하게, 예쁘게 키워주신 우리 엄마, 항상 힘든 상황에서도 부족한 딸에게는 예쁜 마음과 폭넓은 이해심을 심어주시고자 가르침을 주시고, 힘이 들 때마다 짜증을 부리는 못난 딸의 투정을 모두 받아주시며, 한없는 사랑으로 저의 마음을 따뜻하게 채워주시고 헌신적으로 뒷바라지를 해주신 세상에서 제가 가장 사랑하는 저의 어머니께 머리 숙여 깊이 감사를 드리며 이 논문을 바칩니다. 졸업해서 이제는 엄마의 짐을 조금이나마 덜어 드릴 수 있는 딸이 되겠습니다.