

# Aspartic Cathepsin D Endopeptidase Contributes to Extracellular Digestion in Clawed Lobsters *Homarus americanus* and *Homarus gammarus*

Liliana Rojo · Adriana Muhlia-Almazan ·  
Reinhard Saborowski · Fernando García-Carreño

Received: 5 June 2009 / Accepted: 18 January 2010  
© Springer Science+Business Media, LLC 2010

**Abstract** Acid digestive proteinases were studied in the gastric fluids of two species of clawed lobster (*Homarus americanus* and *Homarus gammarus*). An active protein was identified in both species as aspartic proteinase by specific inhibition with pepstatin A. It was confirmed as cathepsin D by mass mapping, N-terminal, and full-length cDNA sequencing. Both lobster species transcribed two cathepsin D mRNAs: cathepsin D1 and cathepsin D2. Cathepsin D1 mRNA was detected only in the midgut gland, suggesting its function as a digestive enzyme. Cathepsin D2 mRNA was found in the midgut gland, gonads, and muscle. The deduced amino acid sequence of cathepsin D1 and cathepsin D2 possesses two catalytic DTG active-site motifs, the hallmark of aspartic proteinases. The putatively active cathepsin D1 has a molecular mass of 36.4 kDa and a calculated pI of 4.14 and possesses three potential glycosylation sites. The sequences showed highest similarities with cathepsin D from insects but also with another crustacean cathepsin D. Cathepsin D1 transcripts were quantified during a starvation period using real-time

qPCR. In *H. americanus*, 15 days of starvation did not cause significant changes, but subsequent feeding caused a 2.5-fold increase. In *H. gammarus*, starvation caused a 40% reduction in cathepsin D1 mRNA, and no effect was observed with subsequent feeding.

**Keywords** Aspartic protease · Cathepsin D · Clawed lobster · Starvation · Protein digestion

## Introduction

Decapod crustaceans evolved complex processes to digest food. After mechanical chopping by the mouth parts, ingested material enters the stomach and mixes with gastric fluid and is hydrolyzed by digestive enzymes derived from the midgut gland (Baker and Gibson 1977). The midgut gland is the site of enzyme synthesis, nutrient absorption, and storage of lipids and glycogen (Glass and Stark 1994; Sánchez-Paz et al. 2007). The digestive enzymes are synthesized by the so-called F-cells (Hu and Leung 2007). They develop into B-cells while accumulating secretion products. The B-cells break down in holocrine fashion and release enzymes into the lumen of the midgut gland tubules (Mikami and Takashima 2000). The gastric fluid passes through the diverticulae into the stomach, where it mixes with food to form chime (Factor and Naar 1990). The hydrolyzed food is absorbed in the lumen of the midgut gland by the R-cells, facilitating assimilation, storage, and resorption (Mikami and Takashima 2000). The gastric fluid of most decapods has a pH between 5 and 6. The gastric fluid of American and European lobsters, however, is slightly more acidic showing pH 4.7 (Baker and Gibson 1977; Brockerhoff et al. 1970; Navarrete del Toro et al. 2006; Omondi and Stark 2001).

L. Rojo · F. García-Carreño (✉)  
Centro de Investigaciones Biológicas del Noroeste (CIBNOR),  
Mar Bermejo 195, Col. Playa Palo de Santa Rita,  
La Paz, BCS 23096, Mexico  
e-mail: fgarcia@cibnor.mx

A. Muhlia-Almazan  
Centro de Investigación en Alimentación y Desarrollo (CIAD),  
Carretera a la Victoria Km. 0.6,  
Hermosillo, Sonora 83000, Mexico

R. Saborowski  
Alfred-Wegener-Institute für Polar- und Meeresforschung (AWI),  
Biologische Anstalt Helgoland,  
P. O. Box 180, 27483 Helgoland, Germany

For catalysis, endopeptidases are classified as serine, aspartic, metallo, and cysteine proteinases (Barrett et al. 1998). In crustaceans, proteolytic enzymes are of major importance in food digestion (Muhlia-Almazán and García-Carreño 2003), breaking down proteins into small peptides and amino acids that are required to fulfill many physiological processes (Neurath 1984). Decapods form a phylogenetically coherent group and possess anatomically similar digestive organs; they express digestive enzymes belonging to the four classes.

In the digestive tract of decapods, serine proteinases, such as trypsin and chymotrypsin (Galgani et al. 1985; Hernández-Cortés et al. 1999; Hernández-Cortés et al. 1997); cysteine proteinases, such as cathepsin L and cathepsin B (Aoki et al. 2004; Le Boulay et al. 1998; Teschke and Saborowski 2005); and metallo-proteinases, such as astacin (Titani et al. 1987; Zwilling and Stöcker 1997) were studied most intensively, while aspartic proteinases remained unnoticed. In clawed lobsters (Homaridae), digestive proteolytic enzymes have been reported since the early 1970s (Brokerhoff et al. 1970; Hoyle 1973), and, unlike other decapods, these proteinases show highest activity at acid pH (Baker and Gibson 1977; Glass and Stark 1994). Cysteine proteinases with optimal activity at acid pH were found in the gastric fluid of American lobster (Laycock et al. 1989) and verified by three cathepsin L cDNAs counterparts isolated from the midgut gland (Laycock et al. 1991). Recently, aspartic proteinases were suggested by Navarrete del Toro et al. (2006), but a specific protein or gene transcript has not yet been identified.

Aspartic proteinases (EC 3.4.23.X) have two aspartic acid residues at the catalytic site. The catalytic apparatus in all aspartic proteinases is virtually the same (Tang and Wong 1987), having no stable covalent acyl or amide intermediate present during catalysis (Davies 1990). Hydrolysis of the peptide bonds catalyzed by aspartic proteases proceeds by general base catalysis, where one water molecule is hydrogen-bonded to the carboxyl groups of both aspartate residues; the water molecule is partly displaced on substrate binding and polarized by one of the catalytic aspartate residues. It then initiates enzymatic catalysis by attacking the scissile carbonyl group bond of the substrate (Coates et al. 2008).

This class of endopeptidases is active at acid pH (Yonezawa et al. 1988). Four major groups of aspartic proteinases have been identified in vertebrates, namely pepsins, cathepsin Ds, cathepsin Es, and renins (Barrett 1979). Mammalian cathepsin D (EC 3.4.23.5) is conventionally regarded as a lysosomal proteinase that acts in intracellular protein turnover. To our knowledge, neither pepsin nor chymosin or cathepsin E activity has been demonstrated in any arthropod species. However, research on cathepsin D proteinases in helminthes, mites, and ticks reveal an extracellular

function for this enzyme (Becker et al. 1995; Tang and Wong 1987). Based on our research (Navarrete del Toro et al. 2006), we hypothesized that cathepsin D might be responsible for the extracellular proteinase activity at acid pH in the gastric fluid of clawed lobsters.

In this study, an aspartic proteinase from the gastric fluid of American and European lobsters was partly purified and identified by mass spectrometry as the active form of cathepsin D1. Then, we detected and amplified the mRNAs of cathepsin D from three tissues and sequenced two cDNAs encoding for two putative isoforms, cathepsin D1, and cathepsin D2. Finally, we assayed the biological function of cathepsin D1 by RT-PCR tissue-specific gene expression and quantitatively evaluated changes in the concentration of cathepsin D1 mRNAs during starvation and feeding episodes.

## Materials and Methods

### Specimens

Live adult American lobsters (*Homarus americanus*) and European lobsters (*Homarus gammarus*) were purchased in Cuxhaven, Germany. The lobsters ranged from 250 to 550 g. In the laboratory, the animals were maintained in aquariums with continuously flowing filtered seawater. The temperature of the water was maintained at  $13\pm0.5^{\circ}\text{C}$  and fed ad libitum with fish or small crustaceans before they were subjected to starvation experiments.

Gastric fluid was withdrawn from lobsters with a disposable 10-ml syringe with the needle replaced by a flexible plastic probe (5 cm long, 3 mm dia). The probe was inserted through the oral cavity into the gastric chamber of lobsters to withdraw at least 1 ml gastric fluid, which was transferred to a 1.5-ml microtube. The gastric fluid was centrifuged for 10 min at  $4^{\circ}\text{C}$  and  $10,000\times g$  to separate solids. The supernatant was removed, freeze-dried, and stored at  $-20^{\circ}\text{C}$  until used. Further assays were carried out with solutions of the freeze-dried gastric fluid ( $50 \text{ mg ml}^{-1}$ ).

### Purification and Identification of Aspartic Proteinases from the American and European Lobsters

An aspartic proteinase from the gastric fluid of both species was purified, based on the affinity of these enzymes for the highly specific aspartic proteinase inhibitor pepstatin A. The gastric fluid was loaded onto an affinity chromatography column containing pepstatin A-agarose (P2032, Sigma-Aldrich Chemical). Briefly, 1 ml of matrix was equilibrated in 50 mM citrate buffer at pH 4.0. Then, 1 ml gastric fluid was loaded onto the column and washed with 5 ml of the same buffer. The enzyme was eluted with 0.5 M Tris-HCl at pH 8.0

containing 1 M NaCl. Then, 0.5-ml fractions were collected, and the protein composition of each fraction was assessed on 12% SDS-PAGE. Proteolytic activity at acid pH was monitored by substrate SDS-PAGE, as described by García-Carreño in 1993 (García-Carreño et al. 1993). Hemoglobin (0.25%) dissolved in glycine-HCl buffer (100 mM, pH 3.0) was used as the substrate. All pH values for buffers were measured at room temperature (25°C).

#### Mass Spectrometry Analyses and N-terminal Sequencing

To identify the mass spectrometric profiles, 20 µg of the protein was first separated by 12% SDS-PAGE. After electrophoresis, gels were stained with Coomassie Blue R-250. Bands matching the size of those showing proteinase activity at low pH were manually cut from the gels. After tryptic digestion, the resulting peptides were subjected to tandem-time of flight (MS/MS-TOF) mass spectrometry.

Additionally, SDS-separated proteins were electroblotted onto polyvinylidene difluoride (PVDF) membranes and used for N-terminal sequencing. N-Terminal sequencing and mass spectrometry were performed at the Stanford Protein Facility, California. The resulting data were compared with the deduced amino acid sequence from cDNA of cathepsin D1 from American and European lobsters.

#### Cathepsin D mRNA Detection in Lobster Tissues

Cathepsin D mRNA was detected by RT-PCR. Muscle, gonad, and midgut gland tissues were dissected from lobsters, and the total RNA was isolated using TRIzol® reagent (15596-026, Invitrogen, Carlsbad, CA, USA). Total RNA concentration was determined spectrophotometrically at 260 nm, and DNA was removed from each sample with DNase I (AMPD-1, Sigma-Aldrich Chemical) following the manufacturer's instructions. One microgram total RNA was reverse-transcribed using a reverse transcription system (A3500, Promega, Madison, WI, USA) using oligo-dT as an anchor primer; cDNA samples from each tissue were used as templates for the following PCR reactions.

A search for sequences encoding cathepsin D proteinases from lobsters available at GenBank yielded nine ESTs from *H. americanus* cathepsin D mRNAs. A comparison among the nine EST outcomes on two discrete sequences corresponded to two aspartic proteinase putative isoenzymes: the cathepsin D isozyme 1(CatD1) (GenBank: CN852987) and cathepsin D isozyme 2 (CatD2) (GenBank: FD483089.1). The two EST sequences share 70% identity in their deduced amino acid sequences (Fig. 1). Based on these sequences, specific primers were designed to amplify short segments of each transcript (CatD1 and CatD2 in Table 1). Using the RT-PCR technique, we analyzed the presence of CatD1 and CatD2 transcripts in different tissues of both species. PCR amplification of cDNA fragments encoding cathepsin Ds was done in a final volume of 25 µl containing 20 pmol of each forward and reverse primer, 1 µl cDNA, 20 nmol dNTPs mixture, 2.5 µl 10× Taq buffer, and 1 U *Taq* DNA polymerase (18067-017, Invitrogen). PCR products were analyzed on 1.5% agarose gels and displayed under UV light after staining with SYBR®Safe DNA gel stain (S33102, Invitrogen).

#### Cathepsin D1 and Cathepsin D2 cDNA Sequence

To amplify the full length of the cDNA of cathepsin D1 and cathepsin D2 gene transcript from American and European lobsters, the 5'/3' rapid amplification of cDNA ends method was used (GeneRacer™ Kit, L1500, Invitrogen). The cDNA was prepared from 1 µg of total RNA isolated from the midgut gland; oligo nucleotide primers were designed to specifically anneal cathepsin D1 and cathepsin D2 (designated in Table 1 as Cat D1F, CatD1R, CatD 2 Race F, and CatD 2 NesRace F, respectively) and used as the internal gene-specific primers; 3' and 5' nested primers included in the kit were used to anchor the cDNA ends. PCR amplifications were carried out for 2 min at 94°C, followed by 30 cycles consisting of 30 s at 94°C, 30 s at 57 to 66°C (depending on primer set used), and 2 min at 68°C. In the last cycle, the extension step at 68°C lasted 10 min. PCR products were cloned using the TOPO-TA cloning kit (K4500-01, Invitrogen) and sequenced for both strands.

**Fig. 1** Comparison of deduced amino acid sequences of two isoforms from American clawed lobster *Homarus americanus* cathepsin D. Cathepsin D1 (GenBank: CN852987), cathepsin D2 (GenBank: FD483089.1). Asterisks show identical or conserved residues in all sequences in the alignment; colons are conservative substitutions; periods are semi-conservative substitutions

FD483089.1-CatD2 CN852987 -CatD1	SHQKGGRICMRHL-IFLVIVALATAELPRIPLHKIKSVRRLQEVDTAVTRAHRKWGNR -----MRSIAALLLLLALAAEELHRIPLKKIEKSR-TLQDLRRTRVFLNHRYG-V *** :*: :****:*** * * * : : . * * * : : . : : :
FD483089.1-CatD2 CN852987 -CatD1	GPMPEPLSNYMDAQYYGPISIGTPPQSFRVVFDTGSSNLWVPSKQCHYTNIACMIHNKYD GSDVIDLDNYEDAQQYGPITIGTPGQGFVDVIFTDGSSNLWIPSEKCFILNLACRLHNRYD . *.* * *****:***** *.* *:*****:*****:*. *:*** :***:***
FD483089.1-CatD2 CN852987 -CatD1	ARKSSTYKKNGTDFAIQYGSGSLSGYLSTDVTAVGSLAVRQQTFAEALSEPGLAFVAAKF STKSSTYIENGTAFDIQYGSGALHGFLSSDNVEMGGVNAMGQTFAEATQE PGLAFIMGKL . :***** :*** * * *****:*** *:***:*. *:.. * ***** .*****: .*:.
FD483089.1-CatD2 CN852987 -CatD1	DGILGMGFDNIAVGTVPTFYNMVKQSLIPAPWFSFYLNRPDSSP----- DGILGMAFTEISVMGIPTVFDTMVAQGAVDQPISFYLNHDVSDMNETLGGEVLVGGSDP *****:*. :*: *:..** .** *. : *:*****:*. *

**Table 1** Sequences of primers used for RT-PCR expression analyses, real-time PCR, and cDNA cloning and sequencing of cathepsin D1 and cathepsin D2

	Primer	Sequence (5'→3')	Target	Location
PCR primers	CatD 1 Forward	5' CTCAGTACTACGGCCCCATC 3'	Cathepsin D1	223–243
	CatD 1 Reverse	5' CCSAGGATGCCGTCTYAACTT 3'	Cathepsin D1	499–519
	CatD 2 Forward	5' AAAATGGGAAATCGAGGAC 3'	Cathepsin D2	168–188
	CatD 2 Reverse	5' TGGATGGCAAATCAGTTCC 3'	Cathepsin D2	399–414
	CatD 2 Race F	5' CATTGGCACTCCCCACAATCCTT 3'	Cathepsin D2	242–265
	CatD 2 NesRace F	5' GGTTCTTGGCTGTGAGGCAACAGA 3'	Cathepsin D2	463–487
qPCR primers	L38 Forward	5' ATTGAACGGGTGAGGTGAAC 3'	L38	ND
	L38 Reverse	5' ATCTGGAGCATCGAACCTTG 3'	L38	ND
	Q-CatD Forward	5' TGGATTCCCTCAGAGAAAGTG 3'	Cathepsin D1	301–320
	Q-CatD Reverse	5' TAGAGAGGAAGCCGTGAAGG 3'	Cathepsin D1	429–449

Location is indicated by nucleotide numbers from the 5'-end of the American lobster cathepsin D1 cDNA. S and Y indicate degenerate sites (S=C or G, Y=C or T)

The BLAST program (Altschul et al. 1997) was used to identify the resulting sequences by comparison and search for similarities in sequences. The deduced amino acid sequences were aligned for comparison using the Clustal X v. 1.81 software (Thompson et al. 1994). A prediction of the entire amino acid sequences was made using Predict-Protein software (Rost et al. 2004) and the SignalP 3.0 Server (Nielsen et al. 1997). This procedure identified both lobster proteins as cathepsin D proteinases and allowed us to infer some of their properties, such as the theoretical isoelectric point, molecular weight, signal peptide cleavage sites, and N-glycosylation sites.

#### Starvation and Re-feeding Experiment

Lobsters were acclimated to the experimental conditions for 7 days. After acclimation, 15 lobsters of each species were starved for 15 days and then fed with fish and small crustaceans to apparent satiety. Three lobsters were sampled at the following times: before starvation (day 1, control group); days 5, 10, and 15 during the starvation period; and 1 day after re-feeding. gonad, muscle, and digestive gland tissue samples were individually dissected from all these specimens. The tissues were preserved in RNAlater solution (AM7024, Ambion, Austin, TX, USA) and stored at  $-80^{\circ}\text{C}$  until used.

#### Quantification of Cathepsin D1 mRNA by qRT-PCR

Changes of cathepsin D1 mRNAs in the midgut gland during starvation and re-feeding were quantified by real-time PCR. The L38 ribosomal protein gene was used as a reference gene. Specific primers were designed to amplify both L38 and CatD1 genes (Table 1). For each tissue, we isolated total RNA, removed genomic DNA, and synthe-

sized cDNA as described in previous sections. Each cDNA sample was diluted 1:2 in sterile, distilled water and used as the template in PCR reactions.

Real-time PCR reaction mixtures contained the same components for both amplified genes; each reaction containing a total volume of 20  $\mu\text{l}$  included 1  $\mu\text{l}$  cDNA (approximately 25 ng of total RNA), 10  $\mu\text{l}$  2× iQ SYBR Green Super Mix, 50 nM of each Q-CatD forward and Q-CatD reverse primers for the cathepsin D1 gene reaction, and L38 forward and L38 reverse for the L38 gene (Table 1), as well as 7  $\mu\text{l}$  sterile, distilled water. Real-time PCR was performed in duplicate for each cDNA sample. Controls without template were included to ensure absence of contaminating DNA. An iQ5 cycler, real-time PCR detection system (170-9750, BioRad) was used with the following procedure: 1 cycle at  $95^{\circ}\text{C}$  for 5 min, 40 cycles at  $95^{\circ}\text{C}$  for 30 s each,  $61^{\circ}\text{C}$  for 65 s, and  $68^{\circ}\text{C}$  for 55 s. The specificity of each primer set was validated by the construction of a melt curve. Melt curves were constructed after the PCR by increasing the temperature from  $60^{\circ}\text{C}$  to  $94.5^{\circ}\text{C}$  at a rate of  $0.3^{\circ}\text{C}$  every 20 s.

PCR efficiency was validated by the construction of a standard curve for each primer pair. The amplification products of each gene were purified with the Illustra™ GFXTM PCR DNA and Gel Band Purification Kit (28-9034-70, GE Healthcare, Buckinghamshire, UK). The purified PCR products of cathepsin D1 and L38 mRNAs were spectrophotometrically quantified, and known concentrations were serially diluted 10-fold in sterile water, and the dilutions were used to generate standard curves for each gene by plotting the average  $C_T$  versus the related log dilutions (ng). Standard curves included data of four serially diluted PCR product samples, from  $25 \times 10^{-3}$  to  $250 \times 10^{-6}$  ng, and they were run at the same time as the experimental samples.

The  $2^{(-\Delta\Delta CT)}$  method was used for the relative quantification of mRNA concentrations of each gene evaluated and the threshold cycle value ( $C_T$ ); PCR amplification efficiencies were directly calculated from the standard curves with the formula,  $E=10^{(-1/\text{slope})}$ .

For a given sample, average  $C_T$  values of L38 were subtracted from the average  $C_T$  values of the cathepsin D1 reactions ( $\Delta C_T$  value). The mean  $\Delta C_T$  of the control group reactions (not starved or day 1) was then subtracted from the  $\Delta C_T$  of each starvation time and re-feeding group ( $\Delta\Delta C_T$ ). The relative mRNA concentrations were calculated by the  $2^{(-\Delta\Delta CT)}$  formula (Livak and Schmittgen 2001).

### Statistical Analysis

To determine the effects of starvation, the data were logarithmically transformed and analyzed by one-way ANOVA followed by Bonferroni multiple comparison test to determine differences among groups (Zar 1984). Statistical significance was set at  $P<0.05$ , using statistical software (Statgraphics Plus v. 5.1, StatPoint, Herndon, VA, USA).

## Results

### Purification and Identification of Aspartic Proteinases from the American and European Lobster

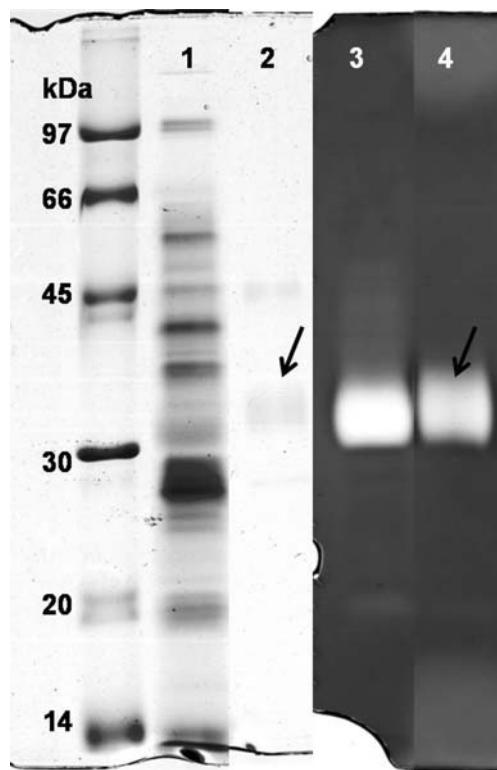
After affinity chromatography, fractions with activity showed two protein bands by SDS-PAGE of ~32 and ~45 kDa; however, the zymography at pH 3.0, with hemoglobin as the substrate, yielded only one band with proteinase activity. The apparent molecular mass of the band with activity was 32 kDa (Fig. 2). The results of MS/MS-TOF mass spectrometry results and the N-terminal sequence of this band matched the deduced amino acid sequence of American and European lobsters cathepsin D1, as described below (Fig. 3), and allowed us to identify it as a cathepsin D and named it cathepsin D1.

### Cathepsin D mRNA Detection in Lobster Tissues

Cathepsin D2 mRNA was detected in all the analyzed tissues from American and European lobster, whereas cathepsin D1 transcripts were found only in their midgut gland (Fig. 4).

### cDNA Encoding Cathepsin D1 and Cathepsin D2 in American and European Lobsters

We sequenced the full cDNA encoding cathepsin D1 and cathepsin D2 from the midgut gland of the American lobster and 90% of the cDNA of cathepsin D1 from the



**Fig. 2** SDS/PAGE and zymogram at pH 3 of the protease fraction purified by pepstatin affinity chromatography from gastric fluid of American lobster. Purified cathepsin D1 (32 kDa) is indicated by arrows. Lane 1 crude extract, lane 2 eluted fraction, lane 3 activity bands of crude extract, lane 4 activity band of eluted fraction

midgut gland of the European lobster. The sequences were submitted to GenBank (accession numbers EU687261 for the American and EU822806 for the European lobster cathepsin D1 mRNA sequences; accession number FJ943775 for American lobster cathepsin D2). Nucleotides and deduced amino acid sequences of American lobster digestive cathepsin D1 are shown in Fig. 3. The cDNA of cathepsin D1 was 1,292 bp. The translation start codon ATG is located at position 40, whereas the stop codon is located at position 1195. Cathepsin D1 cDNA has an ORF extending from position 40 to 1197 that codes for 386 amino acid residues with a predicted molecular mass of 42.6 kDa. The untranslated 3'-region contains 95 bp, which includes an AATAAA polyadenylation signal and a 25-bp polyadenylation (A) tail. A putative signal peptide was identified in the deduced amino acid sequence between residues 1 and 16. N-Terminal sequencing suggests that the active protein starts at amino acid residue 50; this outcome is based on a hypothetical activation peptide cleavage site between residues 17 and 50. Elimination of the activation peptide yielded a mature protein of 336 amino acid residues with a theoretical isoelectric point (pI) of 4.14 and a molecular mass of 36.4 kDa.

**Fig. 3** Nucleotide and deduced amino acid sequence in *Homarus americanus* cathepsin D1 cDNA (GenBank: EU687261). The putative signal peptide is indicated in *bold letters*. A potential asparagine glycosylation residue (designated as N) is marked by *square shapes*. The putative polyadenylation signal (AATAAA) is *underlined*. Aspartic amino acid residues (D) of the catalytic sites are marked by *diamond shapes*. The peptide sequences of the acid enzymatically active band obtained by mass spectrometry analyses are *boxed*. The N-terminal sequence is *underlined* with a dashed line

```

1      agtcacgacacctggccagacactcagcaccaccacaacatgagggtcccttgcgcctt
1      - - - - - - - - - - - - - - - - - M R S L A A A L
61     ctgctactgttggctctagctgccgcagagctccacagaattcctctgaagaaaatcgag
8      L L L L A A A E L H R I P L K K I E
121    aagtgcaggacacttcaggatctccgcgaaccccgagttcttgcataccatcgctacggc
28      K S R T L Q D L R R T R V F L N H R Y G
181    gtcggtagcgatgtaatcgatctcgacaaactacgaggatgtcagttactacgtcccatt
48      V G S D V I D L D N Y E D A Q Y Y G P I
241    accatcgggacacctggacaggattcgatgttatctcgatcatcggttcctccaatctg
68      T I G T P G Q G F D V I F D T G S S N L
301    tggattccctcagagaagtgtttatcctcaacttagctgcgcacttcataaccgctac
88      W I P S E K C [F I L N L A C R] L H N R Y
361    gattctaccaagtcttcaacctacattgagaacggcacagccttcgacatacagtatggc
108    D S T K S S T Y I E N G T A F D I Q Y G
421    tctggtgcccttcacggcttcctcttagcgacacgtggagatggcggtgtaaatggcc
128    S G A L H G F L S S D N V E M G G V N A
481    atggggccagaccttcgctgaggccacacaggagcctgtctcgccatcatcatggtaa
148    M G Q T F A E A T Q E P G L A F I M G K
541    ttagacggcatcctcgaaatggccttcacagagatctcgatggcatccctacagt
168    L D G I L G M A F T E I S V M G I P T V
601    ttgcacaccatggggccccaggcgccgtcgaccagccatcttccttctacctaacc
188    F D T M V A Q G A V D Q P I F S F Y L N
661    catgtatcgacatgaaacactgggaggagctgtgtggcggtcagat
208    H D V S D M N E T L G G E L V L G G S D
721    cccaaaccattacgagggagagttccactacgtgcctgtctccaaatcggtactggcag
228    P N H Y E G E F H Y V P V S K V G Y W Q
781    gtcactgctgaggcgattaaagggtggagacaatgtaaaccggctttgcaccccttgtgaa
248    V T A E A I K V G D N V T G F C N P C E
841    gccatcggtgacactggcacatccctcatcgctggacccaacgcggaggtaaggagatc
268    A I V D T G T S L I A G P N A E V K E I
901    gtacacatgctggaggctacggcttcatcgacgggaaatctcatcagctgccacaag
288    V H M L G G Y G F I A G E Y L I S C H K
961    gtacccgaaatgccagagttcacggtcacccctcaacggcaaggacttcagcatagacgga
308    V P E M P E F T F T L N G K D F S I D G
1021   cctgacccgttcatcgaggacattgaccaaggcaccggagtcaagatttgcattcggtggc
328    P D L V I E D I D P S T G V K I C I V G
1081   atcatggacttcagatggcgagctcgaggcctggatcctgggtatccctcatcgct
348    I M G L Q M G E L E A W I L G D P F I A
1141   gactggataactgagttgcgtggccagaaacgcataaggctcgccaagttccatctaa
368    D W Y T E F D V G Q K R I G F A K S I -
1201   atatttacaatgtggagcatggctttcaataaagaactggaggccttcaggatttc
1261   tactttgaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa

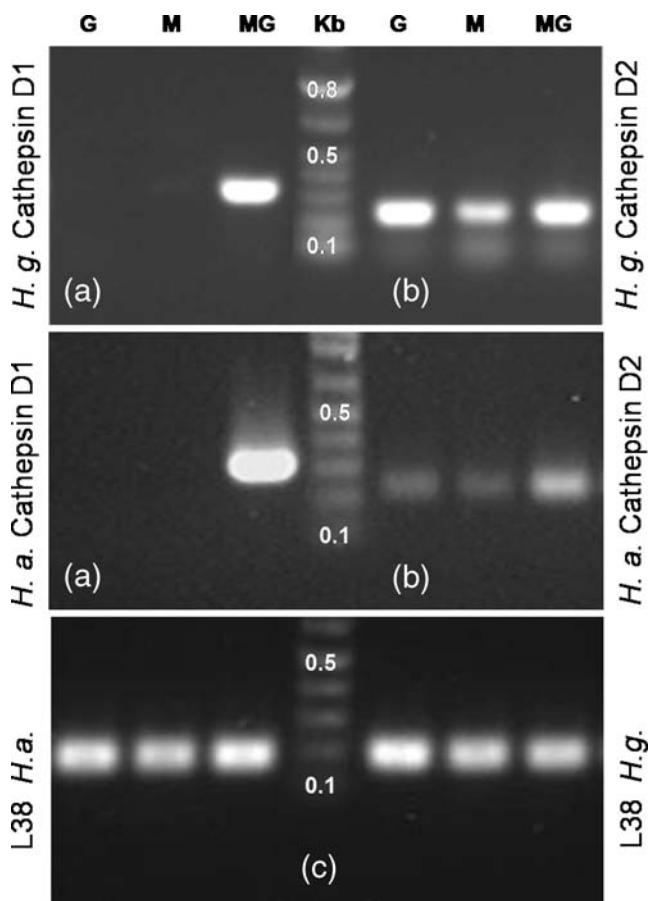
```

### Cathepsin D1 and Cathepsin D2 Deduced Protein Sequence Analysis

The amino acid sequences from American and European lobster cathepsin D1 showed that they are 98% identical. At least three peptide sequences obtained by mass spectrometry matched both sequences (Table 2).

BLASTX analysis of the deduced amino acid sequence (Altschul et al. 1997) revealed that American and European lobster cathepsin D1 and cathepsin D2 share high sequence similarity with arthropod aspartic proteinases. Amino acid residues at homologous positions, in the neighborhood of the active site, were used to identify both enzymes as

aspartic proteinase, specifically, a cathepsin D proteinase. BLASTX analysis demonstrated that 50% of the sequence of American lobster cathepsin D1 is identical to the amino acid sequence of *Apis mellifera* cathepsin D and *Tribolium castaneum* lysosomal aspartic proteinase, 48% of *Haemaphysalis longicornis* aspartic proteinase, and 49% of *Penaeus monodon* cathepsin D (Fig. 5). Remarkably, only 43% of the deduced amino acid sequence of American lobster cathepsin D2 is identical to American lobster cathepsin D1, whereas cathepsin D2 is 83% identical to *P. monodon* cathepsin D, 69% of *T. castaneum* cathepsin D isoform 1, and 64% of *Bombyx mori* cathepsin D (Gu et al. 2006).



**Fig. 4** Expression analyses of cathepsin D1 (a) and cathepsin D2 (b) genes in different tissues of *Homarus gammarus* (*H.g.*) and *Homarus americanus* (*H.a.*), using RT-PCR. Total RNA from gonad (G), muscle (M), and midgut gland (MG) tissues were used to generate first-strand cDNA. Expression of L38 rRNA (c) is shown as an internal control

#### Effect of Starvation on the Concentration of American and European Lobster Cathepsin D1 mRNA

Examination of the real-time PCR melt curves indicated that each PCR reaction produced a specific product. The efficiency of amplification was >94% for each primer set (results not shown). Starvation did not affect the concentration of mRNA of cathepsin D1 in the American lobster midgut gland, but feeding after starvation caused a 2.5-fold increase after 24 h (Fig. 6). Unexpectedly, the concentration of cathepsin D1 mRNA in the European lobster responded in a different manner to starvation, where after 15 days of starvation caused a significant reduction in the amount of cathepsin D1 mRNA to <40% of the control group and did not change with subsequent feeding.

#### Discussion

The exclusive occurrence of mRNA encoding for a cathepsin D1 in the midgut gland and the presence of its codified active

enzyme in the gastric fluid suggest that the enzyme is secreted by the midgut gland, the principal digestive organ in crustaceans, and that it moves to the gastric chamber, where it mixes with food. Because the pH of the gastric fluid is close to the optimum for cathepsin D1 activity, we assume that it must be involved in food protein hydrolysis.

#### General Features of Cathepsins

The historically derived collective name “cathepsins” comprises a set of structurally and functionally incoherent enzymes that facilitate intracellular digestion of protein (Barrett 1998). Cathepsins belong to the aspartic (cathepsin D), serine (cathepsin G), or cysteine (cathepsin L and B) proteinase families and are mostly located in the lysosomes.

In crustaceans, cathepsin activity occurs in the midgut gland of clawed lobsters, shrimp, and crabs (Glass and Stark 1994; Laycock et al. 1989; Navarrete del Toro et al. 2006; Omondi and Stark 2001). Most information about food digestion and cathepsin enzyme activity in crustaceans is available for cysteine proteinases, such as cathepsin L, found in the midgut gland of *Metapenaeus ensis* that performs extracellular food digestion (Hu 2003) and in the European brown shrimp *Crangon crangon*, where a substantial share of the proteolytic activity is caused by cysteine cathepsin L-like proteinases (Teschke and Saborowski 2005). Aoki et al. (2003) found that the mRNA of cathepsin B in the northern shrimp *Pandalus borealis* was almost exclusively found in the midgut gland, which also suggests that the enzyme is secreted to function as a digestive proteinase. In some crustaceans, the digestive vacuoles and lysosome-like vesicles contain cathepsin L, implying extracellular function (Hu and Leung 2007; Lehnert and Johnson 2002) and indicating different ways to hydrolyze food among crustaceans.

In contrast to cysteine proteinases, scarce information is available about the role of aspartic proteinases in extracellular digestive processes in crustaceans. Aspartic proteinases, such as cathepsin D, have been implicated in blood meal digestion of ectoparasitic mites and ticks (Boldbaatar et al. 2006; Hamilton et al. 2003). A first indication for the presence of aspartic proteinases in the gastric fluid of European lobsters was provided by Navarrete del Toro et al. (2006). We followed this lead and proved that lobsters express the cathepsin D1 mRNA

**Table 2** Aspartic protease short amino acid sequences matching with mass spectrometry and deduced amino acid sequences from the cathepsin D1 gene

Peptide	Calculated mass (Da)	Observed mass (Da)
CFILNLACR	1,109.55	1,109.53
LHNRYDSTK	1,133.56	1,133.54
VGYWQVTAEAIK	1,364.72	1,364.68

**Fig. 5** Sequence comparing deduced amino acids of cathepsin D1 with several known aspartic proteases based on primary structure.

Alignments were made with Clustal X (Version 1.81).

*Penaeus monodon* cathepsin D  
(GenBank: ABO10738.1),

*Tribolium castaneum* aspartic protease (GenBank:

protease (GenBank: XP\_966517.1), *Apis mellifera* similar to cathepsin D (GenBank: XP\_302857.2).

(GenBank: XP\_392857.2), *Haemaphysalis longicornis* aspartic proteinase (GenBank:

BAE53722.1). Asterisks represent identical or conserved

residues in all sequences in the alignment; *colons* are conserva-

tive substitutions; *periods* are semi-conservative substitutions. The region surrounding the active site is *boxed*

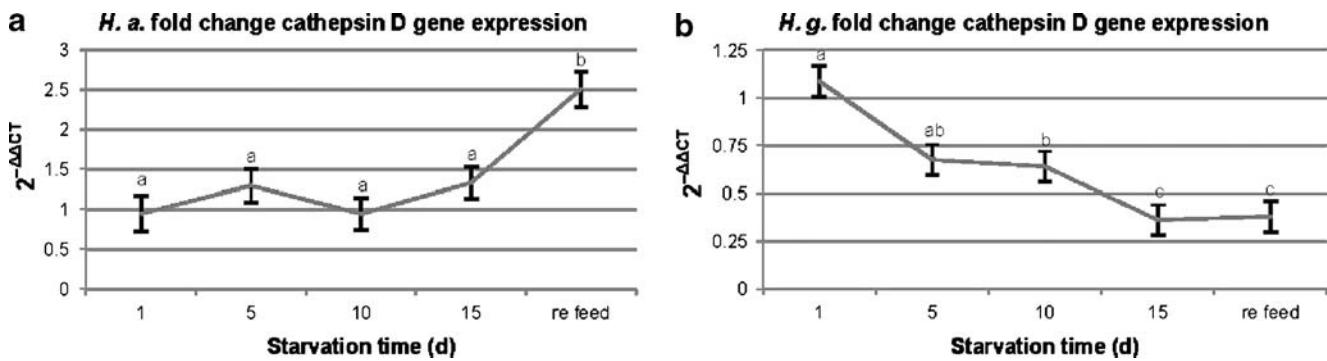
<i>P. monodon</i>		--MKVLLLALVALACADELPRIKLHKFKSARRSLQEVDTAVKVVH--RKWGNKGMPM	54
<i>H. americanus</i> CD2		--MRHLIFLVIVALATA-ELPRIPLHKIKSVRRLQEVDTAVTRAH--RKWGNRGMP	53
<i>T. castaneum</i>		--MLRSSVIFTIVCVLVALANAFVRVPVLYKVKVSARRSLQEVGTHVQQVRR--MRYG-GPTP	53
<i>A. mellifera</i>		--MFHAIIILCFCAFIAIANADVTIRIPLHKIDISRKQFKEYNTIEIYQTH--IFQG-DLPQP	54
<i>H. longicornis</i>		MTSTLSAIIVLLALVGASSALLRVPILKRSQARNLREVGTVLNVTHGWSLGNALGGPP	60
<i>H. americanus</i> CD1		--MRSLAALLLLALAAELHRIPKKIEKSRTLQLDLRRTRVFLNH---RYG-VGSDV	52
<i>H. gammarus</i>		-----IPLKIKEKSRTLQLDLRRTRVFLNH---RYG-VGSDV	32
		: * : . . * * : * : :	.
<i>P. monodon</i>		EPLSNYMDAQYYGPISIGTPQSFRVFTDGSSLWVPSKQCHFTNIACLIHNKYDATKS	114
<i>H. americanus</i> CD2		EPLSNYMDAQYYGPISIGTPQSFRVFTDGSSLWVPSKQCHFTNIACMIHNKYDARKS	113
<i>T. castaneum</i>		EPLSNYDAQYYGPISIGNPPQNFKVVFDTGSSLWVPSKKCHFTNIACLLHNKYDDSQS	113
<i>A. mellifera</i>		EPLSNYLDQAQYYGVVISIGTPQDFRVFDTGSSLWVPSKKCHLTNIACKLHRKYDNTKS	114
<i>H. longicornis</i>		EPLKNYLDQAQYYGDVTLGTPQFVRFVFTDGSSLWVPSKCPFTNIACIMLHHKYNSSKS	120
<i>H. americanus</i> CD1		IDLDNYEDAQYYGPITIGTPQGQFDVIFDTGSSLWIPSEKCFILNLACRLHNRYDSTKS	112
<i>H. gammarus</i>		IDLDNYEDAQYYGPITIGTPQGQFDVIFDTGSSLWIPSEKCFILNLACRLHNRYDSTKS	92
		* . * * ***** : : * . * * * * : * : * * : * : * : * : *	
<i>P. monodon</i>		STYKNGTKFDIQYGSGLSGYLSTDTSVGSVSKDQTFAEAMSEPGLAFVAAKFDGIL	174
<i>H. americanus</i> CD2		STYKNGTDFAIQYGSGLSGYLSTDTSVGSLAVRQQTFAEALSEPGLAFVAAKFDGIL	173
<i>T. castaneum</i>		KTYKNGTDFAIQYGSGLSGFLSTDIVTVGGLKVQQQTFAEAMSEPGLAFVAAKFDGIL	173
<i>A. mellifera</i>		STYKNGTDFAIRYGSGLSGYLSTDVTDIAGMKISDQTFEAEALSEPGLAFVAAKFDGIL	174
<i>H. longicornis</i>		STYAKNGTQEIRYGSGLSGVKEGLSTDVFLGDIRLQGQTFAEI LRESGLIAAKFDGIL	180
<i>H. americanus</i> CD1		STYIENGTAFDIQYGSGLAHGFLSSDNVEMGGVNAMGQTFAEATQEPGLAFIMGKLDGIL	172
<i>H. gammarus</i>		STYIENGTAFDIQYGSGLAHGFLSSDNVEMGGVNAMGQTFAEATQEPGLAFIMGKLDGIL	152
		* . * : * * : * : * : * . : . : * * * * . * . * : * : * : * : *	
<i>P. monodon</i>		GMAYDRIAVDGTPVVFYNNMVNQNVVAPIFSFYLNRDPAAA--EGGELILGGSDPAYYT	231
<i>H. americanus</i> CD2		GMGFNDIAVDGTPVVFYNNMVQKSLIPAPVFSFYLNRDPS --EGGELILGGSDPNYYS	230
<i>T. castaneum</i>		GMAYNRISVDGTPVVFYNNMQNLLVAQPVFSFYLNRDPSAA--QGGEI ILGGSDPNAHY	230
<i>A. mellifera</i>		GMAYSKISVDGTPVVFYNNMVQKGLVPQPVFSFYLNRPDDK--YGGELILGGSDPNHYE	231
<i>H. longicornis</i>		GLGYPQISVLNVPPVFDNMVAQGVAPKPFVFSVYLDRNASDP --NGEVLFGGIDEAHYT	237
<i>H. americanus</i> CD1		GMAFTEISVMGIPTVFDTMVQAGAVDQPIFSFYLNHDVSDMNETLGGEVLVGGSDPNHYE	232
<i>H. gammarus</i>		GMAFTEISVMGIPTVFDTMVQAGAVDQPIFSFYLNHDVSDMNETLGGEVLVGGSDPNHYE	212
		* : . * : * . : * . : * . : * : * : * : * : * : * : * : * : *	
<i>P. monodon</i>		GDFTYVPVDRQGYWQFKMDGLQMNGTTVPCDGGEAIAADTGTSLAAPSEEARLINKKI	291
<i>H. americanus</i> CD2		GNFTYIIPVDRKGYWQIKMDGIQMNGARVPFCCEGGCEIAADTGTSLAAPVEARSINKKI	290
<i>T. castaneum</i>		GDFTYLSDVRQAYWQFKMDSISVGGKNT-FCANGCEAIAADTGTSLIAGPVSEVQGINKAI	289
<i>A. mellifera</i>		GSFTYVPDVDDKGYWQFRMDSIQIG-SDLKVCQGGCEAIAADTGTSLIAGPVKEIEAINKAI	290
<i>H. longicornis</i>		GNITYVPVTRKGYWQFHMGNGVKG-DNATFCNGCAAIAADTGTSLIAGPTEEIHKLNVAI	296
<i>H. americanus</i> CD1		GEFHYPVPSKVGYWQVTAEAIKVGDNVTGFCNP-CEAIVDTGTSLIAGPNAEVKEIVHML	291
<i>H. gammarus</i>		GEFHYPVPSKVGYWQVTAEAIKVGDNVTGFCNP-CEAIVDTGTSLIAGPNAEVKEIVHMM	271
		* . : * : * : . * . : . : * : * : * : * : * : * : * : * : * : *	
<i>P. monodon</i>		GAKPIMGGEWSVDCNLIPLHPTISFVLAGKPFITLEKDYIILR--VSQFGQTTCLSGFIGL	349
<i>H. americanus</i> CD2		GAKPIASGEWSVDCSLIPLHPLKISFVLNGQPFITLEKDYIILK--VSVFGREECVSGFIGL	348
<i>T. castaneum</i>		GATPIVGGEMVDCNLIPNPLIDFTLGGKNTITLEKDYVLR--VAQMGKTICLSGFMGI	347
<i>A. mellifera</i>		GATPIIAAGEAMIDCNSIPNLIPTINFLVGGKSFSLKGEDYVLIK--VTQFGKTVCLSGFMGI	348
<i>H. longicornis</i>		GAAPFMAGEYIVVSCKSIPTLPKINFNLMNGNEFVLEGKDYVLQ--VSQAGIPLCLSGFIGL	354
<i>H. americanus</i> CD1		GGYGFIAGEYLISCHKVPPEMPEFTFTLNGKDFSI DGPDLVIEDIDPSTGVKICIVGIMGL	351
<i>H. gammarus</i>		GGYAFIAGEYLISCHKVPPEMPEFTFTLNGKDFSI DGPDLVIEDIDPNTGVKICICGHHGT	331
		* . : . * : . * : * : * : * : * : * : * : * : * : * : * : * : * : *	
<i>P. monodon</i>		DVPPPMGPIWILGDIFIGRFYTEFDGMGNRVRGFAESK	386
<i>H. americanus</i> CD2		DVPPPMGPLWILGDTFIGRFYTEFDLGNRVRGFAIAK	385
<i>T. castaneum</i>		DIPPPNGPLWILGDFVIGKFYTFEDLGNRVRGFAVAA	384
<i>A. mellifera</i>		DISPPNGPLWILGDFVIGRFYTYTEFDGMGNRVRGFAAK	385
<i>H. longicornis</i>		DVPAPLGPLWILGDFVIGRFYTYTFDRGNDRVRGFAESR	391
<i>H. americanus</i> CD1		QMGEE--LEAWLDPFIADWYTFEDVGQKRIGFAKSI	386
<i>H. gammarus</i>		SDGR--ARG--LDP-----	341

in their midgut glands and secrete the enzyme into the tubule lumen. It finally appears as an extracellular digestive enzyme in the gastric fluid of the gastric chamber.

The mechanism of cathepsin D1 secretion in lobsters is unknown. We think that the model proposed by Hu and Leung (2007) for secretion of cathepsin L in shrimp applies to cathepsin D1 in lobsters. They demonstrate the presence of some small cathepsin L vesicles in the midgut gland lumen by using RNA and protein probes and suggest that

enzyme vesicles result from apocrine (in addition to holocrine) secretion. They suggest that F-cells are the site of enzyme transcription and B-cells are responsible for enzyme secretion. In lobsters, final confirmation can only be provided by detailed histological studies.

There are at least two cathepsin D mRNAs transcribed in clawed lobsters. The D2 form was expressed in three tissues examined in this study. We suggest that it covers a lysosomal function similar to its mammalian counterpart.



**Fig. 6** Changes in cathepsin D1 gene expression in the midgut gland during starvation and feeding as measured by real-time PCR in American lobster (a) and European lobster (b). Cathepsin D1 mRNA concentrations were normalized with the mRNA concentrations of L38 rRNA from control group (no starved lobster). Lobsters were sampled before starvation (day 1); during starvation on days 5, 10, and

15 (end of starvation); and 1 day following re-feeding. The mean of three animals in each group and test are shown. Bars indicate SE of the mean. Differences were determined by one-way ANOVA after logarithmic transformation. Different letters indicate significantly different means from controls for each sample time tested ( $P<0.05$ )

The D1 form was restricted to the midgut gland. The ability of midgut gland cells to secrete cathepsin D1 suggests that clawed lobsters follow a strategy to attain extracellular enzymes with originally intracellular enzymes and, thus, to digest food efficiently under conditions found in their habitats, as observed by Mukhin et al. (2007).

One of the strategies exhibited by poikilothermic animals living at low temperatures to achieve a normal level of proteolysis at temperatures well below that of homeothermic species (Carginale et al. 2004; Smalås et al. 1994) involve their digestive enzymes, which show higher catalytic efficiencies at low temperatures and can be inactivated at relatively low temperatures and mild pH changes (Gerday et al. 2000; Gudmundsdóttir and Pálsson 2005; Leiros et al. 2000; Smalås et al. 1994), and those features make them attractive for their applications in the food, cosmetic, pharmaceutical, and domestic industries (Debashish et al. 2005; Gerday et al. 2000; Haard 1991; Kirk et al. 2002; Schäfer et al. 2005; Shahidi and Janak Kamil 2001). Considering that American and European lobsters live in marine environments where the temperature varies from 0°C to 25°C, depending on season, winds, and tides (Cobb and Phillips 1980); they are by definition, cold-tolerant animals (Crossin et al. 1998; Feller and Gerday 1997; Feller and Gerday 2003), and it is expected that their digestive enzymes will exhibit adaptations similar to other poikilothermic species as described above; however, an extensive study of its biochemical characteristics must be performed to validate our assumption and therefore its applicability.

#### Structural Properties of Lobster Cathepsin D1

Structurally, cathepsins D1 in American and European lobsters are similar to other aspartic proteinases from the pepsin family. Mammalian cathepsin D is glycosylated in Asn residues at positions 67 and 182 (Faust et al. 1985; Nakao et al. 1984).

This carbohydrate portion of cathepsin D seems to play a functional role in lysosomal transportation. It is mediated by the mannose-6-phosphate receptor that recognizes phosphorylated asparagine-linked oligosaccharides of lysosomal cathepsins during biosynthesis (Towatari et al. 1998). According to the deduced primary structure of American lobster cathepsin D1, the molecule contains three potential glycosylation sites located at positions Asn-118, Asn-214, and Asn-258. Two of the glycosylation sites (Asn-118 and Asn-258) are conserved and homologous to positions Asn-70 and Asn-199 of human cathepsin D (Faust et al. 1985) and to positions Asn-67 and Asn-182 of porcine cathepsin D (Shewale and Tang 1984; Takahashi et al. 1983). However, the nature of the carbohydrate side chain and the exact sites remain to be investigated, so the signaling for transportation to lysosomes via glycosylation and, therefore, its intracellular function is yet conjectural.

In mammals, cathepsin D is synthesized as a pre-pro-enzyme. In the first proteolytic post-translational modification, the “pre”-peptide is cleaved off. The signal sequence is usually 20 amino acids long and is removed within the rough endoplasmic reticulum. According to PredictProtein software (Rost et al. 2004), the sequence of American lobster cathepsin D1 shows a well-defined signal peptide, a feature of lysosomal aspartic proteases, such as cathepsins D. In eukaryotes, signal sequences are essential for the efficient and selective targeting of nascent protein chains to the endoplasmic reticulum (Giersch 1989; von Heijne 1990). Since clawed lobster cathepsin D1 bears a signal peptide, this indicates that the molecule is tagged for translocation across membranes for effective intracellular flow.

#### Zymogen—Activation

Zymogens formed by the first post-translational modification migrate to the lysosomes and undergo further proteo-

lytic modification in the acid lysosomal milieu. The activation peptide (40–60 amino acids) is cleaved off by both auto-proteolysis and proteolysis by other, mainly cysteine lysosomal peptidases that yield active cathepsin D (Fusek and Větvička 2005).

Our sequence analyses showed that similar to the mammalian lysosomal counterpart, lobster cathepsin D1 also carries an activation peptide of 34 amino acid residues (Fig. 3). This activation peptide is shorter than in mammalian cathepsin D (44 residues), but similar in size to the digestive aspartic proteinase (36 to 38 residues) in the mosquito (Cho and Raikhel 1992). In mammals, activation of the aspartic proteinase zymogens occurs auto-catalytically at acid pH, and the process may be similar to the activation of pepsinogen. At acid pH, most carboxyl groups of pepsinogens are protonated, which weakens the electrostatic interactions and causes strong conformational changes in the activation peptide. The active site is then exposed, and the enzyme cleaves off its own activation peptide (Kageyama 2002).

Animals expressing digestive aspartic proteinases usually express cysteine proteinases, suggesting a multi-enzyme network and cascade-like digestion of food (Boldbaatar et al. 2006; Delcroix et al. 2006). Cysteine proteinase activity was already reported in the midgut gland of the American lobster (Laycock et al. 1989). The aspartic proteinase cathepsin D1 described in this study and cysteine proteinases seem to contribute to the pool of acid proteinases previously reported by Baker and Gibson (1977), playing an important role in protein digestion in clawed lobsters. The activation mechanism of aspartic proteinases in the gastric fluid of clawed lobsters is still unknown, but given that gastric fluids of lobsters are acid, pH ~4.7 and also contain cysteine proteinases, it may undergo auto-proteolysis along with assisted activation, similar to pepsinogen and lysosomal procathepsin D of mammals (Tang and Wong 1987). Both aspartic and cysteine proteinases may form a multi-enzyme network comparable to that described for platyhelminths and ticks (Delcroix et al. 2006; Sojka et al. 2008).

Cathepsin D in vertebrates can undergo further cleavage to generate an active enzyme. Cleavage at the so-called  $\beta$ -hairpin loop forms a two-polypeptide chain with a N-terminal, 15-kDa light chain and a C-terminal, 30-kDa heavy chain (Yonezawa et al. 1988). In humans, the two chains emerge by cleaving off a small peptide of 7 amino acids between positions 98 through 105 (Fusek and Větvička 2005). Those chains are linked by hydrophobic bonds (Minarowska et al. 2008). Clawed lobster cathepsin D1 has the pro-enzyme sequence (Fig. 3), but lacks a sequence indicating the presence of the  $\beta$ -hairpin loop; this seems to be a common feature of invertebrate cathepsins D (Boldbaatar et al. 2006; Cho and Raikhel 1992; Harrop et al. 1996) and suggests an alternative activation mechanism, yet to be investigated.

## Regulation—Starvation

Starvation can cause acute biological responses in crustaceans, including changes in the composition of digestive enzymes (Sánchez-Paz et al. 2006; Sánchez-Paz et al. 2003). Samples obtained in the bioassay were analyzed to determine if starvation and subsequent feeding can cause significant changes in the concentration of cathepsin D1 mRNAs in the midgut gland of the lobsters. We found that, for both species, expression of cathepsin D1 mRNA is regulated, although the pattern in each species is not the same.

For specimens of American lobster, the concentration of cathepsin D1 mRNA did not vary during starvation; however, feeding after starvation led to a remarkable 2.5-fold increase in the concentration of cathepsin D1 mRNA (Fig. 6). This increase may reflect the physiological demand to digest protein on the first day after feeding, similar to aspartic proteinase genes in sea bass and rainbow trout (Salem et al. 2007; Terova et al. 2007). Regulation of digestive proteinases activity is achieved via transcriptional, post-transcriptional, and/or post-translational control. Changes in the quantity of mRNAs encoding digestive cathepsin D1 in American lobster during starvation and subsequent feeding suggests that regulation of cathepsin D1 could happen during transcription by increasing mRNA synthesis to enhance digestion.

Cathepsin D1 mRNA in European lobster showed a different course with continuous decline in mRNA concentrations during 15 days of starvation. This agrees with observation of trypsin mRNAs, the main digestive enzyme of whiteleg shrimp, where Sánchez-Paz et al. (2003) reported that the amount of trypsin mRNA dropped significantly after 5 days of starvation. Additionally, unlike American lobster, European lobster does not increase the quantity of cathepsin D1 mRNA in re-fed animals; this difference between the two species could be a sampling time artifact in which diagnosis after 1 day of feeding was not enough to detect any change for European lobster. Overall, the different trends in cathepsin D mRNA during the starvation period between the ecologically equivalent species may be caused by a feature of biochemical adaptation to accomplish efficient food digestion in its particular environment and challenges, but the biological implications remain to be assessed.

Studies of the midgut gland ultrastructure in *Penaeus semisulcatus* revealed that waves of E-cell mitoses and F-, R-, and B-cell differentiation were induced after feeding (Al-Mohanna and Nott 1987). The observed changes in cathepsin D1 mRNAs concentration in American lobster after feeding might also be caused by a wave of F/B-cell differentiation induced by feeding, which would result in an increase in cathepsin D1 mRNA. This reflects a larger number of F/B-cells expressing cathepsin D1 mRNA, rather than an increase in the concentration of cathepsin D1 in individual cells.

However, changes in the concentration of mRNA might not necessarily result in physiologically relevant changes in the concentration of the enzyme (Sánchez-Paz et al. 2003).

In summary, this study confirmed the presence of aspartic proteinases in extracellular gastric juice of clawed lobsters and addressed important questions about the site of synthesis and the evolution of proteolytic digestive enzymes in crustaceans. This contributed to understanding of group-specific preferences among decapod crustaceans in expressing different proteinase classes (Navarrete del Toro et al. 2006; Teschke and Saborowski 2005). These results are important in understanding the adaptive options of species from an ecological as well as economic perspective.

**Acknowledgments** We thank M. Navarrete del Toro at CIBNOR for sharing knowledge and invaluable advice on the proteinase experiments and I. Fogel at CIBNOR for valuable comments on this manuscript. L.C.R.A. received a doctoral fellowship from CONACYT.

## References

- Al-Mohanna SY, Nott JA (1987) R-cells and the digestive cycle in *Penaeus semisulcatus* (Crustacea: Decapoda). Mar Biol 95:129–137
- Altschul SF, Madden TL, Schaffer AA, Zhang J, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST a new generation of protein database search programs. Nucleic Acids Res 25:3389–3402
- Aoki H, Nazmul Ahsan M, Shugo W (2003) Molecular cloning and characterization of cathepsin B from the hepatopancreas of northern shrimp *Pandalus borealis*. Comp Biochem Physiol B Biochem Mol Biol 134:681–694
- Aoki H, Ahsan M, Watabe S (2004) Molecular and enzymatic properties of a cathepsin L-like proteinase with distinct substrate specificity from northern shrimp (*Pandalus borealis*). J Comp Physiol B-Biochem Syst Environ Physiol 174:59–69
- Baker PL, Gibson R (1977) Observations on the feeding mechanism, structure of the gut, and digestive physiology of the European lobster *Homarus gammarus*. J Exp Mar Biol Ecol 26:297–324
- Barrett AJ (1979) Cathepsin D: the lysosomal aspartic proteinase. Ciba Found Symp 79:37–50
- Barrett AJ (1998) Cathepsin D: the lysosomal aspartic proteinase. In: Barrett AJ, Rawlings ND, Woessner JF (eds) Handbook of proteolytic enzymes. Academic, San Diego, pp 37–50
- Barrett AJ, Rawlings ND, Woessner JF (1998) Proteolytic enzymes. In: Barrett AJ, Rawlings ND, Woessner JF (eds) Handbook of proteolytic enzymes. Academic, San Diego, pp 801–805
- Becker MM, Harrop SA, Dalton JP, Kalinna BH, McManus DP, Brindley PJ (1995) Cloning and characterization of the *Schistosoma japonicum* aspartic proteinase involved in hemoglobin degradation. J Biol Chem 270:24496–24501
- Boldbaatar D, Sikasunge CS, Battsetseg B, Xuan X, Fujisaki K (2006) Molecular cloning and functional characterization of an aspartic protease from the hard tick *Haemaphysalis longicornis*. Insect Biochem Mol Biol 36:25–36
- Brockhoff H, Hoyle RJ, Hwang PC (1970) Digestive enzymes in the American lobster (*Homarus americanus*). J Fish Res Board Can 27:1357–1370
- Carginale V, Trinchella F, Capasso R, Parisia E (2004) Gene amplification and cold adaptation of pepsin in Antarctic fish: a possible strategy for food digestion at low temperature. Gene 336:195–205
- Coates L, Tuan H-F, Tomanicek S, Kovalevsky A, Mustyakimov M, Erskine P, Cooper J (2008) The catalytic mechanism of an aspartic proteinase explored with neutron and X-ray diffraction. J Am Chem Soc 130:7235–7237
- Cobb JS, Phillips BF (1980) The biology and management of lobsters. Academic, New York
- Cho WL, Raikhel AS (1992) Cloning of cDNA for mosquito lysosomal aspartic protease. J Biol Chem 267:21823–21829
- Crossin G, Al-Ayoub SA, Jury SH, Hunting H, Watson WH III (1998) Behavioral thermoregulation in the American lobster *Homarus americanus*. J Exp Biol 201:365–374
- Davies DR (1990) The structure and function of the aspartic proteinases. Annu Rev Biophys Biophys Chem 19:189–215
- Debashish G, Malay S, Barindra S, Joydeep M (2005) Marine enzymes. Marine Biotechnology I 96:189–218
- Delcroix M, Sajid M, Caffrey C, Lim K, Jan D, Hsieh I, Bahgat M, Dissous C, McKerrow J (2006) A multienzyme network functions in intestinal protein digestion by a platyhelminth parasite. J Biol Chem 281:39316–39329
- Factor JR, Naar M (1990) The digestive system of the lobster, *Homarus americanus*: II. Terminal hepatic arterioles of the digestive gland. J Morphol 206:283–291
- Faust PL, Kornfeld S, Chirgwin JM (1985) Cloning and sequence analysis of cDNA for human cathepsin D. Proc Natl Acad Sci USA 82:4910–4914
- Feller G, Gerday C (1997) Review, psychrophilic enzymes: molecular basis of cold adaptation. Cell Mol Life Sci 53:830–841
- Feller G, Gerday C (2003) Psychrophilic enzymes: hot topics in cold adaptation. Nat Rev Microbiol 1:200–208
- Fusek M, Větvička V (2005) Dual role of cathepsin D: ligand and protease. Biomed Papers 149:43–50
- Galgani F, Benyamin Y, van Wormhoudt A (1985) Purification, properties and immunoassay of trypsin from *Penaeus japonicus*. Comp Biochem Physiol B 81:447–452
- García-Carreño FL, Dimes LE, Haard N (1993) Substrate-gel electrophoresis for composition and molecular weight of proteinases of proteinaceous proteinase inhibitor. Anal Biochem 214:65–69
- Gerday C, Aittaleb M, Bentahir M, Chessa J-P, Claverie P, Collins T, D'Amico S, Dumont J, Garsoux G, Georlette D, Hoyoux A, Lonhienne T, Meuwis MA, Feller G (2000) Cold-adapted enzymes: from fundamentals to biotechnology. Trends Biotechnol 18:103–107
- Giersch LM (1989) Signal sequences. Biochemistry 28:923–930
- Glass HJ, Stark JR (1994) Protein digestion in the European lobster, *Homarus gammarus* (L.). Comp Biochem Physiol 108B:225–235
- Gudmundsdóttir Á, Pálsdóttir HM (2005) Atlantic cod trypsins: from basic research to practical applications. Mar Biotechnol 7:77–88
- Gui Z, Lee K, Kim B, Choi Y, Wei Y, Choo Y, Kang P, Yoon H, Kim I, Je Y, Seo S, Lee S, Guo X, Sohn H, Jin B (2006) Functional role of aspartic proteinase cathepsin D in insect metamorphosis. BMC Dev Biol 6:49
- Haard NF (1991) A review of proteolytic enzymes from marine organisms and their application in the food industry. J Aquat Food Prod Technol 1:17–35
- Hamilton KA, Nisbet AJ, Lehane MJ, Taylor MA, Billingsley PF (2003) A physiological and biochemical model for digestion in the ectoparasitic mite, *Psoroptes ovis* (Acari: Psoroptidae). Int J Parasitol 33:773–785
- Harrop SA, Prociv P, Brindley PJ (1996) Acasp, a gene encoding a cathepsin D-like aspartic protease from the hookworm *Ancylostoma caninum*. Biochem Biophys Res Commun 227:294–302
- Hernández-Cortés P, Whitaker JR, García-Carreño FL (1997) Purification and characterization of chymotrypsin from *Penaeus vannamei* (Crustacea:Decapoda). J Food Biochem 21:497–514
- Hernández-Cortés P, Cerenius L, García-Carreño FL, Soderhal K (1999) Trypsin from *Pacifastacus leniusculus* hepatopancreas:

- purification and cDNA cloning of the synthesized zymogen. *Biol Chem* 380:499–501
- Hoyle RJ (1973) Digestive enzyme secretion after dietary in the American lobster (*Homarus americanus*). *J Fish Res Board Can* 30:1647–1653
- Hu KJ (2003) Molecular cloning and characterization of the cathepsin L gene from the marine shrimp *Metapenaeus ensis*. The University of Hong Kong, China
- Hu KJ, Leung PSC (2007) Food digestion by cathepsin L and digestion-related rapid cell differentiation in shrimp hepatopancreas. *Comp Biochem Physiol* 146B:69–80
- Kageyama T (2002) Pepsinogens, progastricsins, and prochymosins: structure, function, evolution, and development. *Cell Mol Life Sci* 59:288–306
- Kirk O, Borchert TV, Fuglsang CC (2002) Industrial enzyme applications. *Curr Opin Biotechnol* 13:345–351
- Laycock MV, Hirama T, Hasnain S, Watson D, Storer A (1989) Purification and characterization of a digestive cysteine proteinase from the American lobster (*Homarus americanus*). *Biochem J* 263:439–444
- Laycock MV, MacKay RM, Di Fruscio M, Gallant JW (1991) Molecular cloning of three cDNAs that encode cysteine proteinases in the digestive gland of the American lobster (*Homarus americanus*). *FEBS Lett* 292:115–120
- Le Boulay C, Sellos D, Van Wormhoudt A (1998) Cathepsin L gene organization in crustaceans. *Gene* 218:77–84
- Lehnert SH, Johnson SE (2002) Expression of hemocyanin and digestive enzyme messenger RNAs in the hepatopancreas of the black tiger shrimp *Penaeus monodon*. *Comp Biochem Physiol* 133B:163–171
- Leiros HK, Willassen NP, Smalås AO (2000) Structural comparison of psychrophilic and mesophilic trypsin. Elucidating the molecular basis of cold-adaptation. *Eur J Biochem* 267:1039–1049
- Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the  $2^{-\Delta\Delta CT}$  method. *Methods* 25:402–408
- Mikami S, Takashima F (2000) Functional morphology of the digestive system. In: Phillips BF, Kittaka J (eds) Spiny lobsters: fisheries and culture. Blackwell, London, pp 601–610
- Minarowska A, Gacko M, Karwowska A, Minarowski Ł (2008) Human cathepsin D. *Folia Histochem Cytobiol* 46:23–38
- Muhlia-Almazán A, García-Carreño FL (2003) Digestion physiology and proteolytic enzymes of crustacean species of the Mexican Pacific Ocean. In: Hendrickx ME (ed) Contributions to the study of east Pacific crustaceans 2. UNAM, Mexico City, pp 77–91
- Mukhin V, Smirnova E, Novikov V (2007) Peculiarities of digestive function of proteinases in invertebrates—habitants of cold seas. *J Evol Biochem Physiol* 43:476–482
- Nakao Y, Kozutsumi Y, Kawasaki T, Yamashina I, Van Halbeek H, Vliegenthart JFG (1984) Oligosaccharides on cathepsin D from porcine spleen. *Arch Biochem Biophys* 228:43–54
- Navarrete del Toro MA, García-Carreño FL, Díaz LM, Celis-Guerrero L, Saborowski R (2006) Aspartic proteinases in the digestive tract of marine decapod crustaceans. *J Exp Zool* 305A:645–654
- Neurath H (1984) Evolution of proteolytic enzymes. *Science* 224:350–357
- Nielsen H, Engelbrecht J, Brunak S, von Heijne G (1997) Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Eng* 10:1–6
- Omondi J, Stark JR (2001) Studies on digestive proteases from midgut glands of a shrimp, *Penaeus indicus*, and a lobster, *Nephrops norvegicus*. *Appl Biochem Biotechnol* 90:137–153
- Rost B, Yachdav G, Liu J (2004) The PredictProtein server. *Nucleic Acids Res* 32:321–326
- Salem M, Silverstein J, Rexroad CE, Yao J (2007) Effect of starvation on global gene expression and proteolysis in rainbow trout (*Oncorhynchus mykiss*). *Genomics* 8:328
- Sánchez-Paz JA, García-Carreño FL, Muhlia-Almazán A, Hernández-Saavedra N, Yepiz-Plascencia G (2003) Differential expression of trypsin mRNA in the white shrimp (*Penaeus vannamei*) midgut gland under starvation conditions. *J Exp Mar Biol Ecol* 292:1–17
- Sánchez-Paz A, García-Carreño F, Muhlia-Almazán A, Peregrino Uriarte AB, Hernández-López J, Yepiz-Plascencia G (2006) Usage of energy reserves in crustaceans during starvation: status and future directions. *Insect Biochem Mol Biol* 36:241–249
- Sánchez-Paz A, García-Carreño FL, Hernández-López J, Muhlia-Almazán A, Yepiz-Plascencia G (2007) Effect of short-term starvation on hepatopancreas and plasma energy reserves of the Pacific white shrimp (*Litopenaeus vannamei*). *J Exp Mar Biol Ecol* 340:184–193
- Schäfer T, Kirk O, Borchert TV, Fuglsang CC, Pedersen S, Salmon S, Olsen HS, Deinhammer R, Lund H (2005) Enzymes for technical applications. In: Steinbüchel A, Rhee SK (eds) Polysaccharides and polyamides in the food industry: properties, production, and patents. Wiley-VCH, Weinheim, pp 557–618
- Shahidi F, Janak Kamil YVA (2001) Enzymes from fish and aquatic invertebrates and their application in the food industry. *Trends Food Sci Technol* 12:435–464
- Shewale JG, Tang J (1984) Amino acid sequence of porcine spleen cathepsin D. *Proc Natl Acad Sci USA* 81:3703–3707
- Smalås AO, Heimstad ES, Hordvik A, Willlassen NP, Male R (1994) Cold adaption of enzymes: structural comparison between salmon and bovine trypsins. *Protein Struct Funct Genet* 20:149–166
- Sojka D, Franta Z, Horn M, Hajdusek O, Caffrey C, Mares M, Kopácek P (2008) Profiling of proteolytic enzymes in the gut of the tick *Ixodes ricinus* reveals an evolutionarily conserved network of aspartic and cysteine peptidases. *Parasit Vectors* 1:7
- Takahashi T, Schmidt PG, Tang J (1983) Oligosaccharide units of lysosomal cathepsin D from porcine spleen. Amino acid sequence and carbohydrate structure of the glycopeptides. *J Biol Chem* 258:2819–2830
- Tang J, Wong RNS (1987) Evolution in the structure and function of aspartic proteases. *J Cell Biochem* 33:53–63
- Terova G, Rimoldi S, Larghi S, Bernardini G, Gornati R, Saroglio M (2007) Regulation of progastrics in mRNA levels in sea bass (*Dicentrarchus labrax*) in response to fluctuations in food availability. *Biochem Biophys Res Commun* 363:591–596
- Teschke M, Saborowski R (2005) Cysteine proteinases substitute for serine proteinases in the midgut glands of *Crangon crangon* and *Crangon allmani* (Decapoda: Caridea). *J Exp Mar Biol Ecol* 315:213–229
- Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acid Res* 22:4673–4680
- Titani K, Torff HJ, Hormel S, Kumar S, Walsh KA, Rodl J, Neurath H, Zwilling R (1987) Amino acid sequence of a unique protease from the crayfish *Astacus fluviatilis*. *Biochemistry* 26:222–226
- Towatari T, Miyamura T, Kondo A, Kato I, Inoue M, Yano M, Kido H (1998) The structures of asparagine-linked oligosaccharides of rat liver cathepsin L reflect the substrate specificity of lysosomal  $\alpha$ -mannosidase. *Eur J Biochem* 256:163–169
- von Heijne G (1990) The signal peptide. *J Membr Biol* 115:195–201
- Yonezawa S, Takahashi T, Wang X, Wong R, Hartsuck J, Tang J (1988) Structures at the proteolytic processing region of cathepsin D. *J Biol Chem* 263:16604–16611
- Zar JH (1984) Biostatistical analysis. Prentice-Hall, Englewood Cliffs
- Zwilling R, Stöcker W (1997) The astacins: structure and function of a new protein family. Dr Kovac Verlag, Hamburg