# *Complementary Proteomic and Biochemical Analysis of Peptidases in Lobster Gastric Juice Uncovers the Functional Role of Individual Enzymes in Food Digestion*

**Betsaida Bibo-Verdugo, Anthony J. O'Donoghue, Liliana Rojo-Arreola, Charles S. Craik & Fernando García-Carreño**

**Marine Biotechnology**
An International Journal Focusing on Marine Genomics, Molecular Biology and Biotechnology

ISSN 1436-2228
Volume 18
Number 2

Mar Biotechnol (2016) 18:201-214
DOI 10.1007/s10126-015-9681-5


Springer

Mar Biotechnol (2016) 18:201–214
DOI 10.1007/s10126-015-9681-5

ORIGINAL ARTICLE

# Complementary Proteomic and Biochemical Analysis of Peptidases in Lobster Gastric Juice Uncovers the Functional Role of Individual Enzymes in Food Digestion

**Betsaida Bibo-Verdugo[1] · Anthony J. O'Donoghue[2] · Liliana Rojo-Arreola[1,3] · Charles S. Craik[2] · Fernando García-Carreño[1]**

Received: 12 June 2015 / Accepted: 10 November 2015 / Published online: 27 November 2015


**Abstract** Crustaceans are a diverse group, distributed in widely variable environmental conditions for which they show an equally extensive range of biochemical adaptations. Some digestive enzymes have been studied by purification/characterization approaches. However, global analysis is crucial to understand how digestive enzymes interplay. Here, we present the first proteomic analysis of the digestive fluid from a crustacean (*Homarus americanus*) and identify glycosidases and peptidases as the most abundant classes of hydrolytic enzymes. The digestion pathway of complex carbohydrates was predicted by comparing the lobster enzymes to similar enzymes from other crustaceans. A novel and unbiased substrate profiling approach was used to uncover the global proteolytic specificity of gastric juice and determine the contribution of cysteine and aspartic acid peptidases. These enzymes were separated by gel electrophoresis and their individual substrate specificities uncovered from the resulting gel bands. This new technique is called zymoMSP. Each cysteine peptidase cleaves a set of unique peptide bonds and the S2 pocket determines their substrate specificity. Finally, affinity chromatography was used to enrich for a digestive cathepsin D1 to compare its substrate specificity and cold-adapted enzymatic properties to mammalian enzymes. We conclude that the *H. americanus* digestive peptidases may have useful therapeutic applications, due to their cold-adaptation properties and ability to hydrolyze collagen.

**Keywords** *Homarus americanus* · Peptidase · Substrate specificity · Cold-adapted enzyme

Betsaida Bibo-Verdugo and Anthony J. O'Donoghue denotes equal contribution.

**Electronic supplementary material** The online version of this article (doi:10.1007/s10126-015-9681-5) contains supplementary material, which is available to authorized users.

✉ Fernando García-Carreño
fgarcia@cibnor.mx

[1] Centro de Investigaciones Biológicas del Noroeste (CIBNOR), Calle IPN 195, Col. Playa Palo de Santa Rita, La Paz, B.C.S. 23096, Mexico

[2] Department of Pharmaceutical Chemistry, University of California San Francisco, 600 16th Street, San Francisco, CA 94158, USA

[3] Center for Discovery and Innovation in Parasitic Diseases, University of California San Francisco, 1700 4th Street, San Francisco, CA 94158, USA

## Introduction

Crustaceans display a wide range of feeding habits that include filter feeders, scavengers, and active predators; they have evolved a variety of enzymes to digest the complex organic biomolecules in food. In decapod crustaceans, the digestive enzymes are secreted from the midgut gland (or hepatopancreas) and consist mainly of peptidases, glycosidases, lipases, and nucleases. Some of these digestive enzymes have been biochemically characterized (Le Chevalier and Van Wormhoudt 1998; Nilsen et al. 2010; Rivera-Perez et al. 2011; Bibo-Verdugo et al. 2015).

The digestive enzymes of the American lobster *Homarus americanus* have been studied since 1970; biochemical assays using synthetic substrates revealed the presence of peptidase, glycosidase, lipase, and nuclease activities in the gastric juice (Brockerhoff et al. 1970). Since then, proteins with amylase (Wojtowicz and Brockerhoff 1972), chitinase, and chitobiase specificity (Lynn 1990) have been isolated and partially characterized; however, the proteins responsible for these activities have not been determined. Two peptidases capable of

degrading casein and hemoglobin have also been isolated from *H. americanus* gastric juice (Laycock et al. 1989, 1991; Rojo et al. 2010a, b, 2013). N-Terminal sequencing of these proteins identified them as cathepsin L1, a cysteine peptidase (GenBank X63567), and cathepsin D1, an aspartic acid peptidase (GenBank EU687261), respectively. Cathepsin L1 was shown to have sensitivity to E-64, optimum activity at pH 5.0, and a preference for cleaving on the C-terminal side of lysine residues (Laycock et al. 1989). Cathepsin D1 was sensitive to pepstatin A, had optimum activity at pH 4.0, and hydrolyzed an internally quenched fluorescent substrate between a pair of phenylalanine resides (Rojo et al. 2013). When directly compared to bovine cathepsin D, the lobster enzyme showed higher catalytic efficiency below 25 °C and underwent faster loss of activity at 60 °C (Rojo et al. 2013). In addition, cathepsin L1 had lower thermal stability when compared to the cysteine peptidases papain but had constant catalytic efficiency ($k_{cat}/K_m$) between 10 and 60 °C (Laycock et al. 1989). These observations indicate that *H. americanus* peptidases exhibit cold-adapted properties relative to the enzymes isolated from mesophiles (Rojo et al. 2013).

Cold-adapted enzymes are commercially valuable for bioremediation and textile and food production and as additives in detergents (Biggar et al. 2012). In addition, some of these enzymes have unusual substrate specificity (Feller and Gerday 2003). *H. americanus* lives in marine environments where the temperature does not exceed 25 °C (Cobb and Phillips 1980); therefore, an in-depth characterization of gastric juice has the potential to uncover novel enzymes that are likely to be cold-adapted. Gene expression studies have already identified two additional cysteine peptidases (cathepsin L2 and L3) (Laycock et al. 1991) and one additional aspartic peptidase (cathepsin D2) (Rojo et al. 2010a) in the digestive gland of *H. americanus*; however, no biochemical characterization of these enzymes has been performed. In this study, we performed an in-depth proteomic analysis of gastric juice from *H. americanus*. We identified 45 proteins, most of which had putative hydrolytic functions based on similarity to other arthropod enzymes. Some of these proteins could be directly linked to enzymatic activities that have been detected previously. In order to uncover the global proteolytic specificity in gastric juice, we used a multiplex substrate profiling in combination with mass spectrometry (MSP-MS) to determine the contribution of cysteine and aspartic acid peptidases to the overall specificity. In addition, we introduce a novel technique called zymoMSP where gastric juice proteins were separated by semi-denaturant polyacrylamide gel electrophoresis and protein bands were excised for identification and substrate specificity profiling. We found that each cysteine peptidase cleaves a set of unique peptide bonds and that substrate specificity is driven by residues in the S2 pocket. Finally, we used affinity purification to enrich for *H. americanus* cathepsin D1 and thoroughly characterized its specificity and enzymatic properties.

## Materials and Methods

### Enzymes

For our studies, six live adult American lobster *H. americanus* specimens (rostrum to telson length 26.3 ± 4.1 cm) were purchased on August 2013 at a local market in San Francisco, CA. The gastric juice was immediately sampled from lobsters by introducing a probe through the esophagus to reach the gastric chamber. The protein concentration in the juice was 6.7–9.5 mg/mL with a total volume of 2–5 mL (physiological pH ranging from 4.7 to 4.9). Samples were pooled, transferred to a 1.5-mL tube, and centrifuged for 15 min at 10,000×*g* and 4 °C, to remove solids. The supernatant was collected and stored into different tubes at −20 °C until use. Cathepsin D1 was enriched from the gastric juice using pepstatin A-agarose as described by Rojo and collaborators (Rojo et al. 2010b); this method results in cathepsin D1 purified to homogeneity appearing as a single band on a Coomassie-stained SDS electrophoresis gel, with a purification fold of 123 and a yield of 9.3 %. Cathepsin D from bovine spleen was purchased from Calbiochem (219393), and pepsin from porcine gastric mucosa was purchased from Sigma (P-6887).

### Proteomic Analysis of *H. americanus* Gastric Juice

Protein identification in gastric juice was done using peptide sequencing by mass spectrometry. Eight micrograms of gastric protein derived from three lobsters was digested with trypsin following a previously published method (O'Donoghue et al. 2013). Peptides were extracted, desalted using Rainin C18 tips and lyophilized. Mass spectrometry was done with a LTQ Orbitrap XL instrument, equipped with 10,000 psi system nanoACUITY (Waters) ultra-performance liquid chromatography (UPLC) instruments as outlined previously (O'Donoghue et al. 2012). Briefly, reversed phase liquid chromatography was performed using an EZ-Spray C18 column (Thermo, ES800, PepMap, 3 μm bead size, 75 μm × 15 cm). The LC was operated at 600 nL/min flow rate for loading and 300 nL/min for peptide separation over a 65-min linear gradient from 2 to 30 % acetonitrile in 0.1 % formic acid. Survey scans were recorded over 350–1800 *m/z* range, and MS/MS was performed with CID fragmentation on the six most intense precursor ions. Mass spectrometry peak lists were generated using in-house software called PAVA. Expressed sequence tags (ESTs; 29,957) from *H. americanus* were downloaded from GenBank and uploaded to Protein Prospector software v. 5.10.0. Peak list data was searched against the six-frame translation of the ESTs. Peptide

sequences were matched as tryptic peptides with up to two missed cleavages and carbamidomethylated cysteines as a fixed modification. Variable modifications included oxidation of methionine, N-terminal pyroglutamate from glutamine, start methionine processing, and protein N-terminal acetylation. Mass accuracy tolerance was set to 20 ppm for parent and 0.8-Da fragment errors for CID data. At least two non-overlapping peptides had to be observed to positively identify a protein sequence. Protein function was assigned to each EST-derived sequence following a BLASTp homology search.

### Multiplex Substrate Profiling by Mass Spectrometry

Gastric juice, consisting of 22.5 ng/mL of protein, was assayed with an equal molar mixture of 124 tetradecapeptides (500 nM each) in a total reaction volume of 150 μL (O'Donoghue et al. 2012). Each assay contained either 200 μM pepstatin A, 2 μM E-64, or 4 % DMSO. Lobster cathepsin D1 (20 ng/mL) was assayed with the peptide library in the presence of 2 μM E-64 to inhibit contaminating cysteine peptidases. Bovine cathepsin D and porcine pepsin were assayed at 15 and 2 ng/mL, respectively. All enzymes were incubated for 15, 60, 240, and 1200 min and reactions stopped by acidification with concentrated formic acid to a final pH of 2.5. Samples were desalted using C18 tips and analyzed by LC-MS/MS peptide sequencing. The time point at which peptidases showed a similar number of cleavage sites were selected for comparative analysis. This corresponded to 1200 min for lobster cathepsin D1, 60 min for porcine pepsin, and 240 min for bovine cathepsin D.

Mass spectrometry was performed with either the LTQ Orbitrap XL or an LTQ FT instrument, as outlined previously. Data was searched against a database containing the sequences of the 124 14-mer synthetic peptides. For database searching, peptide sequences were matched with no enzyme specificity requirement and variable modifications including oxidation of Trp, Pro and Phe, and N-terminal pyroGlu from Gln. Protein Prospector score thresholds were selected to be minimum protein score of 15, minimum peptide score of 15, and maximum expectation values of 0.1 for "protein" and 0.05 for peptide matches. Cleavage site data was extracted from Protein Prospector using an in-house script called "MSP extractor" software. For comparison of substrate specificity, an iceLogo software was used to generate substrate specificity logos for amino acids at ±4 positions adjacent to the identified cleavage sites.

### Electrophoresis, In-Gel Zymography, and Gel Band Extraction

*H. americanus* gastric juice, consisting of 63 μg of total protein, was incubated with 2 mM DTT for 10 min at room temperature and mixed with Tris-Glycine Native Sample Buffer (Life Technologies). The proteins were separated using a 12 % Tris-Glycine Gel and Tris-Glycine SDS running buffer (Life Technologies). For protein staining, the gel was soaked for 1 h in 30 mL of SimplyBlue™ SafeStain (Thermo Fisher Scientific), and then the staining solution was discarded and replaced with distilled water. For fluorescent in-gel activity assays, the gel was washed in 2.5 % Triton X-100 at room temperature for 30 min, followed by incubation in the dark for 20 min with 30 μM Z-FR-AMC in buffer containing 100 mM sodium acetate pH 4.7, 2 mM EDTA, 2 mM DTT, and 0.05 % Triton X-100. In-gel hemoglobinase assays were performed as described previously (Rojo et al. 2010a). Gels were visualized in a Gel Doc EZ Imager (BIORAD) using the software Image Lab 4.1. Gel bands were excised and soaked in 200 μL of 100 mM sodium acetate at pH 4.7 for 1 h at 4 °C to allow protein diffusion from the gel. The recovered proteins were diluted (1:5) and (1:50) into assay buffer containing an equal molar mixture of 124 tetradecapeptides (500 nM each) in a total reaction volume of 150 μL. Following 60 min of incubation at room temperature, the assay containing a (1:5) dilution of *H. americanus* was subjected to mass spectrometry analysis and data was searched against both the *H. americanus* Uniprot database containing 172 protein entries or the predicted protein sequences derived from ESTs. In this case, peptides were matched with non-specific cleavage sites at each terminus. For the 1 in 50 dilution, the data was searched against a database containing the 124 14-mer synthetic peptides, as outlined previously.

### Thermodynamic Activation Parameters

Aspartic peptidase activity was measured using the fluorogenic substrate 7-methoxycoumarin-4-acetic acid-GKPILFFRLK(DNP)-D-R-amide (Yasuda et al. 1999). Assays contained 1–200 μM of substrate, 50 mM sodium acetate pH 4.0, and 0.01 % Triton X-100. Five microliters of each aspartic peptidase (0.3 to 0.8 nM final) was added to 25 μL of pre-heated substrate. All tubes were incubated for 10, 20, or 30 min in a ThermoStat plus (Eppendorf) set at 5 to 55 °C. Each reaction was terminated by the addition of 170 μL of 5 % TCA and centrifuged for 5 min at 10, 000*g*. One hundred microliters of supernatant was added to each well in a black microplate and fluorescence emission was measured at 393 nm following excitation at 328 nm (Synergy 4, Biotek). Control reactions contained no enzyme and fluorescent intensity was converted to molar amounts of 7-methoxycoumarin-4-acetic acid released by comparison to a standard curve. Peptidase activity at each temperature was defined as the change in nanomolar concentration of product per minute.

The kinetic parameters $k_{cat}$ and $K_m$ were calculated at temperature intervals ranging from 5 to 55 °C. The reaction rate

was measured in triplicate varying the substrate concentration in the range of 1–200 μM. The average of reaction rate was plotted against substrate concentration; $K_m$ and $V_{max}$ were calculated by non-linear regression analysis using the Prism 5 software (Graph Pad Software Inc.). $k_{cat}$ was calculated by dividing $V_{max}$ by the enzyme molar concentration estimated by aspartic peptidase active site titration using pepstatin A as described by Knight (1995). The Arrhenius plot was built using $1/T$ as the abscissa and $\ln k_{cat}$ in the ordinate, $E_a$ was calculated by linear regression using Eq. 1.

$$\ln k_{cat} = \frac{E_a}{RT} + \ln A \quad (1)$$

The thermodynamic parameters of activation were calculated as described by Lonhienne (Lonhienne et al. 2000), using the Eqs. 2–4,

$$\Delta G^{\#} = RT\left(\ln\frac{k_B T}{h} - \ln k_{cat}\right) \quad (2)$$

$$\Delta H^{\#} = E_a - RT \quad (3)$$

$$T\Delta S^{\#} = \Delta H^{\#} - \Delta G^{\#} \quad (4)$$

where $E_a$ is the activation energy of the reaction, $T$ is the absolute temperature, $R$ is the gas constant (8.314 J $mol^{-1}$ $K^{-1}$), A is a constant related to collision frequency, $k_B$ is the Boltzmann constant ($1.3805 \times 10^{-23}$ J $K^{-1}$), $h$ is the Plank constant ($6.6256 \times 10^{-34}$ J s), and $\Delta G^{\#}$, $\Delta H^{\#}$, and $\Delta S^{\#}$ are the free energy, enthalpy, and entropy of activation, respectively.

## Results

### Proteomic Analysis of Gastric Juice

In 1970, Brockerhoff and colleagues performed a set of simple enzyme assays on gastric juice from *H. americanus* and detected peptidase, lipase, nuclease, and glycosidase activity but no identification of the enzymes responsible for these activities was performed (Brockerhoff et al. 1970). In order to identify the enzyme composition of *H. americanus* gastric juice, proteins were digested with bovine trypsin and peptides sequenced using liquid chromatography-tandem mass spectrometry (LC-MS/MS). The genome of *H. americanus* has not been sequenced and only 192 proteins have been deposited in the Uniprot database (May 2015). Therefore, 29,957 available ESTs that were generated by Towle and Smith were translated in silico and used as the *H. americanus* protein database (Towle and Smith 2006). Using a minimum of two non-overlapping peptides to define a protein, a total of 45 unique protein sequences were identified in gastric juice and abundance-ranked based on their intensity sums (Table 1). Of these, 24 proteins were predicted to have enzymatic activity, and all but a thiol reductase had putative hydrolytic functions. These included 11 glycosidases, 6 peptidases, 2 nucleases, 2 amidohydrolases, 1 lipase, and 1 sulfatase (Fig. 1a). Based on sequence comparison with other arthropods, we found that the glycosidases included multiple endo- and exo-acting enzymes, which allowed us to generate a map of how simple sugars can be released from complex plant and animal carbohydrates (Fig. 1b). The endo-glycosidases such as chitinase, cellulase, and α-amylase can digest chitin, cellulose, and starch, respectively. The products of these enzymatic reactions can then be further degraded into simple sugars such as glucose and *N*-acetylglucosamine (GlcNAc) by exo-glycosidases. In addition, putative β-galactosidase and GalNAcase were found that can release simple sugars from glycoproteins. No protein corresponding to a putative β-glucosidase was found in this study; however, β-glucosidase activity was detected in *H. americanus* gastric juice by Brockerhoff and colleagues (Brockerhoff et al. 1970). This enzyme can release glucose by cellobiase during the breakdown of cellulose.

Using peptide sequence coverage as a proxy for estimating relative abundance, our proteomic studies identified four homologous chitinases that are highly abundant proteins in gastric juice. These proteins are likely to be responsible for the individual chitinase activities that were detected and partially purified from *H. americanus* gastric juice (Lynn 1990). In addition, we identified a putative chitin deacetylase in the gastric juice that likely converts chitin to the more soluble chitosan, which can then be degraded by enzymes produced in plants and microorganisms (Zhu et al. 2007). *H. americanus* nucleases and lipases in the gastric juice each have high sequence similarity to proteins from other decapods; however, only the duplex-specific nuclease (GenBank: EX471393.1) has been characterized. Anismova and co-workers have predicted that unlike the "DNA/RNA non-specific β-β-α-metal finger" nucleases from bacteria, the related enzymes from crabs have a much greater preference for double-stranded DNA over single-stranded DNA and RNA (Anisimova et al. 2008). This predicted specificity is consistent with the abundant double-stranded DNase activity detected in *H. americanus* gastric juice (Brockerhoff et al. 1970).

For protein digestion, *H. americanus* gastric juice contained numerous papain family cysteine peptidases being the previously identified digestive enzyme cathepsin L1 the most abundant protease based on peptide sequence coverage. In addition, proteins corresponding to cathepsin L2 and L3 were also detected. Gene expression studies of the aspartic acid peptidases have been performed previously by our group. We determined that cathepsin D1 expression occurs in the midgut tissue but not in muscle or gonads, while cathepsin D2 appears to be stably expressed at low levels in all three.

**Table 1** List of identified proteins in the gastric juice of *Homarus americanus*

| Rank | GenBank ID | Top BlastP protein hit | Accession number for top BlastP hit (residues) |
|---|---|---|---|
| 1 | EY116950.1 | Uncharacterized protein | None |
| 2 | CN949927.1 | Digestive cysteine proteinase 1 (hacl1) | P13277.2 (141-322) |
| 2.1 | CN852959.1 | Digestive cysteine proteinase 1 (hacl1) | P13277.2 (1-243) |
| 3 | FE535164.1 | Chitinase 1 | BAP28983.1 (20-190) |
| 4 | EH401791.1 | Uncharacterized protein | None |
| 5 | EH401809.1 | Chitinase 2 | BAP28983.1 (18-235) |
| 6 | EH401706.1 | Chitinase 3 | BAP28983.1 (22-228) |
| 7 | EH116448.1 | Uncharacterized protein | None |
| 8 | EY291189.1 | Chitin deacetylase | ACB54936.1 (34-220) |
| 9 | EX471196.1 | Chitinase 4 | BAP28983.1 (14-233) |
| 10 | EX827493.1 | Beta-crystallin 1 | ADD38111.1 |
| 11 | FE535048.1 | Beta-galactosidase | AIA09347.1 (37-230) |
| 12 | EX471148.1 | Vitelline membrane outer layer 1-like protein | ACK44247.1 (21-152) |
| 13 | FE840846.1 | Cathepsin D-like protein (hacd1) | ACG70181.1 (170-386) |
| 13.1 | GO271561.1 | Cathepsin D-like protein (hacd1) | ACG70181.1 (67-287) |
| 14 | FF277653.1 | Variable lymphocyte receptor A diversity region | ABO21241.1 |
| 15 | EH035092.1 | Alpha-glucosidase | CAB85963.1 (2-221) |
| 15.1 | EG948746.1 | Alpha-glucosidase | CAB85963.1 (276-494) |
| 16 | EW703171.1 | Hypothetical protein X975_18232 | KFM67210.1 (13-193) |
| 17 | FC556045.1 | Digestive cysteine proteinase 2 (hacl2) | P25782.1 (175-323) |
| 17.1 | CN952065.1 | Digestive cysteine proteinase 2 (hacl2) | P25782.1 (1-223) |
| 18 | EW703171.1 | Uncharacterized protein | None |
| 19 | EG948699.1 | Luciferin 2-monooxygenase non-catalytic subunit | Q9GV46.1 (37-192) |
| 20 | FC071535.1 | Uncharacterized protein | None |
| 21 | EX471393.1 | Duplex-specific nuclease | AAN86143.1 (27-220) |
| 22 | EH034894.1 | Lectin A | ADG85663.1 (3-132) |
| 23 | EX568366.1 | Thiol reductase | AEX07320.1 (1-198) |
| 24 | FD483234.1 | Luciferin 2-monooxygenase non-catalytic subunit | Q9GV46.1 (89-254) |
| 24.1 | FE841074.1 | Luciferin 2-monooxygenase non-catalytic subunit | Q9GV46.1 (41-210) |
| 25 | EW702686.1 | Alpha-N-acetylgalactosaminidase (galnacase) | CAX20733.1 (3-231) |
| 25.1 | CN854381.1 | Alpha-N-acetylgalactosaminidase (galnacase) | CAX20733.1 (129-323) |
| 26 | GO271459.1 | Digestive cysteine proteinase 3 (hacl3) | P25784.1 (170-321) |
| 26.1 | EV781666.1 | Digestive cysteine proteinase 3 (hacl3) | P25784.1 (1-216) |
| 27 | EY290959.1 | Beta-crystallin 2 | ADD38111.1 (18-206) |
| 28 | EH117154.1 | Alpha-amylase | AIJ02078.1 (20-232) |
| 29 | EV782128.1 | Cathepsin C | ADO65979.1 (1-215) |
| 30 | EY290647.1 | Uncharacterized protein | None |
| 31 | EY117386.1 | Cellulase | AAD38027.1 (198-346) |
| 31.1 | EW702563.1 | Cellulase | AAD38027.1 (1-202) |
| 32 | FD483385.1 | Chitinase 5 | BAP28983.1 (18-165) |
| 33 | CN853621.1 | Uncharacterized protein | None |
| 34 | CN854293.1 | Uncharacterized protein | None |
| 35 | FC071608.1 | Arylsulfatase B | KFM79564.1 (8-229) |
| 36 | EX487799.1 | Triacylglycerol lipase | EFX74783.1 (72-245) |
| 36.1 | FE535664.1 | Triacylglycerol lipase | EFX74783.1 (172-384) |
| 37 | FD699760.1 | Uncharacterized protein | None |
| 38 | FD584912.1 | Endoribonuclease | BAM19803.1 (11-182) |
| 39 | EX568277.1 | Uncharacterized protein | None |

**Table 1** (continued)

| Rank | GenBank ID | Top BlastP protein hit | Accession number for top BlastP hit (residues) |
|---|---|---|---|
| 40 | EX487603.1 | S28 serine peptidase | XP_004929850.1 (46-184) |
| 41 | EH116458.1 | Uncharacterized protein | None |
| 42 | EH034776.1 | Beta-N-acetylglucosaminidase (glcnacase) | ACR23316.1 (261-490) |
| 43 | CN852671.1 | Uncharacterized protein | None |
| 44 | FE043895.1 | Neutral ceramidase | XP_008198655.1 (409-614) |
| 45 | EG948890.1 | Lectin D | ADG85659.1 (31-168) |

In this proteomic study, we identified nine peptides from cathepsin D1 but no peptides from cathepsin D2 which supports our previous hypothesis that cathepsin D2 functions in the endo-lysosomal pathway and not in gastric digestion of food (Rojo et al. 2010a). Previous RT-qPCR studies have identified cathepsin C1, C2, and C3 expression in the hepatopancreas of starved lobsters (Clark et al. 2013); however, in this proteomic study only cathepsin C2 was detected in the gastric juice. Three peptides corresponding to a S28 family serine peptidase were identified in the gastric juice (GenBank: EX487603.1). However, further analysis determined that this enzyme was homologous to peptidases encoded by marine ectoparasites, such as *Caligus* and *Lepeophtheirus* species, and therefore is not a lobster enzyme (Burridge et al. 1999).

## Global Analysis of Peptidase Activity in the Gastric Juice

An unbiased substrate profiling assay was used to uncover the global proteolytic substrate specificity or "substrate signature" in gastric juice. This assay consists of a mixture of 124 physiochemically diverse peptides that are each 14 residues in length. Cleavage at any one of the 1612 peptide bonds within these peptides can be readily detected by LC-MS/MS sequencing (Fig. 2a). The complete list of identified cleavage sites is listed in Supplemental file 1. Co-incubation with the gastric juice resulted in 42, 140, 264, and 315 cleavage sites detected after 15, 60, 240, and 1200 min, respectively. The complexity of these hydrolytic events is illustrated in three example peptides where multiple cleavage sites were often detected within each peptide (Fig. 2b). Using iceLogo software (Gillmor et al. 1997), the P4 to P4′ substrate signature was generated using all cleavage sites detected after 60 min incubation. Lobster gastric juice peptidases exhibited a preference for Gly and Ala at P4, Phe, Ser, and Thr at P1 and Ile, norleucine (Nle), and Arg at P1′ (Fig. 2c). In addition, the detected peptidases showed a preference for Trp and Ile at P4′ and a diverse variety of amino acids at P3′. Finally, low tolerance for Pro at P1 and P1′ was observed. Proteomic analysis determined that aspartic acid and cysteine peptidases were the major peptidases in the gastric juice of *H. americanus*. This led us to perform an analysis of proteolytic activity in the presence of the class-specific peptidase inhibitors pepstatin A and E-64 that target aspartic acid and cysteine peptidases, respectively. Cleavage sites that were identified after 60 min in the control assay (DMSO treatment) but appeared at a later time point or never in the inhibitor-treated assay were designated as "sensitive" to these antagonists. From these studies, it was apparent that most of the cleavage sites were sensitive to E-64 treatment (Fig. 2d). In addition, 8 of the 140 cleavage sites were sensitive to pepstatin-A treatment while only 5 cleavage sites remained unchanged after treatment with either inhibitor. Interestingly, 35 of the cleavage sites were sensitive to both pepstatin-A and E-64 indicating an overlap in specificity between to two classes of enzymes. It is likely that in the control assay, aspartic acid and cysteine peptidases are able to cleave many of the same bonds. In the presence of pepstatin A, the overall rate of cleavage slows down but the bonds can still be hydrolyzed by

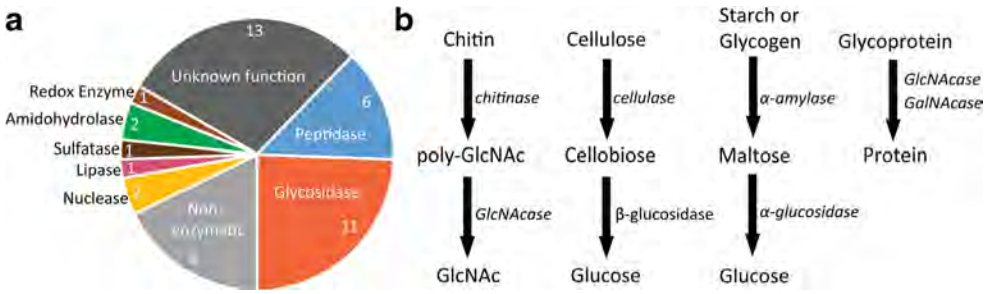



**Fig. 1** Glycosidases and peptidases were the most abundant enzymes in the gastric juice of *H. americanus*. **a** A pie chart showing the number of proteins identified by MS and classified by function. **b** A possible general outline of carbohydrate digestion in the gastric juice of *Homarus americanus*; glycosidases identified in the gastric juice by mass spectrometry are indicated in *italics*

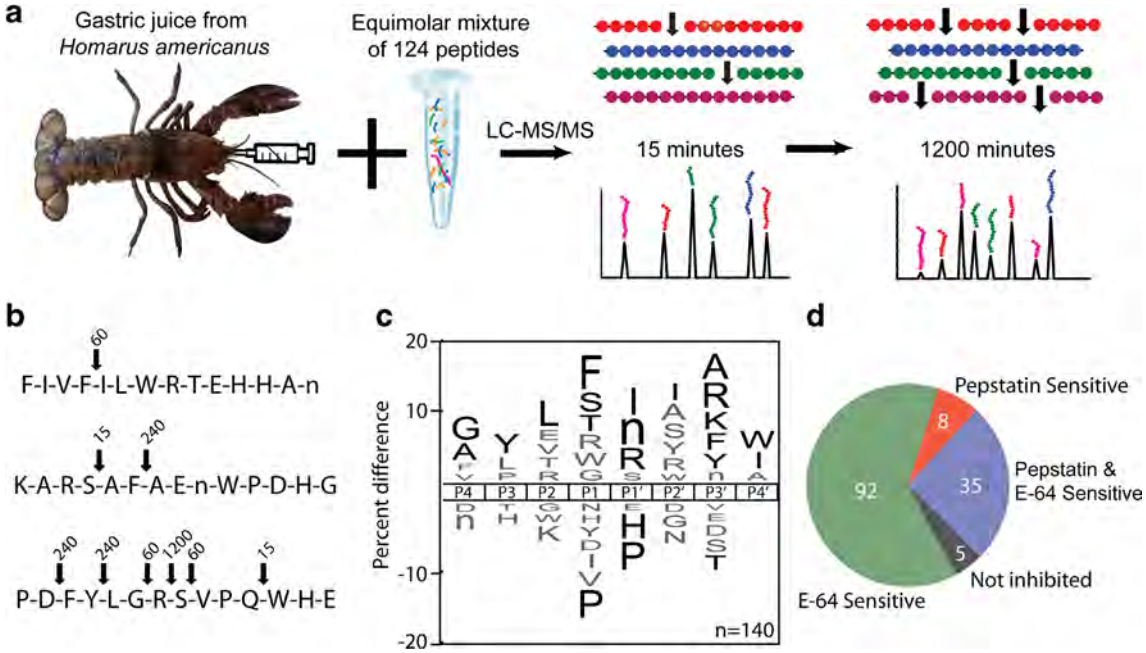



**Fig. 2** Global specificity of gastric juice peptidases from *Homarus americanus* using the MSP-MS assay. **a** An illustration showing the retrieval of gastric juice from *H. americanus* and incubating with a mixture of 124 peptides. At defined time intervals, an aliquot of the assay is removed and the degradation of peptides determined by LC-MS/MS peptide sequencing. **b** An example of three peptides cleaved by peptidases in gastric juice. The time at which cleavage was first detected is indicated in minutes by an arrow. **c** IceLogo generated from the pattern of cleavage events after 1200 min incubation shows the specificity of peptidase activity. Amino acids that are most frequently observed at each position are shown above the axis, and amino acids least frequently observed are shown below the axis. **d** A pie chart showing the number of cleavage sites sensitive to either E-64, pepstatin A or both

the cysteine peptidases. The reverse is true, in that E-64 treatment inhibits the cysteine peptidases but does not affect the action of the aspartic peptidases.

### Identification and Substrate Specificity of Individual Peptidases in the Gastric Juice

Due to complex digestion pattern associated with multiple peptidases acting in unison, we sought to separate the individual *H. americanus* peptidases by gel-based chromatography. Using hemoglobin and the fluorescent substrate, Z-Phe-Arg-AMC, six distinct proteolytically active regions were identified on the zymogram (Fig. 3a). These regions were excised from an unstained gel and proteins were eluted from the gel pieces into assay buffer. An additional band was excised from the gel that corresponded to a region that had no detectable protein or proteolytic activity (gel band #6). This gel band served as a negative control. The eluted proteins from all seven gel pieces were diluted 5-fold and 50-fold into assay buffer containing the 124-member tetradecapeptide library and incubated at room temperature for 15, 60, and 240 min. We have subsequently defined this assay as zymography Multiplex Substrate Profiling (zymoMSP) and lobster gastric juice is the first complex protease mixture to use this technique.

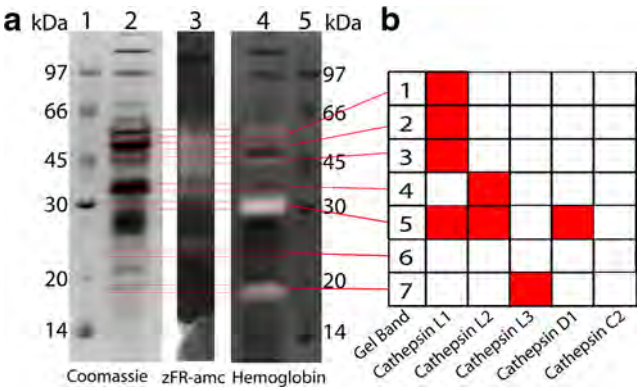



**Fig. 3** Separation of proteolytic active bands from SDS-PAGE gels. **a** Gels were washed with distilled water and treated either with Coomassie stain a fluorescent substrate or hemoglobin as substrate (*lanes 2*, *3*, and *4*). Molecular weight markers are shown in *lanes 1* and *5*. **b** A chart showing the peptidases identified by MS in each band excised from the gel

In the 5-fold diluted assay, excessive peptide cleavage occurred by 15 min and was therefore not informative of peptidase specificity. However, in this reaction, we were able to detect *H. americanus* peptides that were the result of peptidase auto-degradation. Gel bands 1, 2, and 3 contained peptides derived from cathepsin L1 only while gel bands 4 and 7 contained cathepsins L2 and L3, respectively. As expected, no protein was detected in gel band 6. Cathepsin D1 was detected in gel band 5; however, cathepsin L1 and L2 peptides

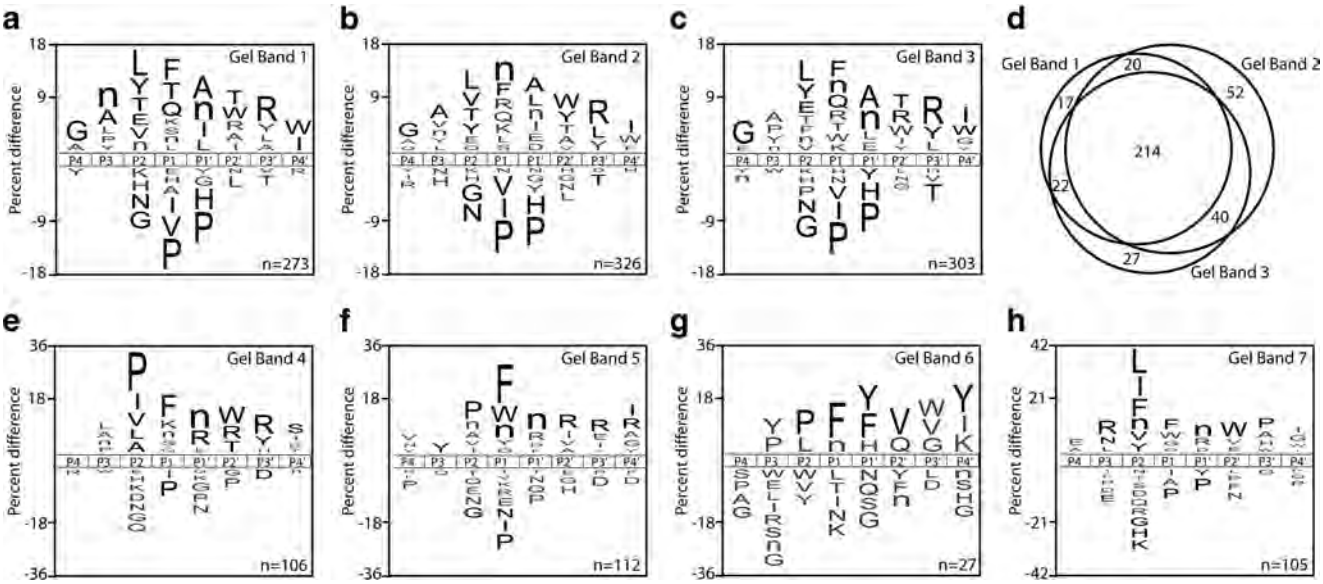



**Fig. 4** IceLogos generated for each gel band show distinct substrate signatures. **a**–**c** Proteolytic signature of three gel bands that contain cathepsin L1. **d** Venn diagram showing the similarity of cleavage sites detected in gel bands 1 to 3. **e**–**h** Proteolytic signature of gel bands 4 to 7 that consist of cathepsins L2 (gel band 4), D1 (gel band 5), and L3 (gel band 7). Gel band 6 is used as a control as there is no detectable cleavage of hemoglobin or fluorescent substrate. Proteomic analysis also detected cathepsins L1, L2, and L3 in gel band 5

were also present (Fig. 3b). No peptides corresponding to cathepsin C2 were detected in any gel band.

In the 50-fold diluted gel band assay, gel bands 1, 2, and 3 (containing cathepsin L1) had similar specificity profiles with a preference for Gly at P4 and Arg at P3′ (Fig. 4a–c). In addition, Leu, Tyr, Thr, and Glu were generally found in the P1 position and Ala and Nle at P1′. Cathepsin L1 has low P1 preference for Ile, Val, and Pro which is consistent with previous studies using *p*-nitrophenyl esters of benzyloxycarbonyl amino acids (Laycock et al. 1989). It is unclear why cathepsin L1 separates into three distinct gel bands; however 214 cleavage sites are shared between the three isoforms (Fig. 4d). Substrate specificity in gel band 4 (cathepsin L2) and gel band 7 (cathepsin L3) is dominated by the presence of hydrophobic residues in the P2 position (Fig. 4e, h). However, cathepsin L2 prefers Pro over other hydrophobic residues while cathepsin L3 has no preference for this amino acid. Based on the proteomic data, gel band 5 contains not only cathepsin D1 but also trace amounts of cysteine cathepsins L1 and L2. Hydrophobic residues are commonly found in the P1 and P1′ position of aspartic peptidases (O'Donoghue et al. 2012; Sojka et al. 2012). Gel band 5 has P1 specificity for Phe, Trp and Nle and P1′ specificity for Nle (Fig. 4f) However, Pro in the P2 position is unusual for this class of peptidases and is likely due to cathepsin L2 activity. Finally, gel band 6 (negative control), cleaved at only 27 sites after 4 h incubation (Fig. 4g), which is likely due to trace amounts of cathepsins L2 and D1 that were undetectable in the proteomic analysis (Fig. 3b).

## Sequence and Specificity Comparison of *H. americanus* Cathepsin L2 and L3 with Other Cysteine Cathepsins

The substrate specificity profile of *H. americanus* cathepsin L3 (gel band 7) was compared to a cathepsin L from the protozoa *Trypanosoma cruz*i, the causative agent of Chagas disease. This peptidase is commonly known as cruzain, and like cathepsin L3 from lobster, it has a distinct preference for hydrophobic residues, other than Pro, in the P2 position (Fig. 5a). Several crystal structures of cruzain have been solved with peptide inhibitors bound in the substrate-binding pocket. These studies have determined that Leu-67, Ala-133, and Leu-157 form the S2 pocket of this enzyme (McGrath et al. 1995). In addition, it is known that Glu-205 rotates out of the S2 pocket upon binding of large amino acids such as phenylalanine (Gillmor et al. 1997). *H. americanus* cathepsin L3 possesses a tryptophan at position 67 (papain numbering) and identical residues at positions 133, 157, and 205 (Fig. 5e). Interestingly, cathepsin L2, which prefers Pro in the S2 pocket, differs from cruzain at positions 67, 133, and 205. The most striking difference is the substitution of the negatively charged Glu-205 to Val. It is possible that such residue does not rotate out of the hydrophobic pocket as readily as Glu making it smaller and less likely to accept bulky P2 residues such as Phe. These conditions may favor the binding of a Pro in the pocket. In collaboration with Jose F. Tort laboratory in Uruguay, our group has profiled the substrate specificity of cathepsin L3 from the sheep liver fluke, *Fasciola hepatica* (Corvo et al. 2013). In this enzyme, Val also replaces Glu in the S1 pocket and we observed a similar P2 substrate specificity preference for Pro and not Phe (Fig. 5d).

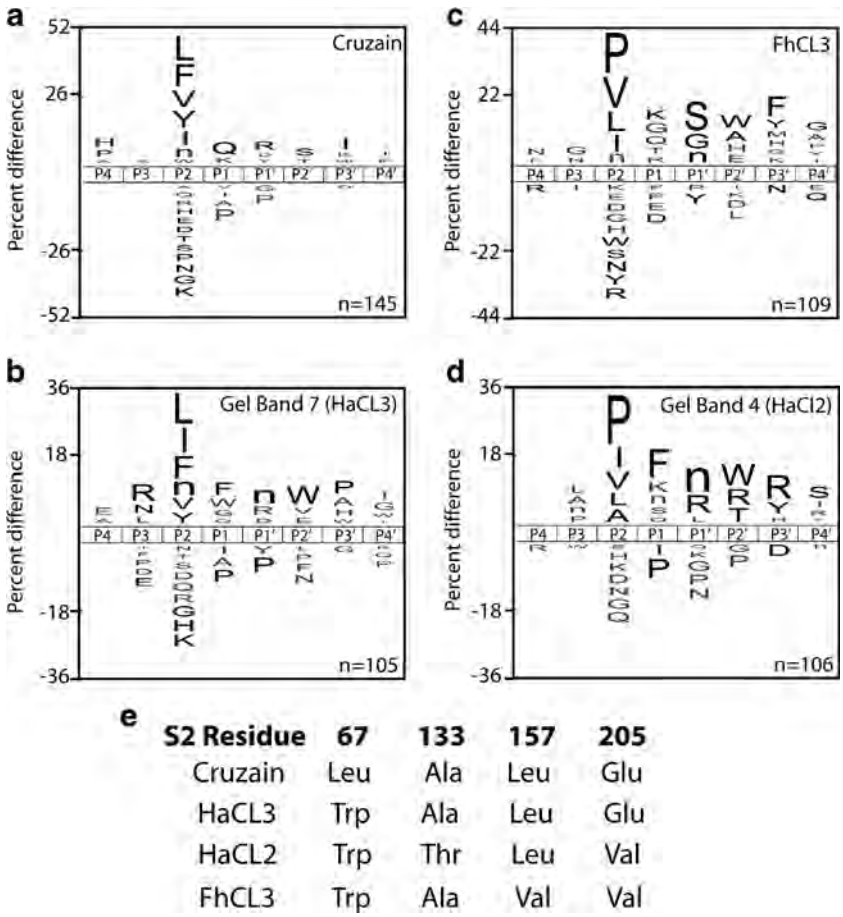



**Fig. 5** Identification of amino acids in the S2 substite of cathepsin L-like peptidases that determine the substrate specificity. The specificity profile of **a** cruzain from *T. cruzi* was directly compared to **b** HaCL3 from *H. americanus* while the specificity of **c** *F. hepatica* cathepsin L3 was compared to **d** HaCL2 from *H. americanus*. **e** Based on structural studies, the amino acids in the S2 subsite of cathepsin L-peptidases that likely play a major role in substrate specificity are *highlighted*

### Substrate Specificity Comparison of *H. americanus* Cathepsin D1 with Mammalian Peptidases

Cathepsin D1 was enriched from *H. americanus* gastric juice using pepstatin A-agarose and used hereafter for specificity and activity analysis. Its substrate specificity was profiled using the MSP-MS assay and directly compared to bovine cathepsin D and porcine pepsin. The specificity of each of these enzymes is largely driven by a preference for hydrophobic residues, particularly Phe, in the S1 pocket (Fig. 6a–c). Bulky residues such as Trp and Tyr are also frequently found at this position in lobster cathepsin D1 assays while porcine pepsin prefers the flexible amino acids, Leu and Nle. Bovine cathepsin D has a more stringent preference for Phe in the P1 position when compared to the lobster and porcine enzymes. Interestingly, all of the studied enzymes have low tolerance for Val or Ile in the P1 position but these residues are often found in the P1′ position. Each peptidase cleaved at 10 to 13 unique sites while 15 cleavage sites were shared by all three enzymes (Fig. 6d). In previous studies, we have used prime and non-prime side specificity profiles to design internally quenched fluorescent substrates for proteolytic assays (O'Donoghue et al. 2015). However, in this study, we deduced that one of the cleavage sites shared by all three aspartic acid peptidases, VIF↓FRL was highly similar to a sequence in a commercially available fluorescent substrate, mca-GKPILFFRLK(dnp). Komei and colleagues have determined that cathepsin D from the hepatopancreas of *Todarodes pacificus* (Japanese common squid) cleaves this fluorescent substrate between the two Phe residues (Komai et al. 2004). We deduced that designing a new fluorescent substrate that would be robustly cleaved by all three aspartic peptidases was not required and therefore this general aspartic acid substrate was used for subsequent biochemical studies.

### Thermodynamic Activation of Cathepsin D1 Aspartic Acid Peptidase from Lobster Gastric Juice

In order to compare the thermodynamic activation of each enzyme, we utilized the general aspartic acid substrate (mca-GKPILFFRLK-dnp) to follow activity. Lobster cathepsin D1 exhibited higher $k_{cat}$ values than bovine and porcine peptidases across the temperature range tested. At low temperature (5–25 °C), lobster cathepsin D1 showed 2-fold and 20-fold higher $k_{cat}$ values than porcine pepsin and bovine cathepsin D, respectively. Activation energy ($E_a$) was calculated by an Arrhenius plot. The lowest values of $E_a$ were obtained for lobster cathepsin D1 (56.4 kJ/mol) followed by porcine

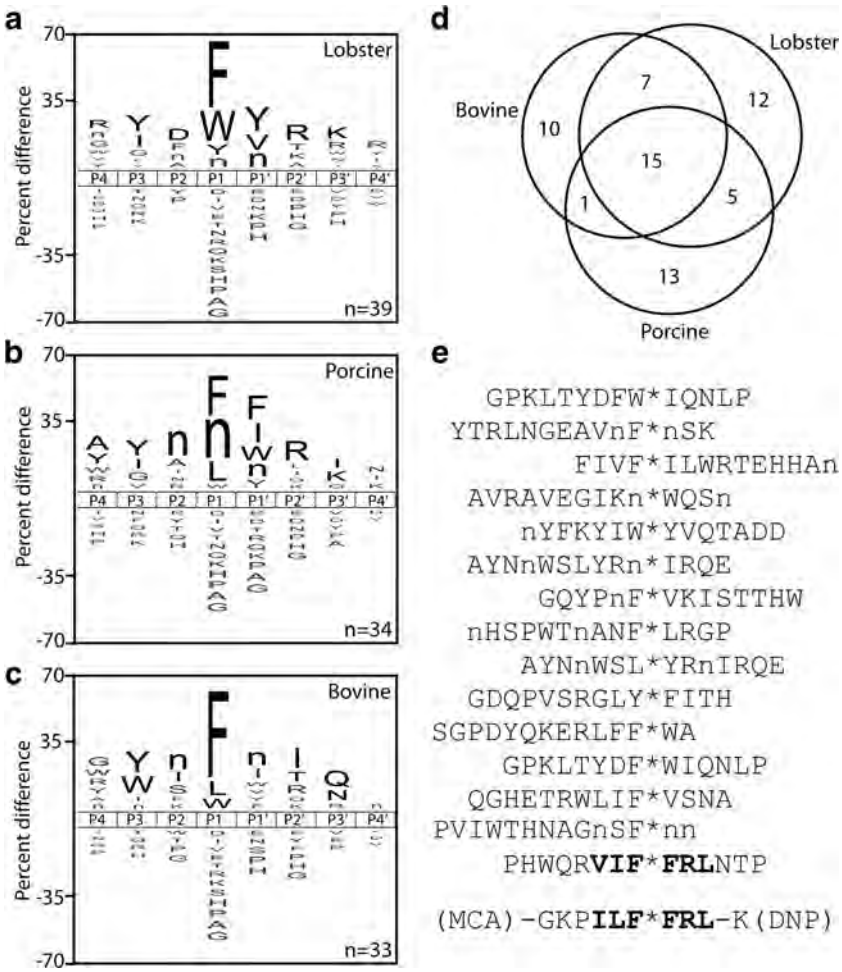


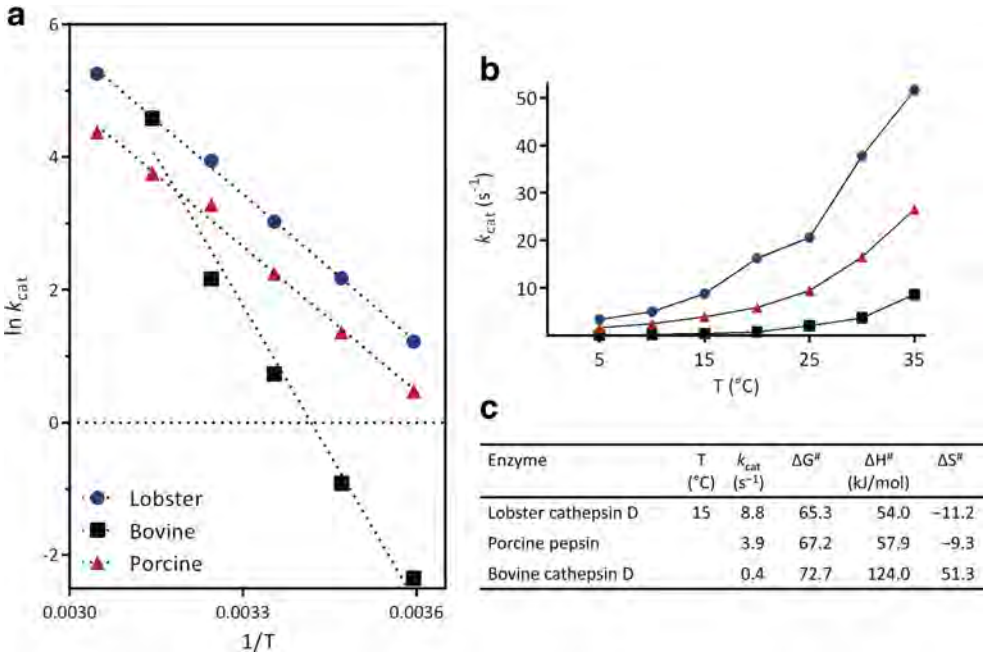

**Fig. 6** Peptide substrate profiles of aspartic peptidases. **a**–**c** IceLogos representing the most frequently (*upper panel*) and least frequently (*lower panel*) observed amino acids in the P4-P4′ sites for lobster cathepsin D1, porcine pepsin, and bovine cathepsin D, respectively. **d** Representation of the number of cleavage sites that are shared and unique for each aspartic peptidase. **e** A list of the cleavage sites shared by the three peptidases. One of these cleavage sites is similar to a commercial substrate for aspartic peptidases (*bold caption*)

pepsin (60.3 kJ/mol), whereas the higher value was obtained for bovine cathepsin D (126.4 kJ/mol) (Fig. 7a, b). Lobster cathepsin D1 also showed lower Gibbs free energy ($\Delta G^{\#}$), enthalpy ($\Delta H^{\#}$), and entropy ($T\Delta S^{\#}$) of activation than bovine and porcine enzymes (Fig. 7c).

## Discussion

The American lobster, *H. americanus*, feeds on a complex diet that includes both plants and animals and therefore is likely to have an equally complex set of digestive enzymes. In this study, we were interested in increasing our understanding of lobster digestion physiology and in identifying hydrolytic enzymes with novel properties that could be exploited for therapeutic and industrial applications. *H. americanus* is the largest known crustacean (by weight) and its gastric juice is easily retrievable and of sufficient quantity for in-depth biochemical studies. Characterization of hydrolytic enzymes in digestive fluids from crustaceans has traditionally utilized simple reporter substrates for detection of enzymatic activity and protein chromatography techniques for isolation. This approach has been largely successful for the biochemical characterization of individual peptidases, glycosidases, lipases, and nucleases, however does not fully explain the function that each enzyme plays within the complex digestion system. Gene expression studies using hepatopancreas tissue shed light on the mRNA levels of enzymes but in the absence of functional assays one can only hypothesize the role that they play.

In this report, we performed an in-depth proteomic study of *H. americanus* gastric juice proteins. In the absence of an available genome sequence, the data files were searched against the six-frame translation of ESTs from *H. americanus*. These ESTs were generated from several lobster tissues that included the hepatopancreas and therefore likely contain many of the mRNA sequences that encode for digestive proteins. Interestingly, many of the proteins identified in the gastric juice such as cathepsins, lectins, and a thiol reductase are typically described as lysosomal proteins (Arunachalam et al. 2000; Turk et al. 2000; Zaidi et al. 2008). Identification of these proteins supports the rapid cell differentiation and apocrine-holocrine secretion hypothesis in crustaceans where large digestive vacuoles containing lysosomal proteins extrude from hepatopancreatic cells (Hu and




| Enzyme | T (°C) | $k_{cat}$ ($s^{-1}$) | ΔG# | ΔH# (kJ/mol) | ΔS# |
|---|---|---|---|---|---|
| Lobster cathepsin D | 15 | 8.8 | 65.3 | 54.0 | −11.2 |
| Porcine pepsin | | 3.9 | 67.2 | 57.9 | −9.3 |
| Bovine cathepsin D | | 0.4 | 72.7 | 124.0 | 51.3 |

**Fig. 7** Thermodynamic activation parameters of lobster cathepsin D1, porcine pepsin, and bovine cathepsin D. **a** Arrhenius plots for temperature dependence of activity of aspartic acid peptidases; lobster cathepsin D1 is compared to porcine pepsin and bovine cathepsin D as reference. **b** The catalytic constant in the temperature range of 5–35 °C. **c** Thermodynamic parameters of activation of the assessed aspartic proteases

Leung 2007). Using both apocrine and holocrine mechanisms, these vacuoles release proteins into the lumen thereby extending enzymatic activity to the extracellular milieu. This hypothesis also explains the acidic nature of Crustacean digestive juice.

Proteomic studies that search against EST or cDNA databases has previously been performed for the analysis of proteins in the digestive fluid of the Venus fly (*Dionaea muscipula*) or the midgut of cattle ticks (*Rhipicephalus microplus*). Both of these species produce a diverse array of digestive enzymes (Kongsuwan et al. 2010; Schulze et al. 2012). Likewise, *H. americanus* secretes numerous digestive enzymes that facilitate the degradation of complex carbohydrates, proteins, glycoproteins, lipids, and nucleic acids. Activity from amylase, chitinase, α-glucosidase, β-galactosidase, GlcNAcase, lipase, nuclease, and peptidases have all been detected in gastric juice from *H. americanus* (Brockerhoff et al. 1970); however, the proteins responsible for these activities have not been elucidated. In this study, we show that glycosidases are the most abundant group of enzymes in gastric juice. These enzymes are capable of degrading starch, cellulose, and chitin, in agreement to *H. americanus* omnivorous lifestyle.

We identified six peptidases in *H. americanus* gastric juice of which four had been previously reported and partially characterized (Laycock et al. 1989, 1991; Rojo et al. 2010b). The role of these individual enzymes in overall protein digestion has not been elucidated, and therefore, we utilized a set of 124 diverse peptides as a synthetic protein source to uncover the substrates specificity. Digestion assays were performed at pH 4.7 to mimic the conditions in *H. americanus* gastric juice. Peptidases in the gastric juice readily degraded the synthetic peptides and this hydrolysis could be largely inhibited by treatment with a cysteine or aspartic acid peptidase antagonist. Previously, this approach was used to dissect the proteolytic activity secreted from a pathogenic fungus (O'Donoghue et al. 2015) and pancreatic ductal adenocarcinoma cells (O'Donoghue et al. 2012). However, *H. americanus* gastric juice is more complex than mammalian and fungal cell secretions and consists of four papain-like cysteine peptidases that are all likely to be sensitive to E-64. Therefore, to dissect the individual activities, proteins were separated on an SDS-PAGE gel and regions showing activity against either zFR-AMC or hemoglobin were excised. Proteomic analysis of the bands indicated that we were able to separate cathepsins L1, L2, L3, and D. It is unclear why cathepsin L1 migrates into three distinct bands but the substrate specificity profile of each isoform is very similar. Unlike cathepsin L1, cathepsins L2 and L3 have a distinct preference for hydrophobic residues in the P2 position. However, these enzymes differ in their ability to accept proline at this position. The amino acids contributing to the S2 pocket of each enzyme were determined based on alignment studies with cruzain, a well-characterized cathepsin L-like peptidase from *Trypanosoma cruzi*.

From an evolutionary standpoint, it is intriguing to observe the diversity in function that these homologous enzymes have achieved. Together, cathepsins L1, L2, and L3 were capable of hydrolyzing 434 peptide bonds after incubation for only 15 min, which amounts to 27 % of all peptide bonds in the library. After 4 h incubation with the peptide library, 42 % of all peptide bonds are cleaved (Supplemental file 1). Interestingly, the three least favored amino acids in the P1 position of cathepsin L1, namely, Ile, Val, and Pro are the most favored residues in the P2 position of cathepsin L2. Therefore,

protein regions that cannot be cleaved by cathepsin L1 will likely be easily hydrolyzed by cathepsin L2. To our knowledge, this synergy in substrate specificity between homologous peptidases has not been previously reported.

Two cysteine peptidases from *Fasciola hepatica* that also differ in their P2 proline selectivity have previously been characterized. Fasciola cathepsin L3 is capable of degrading the proline-rich triple helical regions of collagen while cathepsin L1 cannot (Corvo et al. 2013). We can therefore deduce that *H. americanus* cathepsin L2 is likely to be the enzyme responsible for the collagen-degrading activity in gastric juice (Chen 1991). In other decapods, such as the sand fiddler crab, *Uca pucilator*, the collagen-degrading enzymes have been extensively characterized (Page and Craik 2013). These enzymes are not cysteine peptidases but serine peptidases, called brachyurins, and are optimally active at neutral pH. Enzymes with collagenase activity have important therapeutic applications such as in wound healing (Agren et al. 1992), treatment of Peyronie's disease (Honig 2014), and Dupuytren's contracture (Hurst et al. 2009). Therefore, future studies by our group will assess if *H. americanus* cysteine cathepsin L2 has biochemical characteristics that are superior to the current therapeutic collagenases.

In this study, an in-depth biochemical characterization was performed on cathepsin D1, the only aspartic acid peptidase detected in gastric juice. This enzyme has been isolated previously, and the pH and temperature for optimum activity has been determined (Rojo et al. 2013). This enzyme can readily degrade hemoglobin (Rojo et al. 2010a, 2013) and its substrate specificity profile is similar to cathepsin D isolated from the gut European hard tick, *Ixodes ricinus* (Sojka et al. 2012). Both of these enzymes cleave between a pair of hydrophobic residues, particularly when phenylalanine is on the amino-terminal side of the scissile bond.

Previous studies have shown that *H. americanus* cathepsin D1 has superior enzymatic properties at low temperature (cold-adapted) when compared to bovine cathepsin D (Rojo et al. 2013). Cold-adapted peptidases have been used in dairy manufacturing, as biocatalyst for synthesis of drugs, in environmental bioremediation and in cleaning products (Gerday et al. 2000). In order to directly compare the substrate specificity and thermodynamic activation of "cold-adapted" lobster cathepsin D1 with the mammalian counterpart's bovine cathepsin D and porcine pepsin, we assayed them under identical conditions. Bovine cathepsin D had a preference for phenylalanine and to a lesser extent, leucine in the S1 subsite, while porcine pepsin cleaved after norleucine and leucine residues at a similar frequency to phenylalanine. Interestingly, each enzyme cut 15 of the 124 peptides at the same site. The substrate sequence information generated from these shared cleavage sites validated the use of an internally quenched fluorescent substrate that we and others have previously employed (Yasuda et al. 1999; Delcroix et al. 2006).

In order to achieve metabolic levels comparable to homoeothermic animals, ocean-dwelling poikilothemic animals have evolved cold-adapted enzymes that generally have higher catalytic efficiency and a reduced activation free energy at low temperature, when compared to related enzymes from homoeothermic species (Feller 1996; Stefansson et al. 2010). In this study, lobster cathepsin D1 showed up to 20-fold higher catalytic efficiency between 5 and 35 °C compared to the bovine and porcine enzymes. Although lobster enzyme showed similar activation energy ($\Delta G^{\#}$) than its mammal counterparts, a significantly lower activation enthalpy ($\Delta H^{\#}$) was observed when compared to bovine cathepsin D, meaning that the changes needed to reach the active conformation and stabilization of the enzyme-substrate complex are energetically less costly, those characteristics are considered as adaptive traits of cold-adapted enzymes (Struvay and Feller 2012). Since the structure of lobster enzyme is not known in detail, we can only speculate for explanations regarding this peculiar behavior of lobster cathepsin D1. Although all structural features known to stabilize proteins can be attenuated in cold-adapted enzymes, it is difficult to precise the extent of the differences directly related to the enhancement of the catalytic efficiency (Feller 2003). Perhaps, only a few of the characteristics of cathepsin D1 are related to the optimization of its activation energy, or major structural changes are involved in order to cathepsin D1 attain the same activation energy as the porcine enzyme.

In summary, an in-depth proteomic study identified multiple hydrolytic enzymes in *H. americanus* gastric juice. Peptidase specificity was uncovered using a novel approached that combined zymography with mass spectrometry-based substrate profiling which sheds light on the fundamental aspects of protein digestion in lobster. Future studies will be required to determine if the proteolytic activities and specificities identified here change according to developmental stages, molt cycle, and different habitats or seasons. This approach can be used to investigate the peptidase substrate profile of any complex marine extract and rapidly identify new enzymes with novel properties.

**Acknowledgments** This work was funded by Consejo Nacional de Ciencia y Tecnología (CONACYT Grant No. 177954 to FGC and scholarship No. 277859 to BBV) and by the UCSF Program for Breakthrough Biomedical Research to C.S.C. Mass spectrometry analysis was performed in the Bio-Organic Biomedical Mass Spectrometry Resource at UCSF (A.L. Burlingame, Director) supported by the Biomedical Technology Research Centers program of the NIH National Institute of General Medical Sciences, NIH NIGMS 8P41GM103481. LRA is a CONACYT research fellow at CIBNOR.

**Compliance with Ethical Standards**

**Conflict of Interest** The authors declare that they have no competing interests.

## References


Agren MS, Taplin CJ, Woessner JF, Eaglstein WH, Mertz PM (1992) Collagenase in wound healing: effect of wound age and type. J Investig Dermatol 99:709–714

Anisimova VE, Shcheglov AS, Bogdanova EA, Rebrikov DV, Nekrasov AN, Barsova EV, Shagin DA, Lukyanov SA (2008) Is crab duplex-specific nuclease a member of the Serratia family of non-specific nucleases? Gene 418:41–48

Arunachalam B, Phan UT, Geuze HJ, Cresswell P (2000) Enzymatic reduction of disulfide bonds in lysosomes: characterization of a gamma-interferon-inducible lysosomal thiol reductase (GILT). Proc Natl Acad Sci U S A 97:745–750

Bibo-Verdugo B, Rojo-Arreola L, Navarrete-Del-Toro MA, García-Carreño F (2015) A chymotrypsin from the digestive tract of California spiny lobster, *Panulirus interruptus*: purification and biochemical characterization. Mar Biotechnol 17:416–427

Biggar KK, Dawson NJ, Storey KB (2012) Real-time protein unfolding: a method for determining the kinetics of native protein denaturation using a quantitative real-time thermocycler. Biotechniques 53:231–238

Brockerhoff H, Hoyle RJ, Hwang PC (1970) Digestive enzymes of the American lobster *Homarus americanus*. J Fish Res Board Can 27: 1357–1370

Burridge LE, Haya K, Zitko V, Waddy S (1999) The lethality of Salmosan (*Azamethiphos*) to American lobster (*Homarus americanus*) larvae, postlarvae, and adults. Ecotoxicol Environ Saf 43:165–169

Chen Y (1991) Characterization of semi-purified collagenase fraction from lobster (*Homarus americanus*). Master of Science Thesis, McGill University

Clark KF, Greenwood SJ, Acorn AR, Byrne PJ (2013) Molecular immune response of the American lobster (*Homarus americanus*) to the white spot syndrome virus. J Invertebr Pathol 114:298–308

Cobb JS, Phillips BF (1980) The biology and management of lobsters. Academic, New York

Corvo I, O'Donoghue AJ, Pastro L, Pi-Denis N, Eroy-Reveles A, Roche L, McKerrow JH, Dalton JP, Craik CS (2013) Dissecting the active site of the collagenolytic cathepsin L3 protease of the invasive stage of *Fasciola hepatica*. PLoS Negl Trop Dis 7:e2269

Delcroix M, Sajid M, Caffrey CR, Lim KC, Dvořák J, Hsieh I, Bahgat M, Dissous C, McKerrow JH (2006) A multienzyme network functions in intestinal protein digestion by a platyhelminth parasite. J Biol Chem 281:39316–39329

Feller G (1996) Enzymes from psychrophilic organisms. FEMS Microbiol Rev 18:189–202

Feller G (2003) Molecular adaptations to cold in psychrophilic enzymes. Cell Mol Life Sci 60:648–662

Feller G, Gerday C (2003) Psychrophilic enzymes: hot topics in cold adaptation. Nat Rev Microbiol 1:200–208

Gerday C, Aittaleb M, Bentahir M, Chessa JP, 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

Gillmor SA, Craik CS, Fletterick RJ (1997) Structural determinants of specificity in the cysteine protease cruzain. Protein Sci 6:1603–1611

Honig SC (2014) Intralesional collagenase in the treatment of Peyronie's disease. Ther Adv Urol 6:47–53

Hu KJ, Leung PC (2007) Food digestion by cathepsin L and digestion-related rapid cell differentiation in shrimp hepatopancreas. Comp Biochem Physiol B Biochem Mol Biol 146:69–80

Hurst LC, Badalamente MA, Hentz VR, Hotchkiss RN, Kaplan FT, Meals RA, Smith TM, Rodzvilla J (2009) Injectable collagenase *Clostridium histolyticum* for Dupuytren's contracture. N Engl J Med 361:968–979

Knight CG (1995) Active-site tirtration of peptidases. Methods Enzymol 248:85–101

Komai T, Kawabata C, Amano M, Lee BR, Ichishima E (2004) Todarepsin, a new cathepsin D from hepatopancreas of Japanese common squid (*Todarodes pacificus*). Comp Biochem Physiol B Biochem Mol Biol 137:373–382

Kongsuwan K, Josh P, Zhu Y, Pearson R, Gough J, Colgrave ML (2010) Exploring the midgut proteome of partially fed female cattle tick (*Rhipicephalus* (*Boophilus*) *microplus*). J Insect Physiol 56:212–226

Laycock MV, Hirama T, Hasnain S, Watson D, Storer AC (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 Chevalier P, Van Wormhoudt A (1998) Alpha-glucosidase from the hepatopancreas of the shrimp, *Penaus vannamei* (Crustacea-Decapoda). J Exp Zool 280:384–394

Lonhienne T, Gerday C, Feller G (2000) Psychrophilic enzymes: revisiting the thermodynamic parameters of activation may explain local flexibility. Biochim Biophys Acta Protein Struct Mol Enzymol 1543:1–10

Lynn KR (1990) Chitinases and chitobiases from the American lobster (*Homarus americanus*). Comp Biochem Physiol B 96:761–766

McGrath ME, Eakin AE, Engel JC, Yang AS, Honig B, Fletterick RJ (1995) The crystal structure of cruzain: a therapeutic target for Chagas' disease. J Mol Biol 247:251–259

Nilsen IW, Øverbø K, Havdalen LJ et al (2010) The enzyme and the cDNA sequence of a thermo labile and double-strand specific DNase from northern shrimps (*Pandalus borealis*). PLoS One 5: e10295

O'Donoghue AJ, Eroy-Reveles AA, Knudsen GM, Ingram J, Zhou M, Statnekov JB, Greninger AL, Hostetter DR, Qu G, Maltby DA, Anderson MO, DeRisi JL, McKerrow JH, Burlingame AL, Craik CS (2012) Global identification of peptidase specificity by multiplex substrate profiling. Nat Methods 9:1095–1103

O'Donoghue A, Jin Y, Knudsen G, Perera NC, Jenne DE, Murphy JE, Craik CS, Hermiston TW (2013) Global substrate profiling of proteases in human neutrophil extracellular traps reveals consensus motif predominantly contributed by elastase. PLoS One 8:e75141

O'Donoghue AJ, Knudsen GM, Beekman C, Perry JA, Johnson AD, DeRisi JL, Craik CS, Bennett RJ (2015) Destructin-1 is a collagen-degrading endopeptidase secreted by Pseudogymnoascus destructans, the causative agent of white-nose syndrome. Proc Natl Acad Sci USA 112:7478–7483

Page MJ, Craik CS (2013) In: Rawlings N, Salvesen G (eds) Handbook of proteolytic enzymes, 3rd edn. Elsevier, London, pp 3049–3052

Rivera-Perez C, Garcia-Carreño FL, Saborowski R (2011) Purification and biochemical characterization of digestive lipase in whiteleg shrimp. Mar Biotechnol 13:284–295

Rojo L, Muhlia-Almazan A, Saborowski R, García-Carreño F (2010a) Aspartic cathepsin D endopeptidase contributes to extracellular digestion in clawed lobsters *Homarus americanus* and *Homarus gammarus*. Mar Biotechnol 12:696–707

Rojo L, Sotelo-Mundo R, García-Carreño F, Gráf L (2010b) Isolation, biochemical characterization, and molecular modeling of American lobster digestive cathepsin D1. Comp Biochem Physiol B Biochem Mol Biol 157:394–400

Rojo L, García-Carreño F, de los Angeles Navarrete del Toro M (2013) Cold-adapted digestive aspartic protease of the clawed lobsters

*Homarus americanus* and *Homarus gammarus*: biochemical characterization. Mar Biotechnol 15:87–96

Schulze WX, Sanggaard KW, Kreuzer I, Knudsen AD, Bemm F, Thogersen IB, Brautigam A, Thomsen LR, Schliesky S, Dyrlund TF, Escalante-Perez M, Becker D, Schultz J, Karring H, Weber A, Hojrup P, Hedrich R, Enghild JJ (2012) The protein composition of the digestive fluid from the Venus flytrap sheds light on prey digestion mechanisms. Mol Cell Proteomics 11:1306–1319

Sojka D, Franta Z, Frantova H, Bartosová P, Horn M, Váchová J, O'Donoghue AJ, Eroy-Reveles AA, Craik CS, Knudsen GM, Caffrey CR, McKerrow JH, Mares M, Kopácek P (2012) Characterization of gut-associated cathepsin D hemoglobinase from tick *Ixodes ricinus* (IrCD1). J Biol Chem 287:21152–21163

Stefansson B, Helgadóttir L, Olafsdottir S et al (2010) Characterization of cold-adapted Atlantic cod (Gadus morhua) trypsin I-kinetic parameters, autolysis and thermal stability. Comp Biochem Physiol B Biochem Mol Biol 155:186–194

Struvay C, Feller G (2012) Optimization to low temperature activity in psychrophilic enzymes. Int J Mol Sci 13:11643–11665

Towle DW, Smith CM (2006) Gene discovery in *Carcinus maenas* and *Homarus americanus* via expressed sequence tags. Integr Comp Biol 46:912–918

Turk B, Turk D, Turk V (2000) Lysosomal cysteine proteases: more than scavengers. Biochim Biophys Acta Protein Struct Mol Enzymol 1477:98–111

Wojtowicz MB, Brockerhoff H (1972) Isolation and some properties of the digestive amylase of the American lobster (*Homarus americanus*). Comp Biochem Physiol 42:295–302

Yasuda Y, Kageyama T, Akamine A, Shibata M, Kominami E, Uchiyama Y, Yamamoto K (1999) Characterization of new fluorogenic substrates for the rapid and sensitive assay of cathepsin E and cathepsin D. J Biochem 125:1137–1143

Zaidi N, Maurer A, Nieke S, Kalbacher H (2008) Cathepsin D: a cellular roadmap. Biochem Biophys Res Commun 376: 5–9

Zhu X, Zhou Y, Feng J (2007) Analysis of both chitinase and chitosanase produced by *Sphingomonas* sp. CJ-5. J Zhejiang Univ Sci B 8(11): 831–838