Major acid endopeptidases of the blood-feeding monogenean *Eudiplozoon nipponicum* (Heteronchoinea: Diplozoidae)

LUCIE JEDLIČKOVÁ[1]†, HANA DVOŘÁKOVÁ[1]†, MARTIN KAŠNÝ[1,2], JANA ILGOVÁ[2], DAVID POTĚŠIL[3], ZBYNĚK ZDRÁHAL[3] *and* LIBOR MIKEŠ[1]*

[1] *Department of Parasitology, Faculty of Science, Charles University in Prague, Viničná 7, 128 44 Prague 2, Czech Republic*
[2] *Department of Botany and Zoology, Faculty of Science, Masaryk University, Kamenice 753/5 (A31), 625 00 Brno, Czech Republic*
[3] *RG Proteomics, Central European Institute of Technology, Masaryk University, Brno, Kamenice 753/5 (A26), 625 00 Brno Bohunice, Czech Republic*

(*Received 22 September 2015; revised 30 November 2015; accepted 30 November 2015; first published online 18 February 2016*)

SUMMARY

In parasitic flatworms, acid endopeptidases are involved in crucial processes, including digestion, invasion, interactions with the host immune system, etc. In haematophagous monogeneans, however, no solid information has been available about the occurrence of these enzymes. Here we aimed to identify major cysteine and aspartic endopeptidase activities in *Eudiplozoon nipponicum*, an invasive haematophagous parasite of common carp. Employing biochemical, proteomic and molecular tools, we found that cysteine peptidase activities prevailed in soluble protein extracts and excretory/secretory products (ESP) of *E. nipponicum*; the major part was cathepsin L-like in nature supplemented with cathepsin B-like activity. Significant activity of the aspartic cathepsin D also occurred in soluble protein extracts. The degradation of haemoglobin in the presence of ESP and worm protein extracts was completely inhibited by a combination of cysteine and aspartic peptidase inhibitors, and diminished by particular cathepsin L, B and D inhibitors. Mass spectrometry revealed several tryptic peptides in ESP matching to two translated sequences of cathepsin L genes, which were amplified from cDNA of *E. nipponicum* and bioinformatically annotated. The dominance of cysteine peptidases of cathepsin L type in *E. nipponicum* resembles the situation in, e.g. fasciolid trematodes.

Key words: cysteine peptidase, aspartic peptidase, protease, haematophagous monogenea, cathepsin L, cathepsin B, cathepsin D, fish parasite, common carp.

INTRODUCTION

Monogeneans of the family Diplozoidae are blood-feeding freshwater ectoparasites inhabiting the gills of cyprinid fishes. They can be significantly virulent to their hosts, causing mechanical damage to gill filaments accompanied by a risk of secondary bacterial and mycotic infections or hypochromic microcytic anaemia (Kawatsu, 1978; Buchmann and Brescini, 2006). The alimentary tract of these worms is morphologically adapted to blood uptake, although the mechanisms of blood processing in monogeneans are largely unknown. The foregut is composed of an oral opening with prominent buccal suckers, with an eversible pharynx that leads via an oesophagus into several dead-end side caeca (Smyth and Halton, 1983; Valigurová *et al*. 2011; Konstanzová *et al*. 2015). Both sections of the gut are morphologically and functionally well differentiated. Numerous gland cells with proposed extracellular digestive function open into the foregut lumen. The chemical nature of their secretion is unknown, but it has been expected that they may produce haemolysins or peptidases, thus could be involved in preliminary degradation of complex foodstuff (e.g., blood cells) to smaller components suitable for endocytosis (Smyth and Halton, 1983; Hodová *et al*. 2010; Valigurová *et al*. 2011). The basic principles of the digestive process have been described for a few monogenean species and they were based mainly on ultrastructural and histochemical studies (Jennings, 1959; Tinsley, 1973; Halton and Stranock, 1976; Hodová *et al*. 2010; Valigurová *et al*. 2011; Konstanzová *et al*. 2015). According to these records, the terminal phase of digestion occurs inside specialized types of gut cells within the lysosomal cycle (Smyth and Halton, 1983). From this view, the digestion in diplozoids is more similar to haematophagous mites such as ticks (Sonenshine, 1991), than to other blood-feeding platyhelmints with their extracellular digestion (Dalton *et al*. 2004). The protein part of haemoglobin is broken down into peptides and amino acids (aa); the toxic haem is oxidized to haematin, exocytosed into the gut lumen and regurgitated by the worm through its oral opening (Llewellyn, 1954; Jennings, 1959; Smyth and Halton, 1983; Konstanzová *et al*. 2015).

* Corresponding author. Department of Parasitology, Faculty of Science, Charles University in Prague, Viničná 7, 128 44 Prague 2, Czech Republic. E-mail: mikes@natur.cuni.cz
† These authors contributed equally to this work.

doi:10.1017/S0031182015001808

Current knowledge of monogenean peptidases involved in digestion and other processes is poor. So far, only two reports have been published concerning peptidases of marine monogeneans *Neobenedenia* spp. feeding on epidermal tissues and mucus of their hosts. A gene encoding cysteine peptidase cathepsin L in *Neobenedenia melleni* was characterized and cloned (Rao and Yang, 2007) and some activity of serine peptidases was found in homogenates of adult *Neobenedenia girellae* by zymography (Hirazawa *et al*. 2006). Their functions are unknown.

In general, processing of blood in sanguinivorous helminths and mites relies on an evolutionary conserved network of cysteine and aspartic peptidases (e.g., cathepsins B, L, F, C, legumain and cathepsin D) identified in trematodes – *Schistosoma mansoni, Schistosoma japonicum, Fasciola hepatica* (Caffrey and Ruppel, 1997; Dalton *et al*. 2003; Sajid *et al*. 2003; Caffrey *et al*. 2004; Delcroix *et al*. 2006; Robinson *et al*. 2008), nematodes – *Haemonchus contortus, Ancylostoma caninum, Necator americanus* (Williamson *et al*. 2003*a*, 2004), and in ticks – *Ixodes ricinus, Boophilus microplus* (Renard *et al*. 2000; Sojka *et al*. 2008). We suppose that similar molecular tools for protein hydrolysis are employed inside the digestive tract of monogeneans.

Our study represents the first direct evidence of the presence of acid endopeptidases in a haematophagous monogenean – *Eudiplozoon nipponicum*; these were identified and partially characterized by biochemical, proteomic and molecular methods.

METHODS

Parasites

Living adult worms of *E. nipponicum* were carefully removed from gills of common carps (*Cyprinus carpio*) immediately after slaughter in a commercial facility of Rybářství Třeboň, Plc. The fish originated from localities (ponds) in South Bohemia, Czech Republic.

Sample preparation

Freshly collected worms were repeatedly gently washed in sterilized tap water, placed in Eppendorf tubes, frozen immediately on dry ice and stored at −80 °C. Soluble protein extracts (*sol*PE) were prepared by homogenization of worms in 10 mM acetate buffer pH 5 using motor-driven teflon pellet pestle followed by 3 cycles of sonication on ice (10 W, 30 s) and centrifugation at 16 000 ***g*** for 20 min. Supernatants were collected. Excretory/secretory products (ESP) were obtained by incubation of washed adult live worms in 10 mM phosphate buffer pH 7·4 or 10 mM acetate buffer pH 6 for 3–4 h at room temperature (RT) in Eppendorf tubes. The ability of the ectoparasitic worms to survive for several hours in selected low osmolarity buffers was verified by previous overnight incubation. Full mobility and no obvious damage of the worms were observed. ESP were concentrated on Amicon Ultra filters (MWCO 10 kDa; Millipore). Protein concentration in all samples was measured by Quant-iT Protein Assay Kit (Invitrogen). The samples were stored at −80 °C.

Living adult worms were washed several times in sterile tap water and then placed in TRIzol® Reagent (Invitrogen) for stabilization of RNA and processed or stored at −80 °C. After homogenization, total RNA was extracted following the TRIzol protocol and concentration measured in NanoDrop 1000 (Thermo Scientific). First-strand cDNA synthesis was carried out with 0·3 μg of total RNA using oligo-dT18 primer from SuperScript™ III First-Strand Synthesis kit (Invitrogen).

Peptidolytic activities, pH optima and inhibition assays

Enzyme activities in *sol*PE and ESP were measured with a set of synthetic fluorogenic peptide substrates (Bachem): Z-Phe-Arg-AMC (FR) was used for detection of papain-like cysteine peptidases (acidic pH) and trypsin-like serine peptidases (neutral to basic pH), Z-Arg-Arg-AMC (RR) selectively for cathepsin B and Z-Ala-Ala-Asn-AMC (AAN) for legumain (asparaginyl endopeptidase). Abz–Lys–Pro–Ala–Glu–Phe–Nph–Arg–Leu (KPAEFnFRL), the specific substrate for cathepsin D, was kindly provided by Dr. Martin Horn, Institute of Organic Chemistry and Biochemistry, AS CR Prague (Abz = 2-aminobenzoyl; Nph and nF = nitrophenylalanine). All assays were performed in black 96-well flat bottom plates (Nunc, Thermo Scientific). Samples (1 and 5 μg of protein per well of *sol*PE and ESP, respectively) were pre-incubated (5 min, RT) in various buffers of pH in the range 2–10 for substrates FR, RR, AAN and pH 2–6 for KPAEFnFRL (0·1 M phosphate buffer pH 2; 0·1 M citrate pH 3; 0·1 M citrate/0·2 M phosphate pH 4–7; 0·2 M Tris–HCl pH 8; 0·2 M glycine/NaOH pH 9–10). Final volume was 100 μL. All buffers contained 2 mM dithiothreitol (DTT, Sigma-Aldrich), except for KPAEFnFRL substrate. The reactions were started by addition of particular peptidyl substrate (final concentration 50 μM) in 100 μL of the same buffer. Kinetics of the release of free fluorophors was detected by Infinite M200 flurometer (TECAN) at 28 °C for 60 min in 1 min intervals. Excitation/emission wavelengths for AMC substrates and nF substrate were 355/460 nm and 330/410 nm, respectively. Controls contained equal volume of appropriate buffer instead of the sample. Measurements were performed in triplicates. The quantification of cysteine and aspartic peptidase activities was performed in particular pH optima with FR and KPAEFnFRL substrates, respectively.

https://doi.org/10.1017/S0031182015001808 Published online by Cambridge University Press

Table 1. Sequences of primers used for amplification of *E. nipponicum* cathepsin L genes.

Primer name	Primer sequence
cDNA fragment PCR with degenerate primers	
Cys forward primer[a]	5′-CAA/GGGNCARTGYGGITCNTGC/TTGG-3′
Asn reverse primer[b]	5′-CCANSA/TRTTYTTIACRATCCAA/GTA-3′
RACE–PCR with specific primers	
M-CL1-RACE5′	5′-CCG ACT TGG ATC ACG CCG TAT TGT TGG T-3′
M-CL1-RACE3′	5′-GAC CTT CCA ATG ACC CTG TCG TGG AGA A-3′
M-CL3-RACE5′	5′-GTG CTG CCC CTC TAG TGA TCC TGT TGT-3′
M-CL3-RACE3′	5′-CCA TGG TGT GCT GGT TGT AGG CTA TGG A-3′
GeneRacer™Forw	5′-CGA CTG GAG CAC GAG GAC ACT GA-3′
GeneRacer™Rev	5′-GCT GTC AAC GAT ACG CTA CGT AAC G-3′

[a] **TGC** nucleotide triplet encoding cysteine.
[b] **RTT** nucleotide triplet encoding asparagine.

Results were expressed as nM of substrate cleaved per minute in the presence of 1 μg of *sol*PE total protein.

Peptidase inhibitors (50 μM final concentration, Sigma-Aldrich) were used for inhibition of particular peptidase activities: E-64 (L-*trans*-epoxysuccinyl–leucylamido [4-guanidino] butane), an irreversible inhibitor of papain-like cysteine peptidases; CA-074 (N-(L-3-*trans*-propylcarbamoyloxirane-2-carbonyl)-L-isoleucyl-L-proline]), an irreversible inhibitor of cathepsin B; iCL (Arg–Lys–Leu–Leu–Trp–NH_2), a reversible inhibitor of cathepsin L; pepstatin A (isovaleryl–Val–Val–Sta–Ala–(3S,4S)-4-amino-3-hydroxy-6-methylheptanoic acid), inhibitor of aspartic peptidases. Inhibitors were mixed with samples prior to the addition of substrates and incubated for 15 min. Inhibition tests were performed in pH optimum of peptidase activity for each particular substrate. All measurements were performed in triplicates, repeated at least three times. Values in graphs are expressed as means with standard deviations.

Degradation of haemoglobin

Bovine haemoglobin (10 μg, Sigma Aldrich) diluted in 20 μL of particular buffers in the range of pH 3-6 (see above) containing 2 mM DTT was incubated with 3 μg of ESP or 1·5 μg of *sol*PE proteins for 16 h at 30 °C. Control reactions contained only haemoglobin. The contribution of individual aspartic or cysteine peptidase activities to degradation of haemoglobin was evaluated by inhibition assays at pH 3 and pH 5. Pepstatin, E-64, CA-074 and iCL were used at concentrations of 10 μM. Resulting hydrolysates were mixed with reducing electrophoretic sample buffer and separated by SDS–PAGE in 4–15% gradient precast gels (Bio-Rad) which were stained with Coomassie Brilliant Blue R-250 (CBB) and scanned on GS-800 Calibrated Densitometer (Bio-Rad).

Active site-labelling of cysteine peptidases

Fluorescent Green BODIPY-DCG-04 (Greenbaum *et al*. 2002), an analogue of E-64 inhibitor which binds irreversibly to the active site of papain-like peptidases, was incubated with ESP (15 μg of total protein) for 1 h (20 μM DCG-04, 5 mM DTT, 5 mM $MgCl_2$) at RT in the dark. Controls were incubated for 30 min with cysteine peptidase inhibitors (10 μM) E-64, iCL or CA-074 prior to addition of the probe. SDS–PAGE of labelled samples was performed as described above. Fluorescent signal was recorded on a fluorescence scanner (Molecular Imager FX, Bio-Rad) using excitation/emission wavelengths 488/530 nm. Finally, the gels were stained by Silver Stain Plus™ Kit (Bio-Rad).

Amplification of partial DNA sequences of cysteine peptidases

cDNA obtained by reverse transcription from adult *E. nipponicum* was amplified by PCR using PPP Master Mix (Top-Bio) and degenerate primers (Table 1) designed according to the consensus sequences corresponding to the coding sequences of cysteine peptidases of several parasitic species (especially according to the active site with a cysteine residue – Cys forward primer and with an asparagine residue – Asn reverse primer) (Eakin *et al*. 1990; Heussler and Dobbelaere, 1996; Li *et al*. 2006). The PCR protocol was as follows: one cycle 5 min initial denaturation at 94 °C, then 35 cycles of denaturation at 94 °C for 1 min, primer annealing 45 °C for 1 min, extension at 72 °C for 1 min and finally one cycle 10 min final extension at 72 °C. Control reactions were performed using the same primers and cDNA obtained from gills of a non-infected carp. The amplified gene fragments of the expected length (ca. 500 bp) were electrophoresed and isolated from agarose gel (1·5%) with Gel Extraction Kit (Qiagen), sub-cloned into pCR2·1-TOPO cloning vector (Invitrogen) and transformed into TOP10 *E. coli* (Invitrogen). pCR2·1-TOPO constructs were isolated using Qiaprep Purification Kit (Qiagen) and sequenced with M13 forward and reverse primers (DNA Sequencing Laboratory, Faculty of Science, Charles University in Prague).

https://doi.org/10.1017/S0031182015001808 Published online by Cambridge University Press

Obtained partial DNA sequences were compared with NCBI database using BLASTX (http://blast.ncbi.nlm.nih.gov/).

Rapid amplification of cDNA ends (RACE–PCR)

In order to amplify the 5′/3′ ends and to obtain the entire gene sequences, two pairs of gene-specific primers (M-CL1-RACE5′ + M-CL1-RACE3′, M-CL3-RACE5′ + M-CL3-RACE3′; see Table 1) were designed according to the obtained ca 500 bp sequence fragments (see above). First-strand cDNA for RACE prepared according to manufacturer's instructions of GeneRacer™Kit (Invitrogen) was used as a template. The PCR reaction (25 μL) contained 12·5 μL of the EmeraldAmp PCR master mix (Clontech), 1 μL of the RACE cDNA (100 ng μL^{-1}), 1 μL of the M-CL1/3-RACE5′/3′ (10 μM), 3 μL of the GeneRacer primers Fwd/Rev (10 μM), 7·5 μL of ddH_20. The profile of amplification of 5′/3′ ends was: 1× 94 °C = 120 s, 5 × (94 °C = 30 s + 72 °C = 60 s), 5 × (94 °C = 30 s + 70 °C = 60 s), 35 × (94 °C = 30 s + 60 °C = 90 s + 70 °C = 90 s) and finally 1 × 72 °C = 10 min. PCR products from 5′ and 3′ RACE were cloned into pCR2·1-TOPO cloning vector as described above and submitted for sequencing with M13 forward and reverse primers (DNA Sequencing Laboratory, Faculty of Science, Charles University in Prague).

Sequence analysis

The obtained full-length cDNA sequences were blasted against sequences available in the GenBank database (NCBI) by BLASTX (http://blast.ncbi.nlm.nih.gov/). The presence of a signal sequence was predicted by the SignalP 4·1 Server (http://www.cbs.dtu.dk/services/SignalP/). The theoretical position of a pro-region was determined by multiple sequence alignment of cathepsins L from other organisms: human cathepsin L [GenBank: NP_001903·1], cathepsin L of *Clonorchis sinensis* [GenBank: ABK91809·1], cathepsin L of *C. carpio* [GenBank: BAD08618·1] and cathepsin L of *Neobenedenia melleni* [GenBank: ABK62794·1]. Multiple sequence alignments of aa sequences were carried out using the program Clustal Omega (http://www.ebi.ac.uk/Tools/msa/clustalo/). Molecular mass and theoretical pI of the deduced proteins were determined by the Compute pI/Mw software available at the Expasy site (http://www.expasy.ch/tools/pitool.html). Detection of potential N-glycosylation sites was performed using an online tool at NetNGlyc 1·0 Server (http://www.cbs.dtu.dk/services/NetNGlyc).

Protein identification from 1D gel bands by LC–MS/MS

ESP were separated by SDS–PAGE as described above and stained by CBB. Bands (range 25–40 kDa) were manually excised, destained and incubated each with trypsin (Promega). LC–MS/MS analyses were done using RSLCnano system (Thermo Fisher Scientific) on-line connected to Impact II Q-TOF mass spectrometer with CaptiveSpray nanoBooster ion source (Bruker). Prior to LC separation, tryptic digests were online concentrated and desalted using trapping column (100 μm × 30 mm) filled with 3·5 μm X-Bridge BEH 130 C18 sorbent (Waters). After washing of trapping column (0·1% FA), the peptides were separated (300 nl/min) using Acclaim Pepmap100 C18 column (2 μm particles, 75 μm × 500 mm; Thermo Fisher Scientific) by the 0·1% FA in water : 0·1% FA in 80% ACN gradient program.

MS data were acquired in a data-dependent strategy with 3 s long cycle time. Mass range was set to 150–2200 m/z and precursors were selected from 300 to 2000 m/z. Acquisition speed of MS and MS/MS scans was 2 and 4–16 Hz, respectively. Speed of MS/MS spectra acquisition was based on precursor intensity (low and high absolute thresholds were 10 000 and 100 000 cts, respectively). Mascot (version 2.4·1) MS/MS ion searches were done against in-house database with sequences of expected recombinant proteins. Search results obtained against in-house database were checked for false positive identifications using database search against the whole SwissProt database (version 2014_07). Mass tolerances for peptides and MS/MS fragments were 15 ppm and 0·05 Da, respectively. Oxidation of methionine, deamidation (N, Q) and propionamide (C) as variable modifications and three enzyme miss cleavages were set for all searches.

RESULTS

Activity profiling

The results of fluorometric analyses of peptidolytic activities in *sol*PE and in ESP, pH optima and inhibition tests are shown in Fig. 1. The highest activity in both sample types was attributed to papain-like cysteine peptidases in acid buffers of pH 3–5 (optimum pH 5) with FR substrate (Fig. 1A and B). It dropped down to only ca. 20% in pH 6 and vanished above pH 7. Inhibitors E-64 and iCL suppressed the activity to zero and CA-074 by ca. 25 and 80% in the case of *sol*PE and ESP, respectively (Fig. 1a and b).

With cathepsin B-selective substrate RR, the activity in *sol*PE peaked in pH 5 and it was negligible in pH 6 and above (Fig. 1C). However, in ESP the highest activity appeared in pH 6, dropped to ca. 50% in pH 7 (Fig. 1D). In the case of both sample types, E-64 and iCL inhibited the peptidolytic activities nearly to zero. CA-074 reduced the activities in ESP by ca. 75% and by ca. 95% in *sol*PE (Fig. 1c and d).

The activity of aspartic cathepsin D with was found in the range of pH 2–4 (optimum in pH 3) in *sol*PE of worms in the presence of KPAEFnFRL

https://doi.org/10.1017/S0031182015001808 Published online by Cambridge University Press

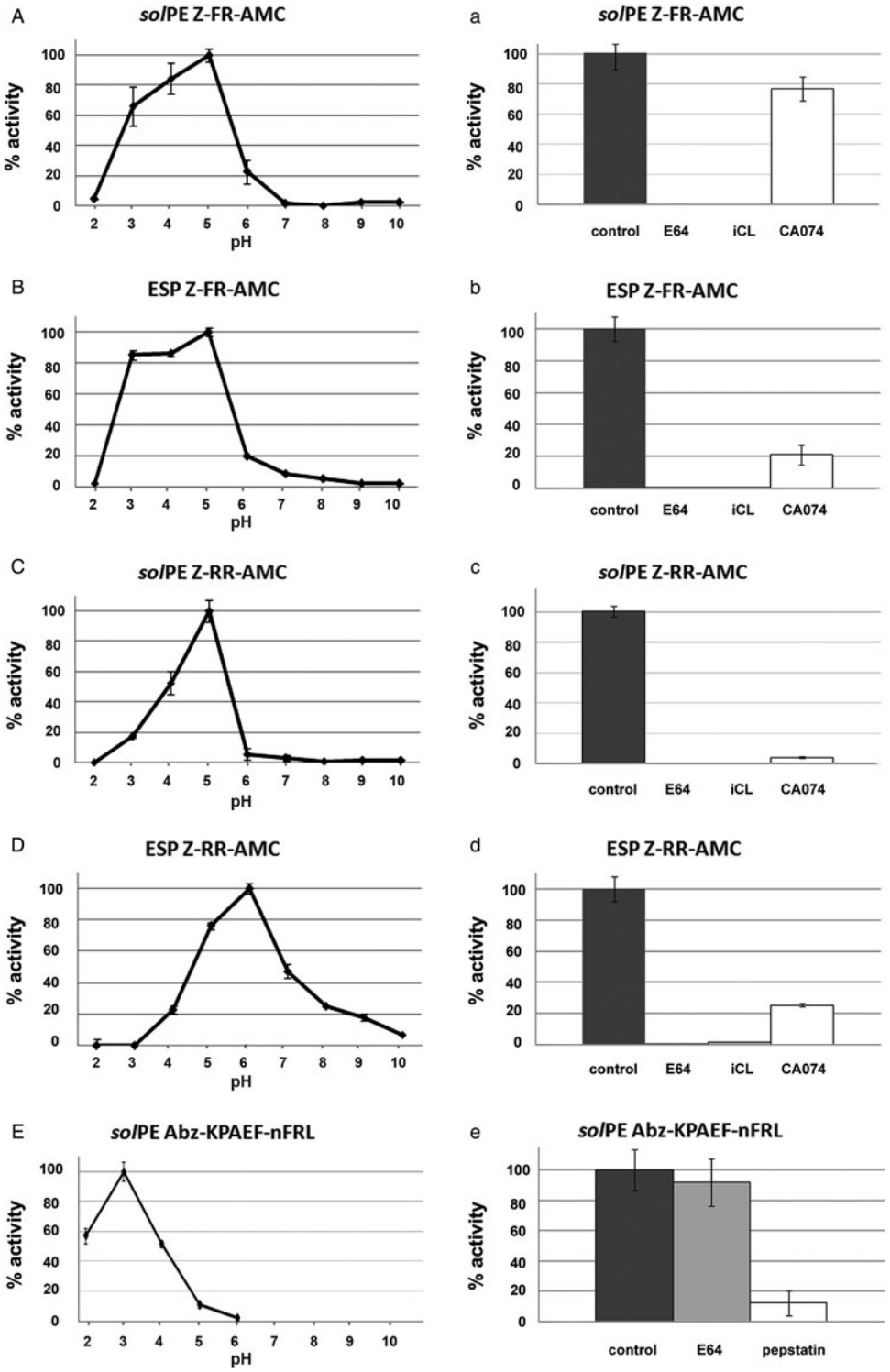

Fig. 1. The pH profiles and inhibition assays of acid endopeptidase activities. The pH optima of peptidolytic activities in soluble protein extracts (*sol*PE) and E/S products (ESP) of *E. nipponicum* and their inhibition at corresponding pH optimum were measured with a spectrum of fluorogenic substrates: (A, a, B, b) Z-FR-AMC; (C, c, D, d) Z-RR-AMC; (E, e) KPAEF-nFRL and with selective peptidase inhibitors (E-64, CA-074, iCL, pepstatin A). The values are expressed as percentage of maximum activity in the sample (100% at pH optimum).

(Fig. 1E); this was inhibited by ca. 90% by pepstatin A (Fig. 1e). In ESP, the activity of cathepsin D was recorded occasionally in some samples (not shown). The ratio of cathepsin L + B to cathepsin D activity in *sol*PE was approx. 4:3 (0·45 *vs* 0·33 nM min^{-1} μg^{-1} of protein). No activity of cysteine asparaginyl endopeptidase (legumain) was detected with AAN substrate in any sample (not shown). There was no difference in activity profiles between the samples of ESP collected at pH 7·4 and pH 6 (not shown). Therefore, a pooled mixture of samples was used for the experiments.

https://doi.org/10.1017/S0031182015001808 Published online by Cambridge University Press

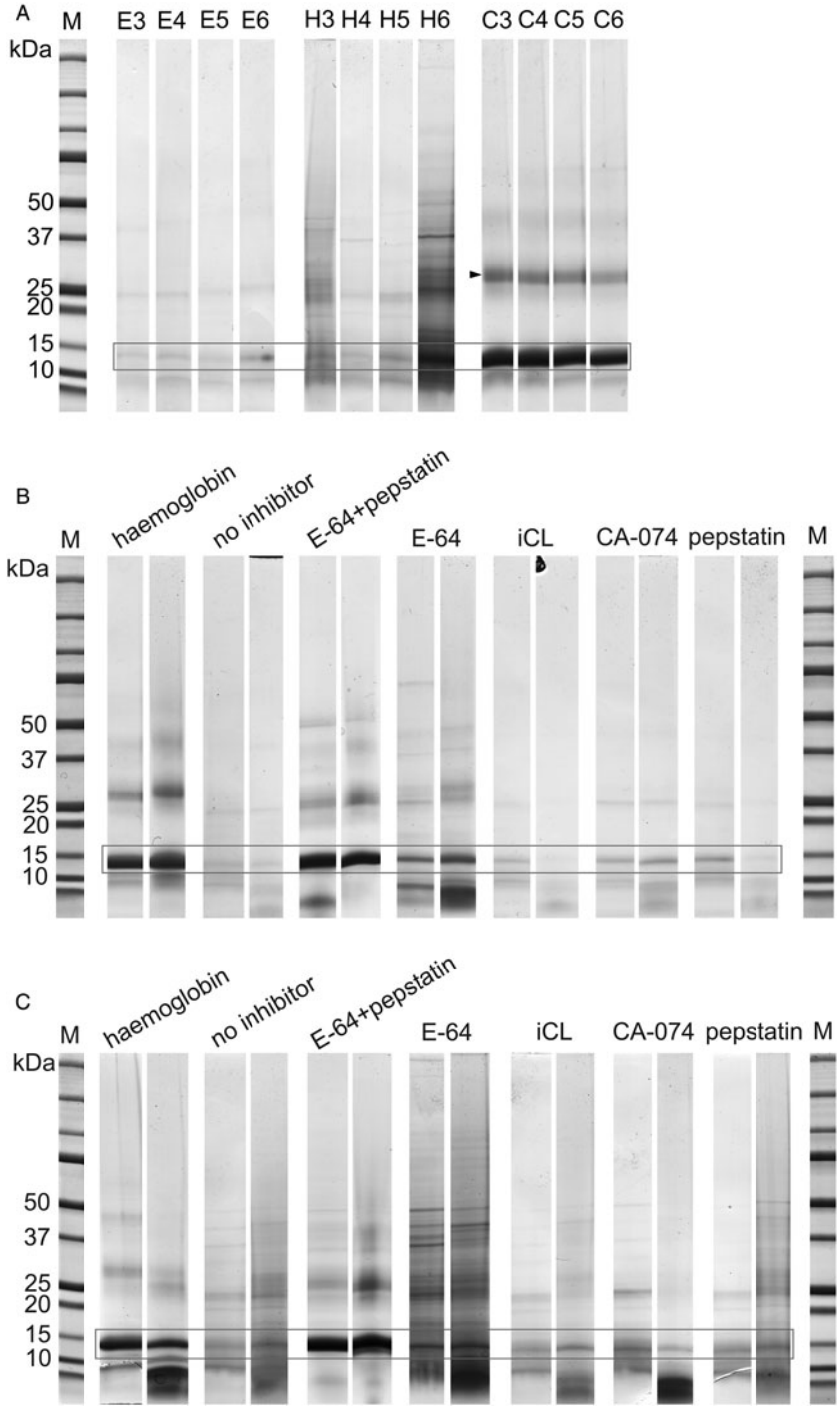

Fig. 2. The pH profile and inhibition of haemoglobin degradation by ESP and *sol*PE. Electrophoretograms of bovine haemoglobin (Hb) after overnight incubation with E/S products and soluble protein extracts of *E. nipponicum*. (A) Degradation in the presence of ESP (lanes E3-E6) and *sol*PE (H3-H6) at pH values 3-6. Controls of Hb in particular pH (C3-C6) contained neither ESP nor *sol*PE. (B) Inhibition of haemoglobinolysis in the presence of ESP in pH 5 and pH 3. (C) Inhibition of heamoglobinolysis in the presence of *sol*PE in pH 5 and pH 3. Headings above each pair of lanes (left lane = pH 5, right lane = pH 3) indicate the type of sample: haemoglobin = controls of Hb without ESP or solPE; no inhibitor = controls of Hb with ESP or *sol*PE; other samples contained inhibitors as labelled. M = markers of molecular size. Bands of Hb monomers are boxed in grey. Arrowhead points to alpha/beta heterodimers of Hb.

Degradation of haemoglobin

ESP as well as *sol*PE efficiently degraded bovine haemoglobin during overnight (16 h) incubation within the pH range of 3–5 and less at pH 6 (Fig. 2A). No substantial degradation occurred above pH 7 (not shown). At pH 5 and 3, haemoglobinolytic activity was markedly (but not completely) inhibited by the general inhibitor of papain-like cysteine peptidases E-64. Lower level of inhibition

https://doi.org/10.1017/S0031182015001808 Published online by Cambridge University Press

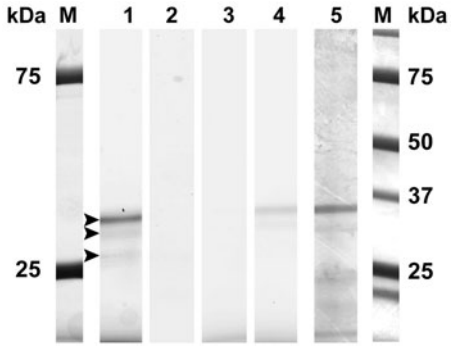

Fig. 3. Active site-labelling of papain-like peptidases in ESP by DCG-04 affinity probe. Electrophoretogram of ESP incubated with fluorescent DCG-04 and inhibitors of particular cysteine peptidases; signal (lanes 1–4) was recorded on a fluorescence scanner, excitation/emission wavelengths 488/530 nm. Lane 1 - control without any inhibitor; lane 2 - E-64; lane 3 - iCL; lane 4 - CA-074; lane 5 - silver staining of the gel. Arrowheads point to the three protein bands labelled by the probe. M = markers of molecular size.

was observed with CA-074, iCL and pepstatin alone; the latter two were less effective at pH 3 in the case of ESP. Complete inhibition at both pH values was reached only when using a mixture of E-64 and pepstatin (Fig. 2B and C). No marked differences in inhibition profiles were observed between the samples containing ESP and *sol*PE.

Active site-labelling of cysteine peptidases

Using the irreversible affinity probe DCG-04, the presence of papain-like cysteine peptidases was demonstrated in the samples of ESP. After SDS–PAGE, three protein bands of molecular sizes between 25 and 35 kDa were demonstrated in the gels. The interaction of DCG-04 with the active sites of cysteine peptidases was efficiently inhibited by the general inhibitor of cysteine peptidases (E-64) and by the inhibitor of cathepsin L (iCL), and only partially by the inhibitor of cathepsin B (CA-074) (Fig. 3).

Amplification of E. nipponicum *cathepsin* L genes

Two amplified DNA fragments (ca. 500 bp) obtained by PCR using *E. nipponicum* cDNA as a template and primers shown in Table 1 were identified as partial sequences of cathepsin L-like cysteine peptidases. In control PCR reactions with DNA template from a non-infected carp and with the same set of primers, a sequence encoding cathepsin L was also amplified. This showed 94·8% sequence identity with cathepsin L gene from *C. carpio* [GenBank: BAD08618·1] and only 55·19 and 49·35% with the fragments amplified from cDNA of *E. nipponicum*.

Applying RACE–PCR on the basis of fragments mentioned above, whole sequences of two *E. nipponicum* cathepsin L genes were obtained, termed EnCL1 and EnCL3. Their nucleotide sequences have been deposited in the NCBI GenBank database under accession numbers [GenBank: KP793605] and [KP793606]. The EnCL1 gene has an open reading frame (ORF) of 954 bp encoding a proenzyme consisting of 317 aa (Fig. 4). Deduced propeptide region contains 93 aa and the catalytic domain (mature enzyme) 224 aa residues. No signal leader sequence has been identified. The calculated theoretical molecular weights of the zymogen/mature protein are 35/24·4 kDa and pI 5·86/6·08. EnCL1 exhibits the highest similarity (53%) to cathepsin L from the Chinese liver fluke *Clonorchis sinensis* [GenBank: ABK91809·1]. The second obtained cathepsin L sequence from *E. nipponicum* was termed EnCL3, as it has the greatest similarity to cathepsin L3 precursor of *Schistosoma mansoni* [GenBank: ABV71063·1]. The ORF of EnCL3 is made up of 1107 bp and encodes a pre-proenzyme of 368 aa (Fig. 4). The sequence is composed of a signal leader sequence (24 aa), an unusually long propeptide (120 aa) and a mature (catalytic) domain (224 aa). Expected Mr/pI values of the zymogen and mature enzyme are 38·0/4·8 and 24·1/4·15, respectively.

The conserved 'ERFNIN' motif typically present in prosequences of various papain-like cysteine peptidases, e.g. cathepsins L, K, S, but not in cathepsin B (Karrer *et al*. 1993), was found in EnCL3; this was slightly modified in the case of EnCL1 (ERFNVN). Another sequence motif, 'GNFD', that may be involved in intra-cellular trafficking and intramolecular processing of some papain-like cysteine peptidases (Vernet *et al*. 1995; Dvořák *et al*. 2009), is modified in the prosequences of EnCL1 and EnCL3 to 'ANLD' and 'TNFD', respectively. The catalytic domains of both enzymes include the typical catalytic triad (EnCL1/3 numbering: Cys_{118}, His_{264}, N_{284}/Cys_{169}, His_{314}, N_{335}) of papain-like cysteine peptidases and the relatively conserved flanking sequences around the catalytic residues (QGQCGS**C**WAFS, LD **H**A/GVL, YWIVK**N**S/TW) (Turk *et al*. 2000) (Fig. 4). No potential N-linked glycosylation sites have been found throughout the sequences of both proenzymes.

The sequence identity between EnCL1 and EnCL3 is 55·21% on the aa level. Identities among *E. nipponicum* cathepsins L and selected aa sequences of cathepsins L from other organisms are summarized in Table 2. EnCL1 shares the highest number of identical aa positions (52·58%) with cathepsin L of *C. sinensis* [GenBank: ABK91809·1] and EnCL3 with cathepsin L3 of *S. mansoni* [GenBank: ABV71063·1] (52·54%). Identities of EnCL1 and EnCL3 with the only published whole sequence of cathepsin L from a monogenean, *N. melleni*, [GenBank: ABK62794·1] are only 44·01 and 41·27%, respectively.

https://doi.org/10.1017/S0031182015001808 Published online by Cambridge University Press

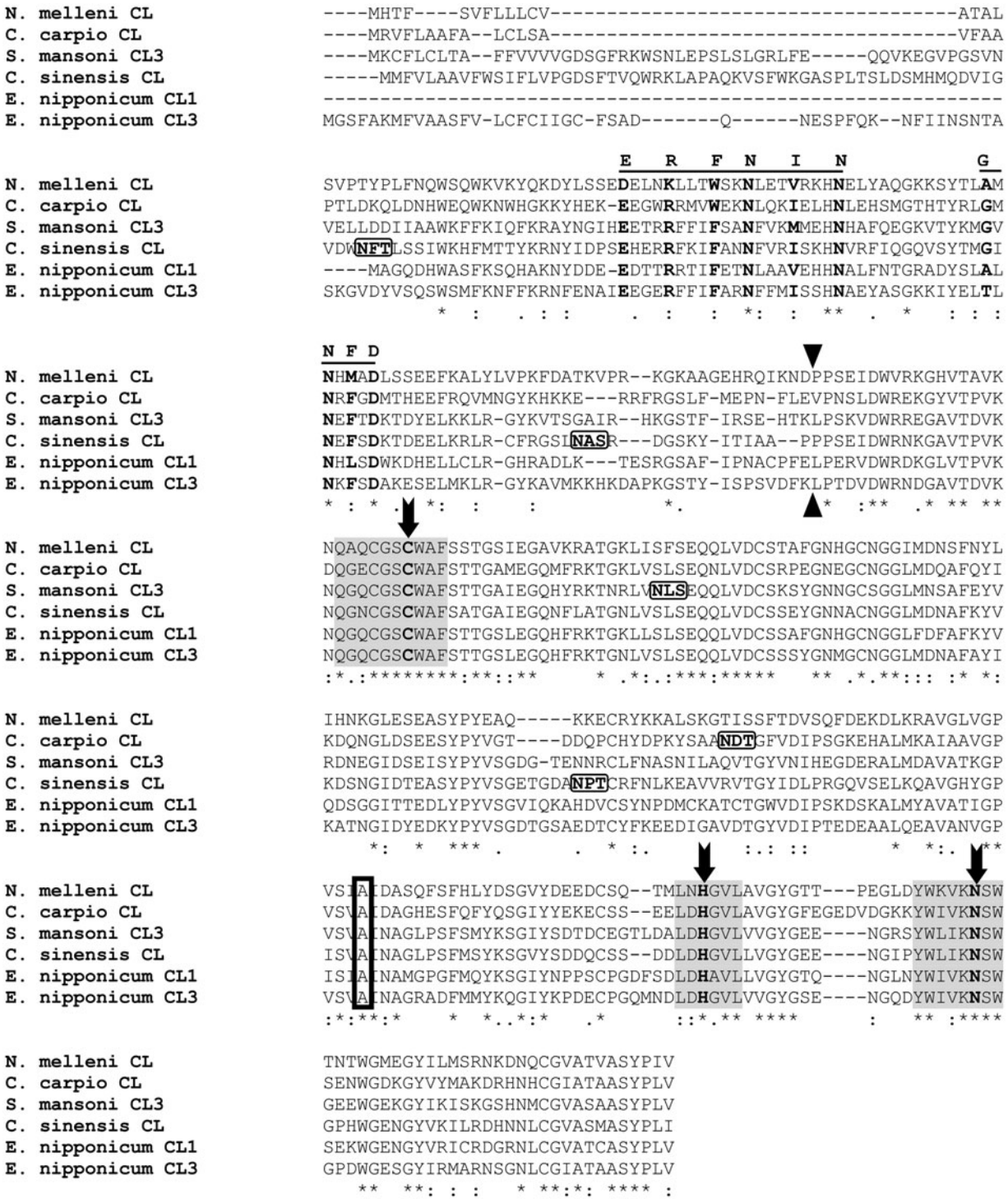

Fig. 4. Multiple sequence alignment of *Eudiplozoon nipponicum* cathepsins L with orthologs from other organisms. *Neobenedenia melleni* [GenBank:ABK62794·1]; *Cyprinus carpio* [GenBank:BAD08618·1]; *Schistosoma mansoni* CL3 [GenBank:ABV71063·1]; *Clonorchis sinensis* [GenBank:ABK91809·1]; *Eudiplozoon nipponicum* CL1 [GenBank: KP793605]; *E. nipponicum* CL3 [GenBank:KP793606]. The position of pro-region cleavage site is marked by arrowheads. ERFNIN- and GNFD-like motifs are in bold and indicated by underlined headings. Catalytic triad of the active site (C, H, N) is marked by arrows. Conserved motifs around active site residues are shaded in light grey. Tripeptides of potential N-glycosylation sites are boxed. Amino residue (A) at the base of the S2 pocket is boxed in black.

Mass spectrometry analysis of ESP of E. nipponicum

Both cathepsins L1 and L3, the genes of which have been amplified, were confirmed in ESP of adult *E. nipponicum* by mass spectrometry. Fifteen peptide sequences from trypsin-digested protein bands matched to the sequence of EnCL1 [GenBank: KP793605] and three peptides matched to EnCL3 [GenBank: KP793606] (Fig. 5). EnCL1

https://doi.org/10.1017/S0031182015001808 Published online by Cambridge University Press

Table 2. Amino acid identities between cathepsins L of *E. nipponicum* and orthologues from other organisms.

	EnCL1	EnCL3
EnCL1	100	55·21
EnCL3	55·21	100
CsCL	52·58	51·12
SmCL3	50·80	52·54
CcCL	50·32	50·30
NmCL	44·01	41·27

Identities expressed in [%]. EnCL1, *Eudiplozoon nipponicum* [GenBank:KP793605]; EnCL3, *E. nipponicum* [GenBank:KP793606]; CsCL, *Clonorchis sinensis* [GenBank:ABK91809·1]; SmCL3, *Schistosoma mansoni* [GenBank:ABV71063·1]; CcCL, *Cyprinus carpio* [GenBank:BAD08618·1]; NmCL, *Neobenedenia melleni* [GenBank:ABK62794·1].

Fig. 5. Identification of cathepsins L1 and L3 in ESP of *E. nipponicum* by mass spectrometry. Each grey line corresponds to a single identified peptide; the sequences of some peptides partially overlap with others. (A) Fifteen tryptic peptides identified in ESP matched to the sequence of cathepsin L1 [GenBank: KP793605]. (B) Three peptides matched to the sequence of cathepsin L3 [GenBank: KP793606].

peptides were identified in nine of ten bands between ca. 24–35 kDa; EnCL3-derived peptides were recorded only above 35 kDa.

DISCUSSION

In contrast to the comprehensive knowledge concerning peptidases from various groups of blood-feeding parasites, only very limited information is available about proteolytic enzymes of haematophagous monogeneans. These neglected helminths in terms of biochemical studies have been thought to digest haemoglobin and other proteins intracellularly in the acidic lysosomes/phagolysosomes within the specialized types of gut cells, unlike other blood-feeding helmints such as schistosomes (Dalton *et al*. 2004), but similarly to ticks (Smyth and Halton, 1983; Sonenshine, 1991). However, it is still not clear whether the first phase of digestion in monogeneans may take place in the extracellular space of the lumen of the gut, with the participation of secreted proteases.

We aimed to identify and partially characterize dominant cysteine and aspartic endopeptidases which can be potentially involved in blood digestion or other processes in the adults of *E. nipponicum*. *sol*PE of worms and ESP were compared in terms of peptidolytic activities. We expected that ESP may be a potential source of peptidases released into the gut lumen (either from gut epithelium or accessory glands that are opened into the front part of the digestive system), and from there excreted outside the worm by continuous regurgitation of the rests of digestion (Smyth and Halton, 1983; Hodová *et al*. 2010; Valigurová *et al*. 2011). However, the presence of peptidases originated in other tissues of the worm cannot be excluded, especially in the case of *sol*PE from whole worms.

Activity profiling revealed that the major peptidase activity can be attributed to the cysteine class, clan CA, family C1 (MEROPS classification). This activity tested with Z-FR-AMC as a substrate at pH 5 was fully inhibited by E-64 (general inhibitor of papain-like cysteine peptidases) and the iCL, but only partially by an inhibitor specific for cathepsin B (CA-074). It is known, that in a reducing environment (presence of DTT), cathepsin L can also be inhibited by CA-074 (Steverding, 2011; Steverding *et al*. 2012). Besides, although iCL (the pentapeptide RKLLW) has been stated as a highly potent iCL (Brinker *et al*. 2000), it was able to diminish significantly the activity of recombinant cathepsin B2 of *Trichobilharzia regenti* in an unrelated previous experiment (Jedličková, unpublished). With the substrate Z-RR-AMC, which is often considered as specific for cathepsin B, the relatively high peptidase activity in ESP was also fully inhibited by E-64 and iCL, but only partially by CA-074 (at pH 6). This brought some confusion on the interpretation of the relative proportion of cathepsin L/B activities.

Cathepsin B- and cathepsin L-like peptidases have been usually distinguished by their ability to degrade the substrate with an arginine in P2 position. Human cathepsin B is able to hydrolyse the substrate, whereas cathepsin L is not – a situation which is not always applicable to orthologous enzymes in other organisms (Sajid and McKerrow, 2002; Choe *et al*. 2006). The substrate specificity of cysteine cathepsins is thought to be determined by interactions in the S2 pocket, particularly the

https://doi.org/10.1017/S0031182015001808 Published online by Cambridge University Press

glutamic acid residue in cathepsin B (Glu_{205} in human CB) and the equivalent alanine residue in cathepsin L (Ala_{205} in human CL) localized at the bottom of the S_2 pocket. The Glu residue can accommodate and stabilize the polar guanidino group of Arg in the substrate, but Ala in this position cannot bind to Arg (Sajid and McKerrow, 2002). This was confirmed for mammalian cathepsins L/B and for peptidases of numerous organisms (including parasites). However, many other parasite peptidases do not possess 'mammalian' residues in this position. For example, the ability to cleave Z-RR-AMC was described for cathepsin L-like peptidase 'cruzain' from *Trypanosoma cruzi* (Gillmor *et al*. 1997) and cathepsin L of *Entamoeba histolytica* (Brinen *et al*. 2000). On the other hand, cathepsin B from *Leishmania major* (Chan *et al*. 1999) and cathepsin B1·4 from *T. regenti* (Dvořák *et al*. 2005) cannot hydrolyse this 'typical' cathepsin B substrate. According to the sequence data of EnCL1 and EnCL3, cathepsins L identified in the ESP of *E. nipponicum* have Ala (Ala_{234} in EnCL1, Ala_{285} in EnCL3) at the bottom of the S2 pocket, and therefore should likely not be able to cleave substrates with Arg in P2 position; however, in the light of the facts mentioned above, this must be verified experimentally with recombinant enzymes.

Cysteine peptidases in both *sol*PE and ESP of *E. nipponicum* exhibited the greatest activity at pH 5 with the substrate Z-FR-AMC; in the presence of Z-RR-AMC the optima were pH 5 and pH 6 for *sol*PE and ESP, respectively. The latter discrepancy could be caused by the presence of different other peptidases in *sol*PE which is a more complex sample in terms of protein composition in comparison to ESP. The results corresponded, e.g. to pH optima measured for cysteine peptidases in secretions of adult *S. mansoni* (Dalton *et al*. 1996); it has been suggested that the activity against Z-FR-AMC at lower pH is predominantly due to cathepsin L, while cathepsin B activity is less significant under these conditions. The presence of at least two different cathepsins L in our samples might explain the high activity in a broader range of acidic pH 3–5. Similarly to our results, the complex sample of protein extract from the gut of *I. ricinus* ticks showed the activity in the presence of the Z-FR-AMC at acidic pH 4-6 (Horn *et al*. 2009). The fact that this substrate was not hydrolysed at pH > 7 suggests that trypsin-like serine peptidases were not present in the samples.

Summarizing the results, we dare to assume that the majority of cysteine peptidase activity in the samples was of cathepsin L-like nature. In addition, a smaller proportion of the activity was cathepsin B-like and sensitive to inhibition by iCL.

To verify the presence of other peptidases in *sol*PE and ESP, substrates suitable for detection of cysteine asparaginyl endopeptidase (legumain) and aspartic cathepsin D were used. Surprisingly, no activity of legumain was detected. Legumain has been found in other blood-feeding helminths as well as ticks (Dalton *et al*. 1995; Caffrey *et al*. 2000; Oliver *et al*. 2006; Sojka *et al*. 2007; Abdul *et al*. 2007). Its function is connected not only with hydrolysis of ingested host proteins, but it also plays an important role in activation of other peptidases by cleaving their prosequences (Sajid *et al*. 2003; Caffrey *et al*. 2004; Sojka *et al*. 2008). There is a certain possibility that this enzyme occurred in our samples as an inactive zymogen or an inactive variant. Significant activity of cathepsin D (ca. ¾ of that of cathepsins L + B) was observed in *sol*PE, whereas only a minute activity was usually recorded in some samples of ESP (not shown). Also cathepsin D acts in the digestive process in other blood-feeding parasites (Brindley *et al*. 2001; Verity *et al*. 2001; Banerjee *et al*. 2002; Williamson *et al*. 2003*b*; Boldbaatar *et al*. 2006; Sojka *et al*. 2012) and we expect its involvement in blood digestion by monogenean parasites, too. The non-presence or low abundance of activities of these two crucial enzymes in some samples could be also possibly explained by alleged instability of their molecules and low production into the gut lumen.

In vitro degradation of haemoglobin by *E. nipponicum* peptidases is optimal at acidic pH, which corresponds with the pH optima of cysteine and aspartic peptidases. The reaction was completely blocked by a mixture of E-64 and pepstatin, less by E-64 and only partially by individual inhibitors. These results clearly confirmed that cysteine cathepsins L + B and aspartic cathepsin D are involved in haemoglobinolysis in *E. nipponicum*.

Affinity labelling with the active-site specific probe DCG-04 confirmed active cysteine peptidases in ESP, proved in three bands (between ca. 25–37 kDa), indicating the presence of various cysteine peptidases or intermediates of their activation by limited proteolysis (the sizes in gel approximately agree with theoretical MWs of proenzymes and mature parts of EnCL1 and EnCL3). The binding of the probe to all three bands was fully inhibited by both E-64 and iCL, whereas CA-074 inhibited only the band of lowest MW. This can be considered as another proof of the presence of both cathepsins L and B in ESP of *E. nipponicum*.

PCR with degenerate primers and RACE-PCR led to the amplification of *E. nipponicum* gene specific DNAs which were sequenced and identified as cathepsin L genes (named EnCL1 and EnCL3) [GenBank: KP793605, KP793606]. The pro-peptide of *En*CL1 and *En*CL3 contains some highly conserved regions typical for cathepsins L. One such a region is a variation of the $EX_3RX_2(V/I)FX_2NX_3IX_3N$ ('ERFNIN') motif (present as 'ERFNVN' in EnCL1) (Karrer *et al*. 1993). The residues in this motif are probably important in stabilizing the globular structure of the propeptide (Coulombe *et al*. 1996; Groves *et al*. 1998).

https://doi.org/10.1017/S0031182015001808 Published online by Cambridge University Press

Another conserved motif is the GXNXFXD ('GNFD') present as 'ANLD' in EnCL1 and as 'TNFD' in EnCL3, with residues supposed to be involved in the control of intramolecular processing of cysteine peptidase precursors and/or participate in folding (Vernet *et al*. 1995; Dvořák *et al*. 2009). Processing is believed to be facilitated by a low pH. When pH is lowered, protonation of Asp_{36} (papain numbering) may cause a conformational change of the cysteine peptidase precursor in which the propeptide is bound less tightly into the active site, making it more susceptible for clip-off (Coulombe *et al*. 1996; Jerala *et al*. 1998). Amino acid differences in these conserved motifs in EnCL1 and EnCL3 should not affect the function.

An asparagine residue, found between the propeptide and mature domain of gut cysteine peptidases – cathepsins from other blood-feeding parasites, e.g. schistosomes and ticks, is responsible for recognition and cleavage by legumain for trans-activation (Sajid *et al*. 2003; Sojka *et al*. 2007). However, this Asn is absent in both EnCL1 and EnCL3, suggesting that pro-EnCL1 and pro-EnCL3 are rather activated by autocatalysis.

The analysis of EnCL1 sequence using SignalP 4·1 software did not indicate the presence of an N-terminal signal sequence necessary for vesicle or extracellular targeting. The absence of such signal sequence has been described for a few cathepsins L from other organisms, e.g. the assassin bug *Rhodnius prolixus* cathepsin L-like protein (Lopez-Ordoñez *et al*. 2001). Additionally, it has been reported that human cathepsin L can be directed to a secretion pathway due to an aa sequence (-SXPXV) located at the carboxy (C) terminus of the enzyme (Chauhan *et al*. 1998). Since the secretion of EnCL1 was supported by our result from mass spectrometry analysis, it seems that it may be realized in an alternative way; a non-signal peptide triggered protein secretion was also predicted by SecretomeP 2·0 software tool (http://www.cbs.dtu.dk/services/SecretomeP/), but just at threshold values of the analysis. So far, it is not clear whether some cryptic (unknown) intramolecular signals are involved in secretion of this enzyme and only speculations can be made on another possibility – extracellular transport in exosomes. On the other hand, EnCL3 possesses a typical N-terminal signal sequence (24 aa). EnCL3 zymogen has an atypical long pro-peptide (120 aa) similarly as the human blood fluke *S. mansoni* SmCL3 (130 aa) [GenBank: ABV71063·1]. It has been suggested that these extensions may have other functions to the proenzyme, for example in protein trafficking or as binding sites for other proteins (Dvořák *et al*. 2009).

Both EnCL1 and EnCL3 have no potential N-linked glycosylation sites in their sequences, similarly to cathepsin L of the monogenean parasite *N. melleni* [GenBank: ABK62794·1]. In general, glycosylation with mannose 6-phosphate has been shown to be an important sorting signal for routing mammalian proteins into lysosomes. However, there are several experimental evidences for mannose 6-phosphate independent trafficking of proteins into lysosomes (Ni *et al*. 2006; Braulke and Bonifacino, 2009). According to this, it would be possible that EnCL occurrence in ESP is not connected with the discharge of phagolysosome residual content into the gut lumen or that the enzyme may be targeted into lysosomes in an alternative manner.

Both cathepsins L were confirmed in *E. nipponicum* ESP in areas of various molecular sizes by means of mass spectrometry. Individual bands may represent partial processing or autolysis of the enzymes. According to the results, it seems that EnCL1 represents an abundant peptidase in ESP of *E. nipponicum*. In addition, the dominant protein spots found in *E. nipponicum sol*PE separated by 2D electrophoresis were also identified by mass spectrometry as EnCL1 (data not shown).

Digestion of proteins in blood-feeding parasites represents a major proteolytic process. We showed that the most abundant haemoglobinolytic endopeptidase activities in *E. nipponicum* belong to the cysteine class, with cathepsin L-like activity predominating over cathepsin B-like activity. Proteomic investigations confirmed this fact. Besides, significant involvement of aspartic cathepsin D was well documented. The possibility of aimed secretion of proteolytic enzymes to the external environment by this ectoparasite of carp gills appears to be purposeless and therefore improbable. From this point of view and based on our results, we believe that cathepsins L, B and D are involved in processing of ingested host's blood in *E. nipponicum*.

ACKNOWLEDGEMENTS

We are grateful to Dr C. Caffrey and Dr D. Greenbaum, Center for Discovery and Innovation in Parasitic Diseases, UCSF, USA, who kindly provided the probe Green BODIPY-DCG-04. The MS part of the work was carried out with the support of Proteomics Core Facility ('CEITEC – Central European Institute of Technology'). We also thank Mr Stanislav Sojka, Rybářství Třeboň, Plc., for arrangements concerning collection of parasites.

FINANCIAL SUPPORT

This research was enabled due to the grants of the Czech Science Foundation (L.M., Grant no. P506/12/1258), (Z.Z., Grant no. P206-12-G151); Grant Agency of the Charles University in Prague (L.J., Grant no. 502313); Masaryk University institutional support (M.K., J.I., Grant no. MUNI/A/1484/2014); Charles University institutional/departmental support (M.K., Grant no. UNCE 204017; L.J., H.D., M.K., L.M., Grant no. PRVOUK P41; L.J., H.D., Grant no. SVV 260202/2015); and the European Regional Development Fund (Z.Z., Grant no. CZ.1·05/1·1·00/02·0068).

REFERENCES

Abdul, A. M., Tsuji, N., Miyoshi, T., Khyrul Islam, M., Huang, X., Motobu, M. and Fujisaki, K. (2007). Characterization of asparaginyl endopeptidase, legumain induced by blood feeding in the ixodid tick *Haemaphysalis longicornis*. *Insect Biochemistry and Molecular Biology* **37**, 911–922.

Banerjee, R., Liu, J., Beatty, W., Pelosof, L., Klemba, M. and Goldberg, D. E. (2002). Four plasmepsins are active in the *Plasmodium falciparum* food vacuole, including a protease with an active-site histidine. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 990–995.

Boldbaatar, D., Sikalizyo Sikasunge, C., Battsetseg, B., Xuan, X. and Fujisaki, K. (2006). Molecular cloning and functional characterization of an aspartic protease from the hard tick *Haemaphysalis longicornis*. *Insect Biochemistry and Molecular Biology* **36**, 25–36.

Braulke, T. and Bonifacino, J. S. (2009). Sorting of lysosomal proteins. *Biochimica et Biophysica Acta* **1793**, 605–614.

Brindley, P. J., Kalinna, B. H., Wong, J. Y., Bogitsh, B. J., King, L. T., Smyth, D. J., Verity, C. K., Abbenante, G., Brinkworth, R. I., Fairlie, D. P., Smythe, M. L., Milburn, P. J., Bielefeldt-Ohmann, H., Zheng, Y. and McManus, D. P. (2001). Proteolysis of human hemoglobin by schistosome cathepsin D. *Molecular and Biochemical Parasitology* **112**, 103–112.

Brinen, L. S., Que, X., McKerrow, J. H. and Reed, S. L. (2000). Homology modeling of *Entamoeba histolytica* cysteine proteinases reveals the basis for cathepsin L-like structure with cathepsin B-like specificity. *Archives of Medical Research* **31**, S63–S64.

Brinker, A., Weber, E., Stoll, D., Voigt, J., Müller, A., Sewald, N., Jung, G., Wiesmüller, K. H. and Bohley, P. (2000). Highly potent inhibitors of human cathepsin L identified by screening combinatorial pentapeptide amide collections. *European Journal of Biochemistry/FEBS* **267**, 5085–5092.

Buchmann, K. and Brescini, J. (2006). *Monogenea*. *In Fish Diseases and Disorders*, Volume 1, 2nd Edn. pp. 297–344. CABI, Wallingford.

Caffrey, C. R. and Ruppel, A. (1997). Cathepsin B-like activity predominates over cathepsin L-like activity in adult *Schistosoma mansoni* and *S. japonicum*. *Parasitology Research* **83**, 632–635.

Caffrey, C. R., Mathieu, M. A., Gaffney, A. M., Salter, J. P., Sajid, M., Lucas, K. D., Franklin, C., Bogyo, M. and McKerrow, J. H. (2000). Identification of a cDNA encoding an active asparaginyl endopeptidase of *Schistosoma mansoni* and its expression in *Pichia pastoris*. *FEBS Letters* **466**, 244–248.

Caffrey, C. R., McKerrow, J. H., Salter, J. P. and Sajid, M. (2004). Blood "n" guts: an update on schistosome digestive peptidases. *Trends in Parasitology* **20**, 241–248.

Coulombe, R., Grochulski, P., Sivaraman, J., Ménard, R., Mort, J. S. and Cygler, M. (1996). Structure of human procathepsin L reveals the molecular basis of inhibition by the prosegment. *EMBO Journal* **15**, 5492–5503.

Dalton, J. P., Hola-Jamriska, L. and Brindley, P. J. (1995). Asparaginyl endopeptidase activity in adult *Schistosoma mansoni*. *Parasitology* **111**, 575–580.

Dalton, J. P., Clough, K. A., Jones, M. K. and Brindley, P. J. (1996). Characterization of the cathepsin-like cysteine proteinases of *Schistosoma mansoni*. *Infection and Immunity* **64**, 1328–1334.

Dalton, J. P., Neill, S. O., Stack, C., Collins, P., Walshe, A., Sekiya, M., Doyle, S., Mulcahy, G., Hoyle, D., Khaznadji, E., Moire, N., Brennan, G., Mousley, A., Kreshchenko, N., Maule, A. G. and Donnelly, S. M. (2003). Fasciola hepatica cathepsin L-like proteases: biology, function, and potential in the development of first generation liver fluke vaccines. *International Journal for Parasitology* **33**, 1173–1181.

Dalton, J. P., Skelly, P. and Halton, D. W. (2004). Role of the tegument and gut in nutrient uptake by parasitic platyhelminths. *Canadian Journal of Zooology*. **82**, 211–232.

Delcroix, M., Sajid, M., Caffrey, C. R., Lim, K. C., Dvořák, J., Hsieh, I., Bahgat, M., Dissous, C. and McKerrow, J. H. (2006). A multienzyme network functions in intestinal protein digestion by a platyhelminth parasite. *Journal of Biological Chemistry* **281**, 39316–39329.

Dvořák, J., Delcroix, M., Rossi, A., Vopálenský, V., Pospíšek, M., Šedinová, M., Mikeš, L., Sajid, M., Sali, A., McKerrow, J. H., Horák, P. and Caffrey, C. R. (2005). Multiple cathepsin B isoforms in schistosomula of *Trichobilharzia regenti*: identification, characterisation and putative role in migration and nutrition. *International Journal for Parasitology* **35**, 895–910.

Dvořák, J., Mashiyama, S. T., Sajid, M., Braschi, S., Delcroix, M., Schneider, E. L., McKerrow, W. H., Bahgat, M., Hansell, E., Babbitt, P. C., Craik, C. S., McKerrow, J. H. and Caffrey, C. R. (2009). SmCL3, a gastrodermal cysteine protease of the human blood fluke *Schistosoma mansoni*. *PLoS Neglected Tropical Diseases* **3**, e449.

Eakin, A. E., Bouvier, J., Sakanari, J. A., Craik, C. S. and McKerrow, J. H. (1990). Amplification and sequencing of genomic DNA fragments encoding cysteine proteases from protozoan parasites. *Molecular and Biochemical Parasitology* **39**, 1–8.

Gillmor, S. A., Craik, C. S. and Fletterick, R. J. (1997). Structural determinants of specificity in the cysteine protease cruzain. *Protein Science: a Publication of the Protein Society* **6**, 1603–1611.

Greenbaum, D., Baruch, A., Hayrapetian, L., Darula, Z., Burlingame, A., Medzihradszky, K. F. and Bogyo, M. (2002). Chemical approaches for functionally probing the proteome. *Molecular & Cellular Proteomics: MCP* **1**, 60–68.

Groves, M. R., Coulombe, R., Jenkins, J. and Cygler, M. (1998). Structural basis for specificity of papain-like cysteine protease proregions toward their cognate enzymes. *Proteins* **32**, 504–514.

Halton, D. W. and Stranock, S. D. (1976). The fine structure and histochemistry of the caecal epithelium of *Calicotyle kröyeri* (Monogenea: Monopisthocotylea). *International Journal for Parasitology* **6**, 253–263.

Heussler, V. and Dobbelaere, D. (1996). Cloning of a protease gene family in Fasciola hepatica using the polymerase chain reaction (PCR). *Schweizer Archiv für Tierheilkunde* **138**, 125–132.

Hirazawa, N., Umeda, N., Hatanaka, A. and Kuroda, A. (2006). Characterization of serine proteases in the monogenean *Neobenedenia girellae*. *Aquaculture* **255**, 188–195.

Hodová, I., Matějusová, I. and Gelnar, M. (2010). The surface topography of *Eudiplozoon nipponicum* (Monogenea) developmental stages parasitizing carp (*Cyprinus carpio* L.). *Central European Journal of Biology* **5**, 702–709.

Horn, M., Nussbaumerová, M., Šanda, M., Kovářová, Z., Srba, J., Franta, Z., Sojka, D., Bogyo, M., Caffrey, C. R., Kopáček, P. and Mareš, M. (2009). Hemoglobin digestion in blood-feeding ticks: mapping a multipeptidase pathway by functional proteomics. *Chemistry & Biology* **16**, 1053–1063.

Chan, V. J., Selzer, P. M., McKerrow, J. H. and Sakanari, J. A. (1999). Expression and alteration of the S2 subsite of the *Leishmania major* cathepsin B-like cysteine protease. *Biochemical Journal* **340**, 113–117.

Chauhan, S. S., Ray, D., Kane, S. E., Willingham, M. C. and Gottesman, M. M. (1998). Involvement of carboxy–terminal amino acids in secretion of human lysosomal protease cathepsin L. *Biochemistry* **37**, 8584–8594.

Choe, Y., Leonetti, F., Greenbaum, D. C., Lecaille, F., Bogyo, M., Brömme, D., Ellman, J. A. and Craik, C. S. (2006). Substrate profiling of cysteine proteases using a combinatorial peptide library identifies functionally unique specificities. *Journal of Biological Chemistry* **281**, 12824–12832.

Jennings, J. B. (1959). Studies on digestion in the monogenetic trematode *Polystoma integerrimum*. *Journal of Helminthology* **33**, 197–204.

Jerala, R., Zerovnik, E., Kidric, J. and Turk, V. (1998). pH-induced conformational transitions of the propeptide of human cathepsin L. A role for a molten globule state in zymogen activation. *Journal of Biological Chemistry* **273**, 11498–11504.

Karrer, K. M., Peiffer, S. L. and DiTomas, M. E. (1993). Two distinct gene subfamilies within the family of cysteine protease genes. *Proceedings of the National Academy of Sciences of the United States of America* **90**, 3063–3067.

Kawatsu, H. (1978). Studies on the anemia of fish - IX. Hypochromic microcytic anemia of crusian carp caused by infestation with a trematode, *Diplozoon nipponicum*. *Bulletin of the Japanese Society of Scientific Fisheries* **44**, 1315–1319.

Konstanzová, V., Koubková, B., Kašný, M., Ilgová, J., Dzika, E. and Gelnar, M. (2015). Ultrastructure of the digestive tract of *Paradiplozoon homoion* (Monogenea). *Parasitology Research* **114**, 1485–1494.

Li, A. H., Moon, S.-U., Park, Y.-K., Na, B.-K., Hwang, M.-G., Oh, C.-M., Cho, S.-H., Kong, Y., Kim, T.-S. and Chung, P.-R. (2006). Identification and characterization of a cathepsin L-like cysteine protease from *Taenia solium* metacestode. *Veterinary Parasitology* **141**, 251–259.

Llewellyn, J. (1954). Observations on the food and the gut pigment of the Polyopisthocotylea (Trematoda: Monogenea). *Parasitology* **44**, 428–37.

Lopez-Ordoñez, T., Rodriguez, M. H. and Hernández-Hernández, F. D. (2001). Characterization of a cDNA encoding a cathepsin L-like protein of *Rhodnius prolixus*. *Insect Molecular Biology* **10**, 505–511.

Ni, X., Canuel, M. and Morales, C. R. (2006). The sorting and trafficking of lysosomal proteins. *Histology and Histopathology* **21**, 899–913.

Oliver, E. M., Skuce, P. J., McNair, C. M. and Knox, D. P. (2006). Identification and characterization of an asparaginyl proteinase (legumain) from the parasitic nematode, *Haemonchus contortus*. *Parasitology* **133**, 237–244.

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

Renard, G., Garcia, J. F., Cardoso, F. C., Richter, M. F., Sakanari, J. A., Ozaki, L. S., Termignoni, C. and Masuda, A. (2000). Cloning and functional expression of a *Boophilus microplus* cathepsin L-like enzyme. *Insect Biochemistry and Molecular Biology* **30**, 1017–1026.

Robinson, M. W., Dalton, J. P. and Donnelly, S. (2008). Helminth pathogen cathepsin proteases: it's a family affair. *Trends in Biochemical Sciences* **3**, 601–608.

Sajid, M. and McKerrow, J. H. (2002). Cysteine proteases of parasitic organisms. *Molecular and Biochemical Parasitology* **120**, 1–21.

Sajid, M., McKerrow, J. H., Hansell, E., Mathieu, M. A., Lucas, K. D., Hsieh, I., Greenbaum, D., Bogyo, M., Salter, J. P., Lim, K. C., Franklin, C., Kim, J.-H. and Caffrey, C. R. (2003). Functional expression and characterization of *Schistosoma mansoni* cathepsin B and its trans-activation by an endogenous asparaginyl endopeptidase. *Molecular and Biochemical Parasitology* **131**, 65–75.

Smyth, J. D. and Halton, D. W. (1983). *The Physiology of Trematodes*. Cambridge University Press, Cambridge, UK.

Sojka, D., Hajdušek, O., Dvořák, J., Sajid, M., Franta, Z., Schneider, E. L., Craik, C. S., Vancová, M., Burešová, V., Bogyo, M., Sexton, K. B., McKerrow, J. H., Caffrey, C. R. and Kopáček, P. (2007). IrAE: an asparaginyl endopeptidase (legumain) in the gut of the hard tick *Ixodes ricinus*. *International Journal for Parasitology* **37**, 713–724.

Sojka, D., Franta, Z., Horn, M., Hajdušek, O., Caffrey, C. R., Mareš, M. and Kopáček, P. (2008). Profiling of proteolytic enzymes in the gut of the tick *Ixodes ricinus* reveals an evolutionarily conserved network of aspartic and cysteine peptidases. *Parasites & Vectors* **1**, 7.

Sojka, D., Franta, Z., Frantová, H., Bartošová, P., Horn, M., Váchová, J., O'Donoghue, A. J., Eroy-Reveles, A. A., Craik, C. S., Knudsen, G. M., Caffrey, C. R., McKerrow, J. H., Mareš, M. and Kopáček, P. (2012). Characterization of gut-associated cathepsin D hemoglobinase from tick *Ixodes ricinus* (IrCD1). *Journal of Biological Chemistry* **287**, 21152–21163.

Sonenshine, D. E. (1991). *Biology of Ticks Volume 1*. Oxford University Press, New York.

Steverding, D. (2011). The cathepsin B-selective inhibitors CA-074 and CA-074Me inactivate cathepsin L under reducing conditions. *Open Enzyme Inhibition Journal* **4**, 11–16.

Steverding, D., Sexton, D. W., Wang, X., Gehrke, S. S., Wagner, G. K. and Caffrey, C. R. (2012). *Trypanosoma brucei*: chemical evidence that cathepsin L is essential for survival and a relevant drug target. *International Journal for Parasitology* **42**, 481–488.

Tinsley, R. C. (1973). Ultrastructural studies on the form and function of the gastrodermis of *Protopolystoma xenopi* (Monogenoidea: Polyopisthocotylea). *Biological Bulletin* **144**, 541–555.

Turk, B., Turk, D. and Turk, V. (2000). Lysosomal cysteine proteases: more than scavengers. *Biochimica et Biophysica Acta* **1477**, 98–111.

Valigurová, A., Hodová, I., Sonnek, R., Koubková, B. and Gelnar, M. (2011). *Eudiplozoon nipponicum* in focus: monogenean exhibiting a highly specialized adaptation for ectoparasitic lifestyle. *Parasitology Research* **108**, 383–394.

Verity, C. K., Loukas, A., McManus, D. P. and Brindley, P. J. (2001). *Schistosoma japonicum* cathepsin D aspartic protease cleaves human IgG and other serum components. *Parasitology* **122**, 415–421.

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

Williamson, A. L., Brindley, P. J., Knox, D. P., Hotez, P. J. and Loukas, A. (2003*a*). Digestive proteases of blood-feeding nematodes. *Trends in Parasitology* **19**, 417–423.

Williamson, A. L., Brindley, P. J., Abbenante, G., Datu, B. J. D., Prociv, P., Berry, C., Girdwood, K., Pritchard, D. I., Fairlie, D. P., Hotez, P. J., Zhan, B. and Loukas, A. (2003*b*). Hookworm aspartic protease, Na-APR-2, cleaves human hemoglobin and serum proteins in a host-specific fashion. *Journal of Infectious Diseases* **187**, 484–494.

Williamson, A. L., Lecchi, P., Turk, B. E., Choe, Y., Hotez, P. J., McKerrow, J. H., Cantley, L. C., Sajid, M., Craik, C. S. and Loukas, A. (2004). A multi-enzyme cascade of hemoglobin proteolysis in the intestine of blood-feeding hookworms. *Journal of Biological Chemistry* **279**, 35950–35957.