FASCIOLA GIGANTICA CATHEPSIN B8, AN ISOFORM PRESENT FROM METACERCARIAL TO MATURE STAGE

Thwet Oo Lwin[1,2], Amornrat Geadkaew-Krenc[1], Sinee Siricoon[3] and Rudi Grams[1]

[1]Graduate Program in Biomedical Sciences, Faculty of Allied Health Sciences, Thammasat University, Pathumthani Province, Thailand; [2]Department of Medical Technology, University of Medical Technology, Mandalay, Myanmar; [3]Thailand Institute of Scientific and Technological Research, Pathumthani Province, Thailand

Abstract. Basic molecular and biochemical properties of a new isoform of tropical liver fluke *Fasciola gigantica* cathepsin B FgCB8 were analyzed. This isoform, also found in *F. hepatica*, most likely represents an evolutionary cathepsin B prototype in trematodes and highly conserved orthologs are present in *Schistosoma mansoni* and *Clonorchis sinensis*. *F. gigantica* cathepsin B FgCB1-7 form a separate clade in phylogenetic analysis and are thought to fulfill specialized roles in infection and/or nutrition processes. FgCB8 was an acidic protease with an ability to undergo auto-activation at pH 5.0 and to sustain maximal enzymatic activity at this pH value for three hours with 50% activity remaining after 24 hours. FgCB8 mRNA was detected from parasite metacercarial to mature stage as well as cecal epithelium of the latter stage. Cross-reactivity of anti-recombinant (r)FgCB8 antiserum to rFgCB5 and human cathepsin B was not detected and immune sera of experimentally-infected rabbits reacted with rFgCB8. Due to a high sequence conservation in trematodes, FgCB8 is not recommended for application as species-specific diagnostic tool.

Keywords: *Fasciola gigantica*, auto-activation, cathepsin B, cysteine protease, proteolytic activity, Trematoda

INTRODUCTION

The peptidase family C1 (papain family), specifically the subfamily A cathepsin L and cathepsin B, are important for infection, protection and nutritional processes in liver flukes, genus *Fasciola* (Dalton *et al*, 2003; Smooker *et al*, 2010). Gene duplication events followed by sequence divergence have created approximately ten cathepsin L and B isoforms each, some of which are highly similar in their sequences making specific detection of distinct members a difficult task (Robinson *et al*, 2008, Robinson *et al*, 2009). In the case of cathepsin B, two clades have been observed in phylogenetic analysis of their amino acid sequences, namely, a clade with members FgCB1-3 and a clade with FgCB4-7 having 63.7-78.5 and 76.7-87.9% identity respectively (Siricoon *et al*, 2015).

FgCB2 and FgCB3 are thought to fulfill an important role in the early infection

Correspondence: Rudi Grams, Graduate Program in Biomedical Sciences, Faculty of Allied Health Sciences, Thammasat University, Phaholyothin Road, Amphoe Klong Luang, Pathumthani 12121, Thailand.
Tel: +66 (0) 2986 9213 ext. 7241;
Fax: +66 (0) 2516 5379
E-mail: rgrams@tu.ac.th

process as demonstrated for their orthologs in *F. hepatica* through mRNA knockdown and inhibitor treatment experiments (McGonigle *et al*, 2008; Beckham *et al*, 2009). These two proteases are stored in metacercariae and released in the newly excysted juveniles (NEJ), but expressions of their genes are downregulated in early juvenile and isotype-specific antibodies detect the proteins only at week-2 post-infection in experimentally infected mice (Sethadavit *et al*, 2009; Chantree *et al*, 2012; Chantree *et al*, 2013). On the other hand, FgCB4, 5, and 7 are not detected in newly excysted juveniles but found in all later stages (Siricoon *et al*., 2015). These isoforms may be adapted for processing of distinct host proteins. Specifically, FgCB5 is the only cathepsin B in *F. gigantica* with a demonstrated carboxy dipeptidyl peptidase activity (Siricoon *et al*, 2015). FgCB1 transcripts are detected in metacercariae, juveniles and mature worms (Meemon *et al*, 2004). *fgcb6* is a transcribed pseudogene carrying two nonsense mutations (Siricoon *et al*, 2015). Pseudogenes may arise following gene duplication because identical gene products are unlikely to provide any advantage and there is no selective pressure to keep both copies functional.

Thus, not only the coding sequences of duplicated cathepsin B genes undergo evolutionary changes to yield proteases with distinct properties, *eg* pH optimum and substrate preferences, but also their regulatory sequences change leading to expression at specific developmental stages. Furthermore, FgCB2 and FgCB3 need to carry signatures for secretion that seem to be lacking in other isoforms. Consequently, analysis of distinct cathepsin B isoforms in *Fasciola* should not only be undertaken to understand their biochemical properties but also to gather information on temporal expression patterns and antigenic properties, which can be important for application of these proteases as diagnostic tools (Tran *et al*, 2019).

A new *F. gigantica* cathepsin B isoform FgCB8 was discovered in a screen of genome data of the genus *Fasciola* available in GenBank as detailed in Materials and Methods section. FgCB8 is considerably different from FgCB1-7 sharing <50% sequence identity to any of them. The study describes molecular and biochemical properties of FgCB8. The findings should assist in the development of diagnosis, vaccines and drugs against liver fluke infection.

MATERIALS AND METHODS

Collection of parasites

Metacercariae (MC), immature (2, 4, 6 weeks old) and mature *F. gigantica* were collected as previously described (Siricoon *et al*, 2012).

Use of experimental animals was approved by Thammasat University Animals Ethics Committee (protocol no. 011/2560).

Molecular cloning and sequence analysis

BLAST searches of *F. hepatica* genomic DNA sequence data with known cathepsin B sequences revealed a DNA sequence of a previously uncharacterized cathepsin B isoform (BN1106_s373B000290 LN627390 *F. hepatica* genome assembly Fhepatica_v1, scaffold scaffold373, whole genome shotgun sequence). DNA sequence of the orthologous gene of *F. gigantica* was found spread across three genomic scaffold clones (GenBank accession nos. MKHB01086990, MKHB01038472 and MKHB01131492). Primers were designed (5′-ATGAGCTTACTGATCTCCAG-3′ and

5′-TTATTGGGGTAATTTTGGCA-3′) to isolate a 1038-bp coding sequence by RT-PCR of total RNA of mature *F. gigantica* (data not shown) and subsequently sequenced (1st BASE Sequencing Asia, Selangor, Malaysia).

Sequences of human cathepsin B (UniProt: P07858) and *F. gigantica* cathepsins B1 to B5 and B7 (UniProt: Q86MW8, Q86MW7, Q86MW6, A0A096ZM60, A0A0U4HGR1, and A0A140F1N7) were culled from NCBI protein database. The encoded sequence of *fgcb6* pseudogene (GenBank: KT781073) was obtained using EMBOSS transeq. Clustal Omega 1.2.0 (Sievers *et al*, 2011) and MUSCLE (Edgar, 2004) were employed to carry out multiple alignments. TEXshade (Beitz, 2000), Dendroscope (Huson *et al*, 2007), EMBOSS 6.6.0 (Rice *et al*, 2000), SignalP 4.1 (Petersen *et al*, 2011), PROSITE database (Sigrist *et al*, 2013) and PhyML version 3.3.20180214 (Guindon *et al*, 2010) were used for formatting alignments, phylogenetic tree construction, sequence editing, sequence analysis, prediction of a signal peptide in FgCB8, detection of potential glycosylation sites, and calculating a maximum likelihood-based tree of the aligned cathepsin B sequences (phyml-d aa-m LG-b-4-v 0.0-c 4-a e-f m-i FgCBs.phy), respectively.

Heterologous expression of FgCB8 in *Escherichia coli* and production of anti-recombinant (r)FgCB8 antiserum

A 765 bp fragment of *fgCB8* cDNA [codon 92 (TTA) to codon 345 (TAA)] was amplified by PCR (primers 5′-*GGATCC*CTTACCGAAAGAGTTTGACG-3′ and 5′-*CTGCAG*TTATTGGGGTAATTTTGG-3′ where sequence in italics indicate *Bam*HI and *Pst*I restriction site respectively) for subsequent insertion into expression vector pQE30 (QIAGEN, Hilden, Germany). Amplicon was inserted into the pGEM-T Easy vector (Promega, Madison, WI) and used for transformation of *E. coli* XL1-Blue. Plasmid DNA extracted from transformants was sent for sequence determination of the insert (1st BASE Sequencing Asia) prior to construction of recombinant pQE30-FgCB8 and heterologous expression in *E. coli* M15. Recombinant (r)FgCB8 was purified by Ni-NTA affinity chromatography as previously described (Chaibangyang *et al*, 2017) and used for production of antisera in three ICR mice as previously described (Chaibangyang *et al*, 2017).

Detection of *fgcb8* transcripts in *F. gigantica*

RNA from metacercariae, 2-, 4-, 6-week-old juveniles and mature *F. gigantica* was extracted using Invitrogen TRIzol extraction kit (Life Technologies, Carlsbad, CA) and treated with DNase (Promega) (1 μg of each stage) at 37°C for 30 minutes and then DNase was inactivated by heating at 65°C for 10 minutes in the presence of 20 mM EGTA. DNase-treated RNA was reverse transcribed using RevertAid™ M-MuLV reverse transcriptase (Thermo Scientific, Vilnius, Lithuania) and the reverse primer described above for 1 hour at 42°C. First strand cDNA was used to amplify the 1038-bp coding sequence of FgCB8 using the primers described above. RT-PCR amplicon was separated by 0.7% agarose gel-electrophoresis, stained with Invitrogen SYBR safe DNA gel stain (Life Technologies, Carlsbad, CA) and visualized under UV illumination. *Fasciola* tubulin (GenBank: AM933581) mRNA was used as positive control and amplified using primers 5′-TGAAGCCTGGGCTCGTTTGGACCA-CAA-3′ and 5′-TTAGTATTCTTCACCCTC-GCCTTCACC-3′ to yield a 200-bp fragment.

Parasite antigen preparation

Crude worm (CW) extract of adult *F. gigantica* was prepared by homogenization

of adult parasites in PBS, pH 7.4 containing 1% Triton X-114, 2 mM PMSF, 5 mM, and 1 mM EDTA). The cleared supernatant was collected after centrifugation at 12,000 × g for 30 minutes at 4°C as previously described (Siricoon *et al*, 2012). Protein concentration was determined by a Bradford assay (Bio-Rad, Richmond, CA) and the CW extract was stored at -20°C until used.

Transformation of *Pichia pastoris* and expression of soluble rFgCB8

A FgCB8 cDNA fragment starting at bp 53 and ending at bp 335 of the coding sequence was amplified by PCR using forward primer 5′-*CTGCAG*CTGAAAATGAACG-3′ containing a natural terminal *Pst*I recognition site and reverse primer 5′-*TCTAGA*CCTTGGGGTAATTTTGGCAAGC-3′ that introduced a terminal *Xba*I endonuclease restriction site (italics). The fragment was inserted into the *Pichia pastoris* expression vector pPICZαB (Invitrogen) to facilitate secretion of the expressed recombinant protein. Transformants were identified by PCR using the primers detailed above and used to express soluble rFgCB8 as previously described (Siricoon *et al*, 2015).

Deglycosylation of yeast-expressed rFgCB8

PNGase-F (New England Biolabs, Ipswich, MA) was used for deglycosylation of rFgCB8 under denaturing conditions following the supplied protocol. In brief, rFgCB8 (15 μg) was mixed with glycoprotein denaturing buffer (0.5% SDS containing 40 mM dithiothreitol) and incubated at 100°C for 10 minutes, chilled on ice, centrifuged 10 seconds, suspended in 50 mM sodium phosphate pH 7.5 containing 1% NP-40 followed by 1 μl (500 units) of PNGase F, and incubated at 37°C for 1 hour. Extent of deglycosylation was assessed by SDS-PAGE and western blotting as described below.

SDS-PAGE, Western blot analysis and ELISA

Parasite CW extract (50 μg), bacterial expressed rFgCB8 (200 ng), deglycosylated yeast expressed rFgCB8 (200 ng) and human cathepsin B from liver (200 ng) (Sigma Aldrich, St Louis, MI) were subjected to 12.5% SDS-PAGE and electro-transferred to nitrocellulose membrane (Amersham Biosciences GE Healthcare Life Sciences, PA) using a Fastblot B33 instrument (Whatman Biometra, Göttingen, Germany) for 1 hour at 60 mA. Membrane was treated with murine primary preimmune and anti-rFgCB8 antisera (1:3,000 dilution), and rabbit primary anti-human cathepsin B (1:2,000 dilution) (Athens Research & Technology, Athens, GA) at 4°C, overnight followed by polyclonal goat alkaline phosphatase-conjugated anti-mouse immunoglobulin (Sigma Aldrich, St Louis, MI) and polyclonal goat alkaline phosphatase-conjugated anti-rabbit immunoglobulin (Dako, Copenhagen, Denmark) secondary antibodies (1:30,000 and 1:1000 dilution respectively) at room temperature for 60 minutes. Immunoreactive protein bands were visualized using chromogenic substrates 5-bromo-4-chloro-3-indolyl-phosphate (BCIP) in conjunction with nitro blue tetrazolium (NBT) (Ameresco LLC, Solon, OH) and electronic images were taken on a Canon scanner (CanoScanLiDE700F). The above listed human cathepsin B, anti-human cathepsin B antiserum and rFgCB5 (our lab, Siricoon *et al*., 2015) were used for comparison in western and ELISA.

Pre-infection sera, 2-, 8-, and 12-week post-infection sera (1:100 dilution) of *F. gigantica*-infected rabbits (Chunchob *et al*, 2010) were employed to detect bacteria-

expressed rFgCB8 and yeast-expressed rFgCB8 (100 ng) by ELISA as previously described (Siricoon *et al*, 2012). Polyclonal goat horseradish peroxidase-conjugated anti-rabbit IgG antibody (1:30,000) (Dako, Copenhagen, Denmark) was used as secondary antibody. Immunoreactive antigen was detected with chromogenic substrate o-phenylenediamine dihydrochloride (OPD, Sigma, MO, USA) at 492 nm (Anthos model 2020, Anthos Labtec Instruments, Eugendorf, Austria). Measurements were carried in duplicate and two-tailed unpaired t-tests assuming the same SD with a significant difference at p-value <0.05 between sample groups employing Prism 8 (www.graphpad.com).

Immunohistochemical detection of FgCB8 in adult *F. gigantica*

Immunohistochemical detection of FgCB8 in adult *F. gigantica* was performed as previously described (Siricoon *et al*, 2012). In short, deparaffinized tissue sections (10 μm) were treated with murine anti-rFgCB8 antiserum (1:500 dilution) overnight at 4°C followed by goat biotinylated anti-mouse secondary antibody (Dako, Copenhagen, Denmark) (1:200 dilution) for 1 hour at ambient temperature. Then washed sections were incubated with avidin–biotin peroxidase (Pierce ABC Staining Kit, Pierce Biotechnology Inc, Thermo Fisher Scientific, Rockford, IL) at ambient temperature for 30 minutes, washed and colorimetric detection performed using aminoethyl carbazole (AEC) substrate solution (Invitrogen). Murine preimmune serum (1:500 dilution) was used as negative control. Sections were mounted in 10% (v/v) glycerol in 10 mM phosphate-buffered saline pH 7.2 and micrographs were examined under an Olympus BX51 microscope (Tokyo, Japan) at 10x magnification and recorded with a PixelLink PL-A686C digital camera (Pixelink, Ottawa, Canada).

Autoprocessing of soluble rFgCB8

Purified soluble rFgCB8 (1.6 μg) was incubated in AMT buffer (100 mM sodium acetate, 100 mM 2-(N-morpholino) ethanesulfonic acid (MES), 200 mM Tris-HCl, and 4 mM EDTA) containing 50 μg/ml dextran sulfate (DS 500K) and 10 mM DTT at pH values of 4.0, 4.5, 5.0, and 5.5 for 0, 0.5, 1.5, 4, and 19 hour(s). Samples were deglycosylated as described above and separated by 12.5% SDS-PAGE under reducing conditions and visualized by Coomassie Blue G-250 staining. Experiments were performed in duplicate.

Activity of treated rFgCB8 (800 ng) was measured by incubating in AMT buffer containing 50 μg/ml DS 500K, 200 mM NaCl and 10 mM DTT at pH values of 3.0, 4.0, 4.5, 5.0, 5.5, 6.0, 6.5, 7.0, 7.5, and 8.0 at 37°C for 2 hours, followed by addition of 10 μm Z-Phe-Arg-AMC (Sigma Aldrich, St Louis, MI). Fluorescence released by substrate hydrolysis was measured at excitation and emission wavelengths of 355 nm and 460 nm, respectively using a Varioskan Flash spectral scanning multimode reader (Thermo Scientific, MA) as previously described (Tarasuk *et al*, 2009).

Stability of rFgCB8 at optimal pH activity was analyzed by incubation of the protein (800 ng) in AMT buffer pH 5.0 containing 50 μg/ml DS 500K, 200 mM NaCl and 10 mM DTT at 37°C for 1, 2, 3, 4, 5 h and 24 hour(s). Enzymatic activity was then measured using 10 μm Z-Phe-Arg-AMC (Sigma Aldrich, St Louis, MI) as described above.

Activity of rFgCB8 against native host protein was assessed by incubation of the protease (500 ng and 1000 ng) with 500 μg/ml of either bovine serum albumin or

human IgG (Sigma Aldrich, St Louis, MI) in AMT buffer containing 50 μg/ml DS 500K, 200 mM NaCl, 10 mM DTT, pH 4.5, 5.0 and 5.5 at 37°C for 4 hours. Reactions were terminated by adding SDS-PAGE sample buffer under reducing condition and heating at 95°C for 5 minutes. Samples then were analyzed by 12.5% SDS-PAGE under reducing condition and stained with Coomassie Brilliant Blue R-250.

RESULTS

Molecular cloning and sequence analysis of *F. gigantica* cathepsin B8

A cDNA encoding an uncharacterized cathepsin B of *F. gigantica* (FgCB8, GenBank: MK472052) was isolated from total RNA of mature parasites based on a bioinformatics analysis of genomic DNA sequence data (Fig 1). FgCB8 showed 43-45% sequence identity to FgCB1-7, while FgCB1-7 had 62-88% identity among each other (Siricoon *et al*, 2015). FgCB8 showed 49% sequence identity to human cathepsin B (UniProt: P07858). This is also evident in the phylogenetic analysis, revealing FgCB8 well separated from FgCB1-7 (Fig 2). Among other trematodes, FgCB8 appeared to be an ortholog of *Schistosoma mansoni* CB2 with 63.7% identity of the preproprotein sequence (75% identity of the mature protein sequence) (Caffrey *et al*, 2002). Extending the phylogenetic analysis to other cathepsin B forms of *S. mansoni* and *Clonorchis sinensis* (see legend of Fig 2 for details) indicated FgCB8 was a highly conserved isoform among these trematodes while FgCB1-7 evolved independently by gene duplication events within the lineage of genus *Fasciola*. Prepro-FgCB8 contains 345 amino acids, with cleavage of the putative signal peptide leading to a 325 aa sequence with a calculated MW of 36.3 kDa. Activation of the proprotein through cleavage of the putative inhibitory proregion would yield a mature 28.1 kDa cysteine protease (254 amino acids). Sequence comparison with human cathepsin B revealed FgCB8 having full conservation of active site residues and cysteine residues forming disulfide bonds (Fig 1). Both human cathepsin B and FgCB5 carry the two adjacent histidine residues in the occluding loop that are required for peptidase activity of the enzyme (Krupa *et al*, 2002; Siricoon *et al*, 2015) while FgCB8 and the other FgCB isoforms have only one histidine residue. PROSITE predicted two putative N-glycosylation sites, the first in the pro-domain at aa position 46-49 (NTTW) and the second in the active protein at aa position 227-230 (NITY).

Expression of *fgcb8* transcripts *in F. gigantica*

RT-PCR analysis demonstrated the presence of *fgcb8* transcripts in metacercariae, in 2-, 4-, and 6-week-old immature and in mature *F. gigantica* with increased transcripts in the adult stage (Fig 3). In support of a previously reported absence of *fgcb5* transcripts in newly excysted juveniles (NEJ) (Siricoon *et al*, 2015) RT-PCR did also not yield *fgcb5* transcripts from metacercariae (data not shown).

Expression of rFgCB8 and immunodetection of FgCB8

The mature form of FgCB8 was heterologously expressed in *E. coli* M15 as an insoluble N-terminal His-tagged recombinant protein. The insoluble protein was purified under denaturing conditions by Ni-NTA affinity chromatography (Fig 4A) and used to produce anti-rFgCB8 antisera in mice. A proprotein form of FgCB8 was also heterologously expressed in *P. pastoris* and secreted as a C-terminal His-tagged protein, which was purified from the medium under

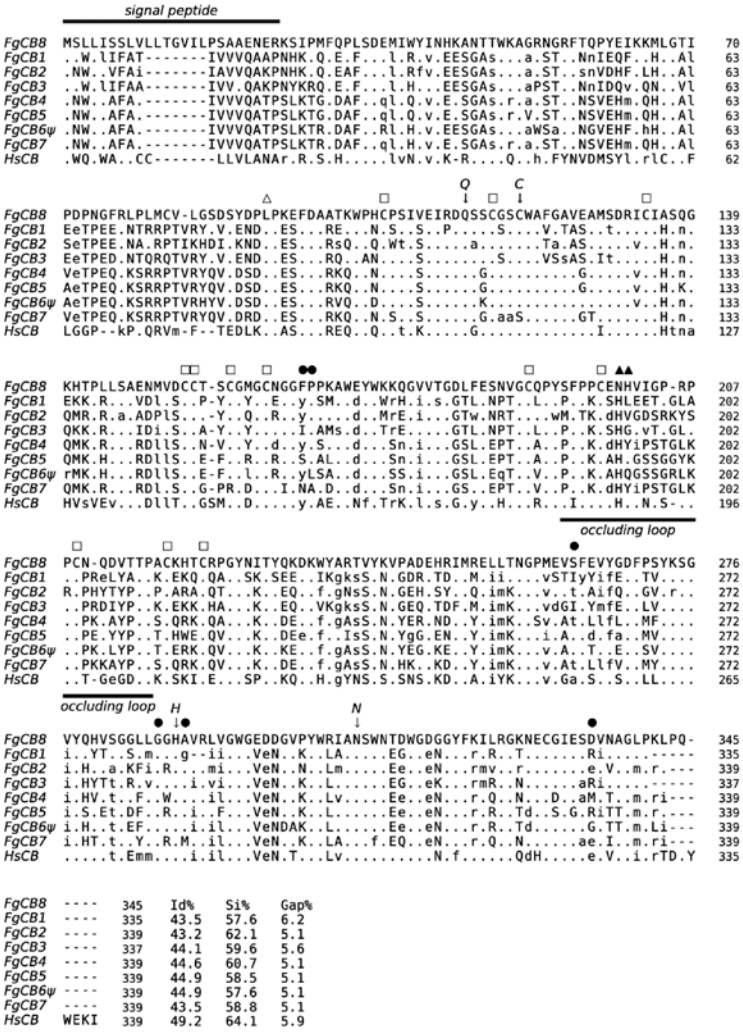

Fig 1-Comparison of *Fasciola gigantica* cathepsin B1 to B8 (FgCB1-8) and human cathepsin B (HsCB) sequences)

FgCB1-7 and HsCB accession numbers are listed in Materials and Methods, the here described FgCB8 was assigned GenBank accession QHT72880. FgCB6 is a conceptual product of a transcribed pseudogene (Ψ) with the nonsense changes at residues 128 and 206 (*) (Siricoon et al, 2015). Features indicated are based on annotations in HsCB. Signal peptide, residues 1-17 in HsCB and 1-15 in FgCB isoforms; N-terminus of mature protease, residue 80 in HsCB and 86 for FgCB isoforms (Δ); active site residues, Q102, C108, H278, and N298; S2 subsite residues 154, 155, 252, 276, 279, and 324; (o), cysteine residues forming disulfide bonds (PDB: 1PBH) (Musil *et al*, 1991); occluding loop, residues 183-205 with H189 and H190 essential for C-terminal exopeptidase activity. Pairwise sequence % identity, % similarity, % gap values of FgCB1-7 and HsCB with FgCB8 are shown at the end of the alignment, calculated using EMBOSS aligncopypair with standard settings. Amino acids identical, similar and non-conserved to FgCB8 are indicated by dots, lower and upper case characters, respectively and gaps introduced for best alignment are indicated by dashes (-).

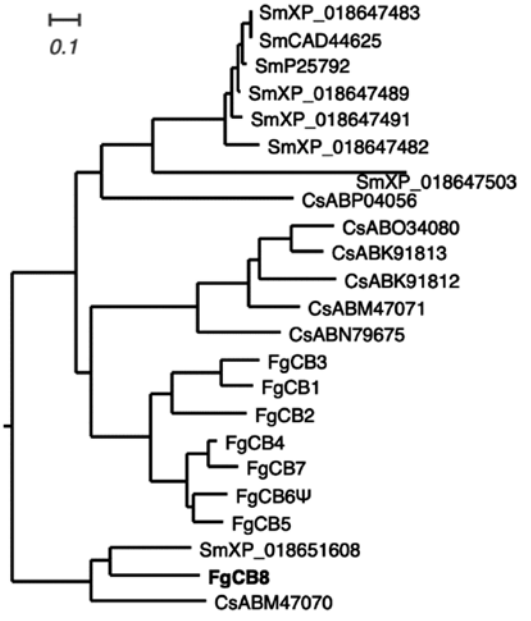

Fig 2-Maximum likelihood phylogenetic tree of cathepsin B sequences of *Fasciola gigantica* (Fg), *Schistosoma mansoni* (Sm) and *Clonorchis sinensis* (Cs) calculated using PhyML

Human cathepsin B was used as outgroup. Log likelihood was -9353.46613. Accession numbers from NCBI protein database are shown for *S. mansoni*, and *C. sinensis*. FgCB1-7 accession numbers are listed in Materials and Methods, FgCB8 was assigned GenBank accession QHT72880.

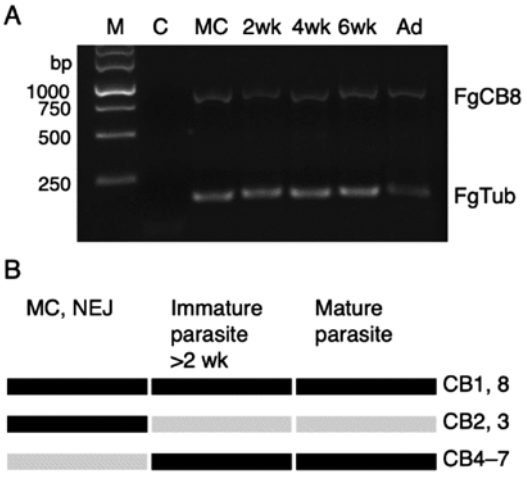

Fig 3-Analysis of expression of FgCB8 at the RNA level by RT-PCR and schematic comparison of RNA expression patterns of FgCB1 to FgCB8 during development

Panel A: Agarose gel showing the resolved *Fasciola gigantica* cathepsin B8 RT-PCR amplicon. First strand cDNA was synthesized from total RNA of metacercariae (MC), 2, 4, 6-week-old juveniles (2wk, 4wk, 6wk) and adults (Ad) and used to amplify a 1,038-bp coding sequence of *fgcb8*. *F. gigantica tubulin* transcript (FgTub) was used as internal control. Panel B: Schematic overview of *F. gigantica* cathepsin B gene expression during development (NEJ: newly excysted juvenile).

native condition by Ni-NTA affinity chromatography (Fig 4B) and used for subsequent functional studies. Yeast-expressed rFgCB8 migrated as a smear of 45-66 kDa due to glycosylation, and following deglycosylation with PNGase-F treatment rFgCB8 demonstrated a sharp band of the expected molecular weight (Fig 4B, lane 2).

Western blotting using murine anti-rFgCB8 antiserum demonstrated reaction with bacteria- and yeast-expressed rFgCB8 and with the mature form of FgCB8 in crude *F. gigantica* worm extract (Fig 5A). *F. gigantica* excretory-secretory products showed a negative western blotting result (data not shown). Murine anti-rFgCB8 antiserum did not cross-detect human cathepsin B (Fig 5A, lane 4) or rFgCB5 (Fig 5B, lane 3). Likewise, anti-human cathepsin B antiserum did not cross-detect rFgCB8 (Fig 5C).

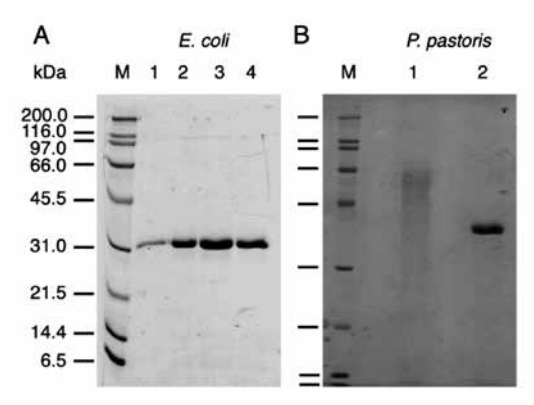

Fig 4-SDS-PAGE of purified recombinant *Fasciola gigantica* cathepsin B8 (rFgCB8) expressed in *Escherichia coli* and *Pichia pastoris*

Panel A: The mature form of FgCB8 was heterologously expressed in *E. coli* M15 as an insoluble N-terminal His-tagged protein and purified under denaturing condition by Ni-NTA affinity chromatography. Elution fractions 1-4 are shown. Panel B: The proform of FbCB8 was heterologously expressed in *Pichia pastoris* as a C-terminal His-tagged protein and was subsequently purified using Ni-NTA affinity chromatography. Lane 1, glycosylated rFgCB8; lane 2, PNGase-F-treated (de-glycosylated) rFgCB8; lane M, protein size markers.

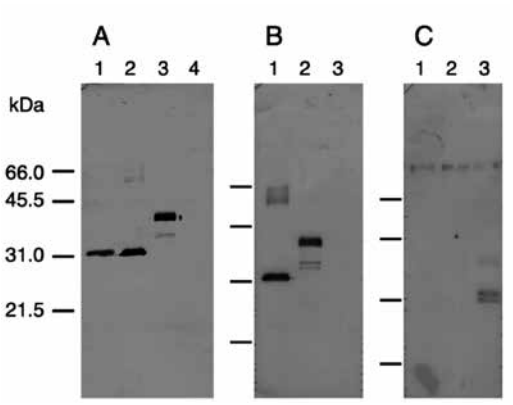

Fig 5-Western blot detection of recombinant *Fasciola gigantica* cathepsin B8 (rFgCB8), rFgCB5, crude worm (CW) extract, and human cathepsin B

Panel A: Proteins were separated by 12.5% SDS-PAGE and subjected to western blotting. Murine anti-rFgCB8 antiserum and goat alkaline phosphatase-conjugated anti-mouse IgG secondary antibody were used for detection. Lane 1: CW extract; Lane 2: bacteria-expressed rFgCB8; Lane 3: deglycosylated yeast-expressed rFgCB8; Lane 4: human cathepsin B.

Panel B: Proteins were subjected to western blotting and detection as described above. Lane 1: bacteria-expressed rFgCB8; Lane 2: deglycosylated yeast-expressed rFgCB8; Lane 3: yeast-expressed rFgCB5.

Panel C. Proteins were subjected to western blotting as described above using rabbit anti-human cathepsin B antiserum and goat alkaline phosphatase-conjugated anti-rabbit IgG for subsequent detection. Lane 1: bacteria-expressed rFgCB8; Lane 2: deglycosylated yeast-expressed rFgCB8; Lane 3: human cathepsin B.

Histoimmunostaining of tissue sections prepared from adult *F. gigantica* with murine anti-rFgCB8 antiserum demonstrated localization of FgCB8 in the gastrodermis (Fig 6).

FgCB8 reactivity to immunosera of rabbits experimentally infected with *F. gigantica*

ELISA using rabbit immune sera (n = 2) from experimental *F. gigantica* infection (Chunchob *et al*, 2010) showed immuoreactivity with bacterial- and yeast-expressed rFgCB8 at week-8 and -12 post-infection (Fig 7). Pre-infection and week-2 post-infection sera were negative. Reactivity towards correctly folded yeast-expressed rFgCB8 was higher than that against bacterial-expressed denatured form.

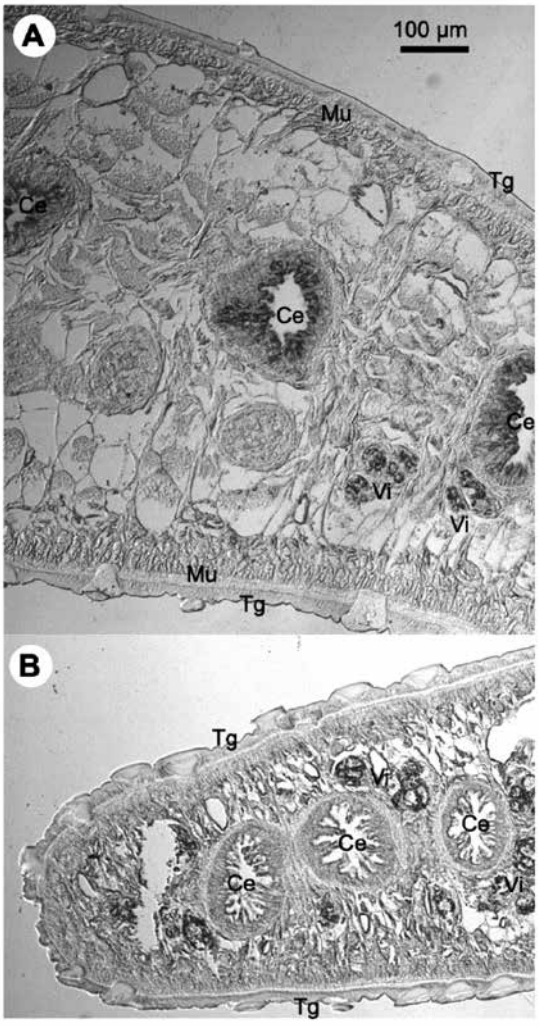

Fig 6-Immunohistochemical detection of *Fasciola gigantica* cathepsin B8 (FgCB8)

Deparaffinized tissue sections of adult specimen (10 µm) were treated with murine anti-recombinant (r)FgCB8 antiserum followed by goat biotinylated anti-mouse secondary antibody, then with avidin–biotin peroxidase and immunoreactive areas visualized with aminoethyl carbazole. Panel A: Treatment with murine anti-rFgCB8 antiserum. Panel B: Treatment with murine preimmune serum.

Ce: cecum; Mu: muscle layer; Tg: tegument layer; Vi: vitellaria

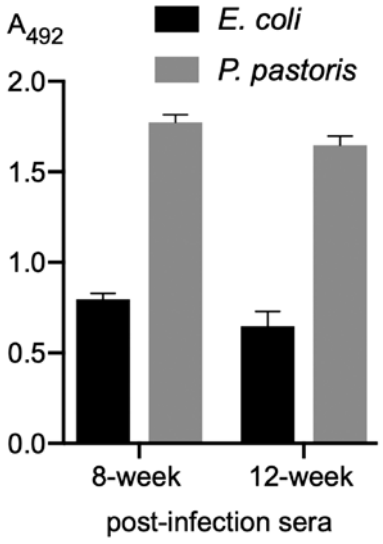

Fig 7-Indirect ELISA of bacteria- and yeast-expressed *Fasciola gigantica* recombinant cathepsin B 8 (rFgCB8)

Bacteria- and yeast-expressed rFgCB8 (100 ng) coated on bottom of wells were treated with rabbit pre-infection serum, and 8-, and 12-week post-*F. gigantica* infection immune sera followed by goat horseradish peroxidase-conjugated anti-rabbit IgG secondary antibody. Immunoreactive antigen was detected with o-phenylenediamine dihydrochloride at 492 nm.

Autoprocessing of soluble yeast-expressed rFgCB8

Glycosylated rFgCB8 heterologously expressed and secreted from *P. pastoris* was used for functional characterization of the protease. Autoprocessing of the proenzyme analyzed at pH 4.0-5.5 in the presence of dextran sulfate demonstrated the enzyme underwent optimal processing to the mature form of ~28 kDa at pH 5.0 and 5.5 (Fig. 8). At pH 4.0 and 4.5 FgCB8 was unstable and could not be detected after 19-hours incubation times. The protein was still glycosylated following

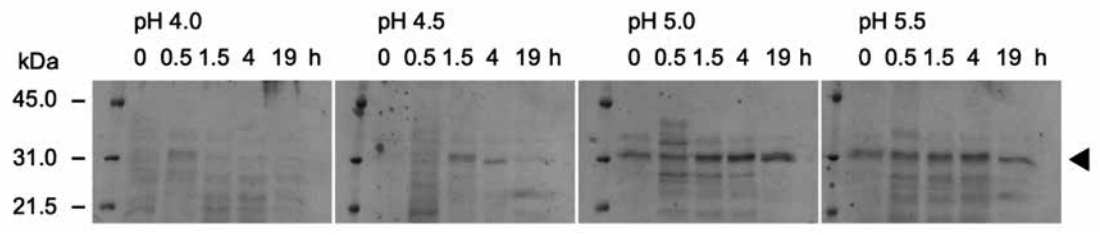

Fig 8- Autoprocessing of yeast-expressed *Fasciola gigantica* recombinant cathepsin B8 (rFgCB8)

Recombinant protein prepared as described in legend to Fig 4 was incubated in the presence of dextran sulfate at the indicated pH values from 0 to 19 hours (h), deglycosylated with PNGase-F, separated by 12.5% SDS-PAGE under reducing condition and visualized by staining with Coomassie Blue G-250. Arrow head indicates processed rFgCB8.

autoprocessing suggesting PROSITE-predicted motif NITY (aa 227 to 230 of prepro-FgCB8) was the glycosylation site in yeast (data not shown).

Activity of soluble yeast-expressed FgCB8

The endopeptidase activity of FgCB8 was analyzed using the substrate Z-Phe-Arg-AMC. In order to determine pH stability, mature FgCB8 was incubated at pH 3.0-8.0 at 37°C for 2 hours. The enzyme showed optimal activity following incubation at pH 4.0-5.0 (Fig 9A). Activity at pH 5.0 was maintained for three hours with 50% activity observed after 24 hours (Fig 9B). FgCB8 was able to digest bovine serum albumin and human IgG at pH 4.5 and 5.5 (Fig 9C).

DISCUSSION

The results demonstrate *F. gigantica* cathepsin B8 belongs to the group of cysteine proteases localized in cecal epithelium of the parasite as has been shown for FgCB2, FgCB3, and FgCB5 (Sethadavit *et al*, 2009; Chantree *et al*, 2012; Siricoon *et al*, 2015). ProFgCB8 was able to autoactivate at acidic conditions comparable to human procathepsin B (Rozman *et al*, 1999). Considering its activity profile it should act as a lysosomal hydrolase and not in an environment with a neutral or higher pH condition. In this regard, it is similar to FgCB5 (Siricoon *et al*, 2015) and different from FgCB2 and FgCB3 that are secreted by newly excysted juveniles (Sethadavit *et al*, 2009; Chantree *et al*, 2012). The high sequence divergence of FgCB8 from to the other cathepsin B isoforms in *F. gigantica* and the phylogenetic analysis demonstrated the former represented an ancestral isoform (ortholog) that has been conserved in trematodes. Possibly, the protease has a house-keeping role as it was found to be present from NEJ to adult stage. RT-PCR analysis showed *fgcb8* expression similar to that reported for *F. hepatica* CB6, 8, 9, and 10, which exhibited intermediate levels of expression in metacercaria and NEJ with higher levels in the later stages (Cwiklinski *et al*, 2015). Thus, FgCB8 and FgCB1 (Meemon *et al*., 2004) are always present throughout the developmental stages, while FgCB2 and 3 are highly abundant in metacercaria and NEJ followed by downregulation in the early immature parasite (Meemon *et al*, 2004; Sethadavit *et al*, 2009; Chantree *et al*, 2012); and expression of FgCB4, 5 and 7 commences after excystation of the larva (Siricoon *et al*, 2015). If FgCB8

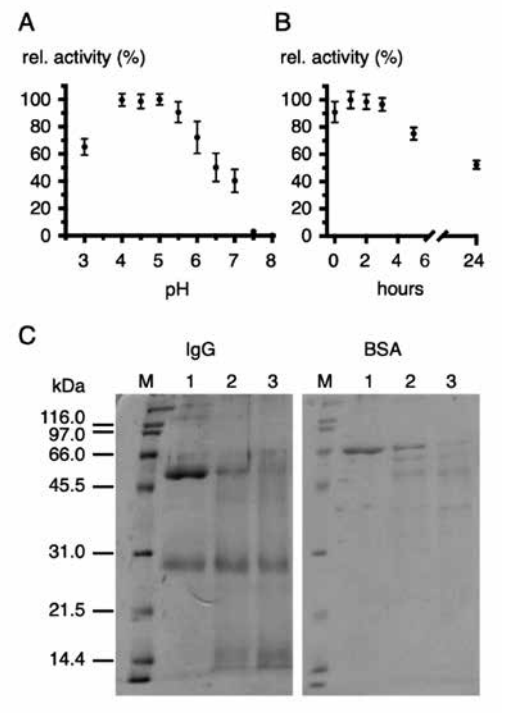

Fig 9-Activity assays of yeast-expressed *Fasciola gigantica* recombinant cathepsin (rFgCB8)

Activity of rFgCB8 (800 ng) prepared as described in legend to Fig 4 was measured using 10 μm Z-Phe-Arg-AMC as substrate. Fluorescence released by substrate hydrolysis was measured at excitation and emission wavelengths of 355 nm and 460 nm. Panel A: Relative activity following incubation at indicated pH values at 37°C for 2 hours. Panel B: Relative activity following incubation at pH 5.0 for the indicated durations at 37°C. Panel C: 12% SDS-PAGE of human IgG and bovine serum albumin (BSA) (500 μg/ml each) following incubation without rFgCB8 (Lane 1), with 500 ng of rFgCB8 (Lane 2), and 1,000 ng of rFgCB8 (Lane 3) at pH 5.0 for 4 hours at 37°C. Gels were stained with Coomassie Blue G-250.

represents the prototype cathepsin B then the other FgCB isoforms should represent evolutionary products of subsequent gene duplication events and may have been maintained because they provide specific advantages in infection and nutritional processes comparable to cathepsin L family in *Fasciola* spp (Irving *et al*, 2003). The phylogenetic tree suggests most of these cathepsin B duplication events happened independently in the lineages of trematode genera leading in each case to a clade of closely related isoforms.

Sequence divergence of FgCB8 was presumably the reason why the generated antiserum showed no cross-reactivity to rFgCB5 and human cathepsin B. In a previous study, antiserum specificity was reported for FgCB2 and FgCB3 that have 63.7% sequence identity (Chantree *et al*, 2013). Nevertheless, cross-reactivity of anti-rFgCB8 antiserum could be expected to the orthologous isoform in other trematodes, e.g. *S. mansoni* CB2 (XP_018651608) with 75% identity and *C. sinensis* CB1 with 69% identity (ABM47070) in the amino acid sequence of the mature protein. This makes FgCB8 less suited as a species-specific diagnostic tool and other isoforms should be selected for this purpose.

FgCB8 was found located in gastric ceca as previously reported for FgCB2, 3, and 5 (Sethadavit *et al*, 2009, Chantree *et al*, 2012, Siricoon *et al*, 2015). SmCB2, the orthologous protein of FgCB8 in *S. mansoni*, was reported as a parenchymal and tegumental protein, which is very unusual and either points to a unique role in this parasite or a problem with the generated antiserum (Caffrey *et al*, 2002).

FgCB8 was detected by sera of *F. gigantica*-infected rabbits as previously reported for FgCB5 (Siricoon *et al*, 2015). This is surprising because FgCB8 was not found in parasite excretory-secretory product and is presumed to work as a lysosomal hydrolase at acidic pH. Still, there must be sufficient leakage

of FgCB8, eventually by lysosomal exocytosis (Saftig and Klumperman, 2009), to stimulate an immune response during infection. Alternatively, antibodies against conserved epitopes from other cysteine proteases cross-react with FgCB8. The results clearly show functioning soluble yeast-expressed FgCB8 was detected at higher sensitivity compared to denatured bacteria-expressed FgCB8. This suggests that the major epitopes are conformational, thereby making it difficult to design FgCB8 peptides with not only high specificity but also sensitivity for application as diagnostic tools. These difficulties have also been observed for FgCB5 peptides (Tran *et al*, 2019). Short peptides might fold differently outside the context of the complete protein due to the absence of neighboring sidechains that affect their folding. It should be noted application of rFgCB8 as a diagnostic tool in natural infections would require a non-glycosylated protein. Unlike sera of naïve experimental animals, many immune sera will cross-detect carbohydrates added during protein processing in yeast and result in false positives, and also access to epitopes on the protein might be blocked by glycosylation (Lee *et al*, 2019).

In conclusion, FgCB8 represents the ancestral cathepsin B in *Fasciola gigantica* and likely in other trematode species, too. It can be thought of as an always available, housekeeping representative of this protease family that has a role as lysosomal hydrolase in the intestinal epithelium of the parasite. Sequence conservation makes it a less suitable target for species-specific application.

ACKNOWLEDGEMENTS

Thwet Oo Lwin was supported by a PhD scholarship for international students from Thammasat University (2015-2017). The work was supported by a grant from the Faculty of Allied Health Sciences, Thammasat University (AHS 1/2561).

REFERENCES

Beckham SA, Piedrafita D, Phillips CI, *et al*. A major cathepsin B protease from the liver fluke *Fasciola hepatica* has atypical active site features and a potential role in the digestive tract of newly excysted juvenile parasites. *Int J Biochem Cell Biol* 2009; 41: 1601-12.

Beitz E. TEXshade: shading and labeling of multiple sequence alignments using LATEX2 epsilon. *Bioinformatics* 2000; 16: 135-9.

Caffrey CR, Salter JP, Lucas KD, *et al*. SmCB2, a novel tegumental cathepsin B from adult *Schistosoma mansoni*. *Mol Biochem Parasitol* 2002; 121: 49-61.

Chaibangyang W, Geadkaew-Krenc A, Vichasri-Grams S, Tesana S, Grams R. Molecular and biochemical characterization of *Opisthorchis viverrini* calreticulin. *Korean J Parasitol* 2017; 55: 643-52.

Chantree P, Phatsara M, Meemon K, *et al*. Vaccine potential of recombinant cathepsin B against *Fasciola gigantica*. *Exp Parasitol* 2013; 135: 102-9.

Chantree P, Wanichanon C, Phatsara M, Meemon K, Sobhon P. Characterization and expression of cathepsin B2 in *Fasciola gigantica*. *Exp Parasitol* 2012; 132: 249-56.

Chunchob S, Grams R, Viyanant V, Smooker PM, Vichasri-Grams S. Comparative analysis of two fatty acid binding proteins from *Fasciola gigantica*. *Parasitology* 2010; 137: 1805-17.

Cwiklinski K, Dalton JP, Dufresne PJ, *et al*. The *Fasciola hepatica* genome: gene duplication and polymorphism reveals adaptation to the host environment and the capacity for rapid evolution. *Genome Biol* 2015; 16: 71.

Dalton JP, Neill SO, Stack C, *et al*. *Fasciola hepatica* cathepsin L-like proteases: biology, function, and potential in the development of first generation liver fluke vaccines. *Int J Parasitol* 2003; 33: 1173-81.

Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 2004; 32: 1792-7.

Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol* 2010; 59: 307-21.

Huson DH, Richter DC, Rausch C, Dezulian T, Franz M, Rupp R. Dendroscope: an interactive viewer for large phylogenetic trees. *BMC Bioinformatics* 2007; 8: 460.

Irving JA, Spithill TW, Pike RN, Whisstock JC, Smooker PM. The evolution of enzyme specificity in *Fasciola* spp. *J Mol Evol* 2003; 57: 1-15.

Krupa JC, Hasnain S, Nagler DK, Menard R, Mort JS. S2′ substrate specificity and the role of His110 and His111 in the exopeptidase activity of human cathepsin B. *Biochem J* 2002; 361: 613-9.

Lee HH, Wang YN, Xia W, et al. Removal of N-linked glycosylation enhances PD-L1 detection and predicts anti-PD-1/PD-L1 therapeutic efficacy. *Cancer Cell* 2019; 36: 168-78.

McGonigle L, Mousley A, Marks NJ, *et al*. The silencing of cysteine proteases in *Fasciola hepatica* newly excysted juveniles using RNA interference reduces gut penetration. *Int J Parasitol* 2008; 38: 149-55.

Meemon K, Grams R, Vichasri-Grams S, *et al*. Molecular cloning and analysis of stage and tissue-specific expression of cathepsin B encoding genes from *Fasciola gigantica*. *Mol Biochem Parasitol* 2004; 136: 1-10.

Musil D, Zucic D, Turk D, *et al*. The refined 2.15 A X-ray crystal structure of human liver cathepsin B: the structural basis for its specificity. EMBO J 1991; 10: 2321-30.

Petersen TN, Brunak S, von Heijne G, Nielsen H. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat Methods* 2011; 8: 785-6.

Rice P, Longden I, Bleasby A. EMBOSS: the European Molecular Biology Open Software Suite. *Trends Genet* 2000; 16: 276-7.

Robinson MW, Dalton JP, Donnelly S. Helminth pathogen cathepsin proteases: it's a family affair. *Trends Biochem Sci* 2008; 33: 601-8.

Robinson MW, Menon R, Donnelly SM, Dalton JP, Ranganathan S. An integrated transcriptomics and proteomics analysis of the secretome of the helminth pathogen *Fasciola hepatica*: proteins associated with invasion and infection of the mammalian host. *Mol Cell Proteomics* 2009; 8: 1891-907.

Rozman J, Stojan J, Kuhelj R, Turk V, Turk B. Autocatalytic processing of recombinant human procathepsin B is a bimolecular process. *FEBS Lett* 1999; 459: 358-62.

Saftig P, Klumperman J. Lysosome biogenesis and lysosomal membrane proteins: trafficking meets function. *Nat Rev Mol Cell Biol* 2009; 10: 623-35.

Sethadavit M, Meemon K, Jardim A, Spithill TW, Sobhon P. Identification, expression and immunolocalization of cathepsin B3, a stage-specific antigen expressed by juvenile *Fasciola gigantica*. *Acta Trop* 2009; 112: 164-73.

Sievers F, Wilm A, Dineen D, *et al*. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* 2011; 7: 539.

Sigrist CJ, de Castro E, Cerutti L, *et al*. New and continuing developments at PROSITE. *Nucleic Acids Res* 2013; 41: D344-7.

Siricoon S, Grams SV, Grams R. Efficient inhibition of cathepsin B by a secreted type 1 cystatin of *Fasciola gigantica*. *Mol Biochem Parasitol* 2012; 186: 126-33.

Siricoon S, Vichasri Grams S, Lertwongvisarn K, Abdullohfakeeyah M, Smooker PM, Grams R. *Fasciola gigantica* cathepsin B5 is an acidic endo- and exopeptidase of the

immature and mature parasite. *Biochimie* 2015; 119: 6-15.

Smooker PM, Jayaraj R, Pike RN, Spithill TW. Cathepsin B proteases of flukes: the key to facilitating parasite control? *Trends Parasitol* 2010; 26: 506-14.

Tarasuk M, Vichasri Grams S, Viyanant V, Grams R. Type I cystatin (stefin) is a major component of *Fasciola gigantica* excretion/secretion product. *Mol Biochem Parasitol* 2009; 167: 60-71.

Tran NTD, Ton Nu PA, Intuyod K, *et al*. Evaluation of a commercial enzyme-linked immunosorbent assay kit and in-house *Fasciola gigantica* cysteine proteinases-based enzyme-linked immunosorbent assays for diagnosis of human fascioliasis. *Am J Trop Med Hyg* 2019; 100: 591-8.