Available online at www.sciencedirect.com
SCIENCE DIRECT
International Journal for Parasitology 35 (2005) 895–910
www.elsevier.com/locate/ijpara

Multiple cathepsin B isoforms in schistosomula of *Trichobilharzia regenti*: identification, characterisation and putative role in migration and nutrition☆

Jan Dvořák[a], Melaine Delcroix[b], Andrea Rossi[c], Václav Vopálenský[d], Martin Pospíšek[d], Miroslava Šedinová[a], Libor Mikeš[a], Mohammed Sajid[b], Andrej Sali[c], James H. McKerrow[b], Petr Horák[a], Conor R. Caffrey[b,*]

[a]*Department of Parasitology, Faculty of Science, Charles University, Viničná 7, CZ 12844 Prague, Czech Republic*
[b]*Department of Pathology, Sandler Center for Basic Research in Parasitic Diseases, University of California San Francisco, Box 0511, HSW 501, San Francisco, CA 94143, USA*
[c]*Departments of Biopharmaceutical Sciences and Pharmaceutical Chemistry, and California Institute for Quantitative Biomedical Research, Box 2240, University of California San Francisco, San Francisco, CA 94143, USA*
[d]*Department of Genetics and Microbiology, Faculty of Science, Charles University, Viničná 5, CZ 12844 Prague, Czech Republic*

Received 7 December 2004; received in revised form 7 February 2005; accepted 21 February 2005

Abstract

Among schistosomatids, *Trichobilharzia regenti*, displays an unusual migration through the peripheral and central nervous system prior to residence in the nasal cavity of the definitive avian host. Migration causes tissue degradation and neuromotor dysfunction both in birds and experimentally infected mice. Although schistosomula have a well-developed gut, the peptidases elaborated that might facilitate nutrition and migration are unknown. This is, in large part, due to the difficulty in isolating large numbers of migrating larvae. We have identified and characterised the major 33 kDa cathepsin B-like cysteine endopeptidase in extracts of migrating schistosomula using fluorogenic peptidyl substrates with high extinction coefficients and irreversible affinity-labels. From first strand schistosomula cDNA, degenerate PCR and Rapid Amplification of cDNA End protocols were used to identify peptidase isoforms termed TrCB1.1–TrCB1.6. Highest sequence homology is to the described *Schistosoma mansoni* and *Schistosoma japonicum* cathepsins B1. Two isoforms (TrCB1.5 and 1.6) encode putatively inactive enzymes as the catalytic cysteine is substituted by glycine. Two other isoforms, TrCB1.1 and 1.4, were functionally expressed as zymogens in *Pichia pastoris*. Specific polyclonal antibodies localised the peptidases exclusively in the gut of schistosomula and reacted with a 33 kDa protein in worm extracts. TrCB1.1 zymogen was unable to catalyse its own activation, but was trans-processed and activated by *S. mansoni* asparaginyl endopeptidase (SmAE aka. *S. mansoni* legumain). In contrast, TrCB1.4 zymogen auto-activated, but was resistant to the action of SmAE. Both activated isoforms displayed different pH-dependent specificity profiles with peptidyl substrates. Also, both isoforms degraded myelin basic protein, the major protein component of nervous tissue, but were inefficient against hemoglobin, thus supporting the adaptation of *T. regenti* gut peptidases to parasitism of host nervous tissue.

Keywords: Cathepsin B; Peptidase; Asparaginyl endopeptidase; Myelin basic protein; *Schistosoma*; *Trichobilharzia*

1. Introduction

Schistosomatid flukes of the genus *Trichobilharzia* are visceral and nasal parasites of birds. *Trichobilharzia regenti* parasitises aquatic birds of the family Anatidae and the freshwater lymnaeid snail *Radix peregra* as definitive and intermediate hosts, respectively (Horák et al., 1998).

☆ The following GenBank accession numbers: AY648119, AY648120, AY648121, AY648122, AY648123 and AY648124.

* Corresponding author. Tel.: +1 415 502 6866; fax: +1 415 514 3165.
E-mail address: caffrey@cgl.ucsf.edu (C.R. Caffrey).

0020-7519/$30.00 © 2005 Australian Society for Parasitology Inc. Published by Elsevier Ltd. All rights reserved.
doi:10.1016/j.ijpara.2005.02.018

On penetration of the avian host by free-swimming larvae (cercariae), *T. regenti* schistosomula adopt an unusual strategy by migrating through the peripheral nervous system, spinal cord, brain and olfactory nerves to the nasal area where, as adults, they lay eggs. Such migration can cause neuromotor disorders and paralysis (Horák et al., 1999; Hrádková and Horák, 2002). Moreover, trichobilharzian cercariae are able to penetrate skin of nonspecific hosts (including humans) causing cercarial dermatitis or 'swimmers itch' (Horák et al., 2002).

Although the penetration of nonspecific hosts does not lead to maturation of the parasite, worms are able to survive for limited periods in their tissues (Horák et al., 2002; Kouřilová et al., 2004). Thus, in laboratory mice, *T. regenti* can migrate through the central nervous system causing varying degrees of tissue degradation depending on host immune status (Kouřilová et al., 2004). Consequently, the potential for damage of the nervous system of mammals, including humans, due to environmental exposure to *T. regenti* should be considered.

Migrating *T. regenti* schistosomula have a well-developed digestive system and host nervous tissue serves as a source of nutrition, whereas, for adult flukes feeding within nasal tissue, the primary nutritive source is blood (Horák et al., 1999). The ability to digest host tissues other than blood for nutrition purposes is unique within the family Schistosomatidae and is presumably mediated by proteolytic enzymes (peptidases). However, as migrating larvae are difficult to obtain, there is no information regarding *T. regenti* peptidases that might function in digestion and/or migration through nervous tissue.

Peptidases elaborated by helminths are implicated in facilitating nutrition, tissue-migration and evasion of the host immune system (Caffrey and McKerrow, 2004). Disruption of peptidase function, therefore, by chemo- (Wasilewski et al., 1996) or immunotherapy (Dalton et al., 2003) is detrimental to both parasite survival and fecundity. For most platyhelminths, Clan CA cysteine peptidases (http://merops.sanger.ac.uk/; Rawlings et al., 2004) orthologous to mammalian cathepsins B, F and L predominate in both extracts and secretions (Caffrey and McKerrow, 2004). In the genus *Schistosoma*, cathepsin B1 is the primary gut-associated activity and is one of a number of peptidases responsible for the digestion of blood proteins (Caffrey et al., 2004 for review). It is reasonable to hypothesise, therefore, that a similar system of orthologous peptidases, facilitating nutrition and migration, exists in *T. regenti*.

Here, we have used fluorogenic peptidyl substrates with high extinction coefficients and irreversible affinity-probes capable of labelling peptidases in limiting amounts of tissue to identify and characterise a predominant cathepsin B-like activity in extracts of migrating *T. regenti* schistosomula. PCR-based cloning strategies were also used to identify multiple isoforms of a *T. regenti* cathepsin B orthologous to the SmCB1 peptidase of *Schistosoma mansoni* (Sajid et al., 2003; Caffrey et al., 2004). Two isoforms were selected for heterologous expression in the yeast *Pichia pastoris* and specific antibodies raised to localise the peptidases in parasite tissues. Protein and peptidyl substrates, and peptidyl inhibitors were used to characterise the specificities of the recombinant peptidases.

2. Materials and methods

2.1. Parasites

Trichobilharzia regenti is maintained in the Prague laboratory by cycling between the freshwater snail, *Radix peregra* and the domestic duck *Anas platyrhynchos f.* domestica. Details of the trichobilharzian life cycle have been described previously (Horák et al., 1998). Ducklings, between 10 and 15 days old, were percutaneously infected with 1000 *T. regenti* cercariae and euthanised 5–6 days p.i. Spinal cord material was squashed and schistosomula isolated over a 150 μm pore wire mesh in Tris-buffered saline (TBS; 20 mM Tris–HCl, 150 mM NaCl, pH 7.8) or RPMI-1640 medium (Gibco BRL; Paisley, UK). Worms were collected after 2–4 h and washed twice in the same buffer or medium. Worms were frozen in liquid nitrogen and stored at −80 °C or lyophilised.

2.2. Preparation of the soluble worm extract

Lyophilised schistosomula (500) were sonicated in 200 μl 50 mM citrate, 100 mM sodium phosphate buffer, pH 6.0, over an ice bath. After centrifugation at 10,000×*g* at 4 °C for 5 min, the supernatant was collected and either used immediately or stored at −80 °C.

2.3. Irreversible active site-labelling of cysteine peptidases in soluble worm extract

The biotinylated Clan CA cysteine peptidase inhibitor, DCG-04 (Greenbaum et al., 2000), was used for active-site labelling. Worm extract (20 μl, at the protein concentration of 0.2 mg/ml; equivalent to 50 schistosomula) was incubated for 1 h with 5 μM DCG-04 in 50 mM citrate, 100 mM sodium phosphate, pH 6.0 containing 2 mM DTT. After SDS-PAGE and electroblotting onto nitrocellulose membrane, the membrane was blocked for 1 h at room temperature in Tris-buffered saline (TBS) containing 5% BSA and 0.05% Tween 20, washed for 3×5 min in TBS containing 0.05% Tween 20 (TBS-T) and then incubated for 30 min in TBS-T containing 2.5 μg/ml avidin-peroxidase (Sigma, Steinheim, Germany). After washing for 3×5 min in TBS-T, the blot was developed in substrate solution (100 mM Tris–HCl, pH 7.6; 0.6 mM 3,3′-diaminobenzidine; Sigma) by adding 0.01% H_2O_2. The reaction was stopped by extensive washing in water.

Alternatively, the fluorescent cysteine peptidase affinity label, BODIPYgreen-DCG-04 (bg-DCG-04;

Greenbaum et al., 2002), was used to visualise cysteine peptidases. Parasite extract or chromatographic fractions thereof (20 µl; see below) were incubated for 1 h with 100 nM bg-DCG-04 in 50 mM citrate, 100 mM sodium phosphate, pH 6.0, containing 2 mM DTT. After SDS-PAGE, gels were washed 3×15 min in water and then placed in a solution of 40% methanol, 10% acetic acid for 2 h. After washing for 3×15 min in water, fluorescence was detected in a Typhoon 8600 Variable Mode Imager (Amersham Biosciences, Piscataway, New Jersey) with excitation and emission wavelengths set to 532 and 555 nm, respectively.

2.4. Measurement of peptidolytic activity and pH optimum with fluorogenic substrates

Cysteine peptidase activity was routinely measured with synthetic peptidyl substrates (Bachem, Torrance, California): Z-Phe-Arg-AMC (benzyloxycarbonyl-phenylalanyl-arginine-7-amido-4-methylcoumarin), which is used to measure the activities of both cathepsins B and L (Barrett and Kirschke, 1981) and Z-Arg-Arg-AMC, which is selective for cathepsin B among lysosomal cysteine endopeptidases (Barrett and Kirschke, 1981). Stocks of 10 mM peptidyl substrates were dissolved in dimethylsulfoxide (DMSO). Assays were performed in black 96-well plates as follows: 20 µl of worm extract (~4 µg) were pre-incubated for 10 min at room temperature in 80 µl 50 mM citrate, 100 mM sodium phosphate, 2 mM DTT, pH 3.0–8.0, containing 300 mM NaCl to minimise variations in ionic strength. The reaction was started by adding 100 µl of the same buffer solution containing 40 µM peptidyl substrate. Release of free AMC was measured at excitation and emission wavelengths of 355 and 460 nm, respectively, in a Labsystems Fluoroskan II fluorometer (Thermo Electron Corporation, West Palm Beach, Florida) or a Spex Fluoromax 3 (Jobin Yvon Horiba, Longjumeau, France) for 20 min at room temperature. Controls contained an equal volume of buffer instead of enzyme. To confirm that cysteine peptidase activity was being measured in worm extracts, 10 µM of the Clan CA cysteine peptidase inhibitor, E-64 (L-*trans*-epoxysuccinyl-leucylamide-(4-guanido)-butane; Sigma), was included in the incubation step prior to addition of substrate solution.

2.5. Inhibition assay to discriminate cathepsin B activity

Two inhibitors were used: CA-074 (N-[L-3-*trans*-propylcarbamoyloxirane-2-carbonyl]-Ile-Pro-OH; Bachem), a selective inhibitor of cathepsin B (Murata et al., 1991), and the general cysteine peptidase inhibitor, E-64. Inhibitor stock solutions were dissolved in DMSO and subsequent serial dilutions prepared in water. Assays were in triplicate for each inhibitor concentration. For each reaction, 20 µl of worm extract (4 µg) were incubated for 10 min at room temperature in 75 µl 50 mM citrate, 100 mM sodium phosphate, 2 mM DTT, pH 5.0, before addition of 5 µl of either inhibitor at one of the selected working concentrations (20, 2, 0.2, 0.02, 0.002, 0.0002 µM). After 15 min, remaining hydrolysis rates were measured with Z-Phe-Arg-AMC as described above.

2.6. Anion-exchange FPLC of T. regenti soluble extract

Anion-exchange FPLC was performed using a Mono Q HR5/5 column (Amersham Biosciences) and an Äkta chromatographic system (Amersham Biosciences). Soluble extract of 5–6-day-old lyophilised *T. regenti* (3000) was prepared in 1 ml 100 mM Bis–Tris buffer, pH 6.0 (see Section 2.2). Sample was loaded onto the column equilibrated with the same buffer. Elution was with a linear gradient of 0–1 M NaCl in 30×1 ml fractions at a flow rate of 1 ml/min. Peptidase activity was monitored using 100 µl of each fraction in 100 µl assay buffer (100 mM Bis–Tris buffer, 2 mM DTT, pH 5.0) containing 20 µM Z-Phe-Arg-AMC or Z-Arg-Arg-AMC (see Section 2.4). Cysteine peptidase activity in column fractions (20 µl) was visualised using bg-DCG-04 and SDS-PAGE as described above.

2.7. Isolation of nucleic acids and amplification of cDNA by PCR with degenerate primers

Worms (1000), isolated from nervous tissue, were stored at −80 °C in the RNA/DNA Stabilization Reagent for Blood/Bone Marrow (Roche, Mannheim, Germany). Parasite total RNA was isolated using the High Pure RNA Tissue Kit (Roche) according to the manufacturer's instructions. Messenger RNA was isolated in a final volume of 20 µl using the mRNA Isolation Kit for Blood/Bone Marrow (Roche). Single-stranded cDNA was prepared using Superscript II Reverse Transcriptase (Invitrogen; Carlsbad, CA) and an oligo $d(T)_{22}$ reverse primer according to the manufacturer's instructions.

Multiple sequence analysis of schistosomal cathepsin B sequences, namely SmCB1 (Klinkert et al., 1989; M21309), SjCB1 (Merckelbach et al., 1994; X70968) and SmCB2 (Caffrey et al., 2002; AJ312106) were used to design degenerate primers for the semi-nested PCR amplification of *T. regenti* cathepsin(s) B. The outer forward, and outer and inner reverse primers were derived from a consensus sequence around the active site Cys and Asn residues, respectively. The forward primer (TrCBdegfrd) was 5′-TTYGGNGCNGTNGARGC-3′ (where N = any nucleotide; R = A/G; Y = T/C). For the reverse primers, the design incorporated different oligonucleotides for each round of PCR to decrease degeneracy. Therefore, outer reverse primers were: TrCBdegrev1, 5′-TCNCCCCARTCNGARTTCCA-3′; TrCBdegrev2, 5′-TCNCCCCARTCRCTRTTCCA-3′ or TrCBdegrev3, 5′-TCNCCCCARTCYTCRTTCCA-3′ and the inner reverse primers were: TrCBdegrev4, 5′-TTNGCDATNAGCCARTANGG-3′ or

TrCBdegrev5, 5′-TTNGCDATYAACCARTANGG-3′ (where D = A/T/G).

PCR was carried out in a final volume of 50 µl and contained 0.5 µl of cDNA, 0.1 mM dNTP, 1 µM of the forward and one of the outer reverse primers, 1×PCR reaction buffer (Roche) and 2.5 U Taq Polymerase (Roche). The amplification profile consisted of an initial denaturation step at 94 °C for 5 min followed by 35 cycles at 94, 45 and 72 °C, each for 1 min. The resulting PCR product was then used as a template (0.5 µl) for the semi-nested reaction, which was performed as described above except for the use of one of the inner reverse primers and an annealing temperature of 50 °C. Final PCR products were separated in 1% agarose gels, excised and purified using the Qiaquick gel extraction kit (Qiagen, Hilden, Germany). Purified DNA was inserted into the pCR2.1-TOPO cloning vector (Invitrogen) and constructs then used for the transformation of chemically competent TOP10 *Escherichia coli* cells (Invitrogen). After propagation of transformed bacterial cultures, plasmid clones were isolated using the Qiaprep purification kit (Qiagen) and sequenced with the M13 forward and M13 reverse primers (UCSF Biomolecular Resource Facility).

2.8. Identification of full-length TrCB1 gene isoforms by Rapid Amplification of cDNA ends (RACE)

To obtain sequence downstream of the cloned TrCB1 PCR fragments, 3′ RACE was performed with 2 gene-specific primers. Briefly, cDNA was reverse transcribed from parasite mRNA using Superscript II reverse transcriptase (Invitrogen) and an oligo d(T)-anchor primer (5′-GACCA CGCGTATCGATGTCGACTTTTTTTTTTTTTTV-3′; Roche) according to the manufacturer's instructions for Superscript II. The first PCR (50 µl final volume) contained 0.5 µl of the cDNA template, 0.4 mM dNTP, 0.3 µM gene-specific outer forward primer (5′-AAGGAGGATT TCCTGGAGCAGCC-3′), 1 µM reverse anchor primer (5′-GACCACGCGTATCGATGTCGAC-3′; Roche), 1×PCR Buffer (MBI Fermentas, Vilnius, Lithuania), 2.5 mM $MgCl_2$ and 2.5 U Taq polymerase (MBI Fermentas). The reaction was performed with an initial denaturation step at 94 °C for 5 min followed by 26 cycles of 94, 48 and 72 °C, each for 1 min. The second (semi-nested) PCR reaction was performed as above with 1 µl of the primary PCR product, 0.3 µM gene-specific inner forward primer (5′-AAATGCGA ACATCATACTACGGG-3′) and 0.3 µM of the same reverse anchor primer at 94 °C for 5 min followed by 36 cycles of 94, 50 and 72 °C, each for 1 min.

To obtain sequence upstream of the cloned TrCB1 fragments, 5′ RACE was performed in semi-nested reactions according to the manufacturer's instructions using the 5′/3′ RACE Kit (Roche). Reverse transcription was done according to the protocol for Superscript II reverse transcriptase (Invitrogen) with a gene-specific reverse primer (5′-TTTCTTCTCTACACCCCATCC-3′). Subsequent PCRs were done first with the supplied anchor forward primer (5′-GACCACGCGTATCGATGTC GAC-3′; Roche) and a gene-specific reverse outer primer (5′-ATTCTAACAGCATGTCCACCG-3′), and then with the same anchor primer and a gene-specific inner reverse primer (5′-AAATCTGAGTGTACTGTGAAAGC-3′). PCRs were carried out as described for 3′ RACE with annealing temperatures of 48 and 55 °C for the first and the second PCRs, respectively.

PCR products from 5′ and 3′ RACE were cloned as described in Section 2.7. and submitted for sequencing (Dept. of Parasitology Sequencing Facility, Charles University). To verify the complete sequence of the TrCB1 gene, PCR was done using a Taq/Pwo polymerase mix (Roche) and primers directed to both the UTR and coding regions. Ten PCR reactions were performed and the products subcloned into the pCR2.1-TOPO vector for propagation in *E. coli*. For each PCR, four clones were randomly selected for sequencing (40 clones in total). Nucleotide sequences were processed by CHROMAS (Technelysium Pty Ltd., Tewantin, Australia) and the DNASTAR package (DNAstar Inc., www.dnastar.com/web/index.php).

2.9. Expression of TrCB1 isoforms in Pichia pastoris

Two TrCB1 isoforms were chosen for expression in *P. pastoris*-TrCB1.1 (the most abundant from PCR screening) and TrCB1.4 (showing the greatest primary sequence divergence among the putatively active peptidases). The starting positions of the TrCB1 pro-regions were predicted with the SignalP software (Nielsen et al., 1997; http://www.cbs.dtu.dk/services/SignalP) and compared with other cathepsin B genes from GenBank. ProTrCB1.1 and proTrCB1.4 sequences were amplified from cDNA using the AccuPrime Pfx DNA polymerase kit (Invitrogen). The same forward primer, proTrCBfrd, was used for both TrCB1 isoforms; 5′-ATACTCGAG*AAAAGA*GAGAATGAGATA CAATTCGAGCC-3′, which contained a *Xho* I restriction site (underlined) and a Kex 2 peptidase cleavage site (in italics). The reverse primers were; proTrCB1rev, 5′-A ATGCGGCCGC*TCA*ATGATGAGGTAATCCACCTGT AAC-3′ and proTrCB4rev, 5′-AATGCGGCCGC*TCA*AT CACGAGGTAGTCCACTTGTAAC-3′. Both reverse primers contained a *Not* I restriction site (underlined) and a termination codon (in italic). Purified PCR products were ligated into the expression vector, pPICZα B, as previously described (Caffrey et al., 2002). Sequences of recombinant vectors were verified at the UCSF Biomolecular Resource Facility.

Protein expression in *P. pastoris* was carried out according to the manufacturer's instructions (Invitrogen; Caffrey et al., 2002). Briefly, electrocompetent *P. pastoris* X33 strain was transformed with 10 µg recombinant vectors. Yeast colonies grown on zeocin-impregnated yeast-peptone-dextrose (YPD-Z) agar plates were picked

for expansion in 15 ml YPD-Z medium and induction in 3 ml of basic minimal methanol (BMM) solution. Peptidase activity in the medium was detected by hydrolysis of Z-Phe-Arg-AMC (see Section 2.4). Those clones expressing most activity were chosen for further expansion and induction in 2 L YPD-Z and 0.5 L BMM, respectively. After induction for 2 days, media were filtered (0.45 μm), lyophilised, rehydrated to 10% of the induction volume and desalted in 50 mM sodium phosphate buffer, pH 6.0 using PD-10 columns (Amersham Biosciences). Solutions were re-lyophilised and stored at −20 °C. The molecular masses of both recombinant enzymes were estimated by a combination of SDS-PAGE on 4–12% precast NuPage gradient gels (Invitrogen) and either affinity-labelling with bg-DCG-04 (see Section 2.3.) or staining with Coomassie Brilliant Blue R-250.

2.10. Production of monospecific sera

Antibodies against the zymogens pro-TrCB1.1 and pro-TrCB1.4 were produced to verify the localisation of the TrCB1 in schistosomula tissues. SDS-PAGE (12% gels) was employed for separation of recombinant proteins (1 mg each). Gels were stained with Coomassie Brilliant Blue and washed in several changes of TBS containing 1% Triton X-100, and then with TBS alone. The fastest migrating protein bands (not hyper-glycosylated) were cut from the gel and homogenised in sterile saline. Two BALB/c mice were injected intraperitoneally without adjuvant three times at 10-day intervals. Seven days after the last injection, mice were euthanised and exsanguinated. After allowing the blood to clot at room temperature for 15 min, sera were removed and stored at −20 °C. Control sera were taken from the same individual mice prior to immunisation.

Alternatively, 4-week-old domestic ducks (*A. platyrhynchos* f. domestica.) were immunised with soluble zymogens. Proteins were mixed with an equal volume of Freund's incomplete adjuvant and administered subcutaneously (150 μg pro-TrCB1.4 per duck) twice at 14-day intervals with an intravenous booster of 50 μg of zymogen in saline 10 days after the second injection. Sera were collected 7 days after the last injection. Control sera were taken from the same individuals prior to immunisation.

2.11. Immunochemistry

For immunoblotting, soluble extract of *T. regenti* schistosomula (300 worms in 50 mM citrate, 100 mM phosphate, pH 6.0) or recombinant TrCB1 zymogens (5 μg) were separated by 10% SDS-PAGE and transferred to nitrocellulose membrane. The membrane was blocked for 2 h in TBS-T with 5% defatted milk at room temperature, then washed 3×5 min in TBS-T and incubated for 1 h with anti-TrCB1.1, 1.4, or pre-immune sera diluted 1:50 in TBS-T. After washing 5×5 min in TBS-T, the membrane was incubated for 1 h with horseradish peroxidase-labelled anti-mouse IgG (Sevac, Czech Republic) or anti-duck IgG (KPL, Gaithersburg, Maryland) at a dilution 1:1000. After washing for 3×5 min in TBS-T, the membrane was developed and stopped as described in Section 2.3.

For immunohistochemistry, 6-day-old schistosomula, isolated from duck spinal cord, were washed twice in TBS, fixed in Bouin's fixative solution containing 5% glacial acetic acid and stored at 4 °C. After soaking in TBS for 4 h, worms were embedded in JB-4 (Polyscience, Niles, Illinois). Embedded material was sectioned at 2.5 μm, placed on glass slides and dried at 60 °C for 5 min. Slides were incubated for 1 h in TBS-T containing 2% BSA, washed briefly in TBS-T and incubated for 1 h at room temperature with pre-immune, duck or mouse serum at a 1:100 dilution in TBS-T. After washing 2×5 min, sections were incubated for 30 min with fluorescein isothiocyanate (FITC)-labelled anti-mouse (Sigma) or anti-duck IgG (KPL) at a dilution of 1:100 in TBS-T. Sections were washed 2×5 min in TBS-T and mounted in TBS-buffered glycerol (4:1). Localisation of TrCB1 expression was observed using an Olympus BX51 microscope.

2.12. Processing of recombinant TrCB1 zymogen isoforms

Both auto- and heterologous-processing were attempted to convert recombinant TrCB1 zymogens to their mature active forms. For auto-processing, lyophilised zymogens (0.5–1 μg/μl) were dissolved in 50 mM citrate, 100 mM sodium phosphate buffer, 2 mM DTT, pH 4.0–7.0, and incubated for different periods at 37 °C with shaking. For heterologous-processing, zymogens were incubated with recombinant *S. mansoni* asparaginyl endopeptidase (SmAE aka. *S. mansoni* legumain; Caffrey et al., 2000) as described for processing of *S. mansoni* pro-cathepsin B1 (Sajid et al., 2003). Activated SmAE (0.1 μg; 10 μl) was mixed with 80 μl of TrCB1 zymogen (1–2 μg/μl) in 50 mM citrate, 100 mM phosphate buffer, 2 mM DTT, pH 4.5–5.0 and adjusted with the same buffer to give a final volume of 200 μl. The solution was incubated for different time periods at 37 °C with shaking. Controls omitted SmAE.

Peptidase activity of processed TrCB1 isoforms was measured using Z-Phe-Arg-AMC as described in Section 2.4. Processing was visualised by SDS-PAGE followed by Coomassie staining or immunoblotting using mouse anti-pro-TrCB1.1 serum. For N-terminal sequencing by Edman degradation, processed intermediates and final forms were resolved by SDS-PAGE and electroblotted onto polyvinyldifluoridine membrane. Proteins were stained in a solution of 0.5% Coomassie Brilliant Blue in 40% methanol then destained in the same solution without Coomassie. Membranes were soaked in distilled water and dried before sequencing (Stanford Protein and Nucleic Acid Biotechnology Facility, Stanford, California).

2.13. S_2 subsite specificity of recombinant TrCB1.1 and 1.4

For S_2 subsite specificity and pH optimum assays, four substrates varying in the P_2 amino acid residue position were employed, namely, Z-Phe-Arg-AMC, Z-Arg-Arg-AMC, Z-Val-Arg-AMC and Z-Leu-Arg-AMC (the latter two were kindly given by Dr Dieter Brömme, Department of Human Genetics, Mount Sinai School of Medicine, New Jersey). Assays (Section 2.4) were carried out with 0.5 nM of either heterologously processed TrCB1.1 or auto-processed TrCB1.4.

2.14. 3D molecular modelling of TrCB1.1 and 1.4

Comparative protein structure modelling by satisfaction of spatial restraints, as implemented in the program MODELLER (Sali and Blundell, 1993) (http://salilab.org/modeller), was used to build three-dimensional models of TrCB1.1 and TrCB1.4. The models relied on two different template structures previously determined by X-ray crystallography: human pro-cathepsin B (Protein Data Bank (PDB) ID: 3pbh; Podobnik et al., 1997) and bovine cathepsin B in complex with the inhibitor NS-134 (PDB ID: 1sp4; Stern et al., 2004). These templates were selected for two reasons: (i) there is relatively high sequence similarity between their catalytic domains and those of TrCB1.1 and TrCB1.4 (sequence identity is ~57%) and (ii) there is a ligand in their active site clefts (the pro-peptide in 3pbh and the inhibitor NS-134 in1sp4). Both models were built with NS-134 in their clefts, mimicking the interaction between NS-134 and bovine cathepsin B.

The substrate molecule, Z-Phe-Arg-AMC, was built into the active site clefts of TrCB1.1 and 1.4 by the program MOLOC (http://www.moloc.ch/) as follows. First, the backbone of the substrate was superimposed on that of the inhibitor NS-134. Second, the substrate backbone was relaxed by a local energy minimisation using the MAB united atoms force field in MOLOC (Gerber and Muller, 1995). Third, sidechains were added to the substrate backbone. Finally, the substrate and neighbouring peptidase sidechains were minimised to obtain the final model. Images of the structures were produced using the CHIMERA package from the UCSF Computer Graphics Laboratory (Pettersen et al., 2004).

2.15. Incubation of recombinant TrCB1 isoforms with protein substrates

Hemoglobinolytic activity was studied with duck, turkey and bovine hemoglobin (Sigma). Hemoglobin solution (50 µl; 0.5 mg/ml in water) in 50 mM citrate, 100 mM phosphate buffer, 2 mM DTT, pH 4.0–7.0, was incubated with recombinant pro-enzymes (1–2 µg) or activated enzymes (0.1–5 µM final) for various times at 37 °C. Control reactions contained buffer instead of enzyme. Digestion of mouse myelin basic protein (mMBP; kindly given by AboaTech, Turku, Finland), isolated according to Maatta et al. (1997), was carried out as described for hemoglobin. The working concentration of mMBP was 0.3 mg/ml. Hydrolysis was visualised by 10% SDS-PAGE and staining with Coomassie Brilliant Blue.

3. Results

3.1. A major cathepsin B activity in extracts of migrating T. regenti schistosomula

Incubation of soluble schistosomula extracts with the biotinylated Clan CA cysteine peptidase-selective inhibitor, DCG-04, identified a prominent 33 kDa peptidase doublet after SDS-PAGE, electroblotting and incubation with streptavidin-conjugated peroxidase (Fig. 1, lane 1). Pre-incubation of extracts with a second cysteine peptidase-selective inhibitor, E-64 (Fig. 1, lane 2) prevented most of the labelling by DCG-04 and, thus, confirmed the identity of

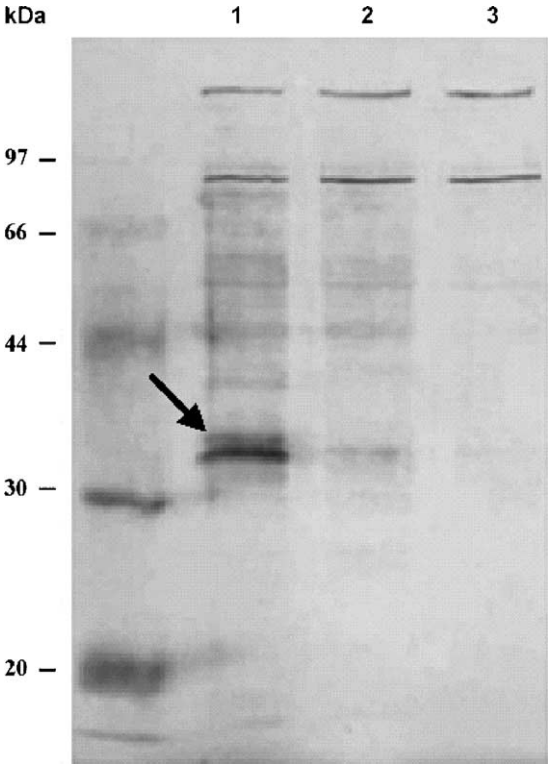

Fig. 1. Identification of Clan CA cysteine peptidase activity in *Trichobilharzia regenti* schistosomula. Soluble worm extract (20 µl; equivalent to 50 schistosomula) was incubated for 1 h at room temperature with 5 µM of the biotin labelled irreversible Clan CA cysteine peptidase inhibitor, DCG-04, then separated by 10% SDS-PAGE and electroblotted to nitrocellulose membrane. The blot was incubated with peroxidase-linked avidin and then developed with 3,3′-diaminobenzidine. Lane 1, sample incubated with DCG-04; lane 2, sample pre-incubated with 10 µM of the Clan CA cysteine peptidase inhibitor E-64 prior to reaction with DCG-04; lane 3, sample heated to 70 °C for 10 min prior to reaction with DCG-04. The arrow indicates a protein doublet at 33–35 kDa specifically labelled by DCG-04.

the protein doublet as a Clan CA peptidase. Prior heating of extracts to 70 °C also eliminated the reaction with DCG-04 (Fig. 1, lane 3), presumably through irreversible denaturation of the peptidase target.

Cysteine peptidase activity in soluble extracts could also be measured with two synthetic peptidyl substrates, Z-Phe-Arg-AMC and Z-Arg-Arg-AMC, commonly used to detect such activity (Fig. 2A). The respective pH optima of substrate hydrolysis were 4.5–5.5 and 5.5–6.5. At pH 5.0, peptidolytic activity was ~80-fold greater with Phe-Arg than Arg-Arg. For both substrates, little activity (<15%) was detected at neutral pH.

The contribution of cathepsin B to the total activity detected in schistosomula extracts was measured by comparing inhibition produced by the general cysteine peptidase inhibitor, E-64, and the selective cathepsin B inhibitor, CA-074 (Fig. 2B). Over the concentration range tested (0.2 nM–20 μM), CA-074 was as potent as E-64. This suggests that most of the cysteine peptidase activity in schistosomula extracts is due to one or more cathepsins B and that the prominent peptidase doublet visualised by

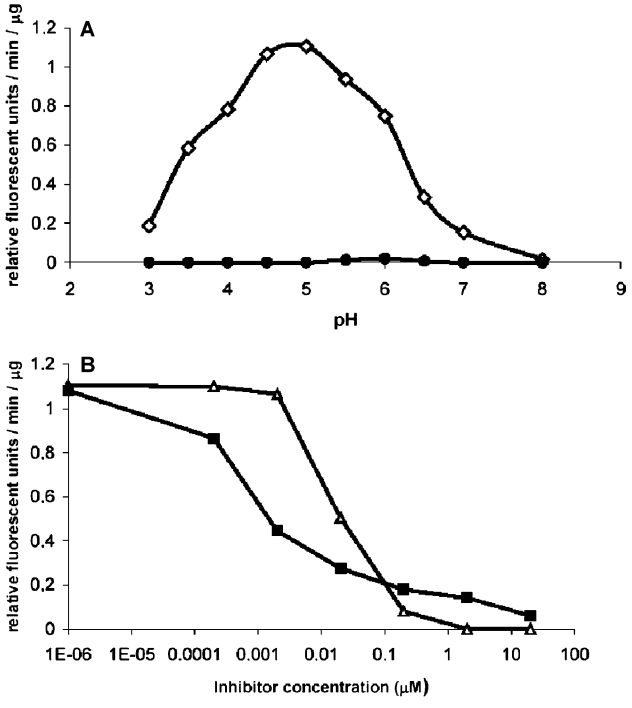

Fig. 2. Quantification of Clan CA cysteine peptidase activity in soluble extract of *Trichobilharzia regenti* schistosomula. (A) pH dependence of protease activity in soluble extracts (20 μl; 4 μg) with the peptidyl fluorogenic substrates Z-Phe-Arg-AMC (■) or Z-Arg-Arg-AMC (●). (B) Discrimination of cathepsin B-like activity. Soluble extract (20 μl; 4 μg) was incubated with different concentrations (0.2 nM to 20 μM) of either the general Clan CA cysteine peptidase inhibitor E-64 (△) or the cathepsin B-selective inhibitor CA-074 (□). Peptidase activity was then measured with 10 μM Z-Phe-Arg-AMC. The similarity of the inhibition curves indicates that cathepsin B activity is the major Clan CA peptidase activity present in soluble extracts. For both A and B, assays were in triplicate and the standard deviation about each mean value was never >5%.

DCG-04 is cathepsin B-like. Mono Q anion exchange chromatography of schistosomula extracts confirmed the presence of a single predominant cathepsin B peptidase as the activity detected in column fractions using the peptidyl substrates Z-Phe-Arg-AMC and Z-Arg-Arg-AMC (Fig. 3A) correlated with the appearance of a fluorescent bg-DCG-04-labelled band at 33 kDa (Fig. 3B).

3.2. Identification of a cathepsin B transcript in schistosomula mRNA

The lack of parasite material made the further detailed characterisation of the cathepsin B-like peptidases difficult (schistosomula of ~500×80 μm are embedded in the spinal cord of infected hosts). Therefore, we pursued PCR-based cloning strategies to identify and characterise the genes responsible for the detected cathepsin B activity. Using template cDNA prepared from mRNA of schistosomula, semi-nested PCR was performed with degenerate primers targeting the relatively conserved amino acid residues around the Cys and Asn active site residues of schistosome cathepsins B (Klinkert et al., 1989; Merckelbach et al., 1994; Caffrey et al., 2002). A single forward primer, combined first with one of three outer reverse primers and then with one of two inner reverse primers, was employed to amplify as many individual *T. regenti* cathepsin B sequences as possible. All combinations of degenerate primers resulted in the amplification of a 550 bp product (not shown). TOPO-cloning and sequencing of this product identified a cathepsin B-like fragment. Rapid amplification of cDNA ends (5′ and 3′ RACE) was performed to complete the open reading frame (ORF) and flanking 5′ and 3′ untranslated regions (UTRs). Finally, the complete cathepsin B gene sequence (~1.1 kb) was amplified from schistosomula cDNA using the 5′ and 3′ UTR-specific primers. The gene was termed TrCB1, in keeping with previously used nomenclature (Caffrey et al., 2004).

3.3. Identification of six isoforms of TrCB1.1

Sequencing of 40 separate full-length TrCB1 clones revealed the presence of six isoforms, which were termed TrCB1.1–1.6. An alignment of the six isoforms is presented in Fig. 4. Of the 40 clones, TrCB1.1 was the most abundant at 68%. The other isoforms were in lower abundance (TrCB1.2-10%; TrCB1.3-5%; TrCB1.4-13%; TrCB1.5 and TrCB1.6-3%).

3.4. Characterisation of T. regenti cathepsins B

The ORFs of all TrCB1 isoforms encode proteins of 342 amino acids each containing a signal leader sequence, a pro-peptide and a mature (catalytic) domain (Fig. 4). The signal leader sequence, predicted by the SignalP 3.0 software contains 21 amino acid residues. The pro-peptide sequence

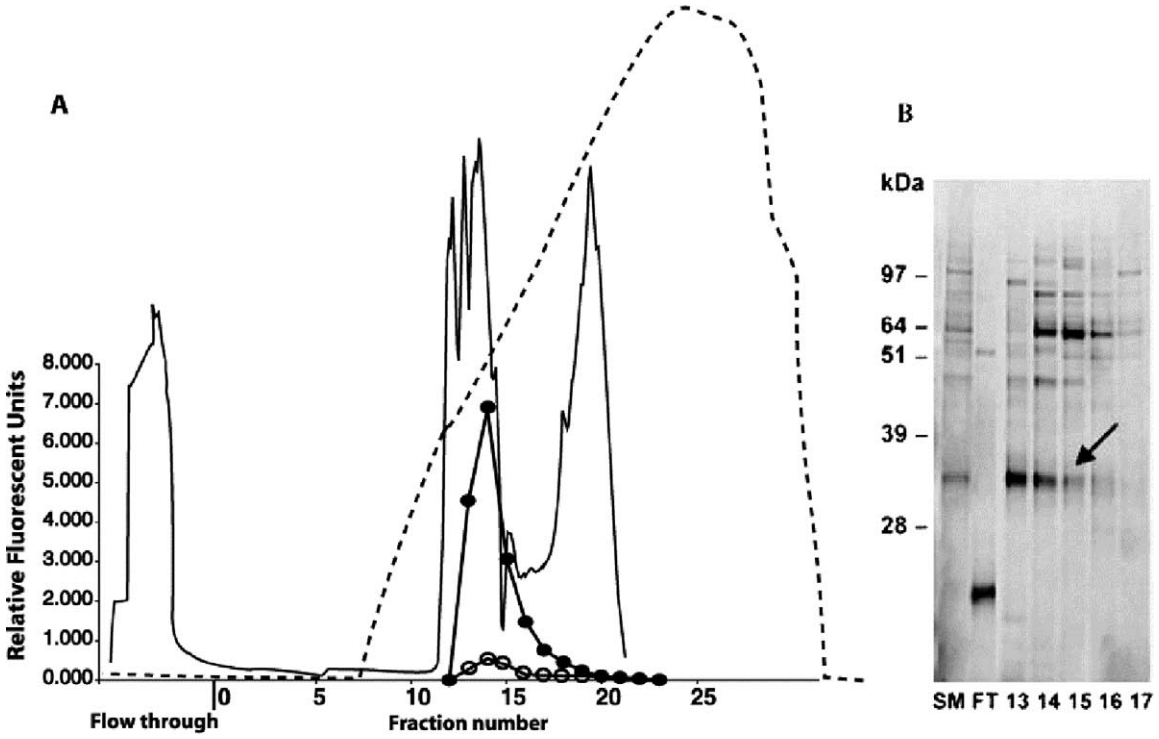

Fig. 3. Fractionation of Clan CA peptidase activity in *Trichobilharzia regenti* schistosomula. (A) Soluble extract (1 ml; from 3000 worms) was separated by Mono Q HR5/5 anion-exchange chromatography in 100 mM Bis–Tris buffer, pH 6.0, over a 0–1.0 M NaCl gradient (dashed curve). The solid curve represents protein detected at A_{280}. Column fractions (20 μl) were analysed for cysteine peptidase activity by hydrolysis of Z-Phe-Arg-AMC (solid circles) and Z-Arg-Arg-AMC (open circles). Activity was detected in fractions 13–19. (B) Column fractions (20 μl) were incubated for 1 h with 100 nM b-gDCG-04 in 50 mM citrate, 100 sodium phosphate, pH 6.0, containing 2 mM DTT and resolved by 4–12% gradient SDS-PAGE. Cysteine peptidases were visualised for fluorescence in a Typhoon 8600 Variable Mode Imager. SM, starting sample material; FT flow-through (material not bound to column); numbers indicate particular fractions resolved as in (A). The arrow indicates a cathepsin B-like peptidase in the fractions labelled by b-gDCG-04.

of 68 residues was predicted from a multiple alignment of cathepsins B: SmCB1.1 and 1.2 (Sajid et al., 2003; AJ506157 and AJ506158), SjCB1 (Merckelbach et al., 1994; X70968), SmCB2 (Caffrey et al., 2002; AJ312106), SjCB2 (AY226984) and human cathepsin B (Chan et al., 1986; M14221). Expected M_r values calculated by PeptideMass (Wilkins et al., 1997; http://www.expasy.org) are 38.7, 36.5, and 28.5 kDa for the full length, zymogen and mature proteins, respectively. With the exception of TrCB1.5, the pro-peptide regions of all TrCB1 isoforms contain an Asn at p64, close to the start of the mature catalytic domain (Fig. 4). Asn at a similar position in pro-SmCB1.1 is known to facilitate zymogen-processing and -activation by *S. mansoni* asparaginyl endopeptidase (SmAE; Sajid et al., 2003).

In the catalytic domain, all 12 Cys residues forming intra-molecular disulphide bridges are conserved between *Trichobilharzia*, schistosome and human cathepsins B (Musil et al., 1991; Caffrey et al., 2002). The catalytic domains of TrCB1.1–1.4 contain 253 amino acid residues, including the conserved catalytic triad residues of Clan CA cysteine peptidases (Turk et al., 2000), namely Cys^{29}, His^{199}, and Asn^{219} (numbering system as shown in Fig. 4). For TrCB1.5 and 1.6, however, the catalytic Cys^{29} is substituted by a Gly, which suggests that any expression product would be inactive. All TrCB1 isoforms contain the 'occluding loop' (Cys^{99}–Cys^{128}) responsible for peptidyl dipeptidase activity of cathepsin B peptidases (Illy et al., 1997). This loop contains either a single His residue at position 109 (TrCB1.6) or two such residues at 109 and 110 (all other isoforms). Other residues responsible for interacting with main chain of the bound substrate (Gln^{23}, Gly^{72}, Gly^{73}, Trp^{221}) are conserved (Turk et al., 2000). Potential N-linked glycosylation sites are found at positions p64 and 95: TrCB1.6 has an additional site at 110.

For full-length TrCB1, the amino acid sequence identity between the putatively active isoforms TrCB1.1–1.4 is 98–99%. Identity between these and the presumed inactive TrCB1.5 and TrCB1.6 isoforms falls to between 84 and 89%. Sequence identities of all TrCB1 isoforms with SmCB1.1 and SjCB1 are between 67 and 71%. Identities with SmCB2 and SjCB2 are 46–50%, and with human cathepsin B, 42–44%.

3.5. Functional expression of recombinant TrCB1.1 and 1.4 zymogens in P. pastoris

TrCB1.1, as the most abundant transcript, and TrCB1.4, as the most divergent among those isoforms expected to be active, were selected for expression as zymogens in

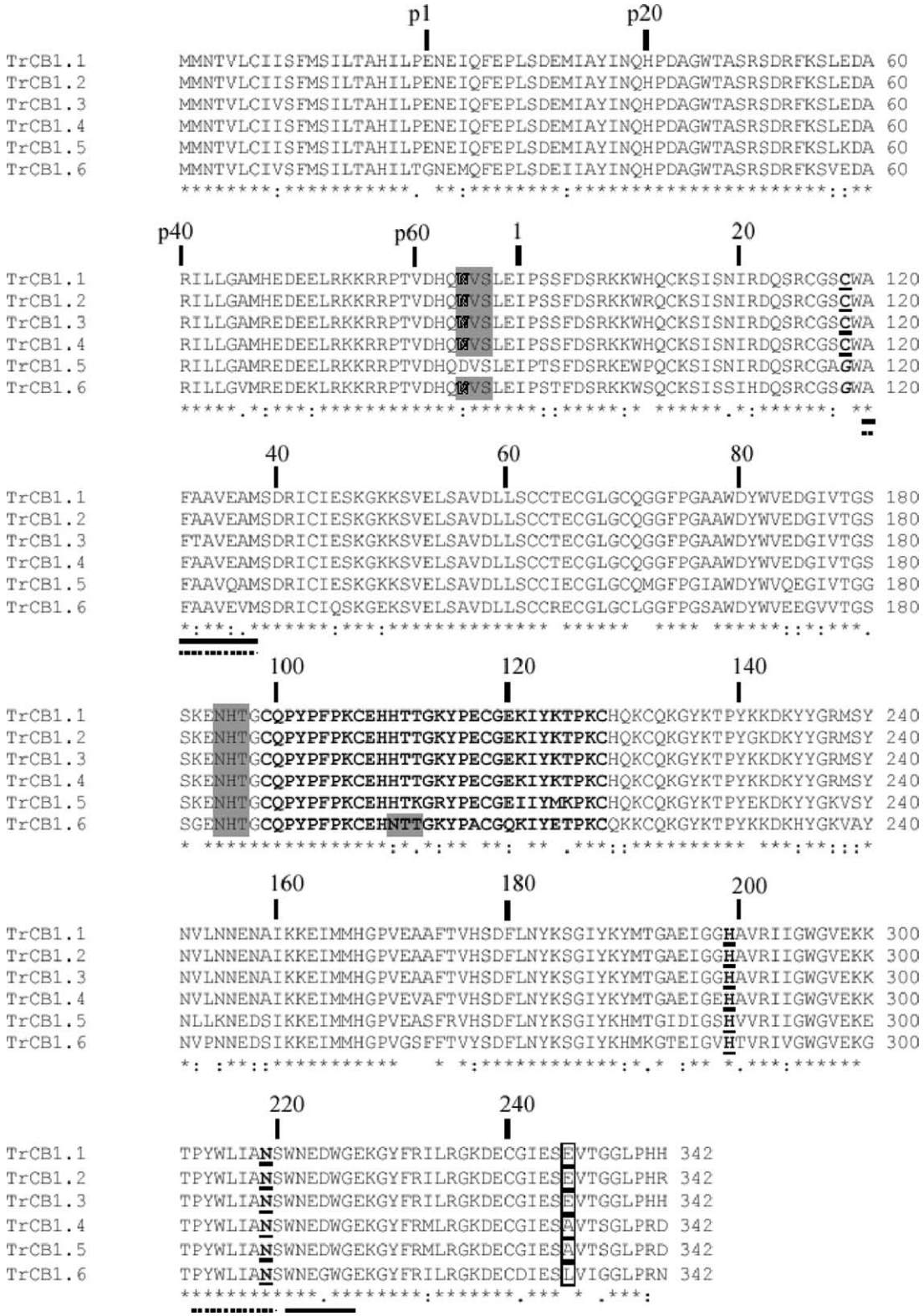

Fig. 4. Multiple sequence alignment of six *Trichobilharzia regenti* cathepsin B isoforms. Above the sequences, the letter 'p' indicates numbering of the propeptide region; the number '1' indicates the start of predicted catalytic domain. The catalytic triad (C, H and N) is marked in bold and underlined. Note that for TrCB1.5 and 1.6, the catalytic C is substituted by a G (bold-italic). The 'occluding loop' unique to cathepsin B enzymes is indicated in bold face. The hypothetical processing site (N) at p64 for *Schistosoma mansoni* asparaginyl endopeptidase is contoured. Potential N-glycosylation sites are shaded in grey and the amino residue at the base of the S_2 subsite is boxed. Solid and dashed lines indicate sequence regions used to construct degenerate primers for first round and second semi-nested PCR reactions, respectively (see Section 2.7).

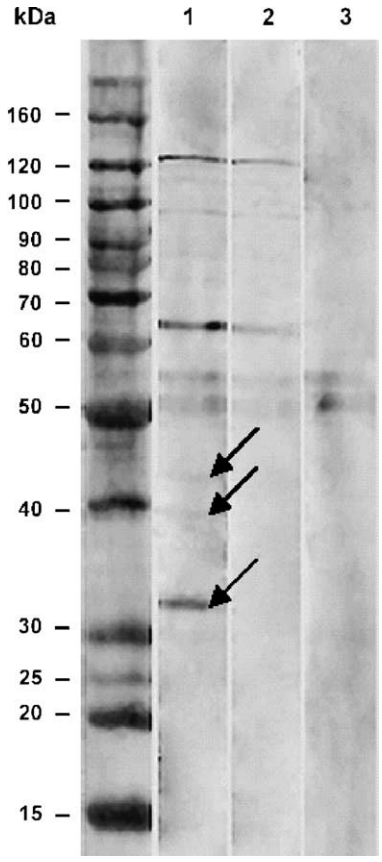

Fig. 5. Duck polyclonal antibodies to pro-TrCB1.1 react with a 33 kDa antigen in *Trichobilharzia regenti* schistosomula extracts. Worm extract (20 μl; from 300 worms) was resolved by 10% SDS-PAGE, electroblotted to introcellulose membrane and incubated with serum from ducks immunised with pro-TrCB1.1 (lane 1), non-immunised mice (lane 2) or with anti-duck IgG alone (lane 3). The reaction with a 33 kDa protein (lane 1, lower arrow) probably represents the fully processed form of TrCB1, whereas the weaker reaction with two bands at 39 and 42 kDa (also arrowed) may represent zymogen forms.

the yeast *P. pastoris*. For both peptidases, SDS-PAGE and Coomassie-staining of *Pichia* induction medium visualised a number of proteins migrating between 38 and 45 kDa (not shown). Amino terminal sequencing by Edman degradation of these proteins yielded the same sequence, E N E I Q F E P, thus confirming their identity as the expected heterologous expression products. The differing molecular masses of these expression products suggests differential glycosylation. Protein yields were in the order of 20–30 and 80–90 mg l^{-1} of induction medium for the TrCB1.1 and TrCB1.4 zymogens, respectively. Both expressed zymogens degraded the peptidyl substrates Z-Phe-Arg-AMC and Z-Arg-Arg-AMC (see below), and reacted with the inhibitors E-64, CA-074 and DCG-04.

3.6. Identification and localisation of TrCB1 in the gut of schistosomula

Mouse and duck polyclonal antisera raised to recombinant TrCB1.1 and 1.4 zymogens reacted with both recombinant enzymes in immunoblots (not shown). With soluble schistosomula extracts, both antisera reacted with a protein species of 33 kDa (Fig. 5, lane 1; duck antiserum only shown). Weak reactions were also detected with two proteins at 39 and 43 kDa. Pre-immune sera and secondary anti-duck IgG did not react (Fig. 5, lanes 2 and 3, respectively). By fluorescence microscopy, mouse and duck polyclonal anti-TrCB1.4 sera reacted exclusively with the lumenal surface of the schistosomulum gastrodermis (Fig. 6A; duck antiserum only shown). No signal was observed in the tegument and parenchyma. A similar reaction to the gut surface was seen with sera collected from ducks 23 days after an experimental infection with 100–200 *T. regenti* cercariae (not shown). Pre-immune serum did not react (Fig. 6B).

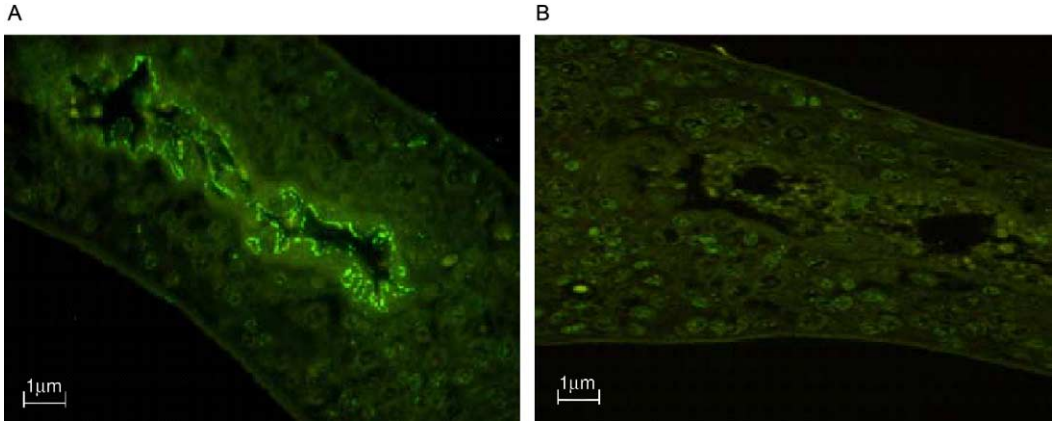

Fig. 6. TrCB1 is expressed by *Trichobilharzia regenti* schistosomula exclusively on the lumenal surface of the gastrodermis. (A) Longitudinal section (2.5 μM) of a 6-day old *T. regenti* schistosomulum incubated with duck anti-pro-TrCB1.4 and then FITC-labelled anti-duck IgG. By fluorescence microscopy, the reaction is most apparent at the lumenal surface of the gastrodermis. (B) Same as in (A), but using pre-immune duck serum. No reaction is visible.

3.7. Auto and heterologous processing of recombinant pro-TrCB1.1 and 1.4

For both TrCB1.1 and 1.4 zymogens, attempts were made at processing and activation using either auto-activation at acid pH or heterologous activation with SmAE (Sajid et al., 2003). Processing was visualised by immunoblotting using mouse anti-pro-TrCB1.1 serum (Fig. 7A) and activation measured with the peptidyl substrate Z-Phe-Arg-AMC (Fig. 7B). Pro-TrCB1.1 (Fig. 7A, lane 1) was partially processed and activated (four-fold) by incubation (up to 12 h) at pH 4.5 in the presence of DTT (Fig. 7A, lane 2; Fig. 7B). In contrast, within 4 h of incubation of the same zymogen with SmAE, major processed products at 32 and 35 kDa (Fig. 7A, lane 3) were detected with a 10-fold increased activity against Z-Phe-Arg-AMC (Fig. 7B).

Pro-TrCB1.4 (Fig. 7A, lane 4), displayed a reversed activation profile. Overnight incubation at pH 4.0 converted the zymogen to products of 32 and 35 kDa (Fig. 7A, lane 5) with an eight-fold increase in peptidolytic activity (Fig. 7B). SmAE failed to significantly trans-process and activate pro-TrCB1.4 (Fig. 7A, lane 6; Fig. 7B) even after overnight incubation.

Repeated attempts to amino-terminal sequence the processed forms (32 and 35 kDa) of both TrCB1 products failed. However, their molecular masses approximate the theoretical M_r values of the mature enzymes with and without a single glycosylation event.

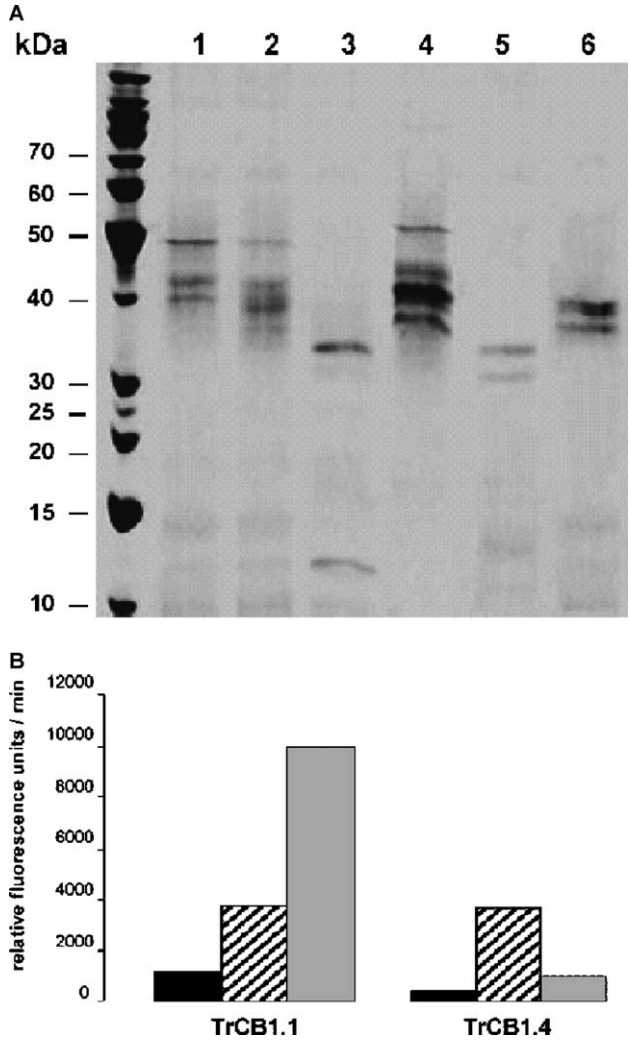

Fig. 7. Processing and activation of recombinant pro-TrCB1.1 and 1.4 in vitro. (A) Processing was visualised by SDS-PAGE (12%) of ~10 μg of both enzymes and immunoblotting using mouse anti-pro-TrCB1.1 serum. Lane 1, pro-TrCB1.1 (doublet at 40 and 42 kDa); lane 2, pro-TrCB1.1 incubated overnight at pH 4.5 and 37 °C; lane 3, pro-TrCB1.1 incubated with *Schistosoma mansoni* asparaginyl endopeptidase (SmAE) for 4 h at pH 4.5; lane 4, pro-TrCB1.4 (triplet at 38, 41 and 43 kDa); lane 5, pro-TrCB1.4 overnight at pH 4.0 and 37 °C; lane 6, pro-TrCB1.4 incubated with SmAE overnight at pH 4.5 and 37 °C. (B) Recombinant pro-TrCB1.1 and 1.4 were subjected to different incubation conditions before their activities were measured at their respective pH optima of 5.5 and 4.0 with Z-Phe-Arg-AMC as substrate. Activities of zymogens are indicated in black. Activities of zymogens incubated for either 4 h at pH 4.5 (TrCB1.1) or overnight at pH 4.0 (TrCB1.4) are indicated by stripes and activities of zymogens incubated with SmAE at pH 4.5 for 4 h (TrCB1.1) are stippled. Assays were in triplicate and the standard deviation about each mean value was never >3%.

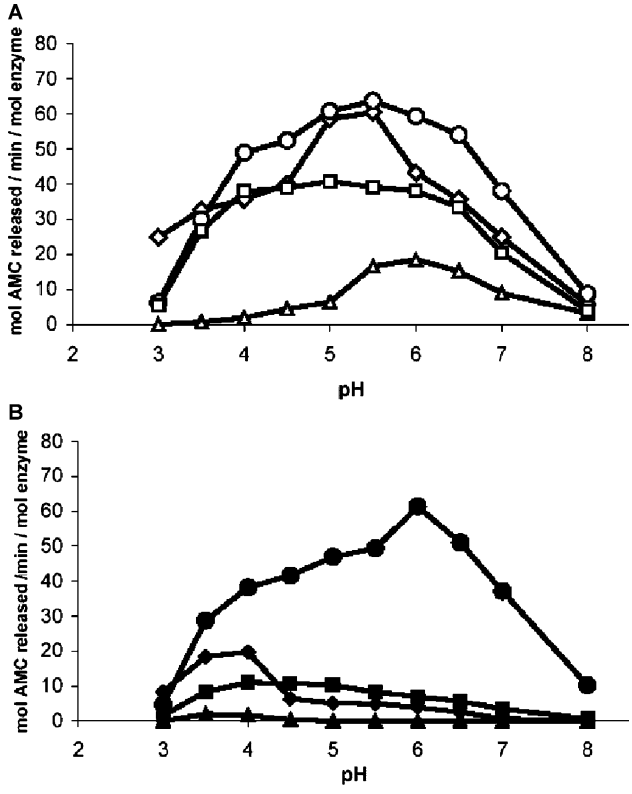

Fig. 8. pH optimum and S_2 subsite specificity of recombinant TrCB1.1 and TrCB1.4. The effect of pH was measured with each of four substrates of the form Z-X-Arg-AMC where X is Phe (diamonds), Arg (triangles), Val (circles) or Leu (squares). (A) TrCB1.1 (0.1 μM) that had been heterologously processed by SmAE and (B) TrCB1.4 (0.1 μM) that had been acid-activated at pH 4.0. The standard deviation about each mean was <5%.

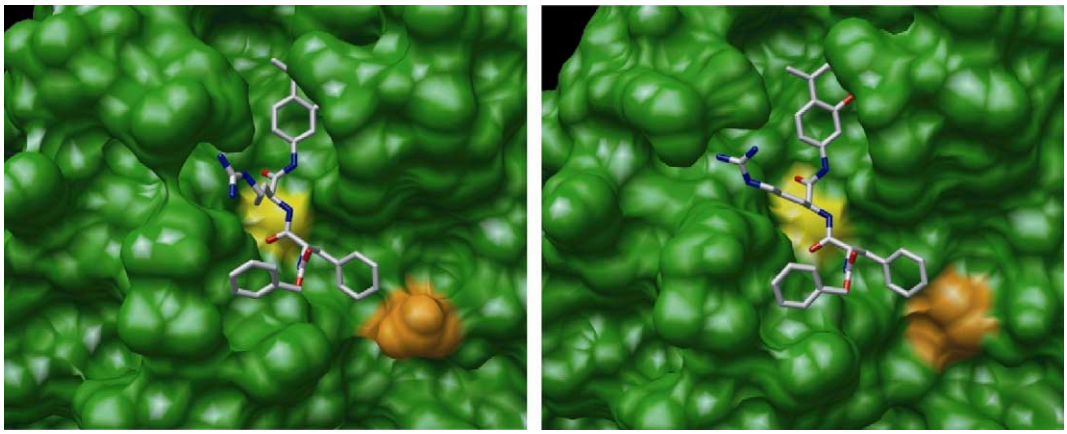

Fig. 9. Comparative protein structure model of TrCB1.1 (left panel) and TrCB1.4 (right) in complex with the substrate Z-Phe-Arg-AMC. The solvent-excluded molecular surface of the active site cleft of TrCB1.4 has a larger S_2 subsite than TrCB1.1 due to the change, $Glu^{245} \rightarrow Ala^{245}$ (indicated in orange). The molecular surface corresponding to the catalytic Cys^{29} is displayed in yellow.

3.8. TrCB1.1 and 1.4 differ in their S_2 subsite specificities

Four dipeptidyl AMC substrates with Arg at P_1 and different residues at P_2 (Phe, Arg, Val and Leu) were employed for testing the S_2 subsite specificity of recombinant processed TrCB1 isoforms. For TrCB1.1, the pH optimum for hydrolysis of substrates with Phe, Leu and Val at P_2 was between 4.5 and 6.5 (Fig. 8A). Val at P_2 was the preferred substrate, but peptides with P_2 Phe and Leu were also efficiently hydrolysed. Hydrolysis of the substrate with Arg at P_2 was least preferred; however, the pH optimum was shifted more to the neutral range than with the other three substrates. In contrast to TrCB1.1, TrCB1.4 had an overriding preference for Val at P_2 with a pH optimum between 5.0 and 6.5 (Fig. 8B). Hydrolysis of substrates with Phe and Leu at P_2 was less than 25% of that with Z-Val-Arg-AMC. Negligible hydrolysis of Z-Arg-Arg-AMC was detected.

3.9. 3D models are consistent with substrate specificity differences at P_2

As expected from the conservative nature of comparative modelling, the three-dimensional models of TrCB1.1 and TrCB1.4 share high overall structural similarity with their two template structures, human pro-cathepsin B and bovine cathepsin B in complex with the inhibitor NS-134. This similarity extends to the binding site where the sequence conservation is much higher than in the rest of the protein (Fig. 9). The only significant difference between the active site clefts in the two modelled isoforms of TrCB1 occurs in the S_2 subsite. Here, the Glu residue at position 245 in TrCB1.1 is replaced by Ala in TrCB1.4. In both models, the S_2 subsite is wide enough to accommodate any of the tested P_2 residues (Phe, Leu, Val, and Arg; Z-Phe-Arg-AMC is depicted in Fig. 9). Nevertheless, the deletion of a carboxylic group from Glu^{245} in TrCB1.1 to obtain Ala^{245} in TrCB1.4 involves a relatively large change in both charge and volume.

3.10. TrCB1.1 and 1.4 degrade mMBP, but hemoglobin is a poor substrate

Activated recombinant TrCB1.1 and 1.4 degraded mMBP within the pH range of 4.0–7.0 (Fig. 10A, only results with TrCB1.4 shown). Neither TrCB1 isoform efficiently degraded duck, turkey or bovine hemoglobin, even after overnight incubation at pH values between 4.0 and 7.0 (Fig. 10B; duck hemoglobin shown).

4. Discussion

Upon penetration of the skin, *T. regenti* schistosomula migrate through the peripheral and central nervous systems of the avian host prior to establishment in the nasal cavity. This is in contrast to other members of the family Schistosomatidae that infect humans and which, after skin-penetration, require the host vascular system to migrate and eventually establish infection. Though previous studies have reported the presence of serine peptidases in free-swimming cercariae of the visceral trichobilharzia, *Trichobilharzia ocellata* (Bahgat and Ruppel, 2002), it is likely that such peptidases are exhausted after entry into the skin and that peptidases actively expressed by schistosomula are required for subsequent migration through the nervous system. A number of gut-associated peptidases (e.g. cathepsins B, C, D, F, and L) are present in schistosomes of humans (Caffrey et al., 2004 for review) and similar enzymes may facilitate alimentation by and migration of *T. regenti* schistosomula.

Biochemical study of migrating *T. regenti* schistosomula is limited by the difficulty in obtaining larvae from nervous tissue. Therefore, we applied recently developed, active site DCG-04 probes that have been employed previously to identify peptidase species in small tissue samples (Greenbaum et al., 2000, 2002). Peptidases were also analysed with selective peptidyl inhibitors and fluorogenic

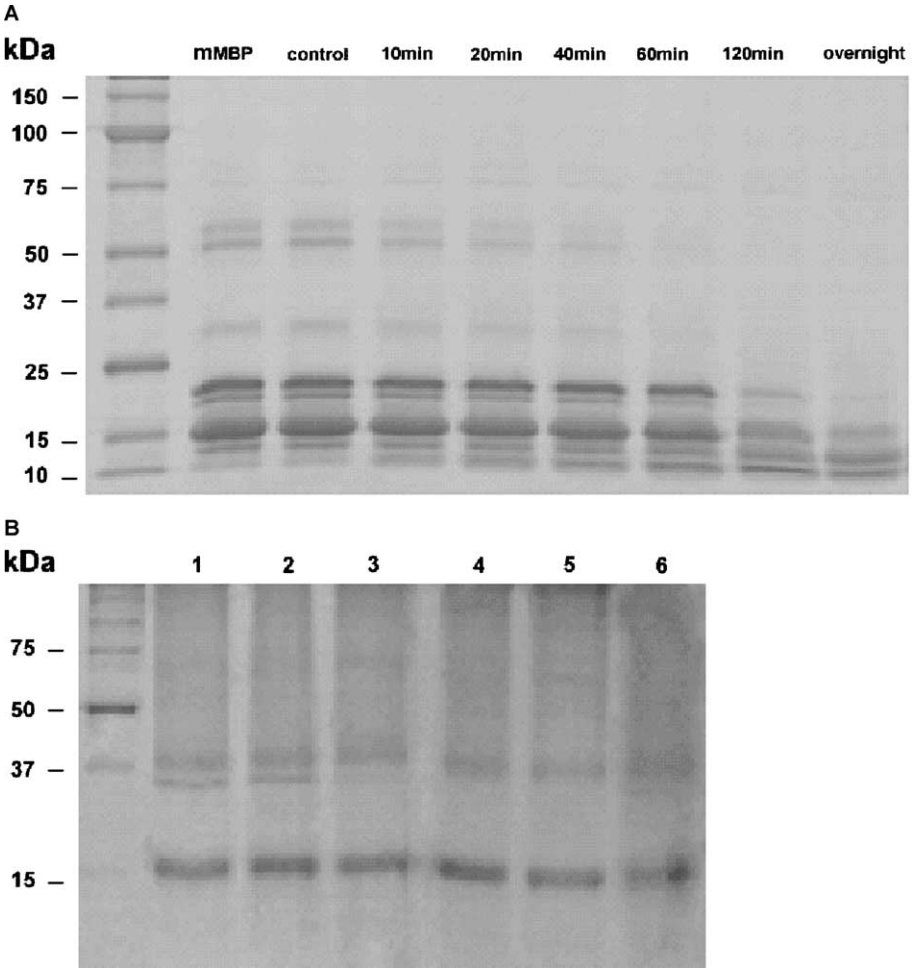

Fig. 10. (A) Time-dependent degradation of mouse myelin basic protein by TrCB1.4. Acid-activated TrCB1.4 (~1 μM) was incubated with 0.3 mg/ml mMBP at 37 °C for 10 min overnight in 50 mM citrate, 100 sodium phosphate, 2 mM DTT, pH 7.0, as indicated. As a control, mMBP was incubated overnight without enzyme under the same reaction conditions. The same time-dependent course of degradation was also observed for heterologously-activated TrCB 1.1 (not shown). (B) Hemoglobin is a poor substrate for either acid-activated TrCB1.4 or heterologously-activated TrCB1.1. Enzymes (~1 μM) were incubated overnight at 37 °C with 0.5 mg/ml duck hemoglobin in 50 mM citrate, 100 sodium phosphate, 2 mM DTT. Lanes 1 and 2, hemoglobin incubated without enzyme at pH 4.5 and 7.0, respectively; lanes 3 and 4, hemoglobin incubated at pH 4.5 with TrCB1.4 and TrCB1.1, respectively; lanes 5 and 6, hemoglobin incubated at pH 7.0 with TrCB1.4 and TrCB1.1, respectively. For both (A) and (B), hydrolysis was visualised by 10% SDS-PAGE and staining with Coomassie Brilliant Blue.

substrates with high extinction coefficients. Using these tools, a major 33 kDa cathepsin B cysteine peptidase was identified in schistosomula extracts. The identification of a major cathepsin B activity in *T. regenti* is comparable with previous findings of predominant cathepsin B enzymes in migratory and adult schistosome parasites of humans (Caffrey et al., 1997; Caffrey and Ruppel, 1997; Sajid et al., 2003), and migratory juveniles of *Fasciola hepatica* (Law et al., 2003) and *Fasciola gigantica* (Meemon et al., 2004). For *Fasciola*, cathepsins B have been proposed to aid nutrition and migration of juvenile parasites by hydrolysis of host-tissues (Law et al., 2003; Meemon et al., 2004) and it is possible that the *T. regenti* orthologues reported here fulfill similar functions.

We also applied a degenerate PCR approach to identify and characterise *T. regenti* cathepsin B genes. Degenerate forward and reverse primers were designed around the conserved Cys and Asn residues, respectively, of Clan CA peptidases with particular reference to the relevant sequence information from three schistosome cathepsins B; SmCB1, SjCB1 and SmCB2. The single PCR product amplified and subsequent 5′ and 3′ RACE reactions allowed the eventual identification of a cathepsin B-like sequence. Nucleotide sequence differences were evident and sequence screening of 40 clones identified six isoforms, termed TrCB1.1–1.6, with TrCB1.1 as the most abundant at 68%.

A multiple sequence alignment of the six TrCB1 isoforms confirmed their orthology to the schistosomal cathepsin B enzymes, SmCB1 and SjCB1, with ~70% sequence identity. In addition, the immunolocalisation of TrCB1 to the gut of *T. regenti* using antisera to recombinant proteins is similar to that observed for SmCB1 (Sajid et al., 2003) and SjCB1 (Caffrey and Ruppel, 1997). Therefore, as has been suggested for the *Schistosoma* orthologues

(Caffrey et al., 2004), the TrCB1 isoforms likely function in the hydrolysis of ingested proteins for nutrition. *Schistosoma* cathepsins B1 are naturally immunogenic and both mouse and human infection sera react strongly with the parasite gastrodermis that corresponds with reactivity to a 31 kDa antigen (Ruppel et al., 1987). Our observation of a strong reaction in the gut of *T. regenti* with duck infection sera, suggests that TrCB1 is likewise naturally immunogenic, perhaps by the periodic regurgitation along with gut contents into host tissues. Moreover, antibodies from sera of naturally infected animals recognise TrCB1.1 and 1.4 recombinant isoforms on immunoblots (Mikeš et al., personal communication). Interestingly, two of the six TrCB1 isoforms, TrCB1.5 and 1.6, possessed a substitution of the catalytic cysteine by a glycine. Consequently, both isoforms are unlikely to be enzymatically active and their function in the biology of the parasite remains unclear. Substitution of the catalytic cysteine residue has been previously reported for the *Schistosoma* cysteine peptidase genes, SjCB1 (Merckelbach et al., 1994) and SmAE (Caffrey et al., 2000), and other putative examples can be found in the *S. mansoni* (Verjovski-Almeida et al., 2003; http://cancer.lbi.ic.unicamp.br/schisto6) and *Schistosoma japonicum* EST databases (Hu et al., 2003; http://schistosoma.chgc.sh.cn). A recent study has shown that enzymes with active site mutations are more the rule than exception in all of the investigated metazoan species (Pils and Schultz, 2004). Such mutations occur in a large number of enzyme families, including the Clan CA peptidases and it was proposed that these inactive enzyme homologues serve to regulate the functions of active enzymes by competing for substrate.

Two catalytically competent isoforms, TrCB1.1 (as the most abundant) and TrCB1.4 (as the most divergent), were expressed as active recombinant peptidases in *P. pastoris*. Both zymogens were secreted as a series of proteins with intact pro-peptide domains and M_r values between 38 and 45 kDa, perhaps due to differential glycosylation of the requisite sites in the pro-peptide and/or mature domain.

Both auto- and heterologous activation protocols were employed to process the TrCB1.1 and 1.4 zymogens to their mature forms. Auto-activation at acid pH has been shown to activate Clan CA mammalian (Turk et al., 2000) and trematode procathepsins (Law et al., 2003; Dalton et al., 2003). For heterologous activation, we used recombinant SmAE (Caffrey et al., 2000), which, like mammalian AE, only cleaves substrates with an available Asn at P_1 (Mathieu et al., 2002). SmAE has been shown to process and activate recombinant pro-SmCB1 at an Asn residue located between the pro-peptide and catalytic domain (Sajid et al., 2003). A similarly positioned Asn site was identified here in the pro-TrCB1 isoforms (except TrCB1.5) 4 amino acid residues upstream of the predicted catalytic domain, thus suggesting the feasibility of heterologous processing by SmAE. Interestingly, pro-TrCB1.1 and 1.4 showed contrasting processing profiles. The former was incapable of complete acid auto-activation, but was heterologously processed by SmAE to final products of 32 and 35 kDa with an increase in specific activity. TrCB1.4, on the other hand, could auto-activate to yield final products of the same molecular weights, whereas processing by SmAE was only partial, without a significant increase in specific activity. For both processed TrCB1 isoforms, the molecular masses of the final products (32 and 35 kDa) are consistent with that found for recombinant activated SmCB1 with and without a single glycosylation event (Sajid et al., 2003). The reasons for the contrary processing profiles of pro-TrCB1 and 1.4 are unclear. It is possible that the Asn residue in pro-TrCB1.1 and targeted by SmAE is more solvent-exposed and, thus, more accessible to hydrolysis. Alternatively, because the same Asn residue is a potential glycosylation site, any preferential glycosylation of pro-TrCB1.4 by *Pichia* would sterically hinder access of SmAE to the zymogen.

Because the S_2 subsite is a major determinant of the overall specificity of Clan CA peptidases (Sajid and McKerrow, 2002), we tested both TrCB1 isoforms with a number of dipeptidyl AMC substrates, varying in Phe, Leu, Val and Arg at P_2, over a range of pH values. TrCB1.1 displayed the P_2 preference of Val > Phe > Leu > Arg at acid pH with activity against Arg increasing at more neutral pH. In contrast, TrCB1.4 displayed the P_2 preference of Val ≫ Leu > Phe > Arg, with essentially no activity against Arg at any pH.

We built comparative 3D models of TrCB1.1 and TrCB1.4 in an effort to rationalise, in molecular terms, at least some of the differences measured in their substrate preferences. The only significant variation in the active site clefts of the two modelled isoforms occurs in the S_2 subsite. The negatively charged Glu^{245} in TrCB1.1 is replaced by a neutral and smaller Ala in TrCB1.4. The corresponding deletion of a carboxylic group from Glu^{245} in TrCB1.1 to obtain Ala^{245} in TrCB1.4 involves a relatively large change in both charge and volume. Thus, the comparative models are consistent with the empirical finding that the substrate preferences at the P_2 position and their pH dependence are different between TrCB1.1 and TrCB1.4. In particular, the models help rationalise the inactivity of TrCB1.4 against Arg at P_2 as follows. Presumably, TrCB1.1 is active against substrates with Arg at P_2 because the negatively charged Glu^{245} at the base of the S_2 subsite forms a productive electrostatic interaction with the positively charged Arg sidechain (Chan et al., 1999; Sajid and McKerrow, 2002). In comparison, the neutral and smaller Ala^{245} in TrCB1.4 may not interact productively with the Arg sidechain.

The overriding AMC substrate preference of TrCB1.4 for Val at P_2 is unusual and, to our knowledge, has not yet been demonstrated for a Clan CA cysteine peptidase. This result can be rationalised based on the modelled molecular interactions between the enzyme and the substrate, particularly at S_2. Here, the mutation from Glu^{245} to Ala does not change the activity against the substrate with Val at P_2, because Val is too small to be shielded from water

molecules by either Glu^{245} or Ala^{245}. In contrast, the decreased activity with substrates containing the larger Phe or Leu residues at P_2 is presumably due to a weakened hydrophobic effect (i.e. greater exposure to interfering water molecules) when Glu^{245} is mutated to the smaller Ala. Further enzymatic and computational studies, involving additional peptidyl substrates and inhibitors examined on their own and in complex with both isoforms in different states (e.g. substrate–enzyme complex, transition state, and enzyme–product complex) are needed to understand the novel specificity of TrCB1.4. For the migrating parasite, the different substrate specificities of TrCB1.1 and 1.4 may offer increased hydrolytic efficiency and specificity against host proteins.

Because *T. regenti* schistosomula utilise a very different migratory route to that of schistosome parasites of humans, we tested two protein substrates for degradation by recombinant activated TrCB1.1 and 1.4: mMBP (the major protein of myelin) and hemoglobin (a major blood protein). Both TrCB1 isoforms degraded mMBP efficiently, however, hemoglobin was poorly hydrolysed. This might reflect an adaptation by the parasite for degradation of the relevant myelin substrate. Hemoglobin, on the other hand, would be seldom encountered by schistosomula during migration through the CNS, being available only in localised vascular damage. In support of this, we have observed that $<5\%$ of the *T. regenti* schistosomula harvested from the CNS contain hematin (black product of hemoglobin-digestion) in the gut. Hemoglobin would only become an important nutritive source after maturation of parasites in the nasal cavity and the commencement of blood-feeding (Horák et al., 1999).

In conclusion, we have identified and characterised six cathepsin B isoforms responsible for the predominant cysteine peptidase activity in migratory schistosomula of the neuropathogenic trematode, *T. regenti*. Clear parallels are evident with the orthologous cathepsins B of *Schistosoma* spp. and invasive juveniles of *Fasciola* spp. in terms of tissue-localisation and inferred biological functions. However, the presence of multiple enzyme isoforms with differing specificities, against both peptidyl and protein substrates, suggest particular adaptations to parasitism of the CNS. These peculiarities are the subject of ongoing research.

Acknowledgements

This work was supported by the Sandler Family Supporting Foundation; the NIAID (AI053247, AI035707), the Welcome Trust—Collaborative Research Initiative Grant (No. 072255; jointly with A.P. Mountford (Dept. of Biology, The University of York, UK)); the Grant Agency of the Czech Republic (GACR 524/03/1263 + GACR 524/04/P082), the Charles University Grant Agency (No. 263/2004/B-BIO/PrF) and the Czech Ministry of Education grants (J13/981131-3 + J13/981131-4). J.D. is a holder of a Czech Literary Fund travel grant and C.R.C. a recipient of a UCSF School of Medicine Research Evaluation and Allocation Committee Award. We gratefully aknowledge AboaTech Ltd. and J.A. Maatta for supplying the mouse myelin basic protein, and Dr Dieter Brömme of the Department of Human Genetics, Mount Sinai School of Medicine for the gift of Z-Leu-Arg-AMC and Z-Val-Arg-AMC. Our thanks to Dr Jan Konvalinka of the Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, for facilitating the initiation of this study.

References

Bahgat, M., Ruppel, A., 2002. Biochemical comparison of the serine protease (elastase) activities in cercarial secretions from *Trichobilharzia ocellata* and *Schistosoma mansoni*. Parasitol. Res. 88, 495–500.

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

Caffrey, C.R., McKerrow, J.H., 2004. Helminth cysteine proteases. In: Barrett, A.J., Rawlings, N.D., Woessner, J.F. (Eds.), Handbook of Proteolytic Enzymes, vol. 2. Elsevier, Amsterdam, pp. 1183–1186.

Caffrey, C.R., Ruppel, A., 1997. Affinity isolation and characterization of the cathepsin B-like proteinase Sj31 from *Schistosoma japonicum*. J. Parasitol. 83, 1112–1118.

Caffrey, C.R., Rheinberg, C.E., Mone, H., Jourdane, J., Li, Y.L., Ruppel, A., 1997. *Schistosoma japonicum*, *S. mansoni*, *S. haematobium*, *S. intercalatum*, and *S. rodhaini*: cysteine-class cathepsin activities in the vomitus of adult worms. Parasitol. Res. 83, 37–41.

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

Caffrey, C.R., Salter, J.P., Lucas, K.D., Khiem, D., Hsieh, I., Lim, K.C., Ruppel, A., McKerrow, J.H., Sajid, M., 2002. SmCB2, a novel tegumental cathepsin B from adult *Schistosoma mansoni*. Mol. Biochem. Parasitol. 121, 49–61.

Caffrey, C.R., McKerrow, J.H., Salter, J.P., Sajid, M., 2004. Blood 'n' guts: an update on schistosome digestive peptidases. Trends Parasitol. 20, 241–248.

Chan, S.J., San Segundo, B., McCormick, M.B., Steiner, D.F., 1986. Nucleotide and predicted amino acid sequences of cloned human and mouse preprocathepsin B cDNAs. Proc. Natl. Acad. Sci. USA 83, 7721–7725.

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

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., Donnelly, S.M., 2003. *Fasciola hepatica* cathepsin L-like proteases: biology, function, and potential in the development of first generation liver fluke vaccines. Int. J. Parasitol. 33, 1173–1181.

Gerber, P.R., Muller, K., 1995. MAB, a generally applicable molecular force field for structure modelling in medicinal chemistry. J. Comput. Aided Mol. Des. 9, 251–268.

Greenbaum, D., Medzihradszky, K.F., Burlingame, A., Bogyo, M., 2000. Epoxide electrophiles as activity-dependent cysteine protease profiling and discovery tools. Chem. Biol. 7, 569–581.

Greenbaum, D., Baruch, A., Hayrapetian, L., Darula, Z., Burlingame, A., Medzihradszky, K.F., Bogyo, M., 2002. Chemical approaches for functionally probing the proteome. Mol. Cell. Proteomics 1, 60–68.

Horák, P., Kolářová, L., Dvořák, J., 1998. *Trichobilharzia regenti* n. sp. (Schistosomatidae, Bilharziellinae), a new nasal schistosome from Europe. Parasite 5, 349–357.

Horák, P., Dvořák, J., Kolářová, L., Trefil, L., 1999. *Trichobilharzia regenti*, a pathogen of the avian and mammalian central nervous systems. Parasitology 119, 577–581.

Horák, P., Kolářová, L., Adema, C.M., 2002. Biology of the schistosome genus *Trichobilharzia*. Adv. Parasitol. 52, 155–233.

Hrádková, K., Horák, P., 2002. Neurotropic behaviour of *Trichobilharzia regenti* in ducks and mice. J. Helminthol. 76, 137–141.

Hu, W., Yan, Q., Shen, D.K., Liu, F., Zhu, Z.D., Song, H.D., Xu, X.R., Wang, Z.J., Rong, Y.P., Zeng, L.C., Wu, J., Zhang, X., Wang, J.J., Xu, X.N., Wang, S.Y., Fu, G., Zhang, X.L., Wang, Z.Q., Brindley, P.J., McManus, D.P., Xue, C.L., Feng, Z., Chen, Z., Han, Z.G., 2003. Evolutionary and biomedical implications of a *Schistosoma japonicum* complementary DNA resource. Nat. Genet. 35, 139–147.

Illy, C., Quraishi, O., Wang, J., Purisima, E., Vernet, T., Mort, J.S., 1997. Role of the occluding loop in cathepsin B activity. J. Biol. Chem. 272, 1197–1202.

Klinkert, M.Q., Felleisen, R., Link, G., Ruppel, A., Beck, E., 1989. Primary structures of Sm31/32 diagnostic proteins of *Schistosoma mansoni* and their identification as proteases. Mol. Biochem. Parasitol. 33, 113–122.

Kouřilová, P., Syrůček, M., Kolářová, L., 2004. The severity of mouse pathologies caused by the bird schistosome *Trichobilharzia regenti* in relation to host immune status. Parasitol. Res. 93, 534–542.

Law, R.H., Smooker, P.M., Irving, J.A., Piedrafita, D., Ponting, R., Kennedy, N.J., Whisstock, J.C., Pike, R.N., Spithill, T.W., 2003. Cloning and expression of the major secreted cathepsin B-like protein from juvenile *Fasciola hepatica* and analysis of immunogenicity following liver fluke infection. Infect. Immun. 71, 6921–6932.

Maatta, J.A., Coffey, E.T., Hermonen, J.A., Salmi, A.A., Hinkkanen, A.E., 1997. Detection of myelin basic protein isoforms by organic concentration. Biochem. Biophys. Res. Commun. 238, 498–502.

Mathieu, M.A., Bogyo, M., Caffrey, C.R., Choe, Y., Lee, J., Chapman, H., Sajid, M., Craik, C.S., McKerrow, J.H., 2002. Substrate specificity of schistosome versus human legumain determined by P1-P3 peptide libraries. Mol. Biochem. Parasitol. 121, 99–105.

Meemon, K., Grams, R., Vichasri-Grams, S., Hofmann, A., Korge, G., Viyanant, V., Upatham, E.S., Habe, S., Sobhon, P., 2004. Molecular cloning and analysis of stage and tissue-specific expression of cathepsin B encoding genes from *Fasciola gigantica*. Mol. Biochem. Parasitol. 136, 1–10.

Merckelbach, A., Hasse, S., Dell, R., Eschlbeck, A., Ruppel, A., 1994. cDNA sequences of *Schistosoma japonicum* coding for two cathepsin B-like proteins and Sj32. Trop. Med. Parasitol. 45, 193–198.

Murata, M., Miyashita, S., Yokoo, C., Tamai, M., Hanada, K., Hatayama, K., Towatari, T., Nikawa, T., Katunuma, N., 1991. Novel epoxysuccinyl peptides. Selective inhibitors of cathepsin B, in vitro. FEBS Lett. 280, 307–310.

Musil, D., Zucic, D., Turk, D., Engh, R.A., Mayr, I., Huber, R., Popovic, T., Turk, V., Towatari, T., Katunuma, N., 1991. The refined 2.15 A x-ray crystal structure of human liver cathepsin B: the structural basis for its specificity. EMBO J. 10, 2321–2330.

Nielsen, H., Engelbrecht, J., Brunak, S., von Heijne, G., 1997. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Protein Eng. 10, 1–6.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., Ferrin, T.E., 2004. UCSF chimera—a visualization system for exploratory research and analysis. J. Comput. Chem. 25, 1605–1612.

Pils, B., Schultz, J., 2004. Inactive enzyme-homologues find new function in regulatory processes. J. Mol. Biol. 340, 399–404.

Podobnik, M., Kuhelj, R., Turk, V., Turk, D., 1997. Crystal structure of the wild-type human procathepsin B at 2.5 A resolution reveals the native active site of a papain-like cysteine protease zymogen. J. Mol. Biol. 271, 774–788.

Rawlings, N.D., Tolle, D.P., Barrett, A.J., 2004. MEROPS: the peptidase database. Nucleic Acids Res. 32, D160–D164.

Ruppel, A., Shi, Y.E., Wei, D.X., Diesfeld, H.J., 1987. Sera of *Schistosoma japonicum*-infected patients cross-react with diagnostic 31/32 kD proteins of *S. mansoni*. Clin. Exp. Immunol. 69, 291–298.

Sajid, M., McKerrow, J.H., 2002. Cysteine proteases of parasitic organisms. Mol. Biochem. Parasitol. 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., Caffrey, C.R., 2003. Functional expression and characterization of *Schistosoma mansoni* cathepsin B and its trans-activation by an endogenous asparaginyl endopeptidase. Mol. Biochem. Parasitol. 131, 65–75.

Sali, A., Blundell, T.L., 1993. Comparative protein modelling by satisfaction of spatial restraints. J. Mol. Biol. 234, 779–815.

Stern, I., Schaschke, N., Moroder, L., Turk, D., 2004. Crystal structure of NS-134 in complex with bovine cathepsin B: a two-headed epoxysuccinyl inhibitor extends along the entire active-site cleft. Biochem. J. 381, 511–517.

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

Verjovski-Almeida, S., DeMarco, R., Martins, E.A., Guimaraes, P.E., Ojopi, E.P., Paquola, A.C., Piazza, J.P., Nishiyama Jr., M.Y., Kitajima, J.P., Adamson, R.E., Ashton, P.D., Bonaldo, M.F., Coulson, P.S., Dillon, G.P., Farias, L.P., Gregorio, S.P., Ho, P.L., Leite, R.A., Malaquias, L.C., Marques, R.C., Miyasato, P.A., Nascimento, A.L., Ohlweiler, F.P., Reis, E.M., Ribeiro, M.A., Sa, R.G., Stukart, G.C., Soares, M.B., Gargioni, C., Kawano, T., Rodrigues, V., Madeira, A.M., Wilson, R.A., Menck, C.F., Setubal, J.C., Leite, L.C., Dias-Neto, E., 2003. Transcriptome analysis of the acoelomate human parasite *Schistosoma mansoni*. Nat. Genet. 35, 148–157.

Wasilewski, M.M., Lim, K.C., Phillips, J., McKerrow, J.H., 1996. Cysteine protease inhibitors block schistosome hemoglobin degradation in vitro and decrease worm burden and egg production in vivo. Mol. Biochem. Parasitol. 81, 179–189.

Wilkins, M.R., Lindskog, I., Gasteiger, E., Bairoch, A., Sanchez, J.C., Hochstrasser, D.F., Appel, R.D., 1997. Detailed peptide characterization using PEPTIDEMASS—a world-wide-web-accessible tool. Electrophoresis 18, 403–408.