Available online at www.sciencedirect.com
ScienceDirect
The International Journal of Biochemistry & Cell Biology 40 (2008) 1389–1398
IJBCB
www.elsevier.com/locate/biocel

Cystatins from filarial parasites: Evolution, adaptation and function in the host–parasite relationship☆

William F. Gregory, Rick M. Maizels *

Institute of Immunology and Infection Research, Ashworth Laboratories, University of Edinburgh, West Mains Road, Edinburgh EH9 3JT, UK

Available online 4 December 2007

Abstract

Cystatins, together with stefins and kininogens, are members of the cystatin superfamily of cysteine protease inhibitors (CPI) present across the animal and plant kingdoms. Their role in parasitic organisms may encompass both essential developmental processes and specific interactions with the parasite's vector and/or final host. We summarise information gathered on three cystatins from the human filarial nematode *Brugia malayi* (*Bm*-CPI-1, -2 and -3), and contrast them those expressed by other parasites and by the free-living nematode *Caenorhabditis elegans*. *Bm*-CPI-2 differs from *C. elegans* cystatin, having acquired the additional function of inhibiting asparaginyl endopeptidase (AEP), in a manner similar to some human cystatins. Thus, we propose that *Bm*-CPI-2 and orthologues from related filarial parasites represent a new subset of nematode cystatins. *Bm*-CPI-1 and CPI-3 share only 25% amino acid identity with *Bm*-CPI-2, and lack an evolutionarily conserved glycine residue in the N-terminal region. These sequences group distantly from the other nematode cystatins, and represent a second novel subset of filarial cystatin-like genes. Expression analyses also show important differences between the CPI-2 and CPI-1/-3 groups. *Bm-cpi-2* is expressed at all time points of the parasite life cycle, while *Bm-cpi-1* and *-3* expression is confined to the late stages of development in the mosquito vector, terminating within 48 h of infection of the mammalian host. Hence, we hypothesise that CPI-2 has evolved to block mammalian proteases (including the antigen-processing enzyme AEP) while CPI-1 and -3 function in the milieu of the mosquito vector necessary for transmission of the parasite.

Keywords: Helminths; Protease inhibitor; Immune evasion; Co-evolution

1. Introduction

The human filarial parasite *Brugia malayi* is an arthropod-borne nematode which causes lymphatic filariasis, an infection afflicting more than 120 million people worldwide (Michael, Bundy, & Grenfell, 1996; Nutman, 1999). Transfer from the mosquito is facilitated through the puncture wound produced during blood feeding after which the infective larvae (L3) migrate through the capillary system to the lymphatic tissues (Scott, 1999). During the following weeks they develop into long-lived adult parasites, releasing thousands of L1 offspring (microfilariae) which migrate to the blood stream awaiting uptake into a mosquito bloodmeal. The mammalian stages of

Abbreviations: AEP, asparaginyl endopeptidase; CPI, cysteine protease inhibitor; EST, expressed sequence tag; SL, spliced leader.

☆ *Nomenclature note*: The convention, established for *Caenorhabditis elegans* and followed in this manuscript, is that genes and mRNA transcripts are denoted by lower case italics, and proteins and protein sequences by upper case Roman type. Each gene is given an italicised two-letter binomial corresponding to the genus/species followed by a three-letter gene designation, and a numeral when required.

* Corresponding author. Tel.: +44 131 650 5511; fax: +44 131 650 5450.
E-mail address: r.maizels@ed.ac.uk (R.M. Maizels).

1357-2725/$ – see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.biocel.2007.11.012

the parasite are long-lived, with adult worms surviving for 5 years or more, generating considerable interest in their ability to suppress host immune effector mechanisms (Maizels & Yazdanbakhsh, 2003; Maizels et al., 2004; Semnani & Nutman, 2004).

The worm's life cycle is punctuated by four obligatory moults, a process involving the synthesis of a new cuticle beneath the existing one, followed by the ecdysis of the old cuticle. This process occurs twice within the mosquito and twice in the human host. In filariae, cysteine proteases have been implicated in the moulting process; small synthetic inhibitors of cysteine proteases arrest moulting after the production of the new cuticle but prior to ecdysis, resulting in morphological abnormalities in the cuticle (Lustigman et al., 1996). Because cysteine proteases control a myriad of other physiological processes, including lysosomal degradation and inflammation, their activity is tightly controlled by a family of specific cysteine protease inhibitors, the cystatins (Abrahamson, Alvarez-Fernandez, & Nathanson, 2003).

Cystatins are protein inhibitors of cysteine proteases found in a wide range of metazoan and plant taxa (Kotsyfakis et al., 2006; Margis, Reis, & Villeret, 1998). On the basis of primary sequence homology the cystatin superfamily has been sub-divided into three families, stefins (family 1), cystatins (family 2) and kininogens (family 3) (Abrahamson et al., 2003). Stefins are predominantly intracellular non-glycosylated proteins that lack disulphide bonds while cystatins possess two disulphide bonds, are generally found in extracellular spaces and may be glycosylated and/or phosphorylated; both family 1 and family 2 inhibitors are low molecular weight (10–15 kDa) proteins. Kininogens are larger plasma proteins containing multiple cystatin-type domains. In this review, we will focus on the family 2 cystatins from nematode organisms, and propose that they fall into three subgroups, two of which contain novel structural features for invertebrate cystatins.

1.1. Identification of filarial nematode cystatins

The nematode phylum includes both free-living organisms, such as *Caenorhabditis elegans*, and numerous parasitic species, including the group of filarial parasites which are the focus of this review. Although the genome of *C. elegans* was the first metazoan to be fully sequenced, nematode cystatins were first recognised among the filarial parasites in which protein- and mRNA-led investigations revealed prominent expression of these inhibitors. In the human filarial nematode *B. malayi*, *Bm-cpi-1* was identified by its high level of expression among transcripts from the infective, mosquito-borne third-stage larvae (L3) (Gregory, Blaxter, & Maizels, 1997), with closest sequence similarity to family 2 cystatins (Fig. 1) Interestingly, cpi-1 cDNA is *trans*-spliced with the 22-nt nematode spliced leader sequence (SL-1) at the 5′ end immediately before the start codon and a 19-aa predicted signal sequence. Antibodies to the 108-aa mature CPI-1 have been used to confirm that the gene product was synthesized and exported to the larval parasite surface, and indeed secreted in vitro by cultured larval parasites (Gregory, unpublished data).

A survey of ESTs from the same parasite confirmed the abundance of *cpi-1* (0.55% of all L3 ESTs), and also identified a related cystatin, designated *Bm-cpi-2* which is similarly *trans*-spliced with SL-1. Antibodies to CPI-2 demonstrated that this cystatin is most associated with the mature, adult worm, in which the protein product can be shown to be located on the parasite surface, as well as in *in vitro* secretions (Gregory, unpublished data). Consistent with its expression by the adult form of the parasite, which is able to survive for many years in an immuno-competent host, CPI-2 appears to have evolved to play a precise role in parasite immune evasion (Maizels, Gomez-Escobar, Gregory, Murray, & Zang, 2001; Manoury, Gregory, Maizels, & Watts, 2001).

Most recently, with the publication of the *B. malayi* genome sequence (Ghedin et al., 2007), we have identified a third cystatin (*Bm-cpi-3*). In parallel, several investigators have reported cystatin homologues from other filarial parasite species. In the case of *Onchocerca volvulus*, the cause of human river blindness, we identified *Ov*-CPI-1 as a homologue from EST data (Gregory et al., unpublished data), as well as *Ov*-CPI-2 which is also termed Onchocystatin (Lustigman, Brotman, Huima, & Prince, 1991; Lustigman, Brotman, Huima, Prince, & McKerrow, 1992; Schönemeyer et al., 2001). Further family members have been described in filarial parasites of rodents, *Acanthocheilonema viteae* (*Av*-CPI), reported to be highly secreted by female adult worms (Hartmann, Kyewski, Sonnenburg, & Lucius, 1997) and *Litomosoides sigmodontis* (*Ls*-CPI) (Pfaff et al., 2002). We discuss below the structural, functional and expression characteristics of these filarial cystatins in relation to homologues from other organisms.

1.2. Analyses of cystatin protein sequences

Although there is a low overall level of sequence identity (14–23%) between the filarial cystatins and the vertebrate members of family 2, all the key structural features of this group are conserved. These features are: (1) an N-terminal signal peptide; (2) a single domain

structure of ∼100 aa; (3) two internal disulphide bonds; (4) an invariant glycine residue within the first 10–15 residues of the mature protein; (5) a central Gln-X-Val-X-Gly motif; (6) a C-terminal Pro-Trp hairpin loop pair. The latter three regions together protrude into the active site cleft of the protease to achieve inhibition (Bode et al., 1988). The filarial cystatins conform with each of these structural patterns, except that only one disulphide bond exists, as the second C-terminal cysteine residue pair that spans the P-W residues is absent (Fig. 1A).

More surprisingly, *Bm*-CPI-1 and CPI-3 are missing the N-terminal glycine residue, with the replacement of a glycine by a serine residue at position 6 in the mature *Bm*-CPI-1 (Fig. 1A). In solution, the lack of any side chain in Gly-6 is thought to confer structural flexibility to the N-terminal trunk of the protein. In complexes with proteases, however, the N-terminus is immobilized by interactions of residues directly preceding the glycine with the substrate-binding sites of cysteine proteases (Bode et al., 1988).

In mammalian cystatins, mutations in residues up to and including the glycine decrease or abolish the ability of the inhibitor to interact with a variety of substrates, as the presence of an amino acid side-chain at this position impedes structural flexibility (Bode et al., 1988; Hall, Dalboge, Grubb, & Abrahamson, 1993; Hall, Hakansson, Mason, Grubb, & Abrahamson, 1995; Shibuya et al., 1995). However, the glycine residue of cystatin C is not necessary for inhibition of the aminopeptidase cathepsin C (dipeptidyl peptidase I) (Hall et al., 1993). In addition, truncation of the N-terminal residues of cystatin C does not affect cathepsin C or cathepsin H inhibition (Abrahamson et al., 1987, 1991). These studies suggest that, despite the substitution of glycine for serine in *Bm*-CPI-1, this product may still function as an inhibitor of a cathepsin C-like enzyme.

In some mammalian cystatins, such as cystatin C, a second inhibitory site has been demonstrated which blocks legumain or asparaginyl endopeptidase (AEP) enzymes (Alvarez-Fernandez et al., 1999). This lies on the opposite side of the protein to the papain-binding site (Fig. 1B), and contains an asparagine essential for legumain inhibition within the motif SND (Alvarez-Fernandez et al., 1999). It is interesting to note that the same motif is present in filarial cystatins of the CPI-2 type (amino acids 83-86 in Fig. 1A). Indeed, *Bm*-CPI-2 has been shown to block the activity of mammalian legumain and the related enzyme, asparaginyl endopeptidase in antigen-presenting cells (Manoury et al., 2001), and site-directed mutagenesis studies have established that, as with the mammalian inhibitor, the asparagine residue of *Bm*-CPI-2 is essential for AEP inhibition but is not required for cathepsin B, L or S inhibition (Murray, Manoury, Balic, Watts, & Maizels, 2005).

A further unusual feature of Bm-CPI-2 and its homologues from 3 other filarial nematodes is the insertion of a short (15–20 aa) N-terminal peptide. Currently, we are unable to assign a functional property to this tract, although one possibility is that it allows the protease inhibitor to interact more effectively with target cells of the immune system. For example, if this inserted sequence were to be recognised by innate receptors of mammalian antigen-presenting cells, the CPI-2 proteins might gain access to intracellular compartments containing target cathepsins and AEP. Experiments are now under way to test this supposition.

1.3. Comparison with C. elegans and other non-filarial nematodes

Two cystatin-like sequences from the free-living nematode *C. elegans* have been readily identified, and designated *Ce*-CPI-1 and -2 (Murray et al., 2005). The predicted proteins are 48% identical to each other and, like the filarial cystatins, are marked by the presence of only one pair of cysteines. The canonical features of an N-terminal Glycine, a central Gln-Val-Val-Ala-Gly, and a C-terminal P-W pair are all present (Fig. 1). Interestingly, on the opposite face to the central QVVAG, *Ce*-CPI-1 encodes SNN, and *Ce*-CPI-2 NNG, at the position where *Bm*-CPI-2 has a functional AEP-inhibiting motif, SND. Testing of recombinant *C. elegans* cystatins, expressed in bacteria, revealed however that neither *Ce*-CPI-1 nor -2 were able to block AEP, although both were effective at inhibiting cathepsin S (Murray et al., 2005). However, Ce-CPI-2 (also named CPI-2a in distinction to a truncated isoform predicted to be nonfunctional) has been shown to be essential in oocyte maturation and fertilization, with *cpi-2* mutant worms being sterile (Hashmi, Zhang, Oksov, Ji, & Lustigman, 2006). Further comparisons between filarial and *C. elegans* cystatins are described below, in the context of immune evasion (Hartmann & Lucius, 2003; Schierack, Lucius, Sonnenberg, Schilling, & Hartmann, 2003).

In addition to *C. elegans*, cystatins have been identified in two non-filarial nematode parasites, both gastrointestinal worms of animal hosts. *Haemonchus contortus*, a prominent gut parasite of sheep, expresses *Hc*-CPI (Newlands, Skuce, Knox, & Smith, 2001), while *Nippostrongylus brasiliensis*, a common model system which infects rodents, expressed *Nb*-CPI (Dainichi et al., 2001).

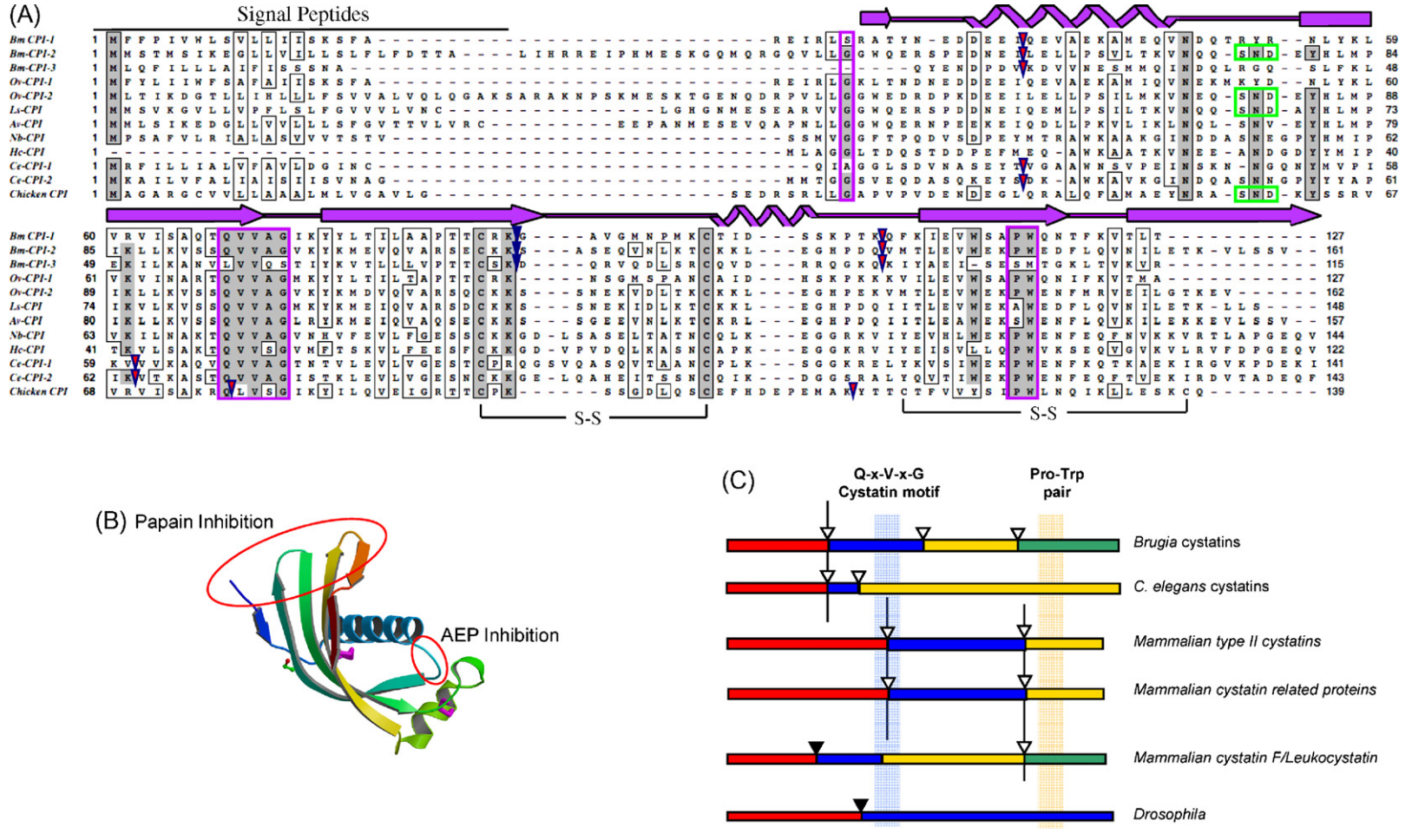

Fig. 1. Comparison of nematode and vertebrate cystatin sequences. (A) Sequence alignment of nematode and vertebrate cystatins. Residues known to be involved in papain binding are in open grey boxes, including the conserved N-terminal glycine residue (numbered 56) absent in *Bm*-CPI-1 and -3. The proposed AEP-binding loop residues (numbered 83–86) are outlined in green boxes. Conserved disulphide bonds are drawn below the alignment. The cystatin-specific flexible loop, QVVAG in *Bm*-CPI-1 and -2, is also shown. Accession numbers for cDNA or proteins sequences are as follows: *Brugia malayi* – *Bm*-CPI-1, U80972; *Bm*-CPI-2, AF015263; *Bm*-CPI-3 (submitted); *Onchocerca volvulus* – *Ov*-CPI-1, AF177194; *Ov*-CPI-2, P22085; *Litomosoides sigmodontis* – *Ls*-CPI, AF229173; *Acanthocheilonema viteae* – *Av*-CPI;, L43053 *Nippostrongylus brasiliensis* – *Nb*-CPI, AB050883; *Haemonchus contortus* – *Hc*-CPI AF035945; *Caenorhabditis elegans* – *Ce*-CPI-1 (K08B4.6), AF100663; *Ce*-CPI-2 (R01B10.1), AF068718; Chicken cystatin C, P01038. (B) Diagram of cystatin structure, illustrating the papain inhibitory loop and AEP inhibitory site on opposite faces of the protein. The diagram is adapted from the original published by Alvarez-Fernandez et al. (1999). (C) Genomic structure of *B. malayi* and *C. elegans* cystatin genes. Intron positions are represented by inverted triangles. Open symbols are in Phase 0, solid symbols in Phase 1. Vertical lines indicate correspondence between introns.

1.4. Analysis of genomic structure

All three *B. malayi* cystatin genes contain three introns varying in size from 120 to 480 bp, which are located at identical positions within the coding sequences (Fig. 1C). The number and position of these introns are atypical for known cystatin genes, as they differ markedly from the structure observed in *C. elegans*. It is especially surprising that intron positions are strictly conserved between three *B. malayi* genes whose coding regions are up to 75% divergent. Thus, while the genomic organisation implies that *B. malayi* cystatin genes have diverged subsequent to the separation of filarial and *C. elegans* lineages, the coding sequence phylogeny suggests the reverse. This may indicate that parasite cystatin protein sequences have undergone rapid evolution in recent times, possibly as an adaptation to the parasitic life cycle.

1.5. Differential expression of filarial cystatins through the life cycle

While EST information gave some indication of filarial cystatin levels in different stages, it is important to verify this with more directed studies. Accordingly, RT-PCR was employed to detect gene expression throughout the parasite's development in both mosquito and mammalian hosts. Gene-specific primer pairs located within exons 1 and 4 of the genes were chosen to ensure discrimination between cDNA and genomic amplification products. PCR from cDNA libraries of the three main stages of *B. malayi* suggested that while *Bm-cpi-2* is constitutively expressed, *cpi-1* and *-3* are more restricted in their expression to the L3 stage.

A more comprehensive insight is afforded by preparing mRNA from each stage of development both during the vector mosquito (Fig. 2A) and also following infection in the mammal (Fig. 2B) (Gregory, Atmadja, Allen, & Maizels, 2000; Murray, Gregory, Gomez-Escobar, Atmadja, & Maizels, 2001). This analysis shows that *Bm-cpi-2* RNA is indeed present in all life cycle stages but *cpi-1* is only detectable in the late stages of parasite development in the vector mosquito and the very early stages of development in the mammal. Twenty-four hours after entry into the mammalian host, *cpi-1* expression is seen to fall below the detection limit of the assay, even after 35 rounds of amplification, with only marginal expression at some later time points. Similarly, *cpi-3* is predominantly expressed in maturing larvae within the mosquito, although transcripts of this gene can be detected from an earlier timepoint, and are evident at later mammalian timepoints, than is the case for *cpi-1*. It is interesting to note that *cpi-1* and *cpi-3* genes are adjacent within the *B. malayi* genome, organised in an inverted head-to-head repeat (Fig. 2C). Between the respective ATG start codons, lie 2.5 kb of intervening sequence which also shows signs of a duplication event. Thus the 500 bp immediately upstream of each gene shows 44% nucleotide identity, suggesting that expression of both *cpi-1* and *cpi-3* may share some common promoter elements.

The distinct developmental expression patterns of the *cpi* genes may give some insight into their respective functions. The restriction of *cpi-1* and *-3* to the mosquito stage indicates that their target protease may be of insect origin rather than mammalian. Initiation of *Bm-cpi*-1 expression coincides with L2 moult and the migration of larvae from the muscle cells of the thorax to the mouthparts of the mosquito (Schacher, 1962). This is also the time at which larvae become infective to the mammalian host. Expression of the gene is high in comparison with β-tubulin and continues until larvae infect the mammalian host. Localization of the protein at the surface of the larvae suggests that it may be involved in protection of the surface cuticle while the larvae are resident within the head and mouthparts of the vector.

In contrast, the continuing production and secretion of *Bm*-CPI-2 throughout the mammalian part of the life cycle implies an active role in parasite maintenance. In the related parasite *O. volvulus*, the homologous gene product is highly expressed in the cuticle of moulting larvae, and indeed synthetic cysteine protease inhibitors are able to arrest moulting altogether (Lustigman et al., 1992; Richer, Hunt, Sakanari, & Grieve, 1993). However, there was no detectable increase in *cpi-2* transcription around the moulting events, as for example observed with collagen genes involved in cuticle synthesis in *C. elegans* (Johnstone & Barry, 1996). In contrast, there is ample evidence that the CPI-2 family are involved in immune interference, as discussed in the following section. The probability remains therefore that CPI-2, at least, continues to be important in the host–parasite relationship after the final moult to the adult stage.

It should also be considered that multiple functions for each cystatin would not be unprecedented. For example, *Bm*-CPI-1 may also be required for protection of larvae from host proteases immediately after infection, or that it plays a regulatory role in the extensive cuticular changes that are initiated by the infection process, including the preparation for moulting. Equally, the expression of CPI-2 in the invertebrate vector is likely to have a functional outcome. In this setting, injection of recombinant *Ov*-CPI-2 was shown to increase

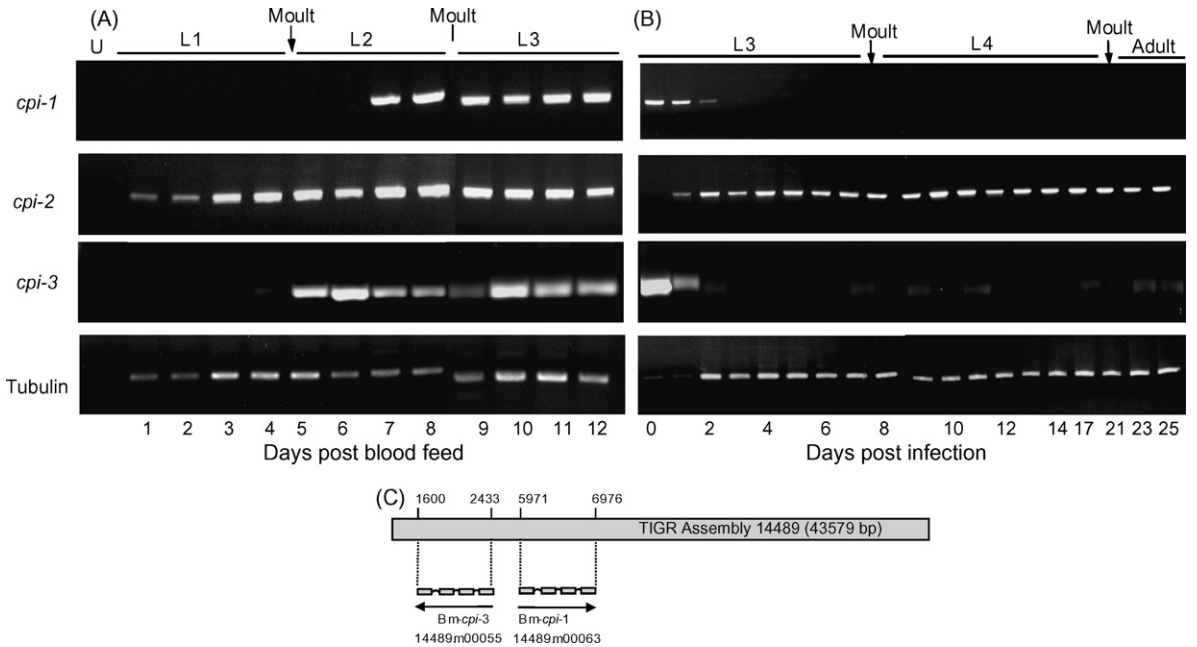

Fig. 2. Expression patterns of *B. malayi* cystatins. (A and B). Stage-specific expression of *Bm-cpi*-1, and -3, compared to constitutive expression of *Bm-cpi-2* throughout the vector stage of the life cycle (A) and during the first 25 days of development in the gerbil (B). L1–L4 denote the larval stages of the parasite. Total RNA was reverse-transcribed with oligo (dT) primer. PCR was then carried out with primers specific for *cpi-1*, *-2*, *-3* or β-tubulin. Products were run on 1% agarose gels and stained with ethidium bromide. Vector stage parasites were obtained from *Aedes aegypti* mosquitoes infected with *B. malayi* by membrane feeding on human citrate-treated blood mixed with peritoneal-derived L1 larvae (termed microfilariae, Mf) to a concentration of 16,000 Mf/ml. Mammalian stage parasites were recovered from gerbils infected intraperitoneally. Adult parasites and Mf were harvested from gerbils 3–12 months after infection. RNA was extracted using either TRIZOLV or RNAZOL B (Biotex Inc.). Total RNA was extracted from individual bloodfed mosquitoes without any attempt to isolate the larvae. For each mosquito 40 μl of first strand cDNA was synthesized using the GeneAmp RNA PCR Kit (Perkin Elmer) with the oligo $d(T)_{16}$ primer. To detect infected mosquitoes each first strand cDNA was amplified with primers specific for *cpi-2*. Total RNA from positive mosquitoes was pooled and PCRs using 5′ and 3′ gene-specific primers for *cpi-1* and *cpi-2* carried out using 5 μl pooled vector stage cDNA or 1 μl mammalian stage cDNA. Each cDNA was also amplified with primers specific for β-tubulin (Gomez-Escobar, Lewis, & Maizels, 1998). 25 μl PCR reactions were performed under standard conditions (94 °C for 1 min, 55 °C for 1 min, 72 °C for 1.5 min, 35 cycles; 72 °C for 10 min) and included 50 μM of each primer, 0.2 mM of each dNTP and 1.25 U Taq polymerase. (C) Diagram of tail-to-tail organisation of cpi-1 and cpi-3 genes in the genome of *B. malayi*, designated as predicted proteins 14489.m00063 and 14489.m00055, respectively.

infection levels of *Onchocerca ochengi* in its blackfly vector, suggesting a role for CPI-2 cystatins in defence against vector proteases (Kläger, Hagen, & Bradley, 1999).

1.6. Filarial cystatins and immune evasion

The cystatins from filarial nematodes have become recognised as one of the major sets of immune evasion molecules produced by these parasites (Hartmann & Lucius, 2003; Maizels, Gomez-Escobar et al., 2001; Maizels, Blaxter, & Scott, 2001; Vray, Hartmann, & Hoebeke, 2002). The first evidence was reported for *A. viteae* CPI-2 (Av17), which directly inhibits murine T cell proliferation *in vitro*, while simultaneously eliciting the production of the suppressive cytokine IL-10 (Hartmann et al., 1997). Interestingly, the *C. elegans* cystatins do not induce IL-10, but instead promote the pro-inflammatory cytokine IL-12 (Schierack et al., 2003).

As with *Av*-CPI, the *O. volvulus* CPI-2 (Ov-17 or onchocystatin) reduces polyclonal and antigen-specific T cell proliferation under conditions which are unaffected by either *C. elegans* cystatin (Schierack et al., 2003). The *Ov*-CPI-2 acts similarly on human T cells, as would be expected from a product of a human parasite, and was found also to inhibit monocyte expression of MHC Class II as well as the co-stimulatory molecule CD86 (Schönemeyer et al., 2001).

This was an important indication that filarial cystatins may target antigen-presenting cell function, which was demonstrated more directly in experiments which showed *Bm*-CPI-2 can block degradation of the tetanus toxoid (TT) protein in vitro by the antigen-processing

enzyme AEP (Manoury et al., 2001). Furthermore, antigen processing was inhibited in whole cell assays, in which the ability of TT-pulsed B cells to stimulate TT-peptide-specific T cell clones was measured, and *Bm*-CPI-2 was also shown to partially inhibit invariant chain breakdown (Manoury et al., 2001). Hence, the parasite inhibitor, known to be secreted from filarial worms, can exert a direct inhibitory effect at key points in the pathway for processing of exogenous antigenic peptides to the immune system.

In the mouse system, filarial cystatins have been found also to influence macrophages. In particular, nitric oxide production is significantly enhanced in IFN-γ-activated macrophages which are exposed to *Av*-CPI (Av17) or *Ov*-CPI-2 (Hartmann, Schönemeyer, Sonnenburg, Vray, & Lucius, 2002) and this effect is independent of the LPS-responsiveness genotype of the murine cells (an important control when testing recombinants derived from bacterial expression systems). Interestingly, enhancement of NO responsiveness

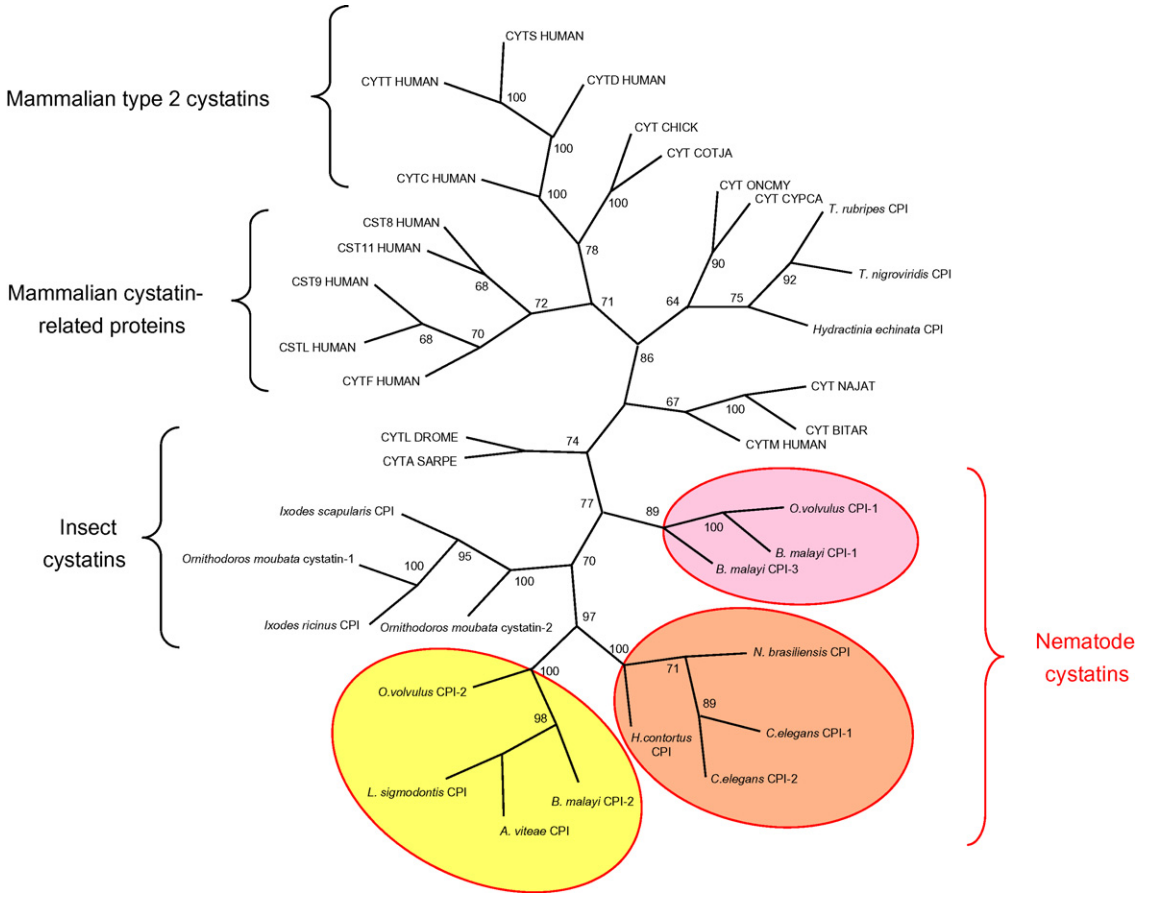

Fig. 3. Unrooted phylogram showing the relationship of nematode cystatins with members of the cystatin, stefin and kininogen families. Numbers on top of the branch lines show the calculated credibility values. GenBank or SwissProt accession numbers are as follows: *B. malayi* CPI-1, P90698; *B. malayi* CPI-2, O16159; *L. sigmodontis* CPI, Q9NH95; *A. viteae* CPI, Q17108; *O. volvulus* CPI-1, Q9U9A1; *O. volvulus* CPI-2, P22085; *N. brasiliensis* CPI. Q966W0; *H. contortus* CPI, O44396; *C. elegans* CPI-1, Q9TYY2; *C. elegans* CPI-2, Q86S25; CST11 HUMAN, Q9H112; CST9 HUMAN, Q9H4G1; CYTA SARPE, P31727; CYTC HUMAN, P01034; CYTD HUMAN, P28325; CYTF HUMAN, O76096; CYTL DROME, P23779; CYTM HUMAN, Q15828; CYTS HUMAN, P01036; CYTT HUMAN, P09228; CYT CHICK, P01038; CYT CYPCA, P35481; CYT NAJAT, P81714; CYT ONCMY, Q91195; CYT BITAR P08935; CYT COTJA P81061; *Ornithodoros moubata* cystatin-1, Q6QZV5_ORNMO; *Ornithodoros moubata* cystatin-2, Q6QD55_ORNMO; *Ixodes scapularis* cystatin 2, CPIQ4PMS6_IXOSC; *Ixodes scapularis* CPI, Q8MVB6_IXOSC; *Ixodes ricinus* CPI, Q86GB6_IXORI; *Hydractinia echinata* CPI, CO539721; *Tetraodon nigroviridis* CPI, CR648555; *Takifugu rubripes* CPI, CA845225. *Phylogenetic analysis*: Cystatin sequences, omitting their signal peptides, were aligned with the aid of the ClustalW function of MacVector with final manual adjustment. The peptide sequences were analysed using the Markov chain Monte Carlo maximum likelihood process as driven by MrBayes v3.0b4. The model for amino acid substitution was not set *a priori* and the Markov chain randomly swapped parameters of both the trees and models. Four chains were run for 1,000,000 generations, with trees saved every 100 generations, and the Bayesian posterior probabilities for nodes in the final consensus tree derived after discarding the first 500 trees (i.e. 50,000 generations) as burn in.

from macrophages does not require intact protease inhibitory activity, and is attributed to a distinct but as yet unidentified part of the cystatin molecule. Indeed, the same properties are also true for mammalian cystatins (Verdot et al., 1996). Finally, a different perspective is provided by studies with the *L. sigmodontis* cystatin, *Ls*-CPI. Recombinant cystatin was infused in vivo via a micro-osmotic pump implanted in the peritoneal cavity. Exposed mice were found with elevated TNF-α responses in peritoneal cells, but lower antigen-specific splenocyte responses and reduced NO levels (Pfaff et al., 2002). These authors also tested Ls-CPI as a vaccine against filarial infection, and reported that immunized mice developed lower frequency of mature infections with circulating microfilariae (Pfaff et al., 2002). This is an intriguing and important result that merits wider assessment in the different experimental systems available.

1.7. Evolutionary analysis

In terms of both sequence similarity and functional capacity for protease inhibition, the set of nematode cystatins described from *C. elegans*, and both filarial and non-filarial parasitic species are clearly members of the cystatin superfamily. Interestingly, all lack the second pair of cysteine residues found in family 2 cystatins, residues that clamp the carboxy terminus to the β-pleated sheet in all other family 2 cystatins. Members of the family 2 cystatins are considered to have evolved from family 1 stefins, which lack cysteine residues, acquiring four cysteine residues in the process (Brown & Dziegielewska, 1997; Müller-Esterl et al., 1985; Rawlings & Barrett, 1990). With the view that disulphide bonds are gained but seldom lost during evolution, nematode cystatins appear to be a fixed intermediate in this evolution.

Within the set of nematode cystatins, however, three evolutionary subgroups can be defined (Fig. 3). The most diverse are represented by *Bm*-CPI-1, which lacks an evolutionarily conserved glycine residue near the N-terminus. This small group (including also *Bm*-CPI-3 and *Ov*-CPI-1, the latter retaining the glycine residue), are not represented in *C. elegans*. As further complete genomes become available (including *H. contortus*), it will be interesting to see if this group remains restricted to vector-borne parasites.

The second subgroup is the most conventional, encoding cystatins relatively similar to the vertebrate homologues other than in the single disulphide bond. Functional studies on recombinant cystatins from *C. elegans* (Murray et al., 2005; Schierack et al., 2003), as well as *N. brasiliensis* (Dainichi et al., 2001), demonstrate that this group encodes functional inhibitors of cathepsin S and L, respectively.

Finally, perhaps the most intriguing subgroup is that typified by *Bm*-CPI-2 and *Ov*-CPI-2, onchocystatin. These inhibitors include an AEP-inhibitory site, in a possible example of convergent evolution, in which the nematode cystatins have acquired a site similar to that evolved in vertebrate cystatin C in order to control AEPs. Moreover, this family are all characterized by an N-terminal extension not observed in any other cystatin sequences; whether this extension plays an active role in the biology of parasite inhibitors remains to be determined.

2. Conclusion

The cystatins are an ancient and conserved family of cysteine protease inhibitors, and their expression among diverse nematode species is not unexpected. It is curious, however, that cystatins are such prominent gene products among parasitic nematodes, and that they have evolved features quite distinct from *C. elegans*, suggesting that cystatins play an essential role in transmission, invasion, and/or immune evasion (Hartmann & Lucius, 2003; Maizels, Gomez-Escobar et al., 2001). Several key features, including the intron conservation between highly divergent *cpi-1* and *cpi-2* genes, and the appearance of an AEP inhibitory site in *cpi-2* genes, indicate that the filarial cystatins, at least, have been subject to rapid evolutionary change since their divergence from a common ancestor with *C. elegans*. Future work will aim at defining the extent to which such changes represent evolutionary adaptations to the parasitic way of life in the arthropod vector or mammalian host.

Acknowledgments

This work was funded by the Wellcome Trust through Programme Grant funding to RMM.

References

Abrahamson, M., Alvarez-Fernandez, M., & Nathanson, C. M. (2003). Cystatins. *Biochem. Soc. Symp.*, *70*, 179–199.

Abrahamson, M., Mason, R. W., Hansson, H., Buttle, D. J., Grubb, A., & Ohlsson, K. (1991). Human cystatin C role of the N-terminal segment in the inhibition of human cysteine proteinases and in its inactivation by leucocyte elastase. *Biochem. J.*, *273*, 621–626.

Abrahamson, M., Ritonja, A., Brown, M. A., Grubb, A., Machleidt, W., & Barrett, A. J. (1987). Identification of the probable inhibitory reactive sites of the cysteine proteinase inhibitors human cystatin C and chicken cystatin. *J. Biol. Chem.*, *262*, 9688–9694.

Alvarez-Fernandez, M., Barrett, A. J., Gerhartz, B., Dando, P. M., Ni, J., & Abrahamson, M. (1999). Inhibition of mammalian legumain by some cystatins is due to a novel second reactive site. *J. Biol. Chem.*, *274*, 19195–19203.

Bode, W., Engh, R., Musil, D., Thiele, U., Huber, R., Karshikov, A., et al. (1988). The 2.0 A X-ray crystal structure of chicken egg white cystatin and its possible mode of interaction with cysteine proteinases. *EMBO J.*, *7*, 2593–2599.

Brown, W. M., & Dziegielewska, K. M. (1997). Friends and relations of the cystatin superfamily—New members and their evolution. *Protein Sci.*, *6*, 5–12.

Dainichi, T., Maekawa, Y., Ishii, K., Zhang, T., Nashed, B. F., Sakai, T., et al. (2001). Nippocystatin, a cysteine protease inhibitor from *Nippostrongylus brasiliensis*, inhibits antigen processing and modulates antigen-specific immune response. *Infect. Immun.*, *69*, 7380–7386.

Ghedin, E., Wang, S., Spiro, D., Caler, E., Zhao, Q., Crabtree, J., et al. (2007). Draft genome of the filarial nematode parasite *Brugia malayi*. *Science*, *317*, 1756–1760.

Gomez-Escobar, N., Lewis, E., & Maizels, R. M. (1998). A novel member of the transforming growth factor-β (TGF-β) superfamily from the filarial nematodes *Brugia malayi* and *B. pahangi*. *Exp. Parasitol.*, *88*, 200–209.

Gregory, W. F., Atmadja, A. K., Allen, J. E., & Maizels, R. M. (2000). The *abundant larval transcript 1/2* genes of *Brugia malayi* encode stage-specific candidate vaccine antigens for filariasis. *Infect. Immun.*, *68*, 4174–4179.

Gregory, W. F., Blaxter, M. L., & Maizels, R. M. (1997). Differentially expressed, abundant *trans*-spliced cDNAs from larval *Brugia malayi*. *Mol. Biochem. Parasitol.*, *87*, 85–95.

Hall, A., Dalboge, H., Grubb, A., & Abrahamson, M. (1993). Importance of the evolutionarily conserved glycine residue in the N-terminal region of human cystatin C (Gly-11) for cysteine endopeptidase inhibition. *Biochem. J.*, *291*, 123–129.

Hall, A., Hakansson, K., Mason, R. W., Grubb, A., & Abrahamson, M. (1995). Structural basis for the biological specificity of cystatin C. Identification of leucine 9 in the N-terminal binding region as a selectivity-conferring residue in the inhibition of mammalian cysteine peptidases. *J. Biol. Chem.*, *270*, 5115–5121.

Hartmann, S., Kyewski, B., Sonnenburg, B., & Lucius, R. (1997). A filarial cysteine protease inhibitor down-regulates T cell proliferation and enhances interleukin-10 production. *Eur. J. Immunol.*, *27*, 2253–2260.

Hartmann, S., & Lucius, R. (2003). Modulation of host immune responses by nematode cystatins. *Int. J. Parasitol.*, *33*, 1291–1302.

Hartmann, S., Schönemeyer, A., Sonnenburg, B., Vray, B., & Lucius, R. (2002). Cystatins of filarial nematodes up-regulate the nitric oxide production of interferon-g-activated murine macrophages. *Parasite Immunol.*, *24*, 253–262.

Hashmi, S., Zhang, J., Oksov, Y., Ji, Q., & Lustigman, S. (2006). The *Caenorhabditis elegans* CPI-2a cystatin-like inhibitor has an essential regulatory role during oogenesis and fertilization. *J. Biol. Chem.*, *281*, 28415–28429.

Johnstone, I. L., & Barry, J. D. (1996). Temporal reiteration of a precise gene expression pattern during nematode development. *EMBO J.*, *15*, 3633–3639.

Kläger, S. L., Hagen, H.-E., & Bradley, J. E. (1999). Effects of an *Onchocerca*-derived cysteine protease inhibitor on microfilariae in their simuliid vector. *Parasitology*, *118*, 305–310.

Kotsyfakis, M., Sá-Nunes, A., Francischetti, I. M. B., Mather, T. N., Andersen, J. F., & Ribeiro, J. M. C. (2006). Antiinflammatory and immunosuppressive activity of sialostatin L, a salivary cystatin from the tick *Ixodes scapularis*. *J. Biol. Chem.*, *281*, 26298–26307.

Lustigman, S., Brotman, B., Huima, T., & Prince, A. M. (1991). Characterization of an *Onchocerca volvulus* cDNA clone encoding a genus specific antigen present in infective larvae and adult worms. *Mol. Biochem. Parasitol.*, *45*, 65–76.

Lustigman, S., Brotman, B., Huima, T., Prince, A. M., & McKerrow, J. H. (1992). Molecular cloning and characterization of onchocystatin, a cysteine proteinase inhibitor of *Onchocerca volvulus*. *J. Biol. Chem.*, *267*, 17339–17346.

Lustigman, S., McKerrow, J. H., Sha, K., Lui, J., Huima, T., Hough, M., et al. (1996). Cloning of a cysteine protease required for the molting of *Onchocerca volvulus* third stage larvae. *J. Biol. Chem.*, *271*, 30181–30189.

Maizels, R. M., Balic, A., Gomez-Escobar, N., Nair, M., Taylor, M., & Allen, J. E. (2004). Helminth parasites: masters of regulation. *Immunol. Rev.*, *201*, 89–116.

Maizels, R. M., Blaxter, M. L., & Scott, A. L. (2001). Immunological genomics of *Brugia malayi*: filarial genes implicated in immune evasion and protective immunity. *Parasite Immunol.*, *23*, 327–344.

Maizels, R. M., Gomez-Escobar, N., Gregory, W. F., Murray, J., & Zang, X. (2001). Immune evasion genes from filarial nematodes. *Int. J. Parasitol.*, *31*, 889–898.

Maizels, R. M., & Yazdanbakhsh, M. (2003). Regulation of the immune response by helminth parasites: cellular and molecular mechanisms. *Nat. Rev. Immunol.*, *3*, 733–743.

Manoury, B., Gregory, W. F., Maizels, R. M., & Watts, C. (2001). *Bm*-CPI-2, a cystatin homolog secreted by the filarial parasite *Brugia malayi*, inhibits class II MHC-restricted antigen processing. *Curr. Biol.*, *11*, 447–451.

Margis, R., Reis, E. M., & Villeret, V. (1998). Structural and phylogenetic relationships among plant and animal cystatins. *Arch. Biochem. Biophys.*, *359*, 24–30.

Michael, E., Bundy, D. A. P., & Grenfell, B. T. (1996). Re-assessing the global prevalence and distribution of lymphatic filariasis. *Parasitology*, *112*, 409–428.

Müller-Esterl, W., Fritz, H., Kellermann, J., Lottspeich, F., Machleidt, W., & Turk, V. (1985). Genealogy of mammalian cysteine proteinase inhibitors. Common evolutionary origin of stefins, cystatins and kininogens. *FEBS Lett.*, *191*, 221–226.

Murray, J., Gregory, W. F., Gomez-Escobar, N., Atmadja, A. K., & Maizels, R. M. (2001). Expression and immune recognition of *Brugia malayi* VAL-1, a homologue of vespid venom allergens and *Ancylostoma* secreted proteins. *Mol. Biochem. Parasitol.*, *118*, 89–96.

Murray, J., Manoury, B., Balic, A., Watts, C., & Maizels, R. M. (2005). *Bm*-CPI-2, a cystatin from *Brugia malayi* nematode parasites, differs from *C.elegans* cystatins in a specific site mediating inhibition of the antigen-processing enzyme AEP. *Mol. Biochem. Parasitol.*, *139*, 197–203.

Newlands, G. F., Skuce, P. J., Knox, D. P., & Smith, W. D. (2001). Cloning and expression of cystatin, a potent cysteine protease inhibitor from the gut of *Haemonchus contortus*. *Parasitology*, *122*, 371–378.

Nutman, T. B. (1999). Lymphatic filariasis. In G. Pasvol & S. L. Hoffman (Eds.), *Tropical medicine: Science and practice* (p. 283). London: Imperial College Press.

Pfaff, A. W., Schulz-Key, H., Soboslay, P. T., Taylor, D. W., MacLennan, K., & Hoffmann, W. H. (2002). Litomosoides sigmodontis cystatin acts as an immunomodulator during experimental filariasis. *Int. J. Parasitol.*, *32*, 171–178.

Rawlings, N. D., & Barrett, A. J. (1990). Evolution of proteins of the cystatin superfamily. *J. Mol. Evol.*, *30*, 60–71.

Richer, J. K., Hunt, W. G., Sakanari, J. A., & Grieve, R. B. (1993). *Dirofilaria immitis:* effect of fluoromethyl ketone cysteine protease inhibitors on the third to fourth-stage molt. *Exp. Parasitol.*, *76*, 221–231.

Schacher, J. F. (1962). Developmental stages of *Brugia pahangi* in the final host. *J. Parasitol.*, *48*, 693–706.

Schierack, P., Lucius, R., Sonnenberg, B., Schilling, K., & Hartmann, S. (2003). Parasite specific immunomodulatory functions of filarial cystatin. *Infect. Immun.*, *71*, 2422–2429.

Schönemeyer, A., Lucius, R., Sonnenburg, B., Brattig, N., Sabat, R., Schilling, K., et al. (2001). Modulation of human T cell responses and macrophage functions by onchocystatin, a secreted protein of the filarial nematode *Onchocerca volvulus*. *J. Immunol.*, *167*, 3207–3215.

Scott, A. L. (1999). Lymphatic-dwelling filariae. In T. B. Nutman & G. Pasvol (Eds.), *Filariasis* (pp. 5–39). London: Imperial College Press.

Semnani, R. T., & Nutman, T. B. (2004). Toward an understanding of the interaction between filarial parasites and host antigen-presenting cells. *Immunol. Rev.*, *201*, 127–138.

Shibuya, K., Kaji, H., Ohyama, Y., Tate, S., Kainosho, M., Inagaki, F., et al. (1995). Significance of the highly conserved Gly-4 residue in human cystatin A. *J. Biochem. (Tokyo)*, *118*, 635–642.

Verdot, L., Lalmanach, G., Vercruysse, V., Hatmann, S., Lucius, R., Hoebeke, J., et al. (1996). Cystatins up-regulate nitric oxide release from interferon-γ-activated mouse peritoneal macrophages. *J. Biol. Chem.*, *271*, 28077–28081.

Vray, B., Hartmann, S., & Hoebeke, J. (2002). Immunomodulatory properties of cystatins. *Cell Mol. Life Sci.*, *59*, 1503–1512.