Current Topics in Biochemical Research
Vol. 14, *1*, 2012

Review

Proteinaceous peptidase inhibitors of the human pathogenic intestinal parasitic protozoa *Entamoeba histolytica* and *Giardia lamblia*

Henning Scholze*

Faculty of Biology, University of Osnabrueck, Biochemistry, Barbarastr. 13, 49069 Osnabrueck, Germany

ABSTRACT

The human pathogenic protozoan parasites *Entamoeba histolytica* and *Giardia lambia*, triggers of intestinal disturbances mainly in developing countries, are equipped with a set of multiple cysteine peptidases used as weapons to attack their victims. Exertion and control of these enzyme activities are absolutely necessary for the survival of these parasites in their inhospitable environments. Generally, peptidase inhibitors are a common tool of many uni- and multi-cellular organisms to regulate endogenous and exogenous peptidase activities. For the regulation of cysteine peptidase activities, either one of two types of proteinacious inhibitors is frequently used: those of the cystatin and those of the chagasin family. Both types of inhibitors, also appearing as multicopy forms, are relatively small and very resistant proteins widespread in nature. Trophozoites of the protists *E. histolytica* and *G. lamblia* also feature these types of specific cysteine peptidase inhibitors, out of which *E. histolytica* is solely equipped with two chagasin-like proteins, whereas *G. lamblia* only contains cystatins, both as single copy forms. Another type of proteinaceous peptidase inhibitors is represented by the group of serpins, usually directed against serine peptidases with particular substrate specificities, but in exceptional cases are also able to inhibit cysteine peptidases. Both protozoan parasites each are equipped with single serpin genes, whose physiological targets have not yet been identified. The significance of these findings is discussed.

KEYWORDS: *Entamoeba histolytica/dispar*, *Giardia lamblia*, cysteine peptidase inhibitors, chagasin, amoebiasin, cystatin, serpin

INTRODUCTION

The *genus Entamoeba* and *Giardia* comprise a group of intestinal parasites primarily colonizing the intestinal tract, but also spending their life in inner organs of their particular target organisms [1, 2]. Their destinations are primates in cases of both *Entamoeba histolytica* and *Entamoeba dispar*, and reptiles in case of *Entamoeba invadens*, whereas *Giardia intestinalis* also affects other mammalian hosts like cats, dogs or birds. *E. histolytica* and *E. invadens* are able to invade and settle tissues and organs of their hosts forming tumour-like abscesses often with deadly effects, whereas the morphologically indistinguishable *E. dispar* solely remains in the human gut as harmless commensal. *G. lamblia* infections, although often proceeding unnoticed, can lead to diarrhea and abdominal pressure and are not fatal in healthy individuals. All these protists pass through a simple life cycle comprising a latent, rigid cyst and a vegetative trophozoite form. The trophozoites spend their life within their host and exclusively proliferate there, and only the cysts

*scholze@biologie.uni-osnabrueck.de

are able to endure the harsher conditions outside for several days and thus represent the infective life stage [3]. However, it is the trophozoites that trigger the well-known symptoms of *amebiasis* and *giardiasis*, and, in case of *E. histolytica*, are responsible for 40,000 thousands incidences with a deadly outcome worldwide [4]. Also the progeny of these parasites is restricted to their trophozoite forms, and this is the reason why attempts as to the rational design of an antiamoebial drug mainly focus on the intervention into fundamental biochemical processes of this life stage.

To get an idea, how the parasites could be hampered to affect their targets, the knowledge of the mechanisms used by the trophozoites to invade the intestinal tract and colonize internal organs of their victims are necessary. Briefly, their life cycles are as follows: After engulfment of the cysts by oral uptake of contaminated food, the parasites pass the stomach more or less unobstructed before arriving at the small intestine, in which the cells excyst [5]. In doing so, altogether eight *Entamoeba* trophozoites are released from a single cyst containing four nuclei, whereas two trophozoites containing two nuclei each per cyst are released in case of *Giardia* [6]. In the intestine the protists proliferate and adhere to the intestinal epithelium triggering resorption disturbances. This impacts water and electrolyte housekeeping processes leading to the well-known symptoms of dysentery. Compared to *Giardia*, the amoebae trophozoites are able to penetrate the intestinal wall faced by epithelial cells and the subjacent *basal lamina*, and thus enter the blood circulation, wherefrom they are able to infest a lot of inner organs including liver, lung and brain. There, the formation of severe tumour-like abscesses can lead to a complete collapse of vital processes with lethal outcome of their victims [7].

The initiation of the pathogenic process caused by both parasites is quite different. For instance, the diplomonad protist *Giardia* adheres to the intestinal wall by a continuous beat of distinct flagella and generates a negative pressure below the contact zone between the ventral disk and the intestinal epithelium. This fixes the trophozoites to the intestinal wall like a suction cup in a physical manner [8]. By contrast, *Entamoeba* trophozoites use a more chemical way for their attack of the human gut [9]. The initial recognition of the intestinal cell surface, at which the amoebae trophozoites may adhere, is realized by their binding to terminal N-acetylglucosamine and N-acetyl-galactosamine residues residing on the surface of the intestinal epithelial cells *via* a specific lectin protruding from their own cell wall. As a next step the amoebae secrete pore-forming proteins (amoebapores) which then insert into the membrane of the target cell thereby forming leakages within this membrane [10]. Consequently, the affected epithelial cell shrinks, dies off, and by this means opens the way for further penetration incidents of the parasite. Afterwards, the trophozoites detach from the epithelial cells by cutting off the terminal carbohydrate anchors keeping hold of them. This may be achieved with the aid of a specific enzyme, a β-N-acetyl-hexosaminidase, which has been found to be secreted by the amoebae [11].

For the subsequent invasion process the *Entamoeba* parasites have to destroy the proteinaceous components of the extracellular matrix, which are broken down with the aid of a bulk of cysteine endopeptidases [12]. The major proteases of *Entamoeba* as well as of *Giardia* belong to the cathepsin B- and L-like types of cysteine peptidases and in turn are relatives of papain [13]. The genome of *E. histolytica* is equipped with a total of 48 different cysteine peptidase encoding genes; the majority of them belong to the C1-papain superfamily (http://pathema.jcvi.org/cgi-bin/Entamoeba/PathemaHomePage.cgi) [14]. Four of these gene products, classified as EhCP-A1, -A2, -A5 and -A7, respectively, account for by far the most of the total peptidolytic activity of the trophozoites under culture conditions and show a broad substrate specificity spectrum towards various protein and peptide substrates [15]. Generally, the knowledge of the subcellular localization of an enzyme permits conclusions as to its functions. The amebic cysteine endopeptidases EhCP-A1 and EhCP-A2 were found to be located in lysosome-like vesicles underlining their role in intracellular protein degradation [16], whereas EhCP-A5 also resides on the cell surface suggesting that further essential physiological tasks of this enzyme are outside the cell [17].

EhCP-A7 is supposed to contribute to erythrophagocytosis, a frequently observed event during infections with *E. histolytica* [15]. However, despite of their important role in the life of *Entamoeba*, the detailed physiological significances of these and other amebic peptidases are not yet completely clear. Still less is known concerning the functions of cysteine proteases in *Giardia*. Although *G. lamblia* does not penetrate the extracellular matrix of its victims, this parasite is equipped with various genes encoding cysteine proteases (see: http://giardiadb.org/giardiadb/), and particular importance for the encystation process during giardial life cycle has been ascribed to a major cysteine protease [18].

The potentially high peptidolytic activity produced by the trophozoites, not least by reasons of self-protection, needs efficient fine-tuning of the corresponding proteins. Up-regulation of the peptidolytic activity within and around a particular organism in most cases is carried out by an increase of the expression of peptidase genes as well as by zymogen activation of the protease precursors by splitting off N-terminal pro-peptides masking the active site of the respective enzyme. Peptidase activity emerging at those sites at which it may be harmful for the cell of its origin can be avoided by specific targeting of the peptidases to their places of action. Down-regulation of enzyme activity is generally less selective, just like disarming of exogenous peptidase activity which, besides their elimination by degrading processes, normally requires the adoption of specific inhibitors. In eukaryotic cells, inhibition of peptidase activity is realized by use of type specific proteinaceous protease inhibitors, whose knowledge has considerably increased in the last decade [19]. To this end, various inhibitor families, which nonetheless have a common mode of action on their target cysteine proteases, are well studied [20]. The fact that nature obviously has found a more general way of protein/protein-interaction to inhibit protease activity reflects the observation that some members of another protease inhibitor, those of the serpins directed towards serine proteases, are also able to inhibit cysteine proteases. However, whereas members of this protease inhibitor class react with their target proteases in an irreversible manner resulting in a complete enzyme inactivation, the two other main lines of naturally occurring cysteine peptidase inhibitors, designated as cystatins and chagasins inhibit their target peptidases reversibly. The docking modes of these inhibitors to their target proteases employing distinct binding loops forming wedge-like structures seem similar [21].

Serpins

Serpins (MEROPS family I4) are named by reasons of their particular properties to effectively inhibit serine proteases (serine protease inhibitors). Well-known representatives of this protein class are involved in important physiological processes such as inflammation or blood coagulation and are triggered by limited proteolysis [22]. Serpins with M_r ranging from 40 to 70 kD are found in all kingdoms of organisms and show a widespread inhibition specificity. For instance, the serpin α1-proteinase inhibitor (α1- antitrypsin) acts as acute phase protein, whereas another serpin, antithrombin, functions as an inhibitor of contact phase activators of the blood coagulation cascade. Three-dimensional structures of a vast number of serpins have been solved thus allowing detailed insights into their particular mode of action. Their well-conserved folding pattern is distinguished by a single exposed β-strand forming a reactive center loop (RCL) which, after contact to the active site of its target protease and subsequent specific cleavage, completely rearranges and together with the bound protease turns down to the middle of the protease molecule resulting in a very effective and irreversible inactivation of the enzyme [23]. The distinct inhibition specificities of the various serpins are mainly determined by those amino acid residues located in the P_1-position of the RCL-loop mimicking the substrate specificity of the protease (Fig. 1). This interaction mode of the inhibitor with its partner peptidase in some way reminds antigen/antibody recognition and binding. Thus, whereas in most cases a well-conserved structural part of a protein molecule is responsible for its particular function, the specific binding of the serpins to their target proteases is primarily assured by the amino acid sequence of a hypervariable region within their polypeptide strands such as a blueprint of the peptide stretch of the substrate to be cleaved. This is reflected

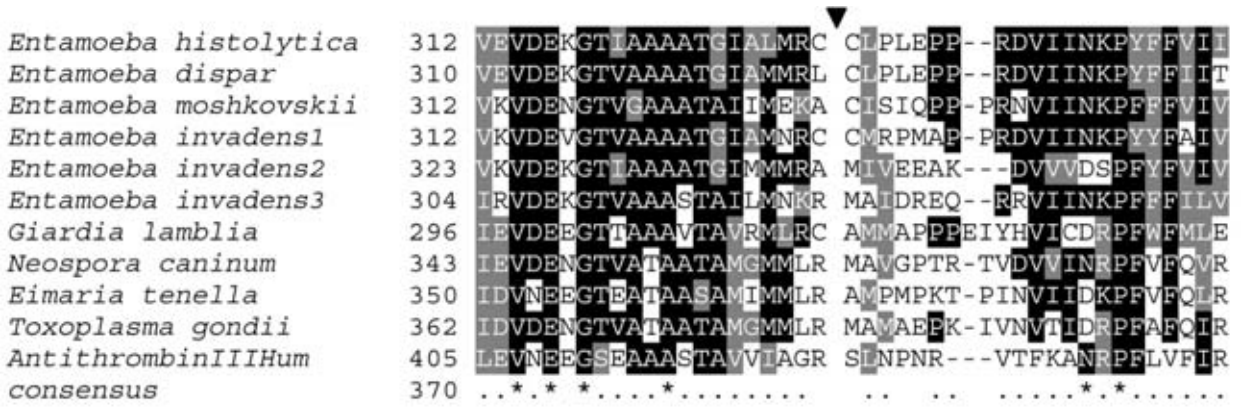

Fig. 1. Sequence alignment of serpins from unicellular organisms together with human antithrombin around their active sites. Multiple alignments were performed using the ClustalW algorithm (http://mobyle.pasteur.fr/cgi-bin/portal.py) [43]. Identical residues are shown in black, strongly similar ones in grey; arrow indicates the peptidolytic cleavage positions in the RCL loops.

in a predominant occurrence of basic and hydrophobic residues in P_1 and P_2 of the serpin mimicking the residues forming the peptide bonds to be favourably cleaved by a series of serine proteases such as trypsin (cleavage after Arg. Lys, His), the blood coagulation factors (cleavage after Arg) or chymotrypsin (cleavage after bulky hydrophobic residues), respectively.

Sequence variabilities of multiple serpin genes in the genome of a given protist in the region of the reactive-center loop reflect the specificity of their protein products. These are distinguished by either large hydrophobic or positively charged residues at the putative P_1- and P_2-positions in their hypervariable regions suggesting that proteases favouring basic residues in their substrates are major targets of these serpins. This applies to serpins of the intracellular parasitic protist *Eimeria tenella* (*phylum Apicomplexa*) as well as to one of the two serpins of the coccidian parasite *Neospora caninum* [23]. The recombinant putative serpin from *Neospora caninum* was found to completely inhibit bacterial subtilisin [24], but showed lower inhibitory capacity towards human neutrophil elastase, animal trypsin, and chymotrypsin, respectively, suggesting differences in effectiveness to inhibit different serine peptidase classes. Two serpins have been found in *Toxoplasma gondii* the causative agent of toxoplasmosis with specificity for trypsin-like serine proteases and are suggested to play a role in the survival of the tachyzoites in the host [25].

The genomes of the human intestinal protists *E. histolytica/dispar* and *G. lamblia* each are equipped with one serpin-encoding gene, the reptilian pathogenic *E. invadens* even contains up to six genes encoding serpin-homologous proteins (http://amoebadb.org/amoeba/) [26]. Compared to the primary sequences of orthologs from higher eukaryotes, these serpins appear to form a phylogenetically distinct clade (Fig. 2A), albeit they all possess the characteristic motifs found in the RCL-loops of other serpins (Fig. 1 and 2B). Genes encoding serine peptidases as potential targets of these serpins have likewise been identified in the genomes of these parasitic protists, but their protein products have not yet been characterized and are supposed to contribute little to the total peptidolytic activity of the trophozoites in culture. Thus, because intracellular targets of the protozoan serpins are difficult to predict, a role of the serpins from these parasites to inhibit host serine proteases seems likely. However, comparison of the amino acid sequences in the RCL loops of these serpins reveal that in cases of *G. lamblia* and the *Entamoeba spp.* arginine/lysine residues reside in the respective P_2 positions (Fig. 1 and 2B). Considering that these amino acids directly correspond to the substrate specificities of the cysteine proteases of these parasites, it also might not be excluded that some of the diverse cysteine proteases of *Entamoeba* and *Giardia* which all favour basic residues in P_1- and/or P_2-positions of their substrates may be targets of these inhibitors.

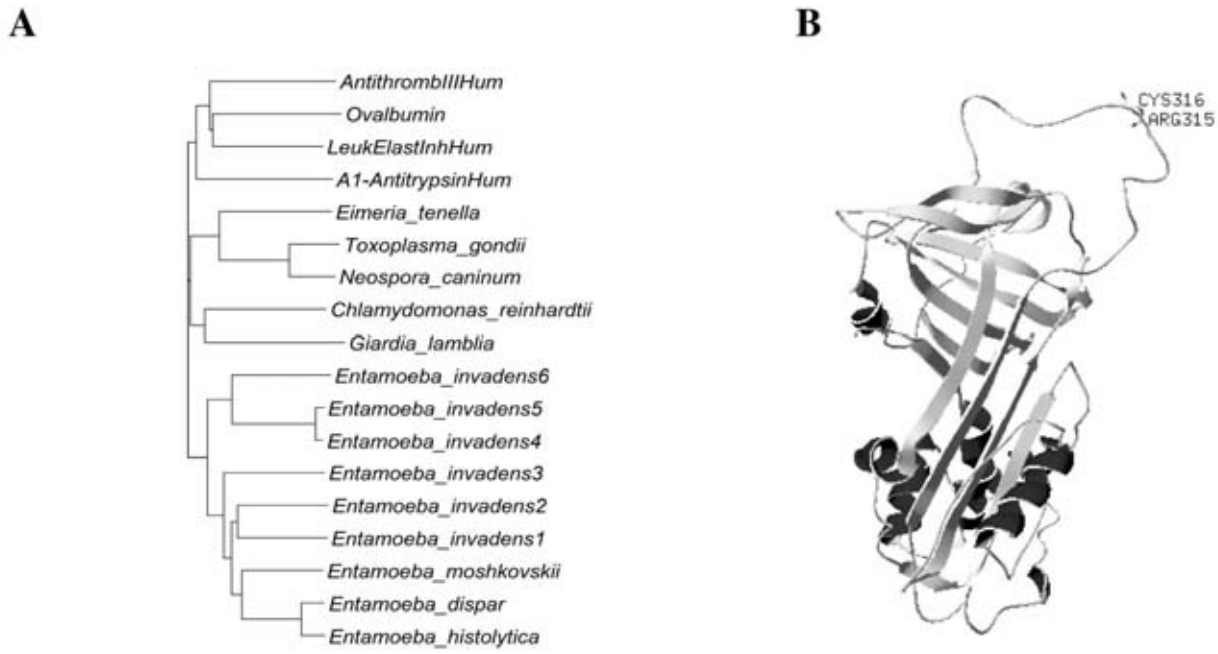

Fig. 2. A. Horizontal cladogram of serpin sequences from vertebrates and several protists. The *Entamoebae* sequences comprise a separate clade. Searches for serpin sequences were performed by BLAST search of the NCBI database (http://www.ncbi.nlm.nih.gov/blast) and by the European pathogen database (http://eupathdb.org/eupathdb/). Multiple alignments of 18 selected serpin homologues were performed using ClustalW algorithm (http://mobyle.pasteur.fr/cgi-bin/portal.py) [43], a dendrogram of these sequences was created by the Neighbour-Joining method [44]. **B.** Structure model of the *Giardia* serpin. Exposed residues in the RCL loop are highlighted. The model was generated with the aid of the Swiss-PdbViewer using the *Manduca sexta* serpin (PDB 1sek) as template [45].

Cystatins

For quite some time, specific proteinaceous inhibitors of cysteine peptidases have been identified in eukaryotes, of which the relatively small cystatins (MEROPS family I25) are the most frequent representatives. These chemically exceptionally sturdy proteins consist of a single cystatin domain of 100-120 amino acid residues and exist as two ancestral lineages, the stefins (type 1-cystatins) and the cystatins (type 2- and type 3-cystatins) (for a detailed review, see: [27]). Compared to the other cystatin types, stefins as solely cytosplasmic proteins have no disulphide bonds and also lack a signal sequence. Cystatin C, a type 2-cystatin, was first isolated from chicken egg white, in which it functions as protector against proteolytic digestion of proteinaceous egg white components [28]. Because of its relatively low molecular mass (13.5 kDa) human cystatin C is used as marker for testing renal functions in medical laboratory diagnostics. Relatives of this class of inhibitors are the fetuins containing two cystatin domains and the kininogens, both being type 3 cystatins (Fig. 3A). Kininogens are multi-domain plasma proteins with three successive cystatin units and are found as low molecular kininogen (HMWK) and as high-molecular weight kininogen (LMWK). Several functions have been attributed to type 3 cystatins: Fetuins mediate transport and availability of a wide variety of cargo substances in the blood stream [29], whereas kininogens as components of the kallikrein-kinin-system are involved in important physiological cascade reactions, such as blood coagulation, blood pressure regulation and inflammatory processes [30].

Cystatins generally inhibit the C1-family (papain-type) and the C13-family (legumain-type) of cysteine peptidases. Along with their primary structures, their polypeptide chains are distinguished by three conserved regions that are important for their action as protease inhibitors. These three

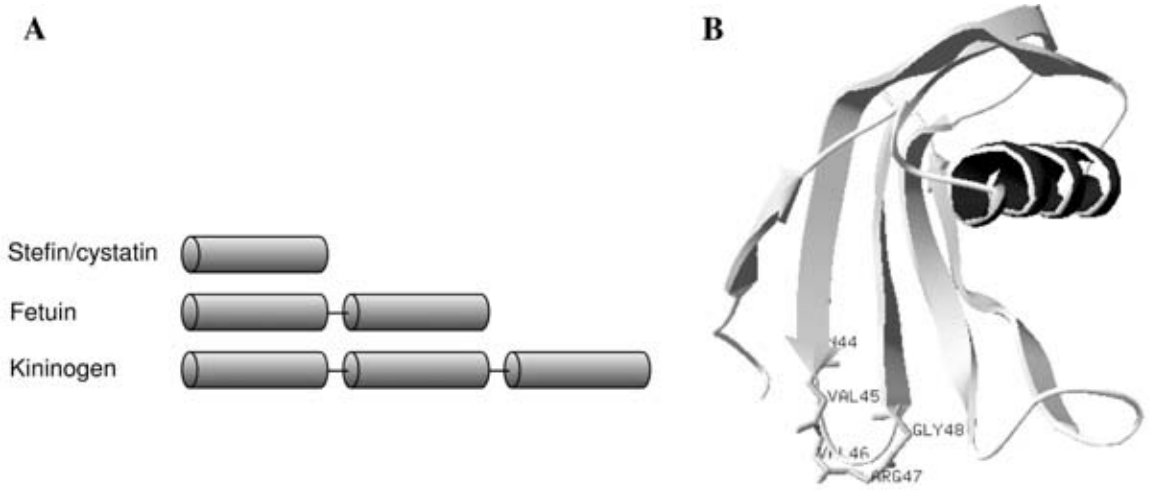

Fig. 3 A. Schematic representation of proteins composed of cystatin-like domains. Each barrel stands for a single cystatin domain. In cases of fetuins and kininogens, any additional non-cystatin domains located at the C terminus have been left out. Note that of the kininogens, only the cystatin domains are shown. **B.** Homology modelling of a giardial cystatin-homologue was done using sialostatin L2 from the tick *Ixodes scapularis* (PDB 3MWZ) as template [46]. The residues of the first hairpin loop are shown.

regions include an N-terminal segment, a highly conserved region with a QXVXG-motif and a small conserved PW(X)-pair in the C-terminal region of the molecule, respectively [31]. The molecule folds to a wedge-like structure with the cone end comprising the protease binding region as elucidated by X-ray crystallography as well as by nuclear magnetic resonance studies using the inhibitor alone and in complex with its target cysteine peptidase (Fig. 3B). The protease binding region of the molecule is built up by the N-terminus together with the two β-hairpin loops formed by the QXVXG-motif (first hairpin-loop) and the PW-pair (second hairpin loop), respectively, and is sterically and chemically complementary to the active site cleft of the target protease [32]. The specific shape of this site and the nature of the amino acid side chains facing it enable a very strong binding of the respective target peptidase to the inhibitor which is expressed by K_i-values being in a picomolar range [33].

Genes encoding proteins of the cystatin superfamily have been identified in two of the three domains of life, namely in *Bacteria* and *Eukarya* including plants, but not yet in *Archaea* [34]. The various cystatin domains found both as single domain and as multi-copy proteins are products of several gene duplications of an anxious cystatin gene during evolution. According to our BLAST searches, the different *Entamoeba spp*., unlike the soil-living amoeba *Dictyostelium discoideum*, obviously do not contain any cystatin genes, whereas the genome of the parasitic protozoon *G. lamblia* is equipped with three genes encoding cystatin homologues. Considering that *Giardia* is supposed to be the earliest diverging eukaryote during evolution, these proteins may represent the most ancient cystatins. However, as predicted from their nucleotide sequences, only one of these giardial cystatins exhibits the canonical protease binding motif, QVVRG, that is part of the central protease binding loop [34]. The arginine residue within this motif correlates with a basic residue in P_2 of a substrate being the preferred split position of the major giardial cysteine protease [18]. This cystatin may help the parasite to regulate the intracellular enzymatic activities of the numerous cysteine proteases expressed by the parasite [12]. The two further open reading frames as disclosed by our extensive screening of the data base of the *Giardia* genome project (http://giardiadb.org/giardiadb/) indeed exhibit significant sequence similarities to this giardial cystatin, albeit the mentioned canonical motif is only rudimentarily discernable in their

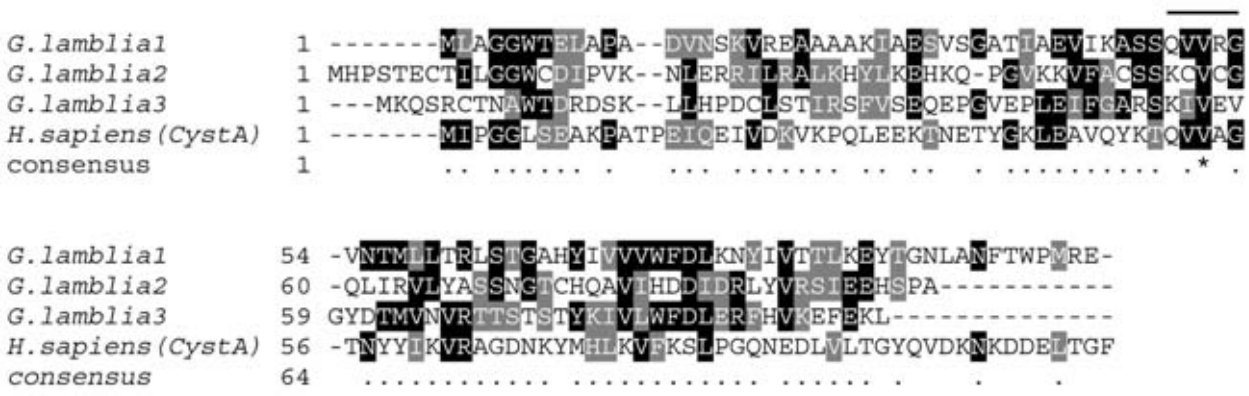

Fig. 4. Sequence alignment of giardial cystatin homologues together with human cystatin A using ClustalW. Identical residues are shown in black, strongly similar ones in grey. Bar above the sequences indicates the first hairpin-loop as the main protease binding motif.

derived primary structures (Fig. 4). All these cystatin-like proteins of *Giardia* have not yet been characterized on a protein level and thus their detailed intracellular functions within the life of *Giardia* are heavily to predict. But although the relevance of these predicted proteins for the life of this parasite is still unknown, the fact that all their derived amino acid sequences do not point to the existence of any signal sequence bears the assumption that these proteins may do their work intracellularly and thus may not directly be involved in the pathogenic processes triggered by the parasite. Beyond that, these ancient cystatins may not necessarily be involved in protease inhibition, but rather fulfil a primordial, still unknown task that requires protein/protein recognition.

Chagasins

Chagasins (MEROPS family I42) comprise a comparatively novel class of peptidase inhibitors evolutionarily not related to those protein molecules already discussed in this review. Chagasins at first have been identified and isolated from the parasitic protozoon *Trypanosoma cruzi*, the causative agent of Chagas disease, which gave the name to this protein family [35]. This inhibitor class, also designated as ICPs (inhibitors of cysteine proteases), represent proteins of molecular sizes similar to the cystatins although being not related to them. In contrast to cystatins, chagasins are found in all kingdoms of organisms, although in eukaryotic organisms they seem to be restricted to single-celled organisms, primarily to parasitic protozoa [36]. Likewise as the cystatins, chagasins are chemically very robust proteins, a property that considerably simplifies their extraction from biological sources [37]. Again like cystatins, chagasin sequences occur both as single- and as parts of multi-domain proteins; multi-copy chagasin domains are also found in single-chain proteins (Fig. 5A). Thus, in the *Archaeon Methanococcus voltae* five chagasin domains altogether constitute a complete polypeptide chain of more than 80 kDa, and two predicted proteins from the aquatic fungus *Allomyces macrogynus* are built up of three chagasin domains. In this diploid phycomycete, another open reading frame encodes a protein having only two consecutive chagasin domains. Derived protein sequences of multimers comprising two chagasin domains have also been found in some *Clostridia spp.* as well as in methanogenic *Archaea*. Further BLAST searches of the protein data base revealed that proteins with single chagasin sequences are found in more than one hundred organisms of all kingdoms, most of them are single domain proteins. However, in some methanogenic *Archaea*, individual chagasin sequences are components of larger proteins containing additional unrelated domains, such as an uncharacterized 10.2 kDa fragment or a β-propeller protein domain as found in *Methanosarcina* or *Methanococcus spp.* Another large protein predicted for the thermophilic bacterium *Pelotomaculum thermopropionicum*, besides its single chagasin sequence, possesses components of a copper

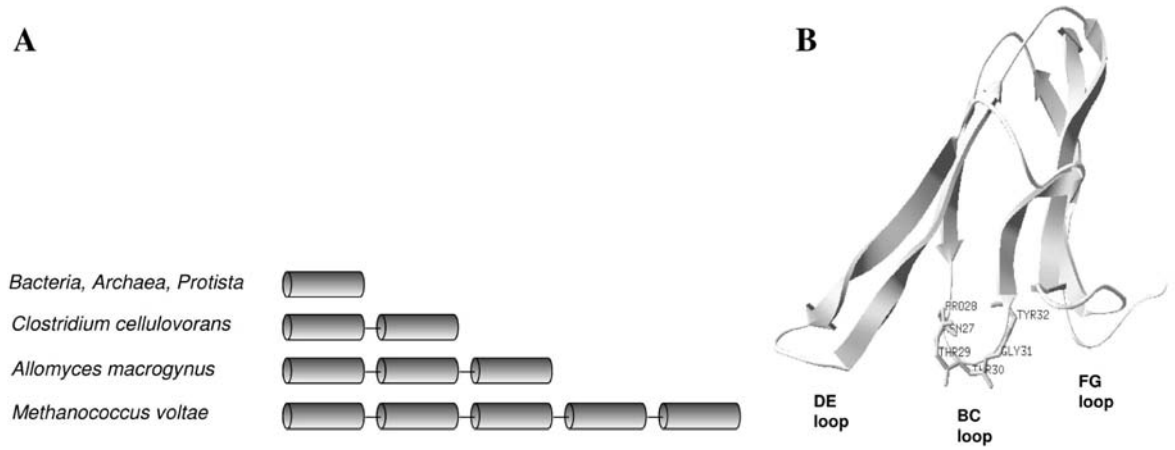

Fig. 5. A. Schematic representation of proteins containing chagasin-like domains. Each barrel stands for a single chagasin domain. Possible additional polypeptide stretches other than chagasins have been left out. **B.** Homology modelling and superposition of EhICP1 and EhICP2 based on the X-ray structure of EhICP2 (PDB 3M88) [40]. The protease contact loops are denoted.

oxidase as well as additional domains of still unknown functions. Similar applies for a protein from the Gram-positive bacterium *Bacillus cereus* and for an extracellular solute binding protein of the thermophilic *Thermotogales bacterium*. All these protein products have not yet been characterized, so nothing can be stated as to their real functions.

As shown in Fig. 6, the primary structures of the individual chagasins in many cases largely diverge, but all of them exhibit three conserved motifs typical for this inhibitor class and thus are suggested to be the result of a horizontal gene transfer [37]. X-ray crystallographic analyses of the single inhibitor as well as of its complex with a target protease revealed that these motifs are embedded in distinct loops that are oriented to the same side of the molecule [38, 39]. These loops named as BC-, DE- and FG-loop, respectively (Fig. 5B), altogether fold to a wedge-like geometry that sterically perfectly fits into the active site of the target cysteine protease. The resulting tight binding hampers the access of a potential substrate resulting in a reversible inhibition of the proteolytic reaction with K_I values in a picomolar range, an inhibitory effect similar to that of cystatins. The exact protease specificities of these inhibitors are not yet clear, but appear mainly to be fixed by the nature of their particular residues in the FG-loops [40].

The closely related protists *E. histolytica* and *E. dispar* each exhibit two genes encoding chagasin-homologs, also referred to as amoebiasins 1 and 2 resp. EhICP1 and EhICP2. In both of these *Entamoeba* species, one of these inhibitor homologs is equipped with an N-terminal signal sequence, required for entering the subcellular sorting pathway and/or secretion of this protein [41]. Recent crystallographic data confirmed that EhICP2, just like the other chagasins whose structures have been elucidated, adopt an immunoglobulin-like folding [40]. The amino acid positions in their FG-loops supposed to be important for their inhibition specificities in each case are occupied by an arginine residue. This basic residue reflects the preferred split position in the P_2-subsite of a substrate and complies well with substrate specificity determinations of the major amebic cysteine proteases [16]. Activity studies in the culture supernatant of the amoebae performed in our laboratory did not provide any indication of a release of cysteine protease inhibitory activity into its environment under culture conditions [42]. Amoebiasins lacking a prosequence are found as expected in the cytosol and thus are supposed to regulate intracellular, cytoplasmic protease activity accidently ended up there. By contrast, ICPs equipped with a prosequence may control the proteolytic processing of other hydrolases in vesicles, or fulfil other, still unknown tasks that need protein/protein recognition

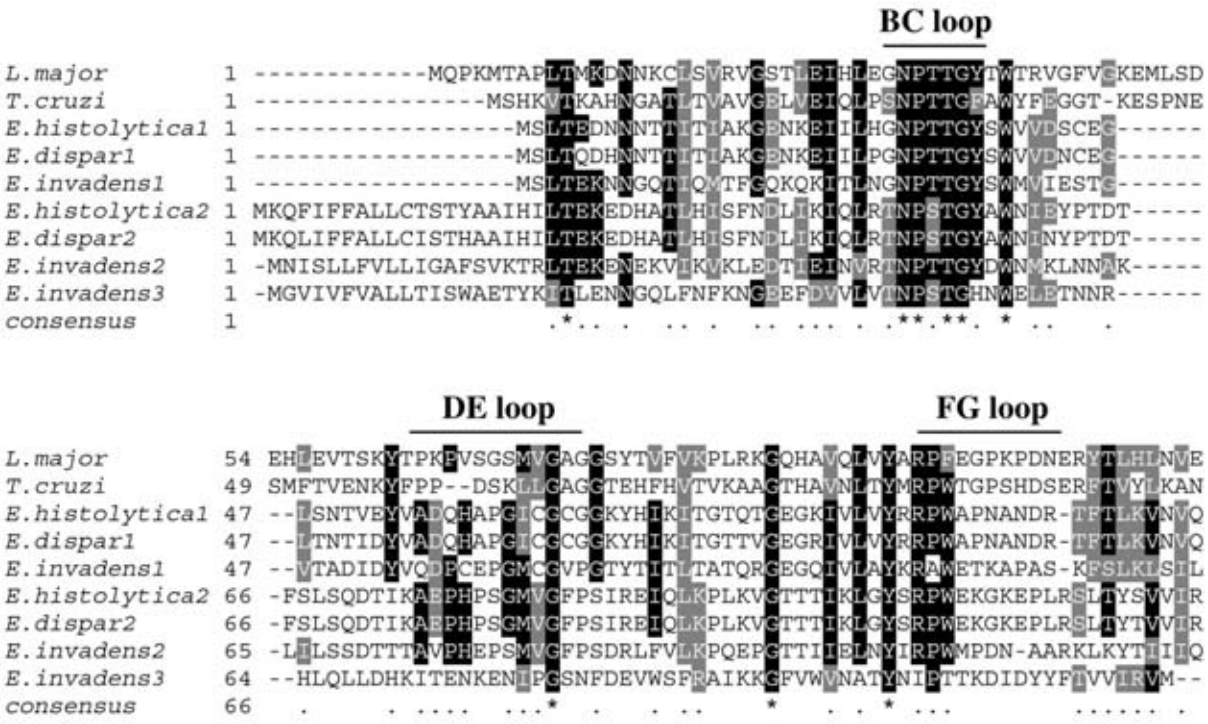

Fig. 6. Sequence alignment of protozoan chagasin-homologues using ClustalW. The three loops contacting the active site of the protease are marked by bars above the sequences. Identical residues are shown in black, strongly similar ones in grey. The last four sequences contain signal sequences with cleavage sites after alanine at positions 17 (*E. histolytica2/dispar2*) and 13 (E. invadens 2/3) predicted by using the Signal P 3.0 server [47].

[41, 42]. These suggestions as to the physiological functions of these inhibitors imply that they are not directly involved in the pathogenesis triggered by invasive *E. histolytica.* This notion is supported by the fact that non-pathogenic *E. dispar* likewise contains two amoebiasin genes strongly similar to those of *E. histolytica*. Although the detailed evidences of these proteins within the life cycle of *Entamoeba* remain unclear, the presence of these inhibitors in pathogenic and non-pathogenic amoebae to a similar extent supports the hypothesis that the own cysteine proteases may be their targets rather than proteases of the host. The reptilian pathogenic *E. invadens* even contains three amoebiasin homologous genes. Of which, however, the sequence of the third one diverges from those of the other two amoebiasins in such an extent, that its attribution to the chagasin class is questionable. Nonetheless, the two actual chagasin genes of *E. invadens* really correspond to the two mentioned chagasin homologues of *E. histolytica/dispar*. The genome of *G. lamblia* obviously does not contain any genes encoding chagasin-homologous proteins. However, the fact that *Giardia* contains at least one cystatin gene, whereas *Entamoeba* does not, raises the assumption that the individual members of both inhibitor classes occupy the same functional niche in these organisms giving rise to the expression of either one of these proteins. This may apply to all organisms because so far no example is known, in which both inhibitor types co-exist.

CONCLUSION

The human pathogenic parasitic protists *Entamoeba histolytica* and *Giardia lamblia* are equipped with different proteinaceous protease inhibitors with partially common properties. Both parasites each contain a single serpin with identical proposed inhibition specificities that surprisingly correspond to the substrates specificities of their major cysteine peptidases. Considering that some serpins are known to also inhibit cysteine proteases it may not be excluded that the frequent cysteine peptidases of these parasites are

also targets of these inhibitors. Additionally, both parasitic protozoa contain members of the well-known proteinaceous cysteine protease inhibitors of the cystatin- and of the chagasin-type, but only either of these members exist in each parasite. Considering that the inhibition mechanisms of chagasins and cystatins are similar, this finding suggests that both inhibitor types are redundantly used in nature. According to that, both inhibitor types also exist as multiple copies in single proteins in some other organisms.

REFERENCES

1. Walsh, J. A. 1986, Arch. Invest. Med., 17 (Suppl. 1), 385.
2. Roberts-Thomson, I. C. 1984, Giardiasis, Tropical and Geographical Medicine, Warren, K. S. and Mahmoud, A. A. F. (Eds.), McGraw-Hill, New York, 319.
3. WHO, 1997, Epidemiol. Bull., Mexico City, Mexico, 13.
4. Martínez-Palomo, A. and Martínez-Báez, M. 1983, Rev. Infect. Dis., 5, 1093.
5. Stanley, S. L. Jr. 2003, Lancet., 361, 1025.
6. Elmendorf, H. G., Dawson, S. C., and McCaffery, J. M. 2003, Int. J. Parasitol., 33, 3.
7. Stanley, S. L. Jr. 2001, Trends Parasitol., 17, 280.
8. Holberton, D. V. 1974, J. Exp. Biol., 60, 207.
9. Ralston, K. S. and Petri, W. A. Jr. 2011, Trends Parasitol., 27, 254.
10. Leippe, M. and Müller-Eberhard, H. J. 1994, Toxicology, 87, 5.
11. Riekenberg, S., Flockenhaus, B., Vahrmann, A., Müller, M. C., Leippe, M., Kiess, M., and Scholze, H. 2004, Mol. Biochem. Parasitol., 138, 217.
12. Sajid, M. and McKerrow, J. H. 2002, Mol. Biochem. Parasitol., 120, 1.
13. Brömme, D. 2001, Curr. Protoc. Protein Sci., 21, Unit 21. 2.
14. Tillack, M., Biller, L., Irmer, H., Freitas, M., Gomes, M. A., Tannich, E., and Bruchhaus, I. 2007, BMC Genomics, 8, 170.
15. Irmer, H., Tillack, M., Biller, L., Handal, G., Leippe, M., Roeder, T., Tannich, E., and Bruchhaus, I. 2009, Mol. Microbiol., 72, 658.
16. Scholze, H. and Tannich, E. 1994, Methods Enzymol., 244, 512.
17. Jacobs, T., Bruchhaus, I., Dandekar, T., Tannich, E., and Leippe, M. 1998, Mol. Microbiol., 27, 269.
18. DuBois, K. N., Abodeely, M., Sakanari, J., Craik, C. S., Lee, M., McKerrow, J. H., and Sajid, M. 2008, J. Biol. Chem., 283, 18024.
19. Rawlings, N. D., Tolle, D. P., and Barrett, A. J. 2004, Biochem. J., 378, 705.
20. Rzychon, M., Chmiel, D., and Stec-Niemczyk, J. 2004, Acta Biochim. Pol., 51, 861.
21. Rigden, D. J., Mosolov, V. V., and Galperin, M. Y. 2002, Protein Sci., 11, 1971.
22. Huntington, J. A. 2011, J. Thromb. Haemost., 9, 26.
23. Roberts, T. H., Hejgaard, J., Saunders, N. F., Cavicchioli, R., and Curmi, P. M. 2004, J. Mol. Evol., 59, 437.
24. Bruno, S., Duschak, V. G., Ledesma, B., Ferella, M., Andersson, B., Guarnera, E. A., and Angel, S. O. 2004, Mol. Biochem. Parasitol., 136, 101.
25. Pszenny, V., Angel, S. O., Duschak, V. G., Paulino, M., Ledesma, B., Yabo, M. I., Guarnera, E., Ruiz, A. M., and Bontempi, E. J. 2000, Mol. Biochem. Parasitol., 107, 241.
26. Riahi, Y., Siman-Tov, R., and Ankri, S. 2004, Mol. Biochem. Parasitol., 133, 153.
27. Turk, V., Stoka, V., and Turk, D. 2008, Front. Biosci., 13, 5406.
28. Barrett, A. J. 1981, Methods Enzymol., 80, 771.
29. Toroian, D. and Price, P. A. 2008, Calcif. Tissue Int., 82, 116.
30. DeLa Cadena, R. A. and Colman, R. W. 1991, Trends Pharmacol. Sci., 12, 272.
31. Bode, W., Engh, R., Musil, D., Thiele, U., Huber, R., Karshikov, A., Brzin, J., Kos, J., and Turk, V. 1988, EMBO J., 7, 2593.
32. Stubbs, M. T., Laber, B., Bode, W., Huber, R., Jerala, R., Lenarcic, B., and Turk, V. 1990, EMBO J., 91, 939.
33. Machleidt, W., Thiele, U., Laber, B., Assfalg-Machleidt, I., Esterl, A., Wiegand, G., Kos, J., Turk, V., and Bode, W. 1989, FEBS Lett., 243, 234.
34. Kordis, D. and Turk, V. 2009, BMC Evol. Biol., 9, 266.

35. Monteiro, A. C., Abrahamson, M., Lima, A. P., Vannier-Santos, M. A., and Scharfstein, J. 2001, J. Cell Sci., 114, 3933.
36. Santos, C. C., Coombs, G. H., Lima, A. P., and Mottram, J. C. 2007, Mol. Microbiol., 66, 991.
37. Rigden, D. J., Monteiro, A. C., and Grossi de Sá, M. F. 2001, FEBS Lett., 504, 41.
38. Figueiredo da Silva, A. A., de Carvalho Vieira, L., Krieger, M. A., Goldenberg, S., Zanchin, N. I., and Guimarães, B. G. 2007, J. Struct. Biol., 157, 416.
39. Ljunggren, A., Redzynia, I., Alvarez-Fernandez, M., Abrahamson, M., Mort, J. S., Krupa, J. C., Jaskolski, M., and Bujacz, G. 2007, J. Mol. Biol., 371, 137.
40. Casados-Vázquez, L. E., Lara-González, S., and Brieba, L. G. 2011, Gene, 471, 45.
41. Sato, D., Nakada-Tsukui, K., Okada, M., and Nozaki, T. 2006, FEBS Lett., 580, 5306.
42. Sarić, M., Irmer, H., Eckert, D., Bär, A. K., Bruchhaus, I., and Scholze, H. 2011, Protist. Epub., ahead of print.
43. Thompson, J. D., Higgins, D. G., and Gibson, T. J. 1994, Nucleic Acids Res., 22, 4673.
44. Felsenstein, J. 1989, Cladistics, 5, 164.
45. Li, J., Wang, Z., Canagarajah, B., Jiang, H., Kanost, M., and Goldsmith, E. J. 1999, Structure Fold. Des., 7, 103.
46. Kotsyfakis, M., Horka, H., Salat, J., and Andersen, J. F. 2010, Mol. Microbiol., 77, 456.
47. Emanuelsson, O., Brunak, S., von Heijne, G., and Nielsen, H. 2007, Nat. Protoc., 2, 953.