Cysteine Proteases and Cell Differentiation: Excystment of the Ciliated Protist Sterkiella histriomuscorum

Eduardo Villalobo,† Clara Moch, Ghislaine Fryd-Versavel, Anne Fleury-Aubusson, and Loïc Morin*

Laboratoire de Biologie Cellulaire 4, Université de Paris-Sud, Orsay, France

Received 4 April 2003/Accepted 5 August 2003

The process of excystment of Sterkiella histriomuscorum (Ciliophora, Oxytrichidae) leads in a few hours, through a massive influx of water and the resorption of the cyst wall, from an undifferentiated resting cyst to a highly differentiated and dividing vegetative cell. While studying the nature of the genes involved in this process, we isolated three different cysteine proteases genes, namely, a cathepsin B gene, a cathepsin L-like gene, and a calpain-like gene. Excystation was selectively inhibited at a precise differentiating stage by cysteine proteases inhibitors, suggesting that these proteins are specifically required during the excystment process. Reverse transcription-PCR experiments showed that both genes display differential expression between the cyst and the vegetative cells. A phylogenetic analysis showed for the first time that the cathepsin B tree is paraphyletic and that the diverging S. histriomuscorum cathepsin B is closely related to its Giardia homologues, which take part in the cyst wall breakdown process. The deduced cathepsin L-like protein sequence displays the structural signatures and phylogenetic relationships of cathepsin H, a protein that is known only in plants and animals and that is involved in the degradation of extracellular matrix components in cancer diseases. The deduced calpain-like protein sequence does not display the calcium-binding domain of conventional calpains; it belongs to a diverging phylogenetic cluster that includes Aspergillus palB, a protein which is involved in a signal transduction pathway that is sensitive to ambient pH.

The exit from dormancy, of which excystment is a particular case, is a widespread process in many eukaryotic and prokaryotic organisms. It is generally accompanied by cellular differentiation, during which numerous intra- and extracellular structures are highly modified, as observed during the germination of fungal conidia and plant seeds (reviewed in references 5, 31, and 39), among others. Cysteine proteases, which are known to be involved in intracellular protein turnover and extracellular matrix remodeling (37), constitute good candidates to take part in this general process. In this paper, we describe the involvement of such proteins in the process of excystment of the ciliate Sterkiella histriomuscorum.

S. histriomuscorum (previously called Oxytricha fallax and O. trifallax) (3) is a free-living ciliated protist from the hypotrich subgroup that is characterized by a highly differentiated pattern of the vegetative cell. This pattern includes a complex oral apparatus and a clustered body ciliumate which are duplicated in an orderly process during cell division (22). The S. histriomuscorum life cycle is interesting in that the vegetative cell undergoes a complete dedifferentiation process when it is deprived of food, leading in about 3 days to a resting stage which is called a cyst. The mature cyst has lost much water, the whole cilature and basal bodies have disappeared, numerous autophagic vacuoles (lysosomes) that were particularly active during encystment are present, and a multilayered cyst wall surrounds the spherical cell, which appears to be metabolically inactive (23). The nature of the cyst wall is unknown, but it may be hypothesized that it includes glycoproteins, as was shown in the case of Parastroystyla sp., a related hypotrich species (29, 44).

In presence of either living food or 0.2% milk powder in water, the excystment process takes place in a few hours, leading through a massive influx of water and the resorption of the cell wall to completely differentiated swimming vegetative cells.

Different taxonomic groups, including both unicellular and multicellular organisms, display encystment-excystment behavior, and two previous studies have shown that cysteine proteases are involved in the excystment process. The first described two distinct proteases that modulate excystation in the case of the trematode Paragonimus westermani metacercariae (9). The second showed that cathepsin B was the protease responsible for the cyst wall degradation in the diplomonad flagellate Giardia (60).

Cysteine proteases constitute a wide superfamily of proteases, including viral, legumain-like, and papain-like proteases. The papain-like group is further divided into families, including calpains and papains, and the papain family includes the bleomycin hydrolase, cathepsin B, and cathepsin L subfamilies (47). Calpains are calcium-dependent cystolic enzymes, while cathepsins are located mainly in lysosomes. Both were previously known to be involved in intracellular protein turnover in mammals, but it appears that they display a much larger array of functions in protists (26, 37, 40).

During a general study of the genes that are differentially expressed in the cyst and vegetative cells, i.e., during the excystment of S. histriomuscorum, we isolated three different cDNAs whose sequences appeared, based on a BLAST analysis, to be highly similar to those of the calpain, cathepsin B,
and cathepsin L family genes. Excytation (but not excytation or cell division) was blocked at a particular differentiating stage when cysteine protease inhibitors were added in the nutrient medium, further confirming the involvement of cysteine proteases in the excytment process. The expression pattern was life cycle stage dependent for both cysteine protease genes. The calpain-like gene is not expressed strongly in cysts and early excysting cells, while cathepsins are expressed more in vegetative cells than in resting cysts and early excysting cells.

The whole sequences of the corresponding minichromosomes were determined, and a comparative analysis of the protein sequences shows an important deletion in the occluding loop of cathepsin B, a cathepsin H signature in the cathepsin L-like sequence, and the absence of the calpain III linker domain and calcium-binding motif in the calpain-like sequence. A phylogenetic study confirmed the cathepsin H identification and the diverging status of the cathepsin B and calpain genes.

**MATERIALS AND METHODS**

**Cell culture, encystment, and excytment.** *S. histrio momusor* strain BA was grown at room temperature in commercial mineral water (Volvic [Volvic, France]) with Tetraselmis sp. or Chlorellas sp. as a food source. Excystment of cells was induced by starvation. Excystment was induced either by adding food (a grown culture of *Aerobacter aerogenes* eubacteria) to the cyst medium or by transferring cysts to a 0.2% milk powder medium or 0.1% H9262 (downstream). To test for excytment, cells were incubated with 10 μM inhibitor in mineral water. To test for excytment, starving vegetative cells were incubated with 10 μM inhibitor in mineral water. To test for excytment, cysts (about 2,000) were incubated into either food medium or the excytment inhibitors were obtained from Sigma and included E64° untreated controls and inhibitor solvent controls were performed. The tested inhibition experiments, dividing vegetative cells were incubated with 10 μM inhibitor in feeding medium. To test for encystment, starving vegetative cells were incubated with 10 μM inhibitor in mineral water. To test for excytment, cysts (about 2,000) were incubated into either food medium or the excytment medium described above with 50 nM to 100 μM inhibitor. Parallel assays of untreated controls and inhibitor solvent controls were performed. The tested inhibitors were obtained from Sigma and included E64° [N-(trans-eposuccinyl)-l-leucine 4-guanidinobutyramide], CA-074° (L-trans-eposuccinyl-Il-PRO-OH-propylamide), CA-074Me (L-trans-eposuccinyl-Il-Pro-OMe-propylamide), MG132 (Z-Leu-Leu-Leu-al), calpain inhibitor II (L- boric acid, 2 mM EDTA [pH 8.3]), and fragments smaller than 400 bp were cut by Villalobo et al. (57). A time-specific display reverse transcription-PCR (RT-PCR) was carried out as described by Adl and Berger (3).

For inhibition experiments, dividing vegetative cells were incubated with 10 μM inhibitor in feeding medium. To test for encystment, starving vegetative cells were incubated with 10 μM inhibitor in mineral water. To test for excytment, cysts (about 2,000) were incubated into either food medium or the excytment medium described above with 50 nM to 100 μM inhibitor. Parallel assays of untreated controls and inhibitor solvent controls were performed. The tested inhibitors were obtained from Sigma and included E64° [N-(trans-eposuccinyl)-l-leucine 4-guanidinobutyramide], CA-074° (L-trans-eposuccinyl-Il-PRO-OH-propylamide), CA-074Me (L-trans-eposuccinyl-Il-Pro-OMe-propylamide), MG132 (Z-Leu-Leu-Leu-al), calpain inhibitor II (N-acetyl-Leu-Leu-Met-al), and lactacystin.

Assessment of cell division, encystment, and excytment was done with a Leitz dissecting microscope. Excytment rate analysis was carried out after an incubation time of 24 h by counting swimming vegetative cells. For cell structure examination, samples were impregnated as described for the silver proteinase method (54); cysts were then either mounted in glycerol for direct observation or dehydrated and mounted in Canada balsam for conservation.

**Differential-display reverse transcription-PCR (RT-PCR)**. A differential-display reverse transcription-PCR (RT-PCR) was carried out by using the QIAquick gel extraction instructions. Nucleic acid hybridization was conducted by standard methods. Clones of interest were isolated, and plasmids were extracted and purified by the conventional alkaline lysis method (48). Sequencing reactions with plasmid DNA were done with an automated DNA sequencer (Vistra; Amersham) and the DNA cycle sequencing kit and Texas Red M13 universal primers (Amersham).

**Inverse PCR.** The ciliate nuclear organization is characterized by the presence of diploid micronuclei, containing the usual diploid chromosomes, and macronuclei which contain expressed genes embedded in minichromosomes. In hypo- trophic ciliates such as *S. histrio momusor*, each minichromosome is a double-stranded DNA molecule with an open reading frame flanked by 5° leader and 3° tailer sequences and specific telomeres; this structure is then duplicated several times in order to produce about 1,900 copies of each sequence (43). This particular organization makes it possible to circulate these minichromosomes and to use the inverse-PCR technique to amplify and clone the whole minichromosome.

Total genomic DNA from *S. histrio momusor* was covalently circularized by ligation, as previously described (57), and then 100 ng of circularized DNA was subjected to PCR in a 50-μl reaction mixture with the Expand long-template PCR system (Boehringer) according to the manufacturer's procedures and with 35 cycles of denaturation, annealing, and polymerization. The PCR products were run on a 1% (wt/vol) agarose gel, extracted and purified as described above, and then cloned in the pGEM-T Easy plasmid.

The sequences of the outward-facing specific primers used to amplify the macronuclear sequences corresponding to each fragment of interest were as follows: for the cathepsin B gene, 5°-TTGGATTTGGGCAACACCTTT-3° (upstream) and 5°-TATACGGGTGTACACCCCA-3° (downstream); for the cathepsin L-like gene, 5°-CGTCITTGAAGCTCAACAGTGC-3° (upstream) and 5°-TGGAGAATGCTTCGACC-3° (downstream); and for the calpain gene, 5°-AAAGTGAACAGTTGAATGC-3° (upstream) and 5°-ATCA GCCAATCTCCCATC-3° (downstream).

**RT-PCR.** Total RNAs (2 μg) from cysts, 1-h excysting cells, and vegetative cells were reverse transcribed by using 3.5 μM random hexamers (pdN6; Amersham) as primers and 200 μM of Moloney murine leukemia virus reverse transcriptase (Eurobio) in a 20-μl reaction mixture containing 40 U of RNasin (Promega). RNA was incubated at 65°C for 5 min prior to the RT step. After 50 min at 65°C, the enzyme was inactivated at 95°C for 5 min. The PCR step was subsequently carried out with 0.1 and 0.3 μl of the RT product, a pair of specific primers, and 1.5 U of Taq polymerase in a 25-μl reaction mixture. The sequences of the forward and reverse primers were, respectively, as follows: for the cathepsin B gene, 5°-TTCCCGACAGATCTGCC-3° and 5°-GGGATGGAAGGAGGCGT-3°; for the calpain-like gene, 5°-TCACATGTTGTTGGT-3° and 5°-GGGTTAAGAGTGGACGT-3°; and for the calpain gene, 5°-AAACACTC CAGGCCTCAAAATG-3° and 5°-AAAGAATTCTCAATCCGGCAT-3°.

**Sequence analysis and reconstruction of phylogenetic trees.** The gapped BlastX program from the National Center for Biotechnology Information was used to find searching protein identities SMART software (34) was used to search domains in the predicted protein sequences. All sequences except *S. histrio momusor* sequences were retrieved from GenBank to constitute a data set of 59 sequences for cathepsin B, 349 sequences for cathepsin-L like proteins, and 79 sequences for calpain. Management, formatting of the sequences for the tree-building programs, and construction of distance trees were carried out with the MUST package (41). Sequence alignments were manually performed with the ED program of the package. Cathepsin B and cathepsin L-like sequence data sets were disjoined in order to optimize the node resolution in each subtree for two reasons: first, they constitute two rather diverging branches in the general tree of cathepsin proteins (4, 60), and second, the number of unambiguously aligned characters in each subset (240 and 229, respectively, for cathepsin B and cathepsin L-like sequences) is much higher than it is in the two combined data sets (181). For legibility reasons, the final analysis retained only representatives of the main clusters in each tree, i.e., 37 sequences for cathepsin B, 72 for cathepsin L, and 41 for calpain. The distance matrix was calculated by considering all amino acid differences without weighting the transition probabilities. This matrix served as a basis for phylogenetic reconstructions by the neighbor-joining procedure (46). A 1,000-bootstrap resampling analysis was carried out with the Nboot program. Parsimony and maximum likelihood analyses were conducted for 100 bootstrap resamplings with the PROTPARS and PROML programs from the PHYLIP package (14); for the latter analysis, the Jones-Taylor-Thornton model was used with a constant rate of change among sites.

**Nucleotide sequence accession numbers.** The nucleic acid and deduced amino acid sequences of the cathepsin L-like, and A gene have been deposited in GenBank under accession numbers AY204512, AY204513, and AY204514, respectively.
RESULTS

Isolation of the cysteine protease molecules. From the differential-display RT-PCR gels, we selected several bands that appeared at first to be much more expressed in vegetative and excysting cells than in resting cysts. Subsequent cloning and insert sequencing led to the identification, as determined from BLAST searches, of putative genes of cathepsin L-like and calpain-like proteins for two of the isolated clones. From the 2-h time-specific excysting library, a crude analysis by random sequencing of several clones allowed us to recover a particular clone which was found several times and whose insert appeared to encode a putative cathepsin B protein.

The sequences of these short cDNA inserts (about 400 bp) were used as anchoring sequences to design primers that were utilized with the inverse-PCR technique in order to clone the minichromosomes bearing the cathepsin B, the cathepsin L-like, and the calpain-like genes.

Effects of cysteine protease inhibitors. In order to confirm the involvement of cysteine protease molecules in the excystment process, the effects of several cysteine protease inhibitors were tested with growing, excysting, and excysting cells. We used 1-trans-epoxysuccinyl peptides such as E64, CA-074, and its membrane-permeable derivative CA-074Me; peptide aldehydes such as MG132 (Cbz-Leu-Leu-leucinal) and calpain inhibitor II (N-acetyl-Leu-Leu-Met-aldehyde); and a natural inhibitor with a different structure, lactacystin. No inhibitor, whatever concentration was used (up to 1 mM), was able either to prevent or to delay cell division in a culture of growing cells or to block the excystment process in a nutrient-depleted medium. In contrast, E64, calpain inhibitor II, and MG132 were able to block the excystment at a specific stage after the onset of the cell differentiation process. Efficient inhibitor concentrations were determined by counting swimming vegetative cells, leading to a value of 10 μM for complete inhibition (Fig. 1). Cell observation during the experiments suggests that excystation started normally, as we witnessed a change of cyst refringency and the onset of contractile vacuole beating. Moreover, blocked cysts that were fixed and protargol impregnated display the ciliary “anlagen” as a whole on the cell cortex; that is, the basal body assembly is almost finished (Fig. 2). Washing the blocked cells with fresh medium lacking the inhibitor reactivated the excystment process and led to normally dividing cells. These results clearly show that the active inhibitors are responsible for a specific blockage of the excystment process and that the inhibition occurs at a late stage of excystation.

E64 is a nonspecific inhibitor of cysteine proteases, calpain inhibitor II, and MG132 are relatively nonspecific inhibitors of cysteine proteases and proteasome (33), and CA-074 and CA-074Me are specific inhibitors of cathepsin B (53); lactacystin is a specific inhibitor of the proteasome (33). The excystment process involves cysteine protease activities (apparently with the exception of cathepsin B) and is not dependent of the proteasome activity.

Expression pattern. The RT-PCR patterns observed for the three cysteine proteases are displayed in Fig. 3. Cathepsin B mRNA is strongly expressed in vegetative growing cells which are actively dividing, while a relatively low expression level is observed in cysts and in 1-h excysting cells. Cathepsin L-like mRNA displays a pattern which appears as similar to that of cathepsin B, except that the expression level in vegetative cells is hardly greater than in cysts and 1-h excysting cells. Calpain mRNA is apparently absent in cysts and 1-h excysting cells but is strongly expressed in growing cells. These results suggest that the three cysteine proteases are active during the normal vegetative cell growth and that their expression, at least for cathepsin B and calpain, is upregulated during the excystment process; moreover, this expression is increased at least after 1 h of excystment induction, suggesting that upregulation occurs at a late stage after the onset of excystation, in agreement with the above observations concerning the effects of cysteine protease inhibitors.

Structure and phylogeny. The minichromosome bearing the cathepsin L-like gene is around 1.3 kbp long, as confirmed by Southern blot experiments (data not shown). Both ends of the chromosome display extremely AT-rich regions (82% at the 101-bp-long 5’ end and 79% at the 109-bp-long 3’ end) surrounding a central region of 1,101 bp which is 60% AT rich and probably corresponds to the transcribed gene. The poly(A) tract of the mRNA is inserted after the position 1298, 85 bp downstream from the coding sequence and very close to the 3’ telomere sequence. The translated amino acid sequence, using the ciliate nuclear genetic code (TAA and TAG codons code for Gln), discloses a single open reading frame with six TAA codons and 1 TAG codon and a putative methionine start codon located at the end of the 5’ AT-rich region of the sequence. The sequence is 366 amino acids long, and a preliminary BLAST analysis shows a strong similarity with cathepsin L-like proteins.

The deduced protein displays the usual structural characteristics of cathepsins, with a strongly conserved active site involv-
FIG. 2. Excystment process (A to F) and effect of the E64 inhibitor (G) as observed on cells after silver proteinate impregnation and mounting in Canada balsam (B to F). Bars, 10 μm. (A to F) Excystment process. The resting cyst (A) is devoid of ciliature and enclosed in a wall (arrowheads). The macronucleus (Ma) is visible in the center of the cell. After induction of excystment, the first cortical event detected by silver proteinate impregnation is the appearance of a field of basal bodies (arrows in panel B). Thousands of basal bodies are quickly assembled and organized into anlagen (C and D), which will differentiate into specific ciliary structures of the young (E) and interphase (F) cell. Several of these anlagen are recognizable very early, and their maturation is easily monitored through the whole process. Adoral membranelles (AM) (or polykineties) and paroral membrane (P) (or haplokineties) represent the two sides of the oral ciliature in the interphase cell (F). The AM differentiate from the field of basal bodies already seen in panel B. This field enlarges (C); basal bodies are progressively organized into transversal rows (D) characteristic of the cell pattern (E and F). The paroral ciliature (P) of the vegetative cell (F) differentiates from a streak of basal bodies situated on the right of the AM anlage (C to E). The ventral cirri (C), which appear as dark points in the interphase cell (F), differentiate from streaks of basal bodies (C) which segment into individual cirri (D) and become patterned on the cortex as cell enlarges (E and F). The specific pattern of the interphase cell may be clearly recognized as the cell is still enclosed in its wall (arrowheads in panel E). During the process, the macronucleus elongates (D) and separates into two parts (E and F). (G) Effect of the E64 inhibitor. Inhibition of the excystment occurs as soon as the oral apparatus (AM and P) as well as those of the cirri (C) are visible. The macronucleus (Ma) has not separated into two parts.
A codon located at the end of the 5' AT-rich region of the sequence. The sequence is 294 amino acids long, and a preliminary BLAST analysis shows a strong similarity with cathepsin B-like proteins.

The protein displays a conserved active site identical to that observed in cathepsin H (Fig. 4B); the propeptide region is 75 amino acids long, much shorter than the cathepsin H propeptide region. A signal peptide (MFKLVIIGTIVAVAVH) was identified at the amino-terminal end of the molecule, here also following the bacterial rule for signal peptidase I cleavage. The mature peptide displays an important deletion in the so-called occluding loop, which is involved in the dipeptidyl carboxypeptidase activity of the enzyme; this deletion is also observed in the cathepsin B sequence of Giardia (Fig. 7), suggesting that both enzymes display solely the endopeptidase activity. Similar deletions are observed in plant cathepsin Bs and in some cathepsins from Caenorhabditis elegans. Only one putative N-glycosylation site was found in the propeptide domain, suggesting a relatively low-affinity binding with mannose-6-phosphate receptors and the possibility that this cathepsin B could be mostly excreted (12).

The neighbor-joining phylogenetic tree of cathepsin B proteins displays two paraphyletic clusters (Fig. 8). The S. histrio-muscorum sequence is strongly clustered with Giardia cathepsin B, thus confirming the shared structural signature of the occluding loop; both molecules are themselves clustered with a monophyletic subgroup involving particular cathepsins, either isolated as excreted tubulointerstitial nephritis antigens (Homo, Mus, and Oryctolagus) or deduced from genome sequencing (Caenorhabditis). Altogether, these molecules appear to be related to the plant cathepsins, as opposed to the more conventional animal cathepsin B clustering.

The minichromosome bearing the calpain-like gene is around 3.1 kbp long, as confirmed by Southern blot experiments (data not shown); the variations in the percent GC contents define three regions: a 5'-terminal region (393 bp) and a 3'-terminal region (133 bp) that are extremely AT rich (76 and 80%, respectively) and a central region of 2,511 bp which is 64% AT rich. A single open reading frame is included in the central region, with 7 TAG and 15 TAA codons, and a putative methionine start codon is located at the end of the 5'
AT-rich terminal region. The protein sequence is 836 amino acids long, and a preliminary BLAST analysis shows a strong similarity with the large subunit of calpain-like proteins. The protein displays a cysteine protease domain located at the amino-terminal part of the molecule (Fig. 4C); the conserved active site of this domain is identical to those of cathepsins B and H. There is no signal peptide, suggesting a cytosolic location of the protein. Conventional calpains display a calpain III linker domain followed by a calcium-binding domain at the carboxy-terminal end of the molecule (51), but no such domain was found in the *S. histriomuscorum* sequence. However, an internal repeat was observed in the carboxy-terminal region, and the first repeat displays good similarity with the ciliate *Eufolliculina uhligi* calcium/calmodulin-dependent protein kinase catalytic domains X and XI (36).

The neighbor-joining phylogenetic tree of calpain protein also displays paraphyletic clusters, among which are represented the known mammalian sequences (52) (Fig. 9). The position of the *S. histriomuscorum* sequence is poorly defined: in the neighbor-joining distance analysis, it is related to the *Drosophila melanogaster* sol (small optic lobes) protein, and in the parsimony analysis, it is related to plant calpains (63.7% bootstrap value); in the maximum likelihood analysis, no significant clustering is observed. Nevertheless, whatever the reconstruction method that was used, it clearly belongs to a larger phylogenetic group including plants, trypanosomes, fungal *palB*, and animal calpain 7, to the exclusion of the more conventional animal calpains. Sequences which so far have been more precisely characterized in this group (*sol*, *palB*, *cap5.5*, *capn7*, and *Dek1*) do not display the calcium-binding domain IV of conventional calpains.

**DISCUSSION**

Inhibition experiments indicate the involvement of cysteine proteases during late stages of the excystment process. These experiments clearly show that inhibitor targets, i.e., cysteine proteases and/or proteasome, are specifically involved in the excystment process. MG132 and calpain inhibitor II are known to block proteasome, calpain, and cathepsin activities (33), while E64 is a general inhibitor of cysteine proteases (6, 8, 42). Treating cysts with either MG132 or E64 leads to an excystment arrest, a result that clearly suggests that the proteasome is not directly involved in the excystment process. This indication is further reinforced by the fact that lactacystin (a specific inhibitor of proteasome) has no effect on the excystment process. However, this negative effect must be analyzed cautiously: while lactacystin is known to be active in vivo in cultured cells (19, 33), we cannot completely rule out the possibility that it might be actively pumped out of the cell, as is the case with yeast cells (32). Nevertheless, the results suggest that cathepsins and/or calpains are the main proteases involved in the excystment process.

The apparent absence of an effect of CA-074 and CA-074Me is rather surprising, since it was shown previously that cathepsin B is involved in the cyst wall breakdown of *Giardia* cysts (60). We may hypothesize that these inhibitors are unable either to enter the cell or to link their target, due to structural particularities. The first hypothesis can be rejected because the process of cyst wall breakdown implies extracellular secretion of cathepsin B, which then would be directly accessible to the inhibitor; this would result in an incomplete cell wall breakdown. On the other hand, *S. histriomuscorum* and *Giardia* molecules do not display the occluding loop which is specific to conventional cathepsin B proteins. Because histidine residues from this loop are essential for the inhibition specificity of cathepsin B, which then would be directly accessible to the inhibitor; this would result in an incomplete cell wall breakdown. On the other hand, *S. histriomuscorum* and *Giardia* molecules do not display the occluding loop which is specific to conventional cathepsin B proteins. Because histidine residues from this loop are essential for the inhibition specificity of cathepsin B, which then would be directly accessible to the inhibitor; this would result in an incomplete cell wall breakdown. Finally, the absence of the occluding loop implies that cathepsin B works as an endopeptidase (the dipeptidyl carboxypeptidase activity is absent) in the process of cyst wall breakdown.

The blockage of excystment occurs at a specific stage: all ciliary anlagen are present in blocked cysts. Many arguments
FIG. 6. Neighbor-joining phylogenetic tree of 72 cathepsin L-like sequences, including 229 amino acids. GenBank accession numbers are indicated when no precise identification was published and when several sequences were available for one species. Bootstrap values at the nodes are noted (neighbor-joining/parsimony/maximum likelihood, respectively); values under 50% have been discarded and are indicated by dots. The scale bar indicates 2.6834% of amino acid substitutions.

2.6834 species, 229 characters
suggest that this stage could correspond to the commitment point of morphogenesis. In ciliates, whatever cycle is concerned (division, sexual cycle, regeneration, or excystment), the morphogenesis always proceeds through the same chronology: a new ciliature is assembled in the form of ciliary anlagen at specific locations on the cortex, and then all of these ciliary elements move on the cortex towards their specific location. All of these stages are well known during the fission cycle in many species (reviewed in references 18 and 20). Detailed studies performed with Paramecium (1) and Tetrahymena (17) have demonstrated that, as in metazoans, ciliates become committed to division at a specific stage; this stage is characterized by the presence of a complete set of ciliary organelles (1) and is thought to result from the classical cyclin/kinase cascade involving the presence of a complete set of ciliary organelles (1) and is thought to result from the classical cyclin/kinase cascade.

Structural and phylogenetic analyses suggest putative functions for the isolated cathepsins and calpain. The above discussion strongly suggests that cathepsin B is involved in the cyst wall breakdown process; while the mRNA differential-display technique may produce false-positive results, the isolated cathepsin H and calpain-like genes are not the conventional and abundant cathepsin L and capn1 (µ) or capn2 (m) calpain genes that would have been isolated in a cDNA random screening. We thus hypothesize that the three isolated molecules are the main inhibitor targets involved in the excystment process. If we consider their gene expression patterns, only the cathepsin B and calpain genes display a strong differential expression between the cyst and growing cell stages. Vegetative cells divide and encystment proceeds, apparently without any problem, whatever inhibitor has been added. This observation is similar to that for Giardia, in which cathepsin B is essential for the excystation to be achieved and the corresponding gene is strongly expressed in vegetative cells (60). This is not the case of Entamoeba, where lactacystin blocks in vivo the excystment process (19). As was previously reported for Saccharomyces, we assume that during the vegetative growth, either the inhibitor is actively pumped out of the cell or inhibition is not complete (15, 32).

Cathepsin B, which we know enters the secretory pathway and could be mostly excreted, would be involved in the breakdown of the cyst wall, in a way similar to that observed for Giardia. It is noteworthy that the cluster most related to the S. histriomuscorum cathepsin B sequence is made up of sequences of tubulointerstitial nephritis antigen, a protein which is accumulated in the basement membrane and is able to promote the adhesion of epithelial cells (28, 30), i.e., a protein which is excreted and involved in cell wall remodeling. Concerning the related plant cluster, a recent study describes a molecule involved in cell wall remodeling by cleavage of cell wall structural proteins in Phleum pratense (24).

Cathepsin H also enters the secretory pathway and should be directed mostly to lysosomes through the mannose-6-phosphate receptor pathway. However, some of it also could be excreted in order to take part in the cell wall breakdown process. Cathepsin L-like proteins are involved in intracellular protein degradation throughout their lysosomal activity; outside lysosomes, their activity contributes to extracellular matrix remodeling and apoptosis (56). Cathepsin H expression is increased in several cancer diseases and could be involved in the degradation of extracellular matrix components, leading to cell proliferation, migration, and metastasis (59). On the other hand, plant cathepsin H-like proteins are strongly expressed during leaf senescence and seedling germination, but also following water stress (7, 21), and are regulated by the plant hormones gibberellic acid and abscissic acid (45). These data suggest that secreted S. histriomuscorum cathepsin H could participate in the cyst wall breakdown, even if there is no strong differential expression of this molecule during the excystment process.

The calpain-like molecule does not display EF hands (which are characteristic of the calcium-binding domains) in its carboxy-terminal part, but an internal repeat similar to a calcium/calmodulin-dependent protein kinase domain has been found in this region. In the case of the conventional m-calpain, only one EF hand of the calcium-binding domain is strictly required for calcium-induced activation (13, 27). Moreover, the calpain-like tra-3 molecule displays a calcium-dependent protease activity, but lacks the EF hand domains (50). Although there is no formal evidence, these observations suggest that calcium ions could be involved in some way in the activation of the isolated S. histriomuscorum calpain. From other experiments, we also know that treatment of ionophore-permeabilized cysts with calcium ions is able to set off the excystment process,
FIG. 8. Neighbor-joining phylogenetic tree of 37 cathepsin B sequences, including 248 amino acids. GenBank accession numbers are indicated when no precise identification was published and when several sequences were available for one species. Bootstrap values at the nodes are noted (neighbor-joining/parsimony/maximum likelihood, respectively); values under 50% have been discarded and are indicated by dots. The scale bar indicates 3.0818% of amino acid substitutions.

3.0818  37 species, 248 characters

FIG. 8. Neighbor-joining phylogenetic tree of 37 cathepsin B sequences, including 248 amino acids. GenBank accession numbers are indicated when no precise identification was published and when several sequences were available for one species. Bootstrap values at the nodes are noted (neighbor-joining/parsimony/maximum likelihood, respectively); values under 50% have been discarded and are indicated by dots. The scale bar indicates 3.0818% of amino acid substitutions.
3.9715  41 species, 268 characters

FIG. 9. Neighbor-joining phylogenetic tree of 41 calpain sequences, including 268 amino acids. GenBank accession numbers are indicated when no precise identification was published and/or when several sequences were available for one species. Bootstrap values at the nodes are noted (neighbor-joining/parsimony/maximum likelihood, respectively); values under 50% have been discarded and are indicated by dots. The scale bar indicates 3.9715% of amino acid substitutions.
without any addition of some other nutrient (R. Lescasse, G. Fryd-Versavel, and A. Fleury, unpublished data). Thus, a release of calcium ions in the cyst cytoplasm could be a signal for calpain-like activation. In other respects, RT-PCR experiments have shown that the S. histrio-muscum calpain-like molecule is not expressed during the first stages of the excystment process, suggesting that a signal cascade might be involved in the gene expression. Phylogenetic considerations suggest a relationship with the sol protein from D. melanogaster and with the pαB protein from Aspergillus. While the exact function of the sol protein is unknown, it has been shown that pαB belongs to a signal transduction cascade which is sensitive to ambient pH (10, 51). At alkaline pH, the transcription factor pacC is proteolytically processed to a functional form which activates all-trans retinoic acid receptor gamma and chromosomal localization of a human tubulointerstitial nephritis antigen (dek1) gene required for aleurone cell development in the endosperm of P. amygdalus. Tree Physiol. 2083–2092.

REFERENCES

1. Adl, S. M., and J. D. Berger. 1991. Timing of oral morphogenesis and its relation to commitment to division in Paramecium tetraurelia. Exp. Cell Res. 212:497–504.

2. Adl, S. M., and J. D. Berger. 1996. Commitment to division in ciliate cell cycles. J. Eukaryot. Microbiol. 43:77–86.

3. Adl, S. M., and J. D. Berger. 1997. Timing of life cycle morphogenesis in synchronous samples of Sterkiella histrio-muscum. I. The vegetative cell cycle. Eur. J. Protistol. 33:99–109.

4. Berti, P. J., and A. C. Storer. 1995. Alignment/phylogeny of the papain superfamily of cysteine proteases. J. Mol. Biol. 246:273–283.

5. Bewley, J. D. 1997. Seed germination and dormancy. Plant Cell 9:1055–1066.

6. Boudjelal, M., Z. Wang, J. J. Voorhees, and G. J. Fisher. 2000. Ubiquitin/proteasome pathway regulates levels of retinoic acid receptor gamma and retinoid X receptor alpha in human keratinocytes. Cancer Res. 60:2247–2252.

7. Campalans, A., M. Pages, and R. Messegueur. 2001. Identification of differentially expressed genes by the CDNA-AFLP technique during dehydration of almond (Prunus amygdalus). Tree Physiol. 21:633–643.

8. Castilleja, A., N. E. Ward, R. B. Epstein, A. P. Kudelka, D. M. Gershenson, C. L. Efferson, C. A. O’Brien, and C. G. Ioannides. 2002. Treatment with HER-2 phosphorylation agonists enhance tumor ability to stimulate epidermal growth factor receptor. Oncol. Rep. 9:49–49.

9. Chong, Y. B., Y. Kong, I. J. Joo, S. Y. Cho, and S. Y. Kang. 1995. Excystment of Paragonimus westermani metacercariae by endogenous cysteine protease. J. Parasitol. 81:137–142.

10. Denison, S. H., M. Orejas, and H. N. Arst, Jr. 1995. Signaling of ambient pH in Aspergillus involves a cysteine protease. J. Biol. Chem. 270:28519–28522.

11. Dicz, E., J. Alvaro, E. A. Espeso, L. Rainbow, T. Suarez, J. Tiliburn, H. N. Arst, Jr., and M. A. Penalva. 2002. Activation of the Aspergillus PacC zinc finger transcription factor requires two proteolytic steps. EMBO J. 21:1350–1359.

12. Dong, J. M., and G. G. Sahagian. 1990. Basis for low affinity binding of a lysosomal cysteine protease to the cation-independent mannose 6-phosphate receptor. J. Biol. Chem. 265:4200–4217.

13. Butt, P. J., S. S. Arthur, P. Grochulski, M. Cygler, and J. S. Elce. 2000. Roles of individual EF-hands in the activation of m-calpain by calcium. Biochem. J. 348:37–43.

14. Grimes, G. W. 2001. PHYLIP (phylogeny inference package), version 3.6 (alpha 2). University of Washington, Seattle.

15. Fleming, J. A., E. S. Lightcap, S. Sadis, V. Thordoldsen, C. E. Bulawa, and R. K. Blackman. 2002. Complementary whole-genome technologies reveal the cellular response to proteasome inhibition by PS-341. Proc. Natl. Acad. Sci. USA 99:1461–1466.

16. Fizzy, A., and G. Fryd-Versavel. 1990. Précisions sur la dynamique morphogénétique de Keranopsis muscica (cilié hypotrich). J. Protozool. 37: 1–10.

17. Frankel, J. 1962. The effects of heat, cold, and p-fluoro-phenyl-alanine on morphogenesis in synchronized Tetrahymena pyriformis GL. C. R. Trav. Lab. Carlsb. 331:52–56.

18. Frankel, J. 1989. Pattern formation: ciliate studies and models. Oxford University Press, New York, N.Y.

19. Gonzalez, J., G. Bai, U. Frevert, E. J. Corey, and D. Eichinger. 1999. Proteasome-dependent cyst formation and stage-specific ubiquitin mRNA accumulation in Entamoeba invadens. Eur. J. Biochem. 260:897–904.

20. Grasse, P. P. 1994. Traité de zoologie. Anatomie, systématique, biologie, tome II. Infusoriens ciliés. Masson, Paris, France.

21. Griffiths, C. M., S. E. Hosken, D. Oliver, J. Chojek, and H. Thomas. 1997. Sequencing, expression pattern and RFLP mapping of a mesencephalon-enhanced cDNA from Zea mays with high homology to oryzain gamma and aleurain. Plant Mol. Biol. 34:815–821.

22. Grimes, G. W. 1972. Cortical structure in nondiving and cortical morphogenesis in dividing Oxytricha fallax. J. Protozool. 19:428–445.

23. Grimes, G. W. 1973. Differentiation during encystment and excystment in Oxytricha fallax. J. Protozool. 2092–1014.

24. Grobe, K., M. Poppelmann, W. M. Becker, and A. Petersen. 2002. Properties of group I allergens from grass pollen and their relation to cathepsin B, a member of the C1 family of cysteine proteases. Eur. J. Biochem. 269:2883–2892.

25. Hashimoto, K. 1963. Formation of cilium in excystment and induced re-encystment of Oxytricha fallax Stein. J. Protozool. 10:156–166.

26. Hertz-Fowler, C., K. Ersfeld, and K. Gull. 2001. CAP 5.5, a life-cycle-regulated, cytoskeleton-associated protein is a member of a novel family of calpain-related proteins in Trypanosoma brucei. Mol. Biochem. Parasitol. 115:225–34.

27. Hosfield, C. M., T. Molodaveu, P. L. Davies, J. S. Elce, and Z. Jia. 2001. Calpain mutants with increased Ca2+ sensitivity and implications for the role of the C(2)-like domain. J. Biol. Chem. 276:7040–7047.

28. Ikneda, M., T. Takeamura, S. Hino, and K. Yoshoka. 2000. Molecular cloning, expression and chromosomal localization of a human tubulointerstitial nephritis antigen. Biochem. Biophys. Res. Commun. 268:225–230.

29. Izquierdo, A., A. Martin-Gonzalez, and J. C. Gutierrez. 2000. Resting cyst wall glycoproteins isolated from two copodid ciliates are glycan-rich proteins. Cell. Biol. Int. 24:115–119.

30. Kalfa, T. A., J. D. Thull, R. J. Butkowski, and A. S. Charonis. 1994. Tubulo-lointerstitial nephritis antigen interacts with laminin and type IV collagen and promotes cell adhesion. J. Biol. Chem. 269:1654–1659.

31. Koornneef, M., L. Bentink, and H. Hilhorst. 1993. Seed dormancy and germination. Curr. Opin. Plant Biol. 5:33–36.

32. Lee, D. H., and A. L. Goldberg. 1996. Selective inhibitors of the proteasome-dependent and vacuolar pathways of protein degradation in Saccharomyces cerevisiae. J. Biol. Chem. 271:27280–27284.

33. Lee, D. H., and A. L. Goldberg. 1998. Proteasome inhibitors: valuable new tools for cell biologists. Trends Cell. Biol. 8:397–403.

34. Letunic, I., L. Goodstadt, N. J. Dickens, T. Doerks, J. Schultz, R. Mott, F. Ciccarelli, R. R. Frankel, J. D. Berger. 2002. Recent improvements to the SMART domain-based sequence annotation resource. Nucleic Acids Res. 30:242–244.

35. Lid, S. E., D. Gruis, R. Jung, J. A. Lorentzen, E. Ananiev, M. Chamberlin, N. N. M., and R. Meeley, S. Nichols, and O. A. Olsen. 2002. The defective killer 1 (dek1) gene required for aleurone cell development in the endosperm of
maize grains encodes a membrane protein of the calpain gene superfamily. Proc. Natl. Acad. Sci. USA 99:5460–5465.
36. Markmann-Mulisch, U., B. Reiss, and M. Mulisch. 1999. Cell type-specific gene expression in the cell cycle of the dimorphic ciliate Eufolliculina ulhigi. Mol. Gen. Genet. 262:390–399.
37. McGrath, M. E. 1999. The lysosomal cysteine proteases. Annu. Rev. Biophys. Biomol. Struct. 28:181–204.
38. Nothwehr, S. F., and J. I. Gordon. 2000. Substrate specificity of bovine calpain B and its inhibition by CA074, a novel cysteine protease inhibitor. J. Biol. Chem. 275:6131–6137.
39. Osherov, N., and S. M. Gregory. 1991. Novel epoxysuccinyl peptides. A selective inhibitor of cathepsin B, in vivo. FEBS Lett. 280:311–315.
40. Prescott, D. M. 1994. The DNA of ciliated protozoa. Microbiol. Rev. 58:233–267.
41. Rios, R. M., J. Martin, A. Torres, and C. Fedriani. 1992. Production and characterization of three polyclonal antibodies raised against cyst wall proteins of a hypotrichous ciliate. J. Protozool. 39:584–588.
42. Rogers, J. C., D. Dean, and G. R. Heck. 1985. Alcurain: a barley thiol protease closely related to mammalian cathepsin H. Proc. Natl. Acad. Sci. USA 82:6512–6516.
43. Saitou, N., and M. Nei. 1987. The neighbour-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4:406–425.
44. Sajid, M., and J. H. McKerrow. 2002. Cysteine proteases of parasitic organisms. Mol. Biochem. Parasitol. 120:1–21.
45. Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
46. Santella, L., K. Kyozuka, L. De Riso, and E. Carafoli. 1998. Calcium, protease action, and the regulation of the cell cycle. Cell Calcium 23:123–130.
47. Sokol, S. B., and P. E. Kewabara. 2000. Proteolysis in Caenorhabditis elegans sex determination: cleavage of TRA-2A by TRA-3. Genes Dev. 14:901–906.
48. Sorimachi, H., S. Ishiura, and K. Suzuki. 1997. Structure and physiological function of calpains. Biochem. J. 328:721–732.
49. Sorimachi, H., and K. Suzuki. 2001. The structure of calpain. J. Biochem. 129:653–664.
50. Towatari, T., T. Nikawa, M. Murata, C. Yokoo, M. Tamai, K. Hanada, and N. Katunuma. 1991. Novel epoxysuccinyl peptides. A selective inhibitor of cathepsin B, in vivo. FEBS Lett. 280:311–315.
51. Tuffrau, M. 1967. Perfectionnement et pratique de la technique d’imprégnation au protargol des Infusoires ciliés. Protistologica 3:91–98.
52. Turk, B., D. Turk, and V. Turk. 2000. Lysosomal cysteine proteases: more than scavengers. Biochim. Biophys. Acta 1477:98–111.
53. Turk, D., M. Podobnik, T. Popovic, N. Katunuma, W. Bode, R. Huber, and V. Turk. 1995. Crystal structure of cathepsin B inhibited with CA030 at 2.0-A resolution: a basis for the design of specific epoxysuccinyl inhibitors. Biochem. 34:4791–4797.
54. Villalobo, E., C. Moch, R. Perasso, and A. Baroin-Tourancheau. 2001. Searching for excystment-regulated genes in Sterkella histriomuscorum (Ciliophora, Oxytrichidae): a mRNA differential display analysis of gene expression in excysting cells. J. Eukaryot. Microbiol. 48:382–390.
55. Von Heijne, G. 1990. The signal peptide. J. Membr. Biol. 115:195–201.
56. Waghray, A., D. Keppler, B. F. Sloane, L. Schuger, and Y. Q. Chen. 2000. Invariant chain controls H2-M proteolysis in mouse splenocytes and dendritic cells. J. Exp. Med. 191:1057–1062.
57. Wagh, V., L. Alvarado, N. D. Rawlings, J. C. Engel, C. Franklin, and J. H. McKerrow. 1997. A primitive enzyme for a primitive cell: the protease required for excystation of Giardia. Cell 89:437–444.
58. Yamamoto, A., K. Tomoo, T. Harata, M. Murata, K. Kitamura, and T. Ishida. 2000. Substrate specificity of bovine cathepsin B and its inhibition by CA074, based on crystal structure refinement of the complex. J. Biochem. 127:635–643.