JCI The Journal of Clinical Investigation

Cloning of a virulence factor of Entamoeba histolytica. Pathogenic strains possess a unique cysteine proteinase gene.

S Reed, … , P Hagblom, F Gillin

Research Article

Cysteine proteinases are hypothesized to be important virulence factors of Entamoeba histolytica, the causative agent of amebic dysentery and liver abscesses. The release of a histolytic cysteine proteinase from E. histolytica correlates with the pathogenicity of both axenic strains and recent clinical isolates as determined by clinical history of invasive disease, zymodeme analysis, and cytopathic effect. We now show that pathogenic isolates have a unique cysteine proteinase gene (ACP1). Two other cysteine proteinase genes (ACP2, ACP3) are 85% identical to each other and are present in both pathogenic and nonpathogenic isolates. ACP1 is only 35 and 45% identical in sequence to the two genes found in all isolates and is present on a distinct chromosome-size DNA fragment. Presence of the ACP1 gene correlates with increased proteinase expression and activity in pathogenic isolates as well as cytopathic effect on a fibroblast monolayer, an in vitro assay of virulence. Analysis of the predicted amino acid sequence of the ACP1 proteinase gene reveals homology with cysteine proteinases released by activated macrophages and invasive cancer cells, suggesting an evolutionarily conserved mechanism of tissue invasion. The observation that a histolytic cysteine proteinase gene is present only in pathogenic isolates of E. histolytica suggests that this aspect of virulence in amebiasis is genetically predetermined.

Find the latest version:

https://jci.me/116359/pdf

Cloning of a Virulence Factor of *Entamoeba histolytica*

Pathogenic Strains Possess a Unique Cysteine Proteinase Gene

Sharon Reed,* Jacques Bouvier,‡ Anna Sikes Pollack,‡ Juan C. Engel,‡ Margaret Brown,‡ Ken Hirata,*
Xuchu Que,* Ann Eakin,‡§ Per Hagblom,‖ Frances Gillin,* and James H. McKerrow*¶

*Departments of Pathology and Medicine, University of California, San Diego, California 92103-8416; Departments of ‡Pathology and
§Pharmaceutical Chemistry, University of California, San Francisco, California 94143; the ‖Department of Microbiology, University of
Uppsala, Uppsala, Sweden; and the ¶Department of Veterans Affairs Medical Center, San Francisco, California 94121*

Abstract

Cysteine proteinases are hypothesized to be important viru-
lence factors of *Entamoeba histolytica*, the causative agent of
amebic dysentery and liver abscesses. The release of a histoly-
tic cysteine proteinase from *E. histolytica* correlates with the
pathogenicity of both axenic strains and recent clinical isolates
as determined by clinical history of invasive disease, zymodeme
analysis, and cytopathic effect. We now show that pathogenic
isolates have a unique cysteine proteinase gene (ACP1). Two
other cysteine proteinase genes (ACP2, ACP3) are 85% identi-
cal to each other and are present in both pathogenic and non-
pathogenic isolates. ACP1 is only 35 and 45% identical in se-
quence to the two genes found in all isolates and is present on a
distinct chromosome-size DNA fragment. Presence of the
ACP1 gene correlates with increased proteinase expression and
activity in pathogenic isolates as well as cytopathic effect on a
fibroblast monolayer, an in vitro assay of virulence. Analysis of
the predicted amino acid sequence of the ACP1 proteinase gene
reveals homology with cysteine proteinases released by acti-
vated macrophages and invasive cancer cells, suggesting an
evolutionarily conserved mechanism of tissue invasion. The ob-
servation that a histolytic cysteine proteinase gene is present
only in pathogenic isolates of *E. histolytica* suggests that this
aspect of virulence in amebiasis is genetically predetermined.
(*J. Clin. Invest.* 1993. 91:1532–1540.) Key words: amebiasis •
proteinases • cathepsins • pathogenicity • cytopathic

Introduction

Entamoeba histolytica infects more than 500 million people
worldwide (1). Almost from the time of its discovery, it was
observed that although *E. histolytica* most often causes mild or
asymptomatic infections, ~ 10% of patients develop severe
dysentery and life-threatening invasive and extraintestinal dis-
ease (2). Whether these distinct clinical courses represent in-
fection by two different species of *Entamoeba* (2) or, con-
versely, whether any strain of *E. histolytica* can potentially
cause disease in the presence of certain environmental factors
has been an area of active debate (3, 4). Distinct isoenzyme
patterns (zymodemes) (5) and differences in restriction frag-
ment–length patterns between pathogenic and nonpathogenic

strains (6–10) strongly support the premise that the potential
to cause invasive disease is genetically determined. If virulence
is genetically determined, we would expect a gene encoding a
virulence factor that correlated with pathogenicity to be pres-
ent only in pathogenic strains. We tested this hypothesis using
the ameba cysteine proteinase.

Factors associated with the virulence of *E. histolytica* in-
clude surface lectins (11, 12), cytolytic ion channel–forming
proteins (13–15), phospholipases (11), and proteinases (10,
16–18). The major proteinase of *E. histolytica* is a cysteine
proteinase biochemically similar to cathepsin B (17–20). Ex-
perimental evidence supporting the role of this cysteine pro-
teinase in the pathogenesis of amebiasis includes its ability to
degrade fibronectin, collagen, and basement membrane matrix
and to activate the third component of complement (17–19,
21). Purified proteinase reproduces the cytopathic effect of
pathogenic amebae (17–20), and a specific irreversible inhibi-
tor of the proteinase prevents destruction of cell monolayers by
live axenic amebae (22). Increased expression and excretion of
the proteinase correlates with virulence of both axenic labora-
tory strains of *E. histolytica*, and fresh clinical isolates (16, 19,
20). Antibodies to the proteinase were detected in 83% of pa-
tients with invasive disease but were not detected in patients
with noninvasive infections (16).

Eukaryotic cysteine proteinases have long been recognized
as particularly potent proteinases but were best known for their
role in intracellular protein digestion (23). More recently, evi-
dence has accumulated that a subset of invasive cancers excrete
active or activatable cysteine proteinases. A cathepsin L–like
cysteine proteinase is the major excreted protein of trans-
formed fibroblasts (24–26). Levels of cathepsin B were found
to correlate with the metastatic potential of melanoma vari-
ants (24).

To determine whether the cysteine proteinase released by
pathogenic *E. histolytica* is structurally related to those re-
leased by cancer cells and activated macrophages, we used poly-
merase chain reaction to identify and amplify genes encoding
the cysteine proteinases of *E. histolytica* using primers based
on conserved structural motifs of eukaryotic cysteine protein-
ases (27). We identified and sequenced three cysteine protein-
ase genes and now show that one gene is unique to pathogenic
isolates of *E. histolytica*.

Methods

Strains of E. histolytica. Axenic strain HM-1:IMSS was obtained from
the American Type Culture Collection (Rockville, MD) and cultured
in TYI-S-33 media (28). Nine pathogenic and 12 nonpathogenic clini-
cal isolates were cultured directly from stools or liver abscesses into
Robinson's media as previously described (29). Three strains (SAW
760, SAW 1453, and SAW 1734) were a gift from Peter Sargeaunt
(London School of Hygiene and Tropical Medicine), and five strains

---

(FAT 957, FAT 967, FAT 973, FAT 1010, and FAT 1014) were a gift from T.H.F.G. Jackson (Research Institute for Diseases in a Tropical Environment, Durban, South Africa). One strain was cultured from a patient in a refugee camp in Costa Rica (REF 291). The other 12 strains (pathogenic: SD 4, 53, 92, 135, and 136; nonpathogenic: SD 11, 107, 116, 130, 137, 143, and 147) were isolated by the Microbiology Laboratory at the University of California Medical Center, San Diego. Strains were assigned to zymodemes by the method of Sargeaunt et al. (5).

DNA and RNA purification. DNA was isolated by a modification of the method of Huber et al. (30). Cell pellets from clinical isolates were first incubated with proteinase K (1 mg/ml) for 2 h at 56°C. Total RNA was isolated from 10^7 to 10^8 trophozoites according to the method of Chomczynski and Sacchi (31).

PCR amplification and sequencing. Degenerate oligonucleotide primers based on active site sequences conserved in all eukaryotic cysteine proteinases were used to amplify cysteine proteinases from genomic HM-1 strain DNA by the PCR under conditions previously described (27). A new 5' primer (GCC GAA TTC GCT GCT CCA GAA TCA GTT GAT TGG AGA) based upon the amino-terminal sequence of ACP1 was used with the original 3' primer to amplify and isolate the two additional proteinase gene fragments (ACP2 and ACP3). The 5' sequence of the mature ACP2 proteinase was then confirmed using a new 3' primer (AAA GGA TCC ACA TGA TCC GCA TTG TC/GC TT).

The remaining sequence of ACP2 was obtained using an *E. histolytica* cDNA library in λgt11 (strain H-302:NIH; a gift from Dr. Bruce Torian) (32). DNA was purified from bacteriophage plaques by bind-

ing to DEAE-cellulose (DE52; Whatman Inc., Clifton, NJ). The 5' sequence was obtained by PCR using a λgt11 reverse sequencing primer (Promega Corp., Madison, WI) and a second primer corresponding to a seven–amino acid sequence at residues 18–25 of the mature ACP2 enzyme (5'-AGA GTC GAC TGT ATA ACA TGA TCC ACA TTG TCC TTG ATC). After PCR amplification, a 321-bp fragment that encoded the entire pro-sequence was obtained. The carboxy-terminal sequence was similarly obtained by PCR using the forward sequencing primer for λgt11 (antisense) and a sense primer made to the same seven–amino acid sequence. Sequencing was performed by the dideoxy method with the Sequenase Kit (U.S. Biochemical Corp., Cleveland, OH) or by automated dye terminator sequencing at the Biomolecular Resource Center, University of California at San Francisco, using an ABI373A instrument.

Genomic library. An *E. histolytica* genomic library of EcoRI*-digested DNA fragments was prepared in Lambda ZAP (Stratagene, Inc., San Diego, CA). The ACP1 probe (452 bp) was labeled by the random primer method (Bethesda Research Laboratories, Gaithersburg, MD) and used to screen 2 × 10^6 plaques. Five clones were selected after tertiary screening and sequenced.

Southern and Northern analysis. For Southern blot analysis, 5–10 μg of DNA was digested with EcoRI and subjected to electrophoresis on 0.8% GTG agarose gels (FMC Corp., Rockland, ME) and transferred to Gene Screen Plus nylon membranes (New England Nuclear Research Products, Boston, MA). Hybridizations were performed under high stringency by the method of Church and Gilbert (33) in 7% SDS, 0.5 M $NaHPO_4$, pH 7.2, 1 mM EDTA, and 1% bovine serum albumin at 65°C. Ethidium bromide staining of the gel and hybridiza-
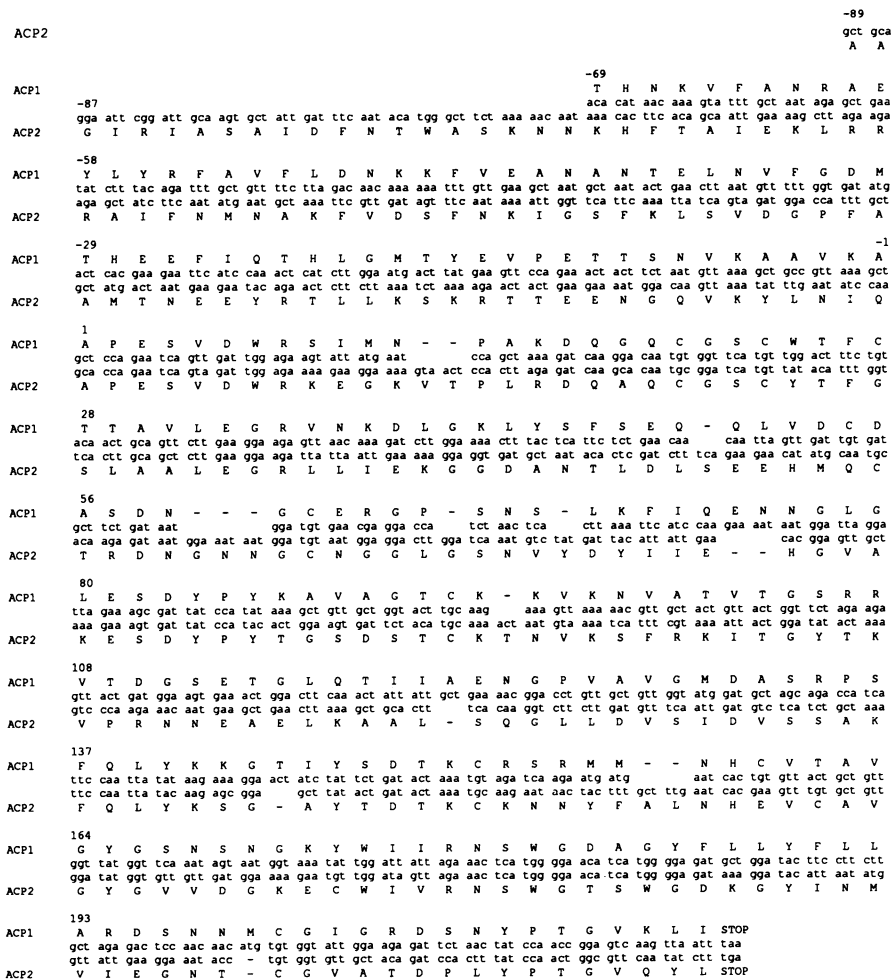
Figure 1. Nucleic acid and predicted protein sequences of cysteine proteinase genes from *E. histolytica* HM-1 strain. ACP1 and ACP2 are shown. The ACP3 gene fragment was identical to a cDNA sequence (cEh-CPp) previously published (10); the amino acid sequence of ACP3 is shown in Fig. 6. Sequence data for ACP1 is from a genomic clone and includes 207 bp of proenzyme sequence 5' to the amino terminus of the mature proteinase. Residue +1 is the amino terminus of mature proteinase.
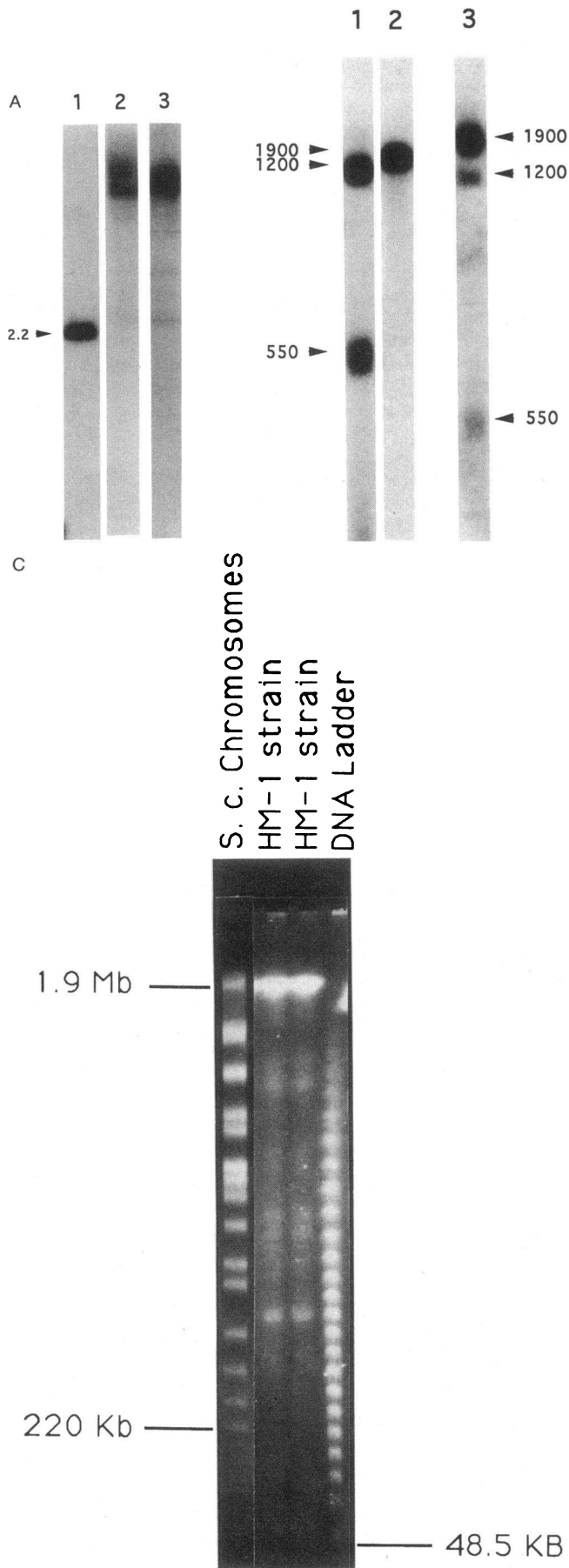
tion of the blot with ^{32}P-labeled *E. histolytica* actin cDNA (34), a gift of Dr. Isaura Meza, confirmed the presence of equivalent amounts of DNA from pathogenic and nonpathogenic strains. For Northern analysis, 5–10 µg of total RNA was subjected to electrophoresis on a 1% agarose–2.2 M formaldehyde gel in 3-(*N*-morpholino)propanesulfonic acid buffer and transferred to Gene Screen Plus nylon membranes. Hybridizations were performed under identical conditions.

Ribonuclease protection assay. Antisense RNA transcripts were synthesized from Bluescript (pKS) vectors containing the ACP1 gene cut with SmaI and ACP2 cut with EcoRI to terminate transcription of the insert. α-[^{32}P]CTP–labeled transcript was synthesized with the T3 bacteriophage polymerase according to the Maxiscript kit instructions (Ambion, Inc., Austin, TX) and the DNA template was digested with RNase-free DNase. Aliquots of 5–10 µg of total RNA were added to a fourfold molar excess of labeled probe and hybridized at 42°C overnight using the RPAII kit (Ambion, Inc.). Unhybridized RNA was digested with RNase A, RNA hybrids were precipitated and pelleted, and the fragments were separated on an 8 M urea, 5% acrylamide gel. The intensity of the resulting bands on autoradiographs was compared by scanning with a Quikscan (Helena Laboratories, Beaumont, TX).

Field inversion gel electrophoresis (FIGE). Trophozoites were harvested from early stationary phase cultures and washed twice by centrifugation (10 min at 200 *g*) in TSE buffer (100 mM NaCl, 50 mM EDTA, 20 mM Tris base, pH 8.0). An equal volume of melted 1.2% InCert Agarose (FMC Corp.) in TSE was added to the parasite pellet. The mixture was poured into 63-µl wells of a Hexa-A-Field agarose plug mold (Bethesda Research Laboratories) and allowed to gel at 4°C. The agarose blocks were transferred to a tube containing 10 ml of TSE plus 1% *N*-lauroylsarcosine, 1% Nonidet P-40, and incubated for 3 h with gentle agitation at 4°C. The buffer was then replaced with fresh TSE, 1% *N*-lauroylsarcosine (TSE-Sarkosyl) and stored at 4°C for a minimum of 18 h. 5 h before use the agarose plugs were transferred to a new tube containing TSE-Sarkosyl with 2 mg/ml Proteinase K and incubated at 45°C.

FIGE was performed in a horizontal electrophoresis unit (model HE 100 SuperSub Hoefer, Scientific Instruments, San Francisco, CA) connected to a PC 750 pulse controller (model PC 750; Hoefer Scientific Instruments). The agarose plugs were loaded into wells of a 1.2% agarose gel (Seakem agarose; FMC Corp.) made with 1.0× TBE buffer (89 mM Tris base, 89 mM boric acid, 2 mM EDTA, pH 8.0). DNA molecules were separated using two cycles at a constant voltage of 125 V at 8°C in 1× TBE buffer. The first cycle was run with a beginning forward pulse of 2.4 s, a reverse pulse of 0.8 with a 1.6 ramp value for 22–24 h. The second cycle lasted 20–24 h with an initial forward pulse of 3.6 s and 1.2 reverse pulse with a 1.6 ramp value. Lambda DNA ladder and *Saccharomyces cerevisiae* chromosomal DNA (FMC Corp.) were included to estimate the molecular size of the *E. histolytica* chromosomes. The DNA was depurinated with 0.2 N HCl for 10 min and denatured with 0.5 N NaOH for 45 min before it was transferred to nylon membranes. Hybridizations were performed in 30% formamide, 6× SSPE (sodium chloride, sodium phosphate, EDTA), 5× Denhardt's, 0.2% SDS, and 100 µg/ml transfer RNA (tRNA) at 45°C overnight. The filters were washed as described in Sambrook et al. (35) at a final temperature of 68°C for 30 min.

Proteinase purification and peptide sequencing. HM-1 trophozoites were washed and suspended at a concentration of 10^7/ml in PBS and

Figure 2. (*A*) Southern blot of ACP1, ACP2, and ACP3 with HM-1 strain DNA. Southern hybridization of EcoRI-digested DNA from *E. histolytica* axenic strain HM-1 with a probe corresponding to the ACP1 gene (lane *1*), ACP2 gene (lane *2*), and ACP3 gene (lane *3*). (*B*) Southern blot after FIGE of DNA from HM-1 strain hybridized with ACP1 and ACP2 genes. Lane *1*, ACP1; lane *2*, ACP2 hybridized to same blot; lane *3*, both probes simultaneously hybridized to second blot. ACP3 gave identical pattern to ACP2 (not shown). Note unique location of ACP1 gene copy on < 550-kb "chromosome." (*C*) Ethidium bromide–stained FIGE gel used for Southern blot transfer in *B*.
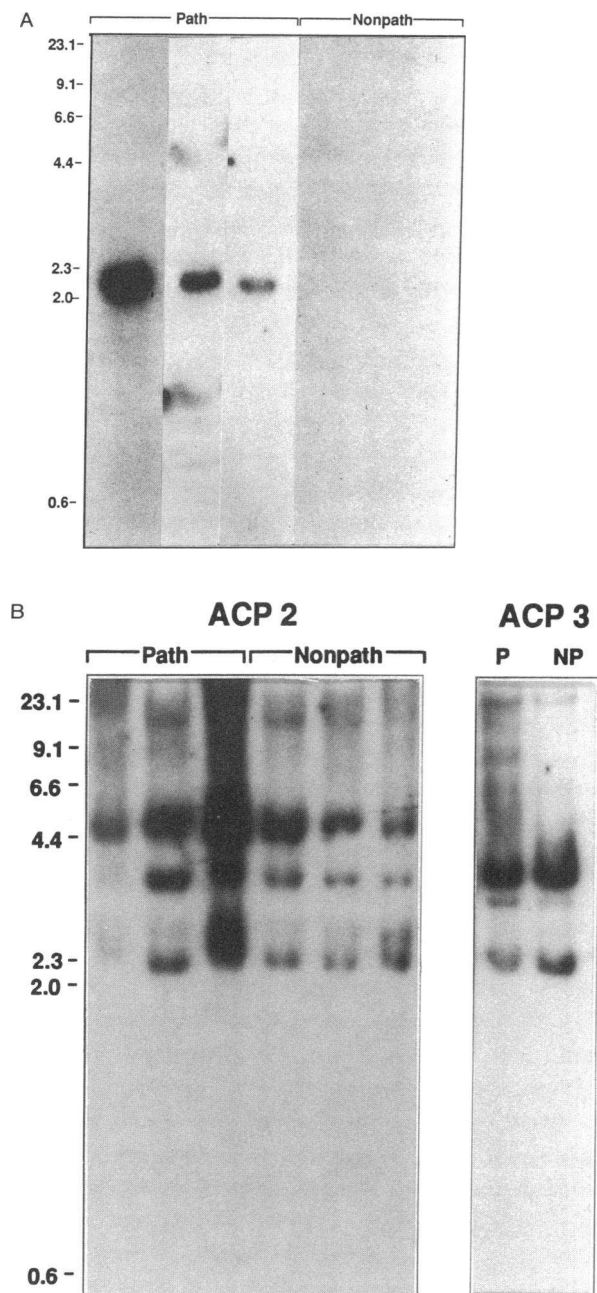
Figure 3. (*A*) Southern blot of ACP1 cysteine proteinase gene with DNA of clinical isolates. Southern hybridization of EcoRI-digested DNA from pathogenic (*Path*) and nonpathogenic (*Nonpath*) clinical isolates with ACP1 at 65°C. Examples of three of nine pathogenic and three of nine nonpathogenic isolates are shown. (*B*) Southern blot of ACP2 and ACP3 cysteine proteinase gene with DNA of clinical isolates. Southern hybridization of EcoRI-digested DNA with ACP2 (three of six pathogenic and three of five nonpathogenic strains shown) and ACP3 (one of three pathogenic and one of three nonpathogenic strains shown) at 55°C.

incubated for 3 h at 37°C. Cysteine proteinase activity was identified and the enzyme purified from the supernatant by fast protein liquid chromatography as previously described (17). Both the high molecular mass (56 kD) and low molecular mass (27 kD) forms of the proteinase were purified. The purified proteinase was subjected to electrophoresis by SDS-PAGE on a 10% gel and transferred to a polyvinylidenedifluor-

ide membrane (Immobilon; Millipore Corp., Bedford, MA) by the method of Matsudaira (36). Peptide sequencing of the amino terminus was performed at the Biomolecular Resource Center at University of California at San Francisco, using a gas phase sequencer (Applied Biosystems Inc., Foster City, CA).

Monolayer assay for cytopathic effect. Trophozoites were purified from xenic media as previously described (16) and resuspended in PBS containing 20 mM cysteine, 0.15 mM $CaCl_2$, and 0.5 mM $MgCl_2$ at a concentration of 10^7/ml. After incubation for 3 h at 37°C, the supernatants were separated and passed through a 0.2-μM filter to remove remaining bacteria from the medium. This procedure did not affect the proteinase activity but was necessary to remove any residual bacterial flora, which is different for each clinical isolate. Proteinase activity was quantified by the cleavage of a synthetic peptide substrate, Z-Arg-Arg-AMC (benzyloxycarbonyl-arginine-arginine-4-amino-7-methylcoumarin) (Enzyme Systems Products, Livermore, CA), as previously described (17), and recorded as the initial velocity of cleavage of the fluorescent 4-amino-7-methylcoumarin/10 μl. Proteinase activity was inhibited by preincubating with Z-Phe-Arg-CH_2F, an irreversible cysteine proteinase inhibitor (10 μM for 30 min at 37°C), which has no effect on the viability of the monolayer (Enzyme Systems Products). 24-well culture plates were seeded with 2×10^5 cells from a foreskin epithelial cell line (HFS 1) and incubated overnight in DME with 10% fetal calf serum in CO_2. The medium supplemented with serum was removed and the monolayer washed twice with MEM without sera. Amebic supernatants (100 μl) were added in triplicate to 400 μl of MEM and incubated for 3 h at 37°C with CO_2. Detached cells were removed by washing twice with PBS. The remaining cells were fixed (4% Formalin) and quantified by staining with methylene blue dye (0.1% in 0.1 M borate, pH 8.7) and measuring the extracted absorbance at 660 nM (22). Cytopathic effect (CPE)[1] was calculated as the absorbance of control wells (MEM alone) minus the absorbance of sample wells divided by control × 100%. To control for variations between different experiments, the values were standardized against the CPE caused by strain HM-1, which was taken as 1.00 (relative CPE).

Results

To isolate the cysteine proteinase genes of *E. histolytica*, we amplified cysteine proteinase gene fragments by PCR using primers based on conserved structural motifs identified in the eukaryotic cysteine proteinase family (27). A 450-bp fragment was initially amplified from DNA isolated from the pathogenic axenic strain HM-1, as previously reported (27). The 450-bp fragment was subsequently used to isolate a genomic clone (ACP1) containing the entire coding region of the mature proteinase (Fig. 1). A second set of primers based on the amino terminus of this gene and the region around the active site asparagine was then used to amplify two 550-bp genomic fragments of the second and third cysteine proteinase genes (ACP2 and ACP3) (Fig. 1). The 5' end of ACP2, representing the amino terminus of the mature proteinase, was then confirmed by sequence from genomic DNA using a 3' to 5' primer based on adjacent downstream sequence (see Methods). The 5' pro sequence and the carboxy-terminal sequence of ACP2 were obtained by PCR from DNA purified from a cDNA library (see Methods). The sequence of the ACP3 fragment was found to be identical to the sequence of a cDNA, cEh-CPp, previously published by Tannich et al. (10).

To confirm that these genes encoded *E. histolytica* protein-

---

1. *Abbreviations used in this paper:* CPE, cytopathic effect; FIGE, field inversion gel electrophoresis.

ases and were expressed in the HM-1 strain, we purified proteinase from the culture supernatant of this strain. Microsequencing of the first eight amino acids of the amino terminus of the purified 27-kD enzyme revealed heterogeneity at positions 3 and 4 in an approximately equimolar ratio (2.1:1.8), A-P-E/K-S/A-V-D-W-R. One sequence (APES . .) was identical to that predicted from ACP1 and ACP2 and also identical in seven out of eight residues with the amino-terminal sequence of a cysteine proteinase (called "histolysin") purified by Luaces and Barrett (18) from HM-1 trophozoites. The second sequence (APKA . . .) was identical to the first eight residues of the cEh-CPp proteinase predicted from a cDNA clone isolated from HM-1 amebae by Tannich et al. (10). The nucleotide sequence of cEh-CPp is identical to ACP3.

Southern blot analysis showed that the ACP1 gene fragment hybridized with a single restriction fragment in EcoRI (2.2 kb)- (Fig. 2 A) or BglII (3 kb)- (not shown) digested HM-1 genomic DNA whereas ACP2 and ACP3 each hybridized to two larger (10 and 12 kb, respectively) EcoRI fragments (Fig. 2 A). Southern blot analysis after FIGE of HM-1 strain (Fig. 2 B) showed that the ACP1 gene fragment hybridized with two DNA bands of a relative size of ~ 1 Mb and 550 kb. In contrast, the ACP2 and ACP3 gene fragments hybridized with a single DNA band of relative size 1.9 Mb. FIGE-derived Southern blots hybridized with mixtures of ACP1 and ACP2 or ACP3 gene fragments (Fig. 2 B, lane 3) showed three hybridization sites: 1.9 Mb corresponding to ACP2/ACP3 genes, and 1 Mb and 550 kb corresponding to the ACP1 gene as described above. These data suggested there was a distinct chromosomal location for the ACP1 gene versus ACP2 and ACP3. The detection of two sites of hybridization of the ACP1 probe on FIGE suggested more than one gene copy was present. The presence of multiple gene copies was also supported by repeating Southern blot analysis under conditions of varying EcoRI digestion (0.1, 0.5, 2, 4, and 8 U). Six hybridization bands were detected ranging in six regular increments from 2.2 to 10 kb, suggesting multiple gene copies are present, at least six of which may be in tandem. This pattern was reminiscent of the multiple tandem copies of the cysteine protease gene of *Trypanosoma cruzi* (37).

To determine whether the presence of any of these genes was correlated with amebic virulence, DNA was extracted from nine pathogenic and nine nonpathogenic clinical isolates classified by their zymodeme patterns and the clinical syndrome of the patients (5). In Southern blot analysis of EcoRI-digested DNA from all nine pathogenic isolates, ACP1 hybridized to a 2.1–2.3-kb fragment, similar to that seen in the digest of HM-1 DNA. It did not hybridize to any DNA fragment from any of the nine nonpathogenic isolates (Fig. 3 A). A portion of the cysteine proteinase gene from one pathogenic clinical isolate (SAW 1453) was amplified by PCR. The 450-bp amplified sequence from this isolate was identical to that of the HM-1 axenic strain (ACP1) shown in Fig. 1. However, there was slight but detectable variation in size (2.1–2.3 kb) or intensity of the EcoRI fragment in DNA from other pathogenic isolates, suggesting the presence of some sequence heterogeneity at this locus among different isolates.

In contrast, ACP2 hybridized to 2.2-, 4.0-, and 5.5-kb fragments and ACP3 hybridized to 2.2-, 2.9-, and 3.6-kb fragments in EcoRI-digested DNA from six pathogenic and five nonpathogenic clinical isolates (Fig. 3 B). Because complete genomic clones of ACP2 and ACP3 were not available, it is difficult to
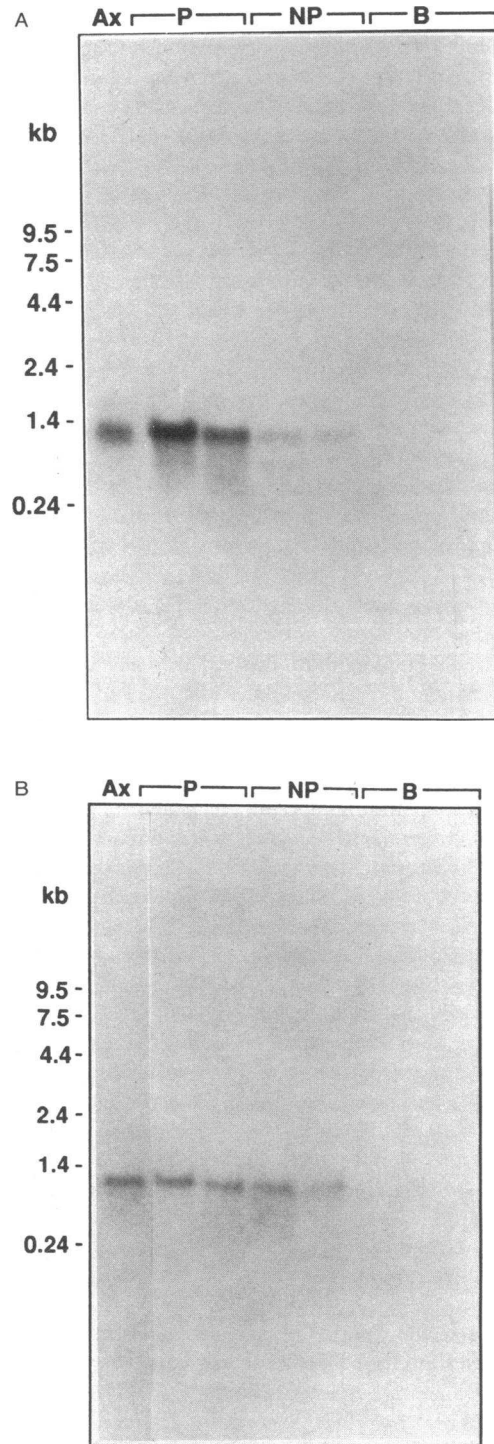
Figure 4. (*A*) Northern blot with ACP1. Northern blot of pathogenic (*P*), nonpathogenic (*NP*), and bacteria (*B*) RNA hybridized with ACP1 gene probe. Examples of three of five pathogenic and two of five nonpathogenic isolates are shown. (*B*) Northern blot with ACP2. Same blot as *A* with ACP2 gene as probe. (*C*) RNase protection assay. Total RNA from one axenic (HM-1; *Ax*), three pathogenic (*Path*), and four nonpathogenic (*Nonpath*) strains was hybridized with antisense RNA transcripts from ACP1 and ACP2.

assign specific map positions to the fragments in Fig. 3 *B*. It is possible some may be the result of partial digestion by the restriction enzyme. None of the ameba gene probes hybridized to

C

Ax ┌──────Path──────┐ ┌──────Nonpath──────┐

ACP 1 — 1.2 kb

ACP 2 — 1.2 kb

Figure 4. (Continued)

DNA from the bacteria in the cultures of clinical isolates. Northern blots probed with ACP1 revealed a strong signal at 1.2 kb with RNA from HM-1 and five of five pathogenic strains, but only a faint band in five of five nonpathogenic 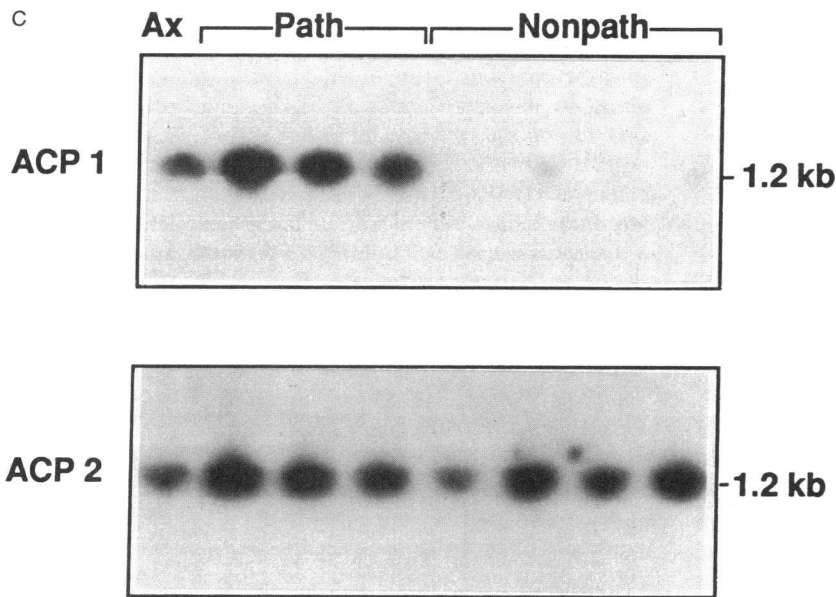strains (Fig. 4 *A*). When the same blot was hybridized with the ACP2 probe, a band of equivalent intensity and identical molecular weight was observed with all strains. The ACP3 probe also hybridized to a band of identical size in all three pathogenic and four nonpathogenic RNA samples tested (not shown). These Northern blots suggested that ACP2 and ACP3 were expressed in all isolates whereas ACP1 was expressed only in pathogenic isolates but crosshybridized weakly with the other genes (faint bands in nonpathogenic lanes in Fig. 4 *A*). To confirm this result with a more sensitive measure of potentially low abundance mRNAs, ribonuclease protection assays were performed using antisense RNA transcripts synthesized from ACP1 and ACP2. A 1.2-kb band was detected in all isolates with the ACP2 probe whereas the ACP1 probe hybridized strongly only with transcripts from the pathogenic isolates (Fig. 4 *C*).

To directly correlate the presence of the ACP1 gene in pathogenic amebae with enhanced proteinase activity and a quantitative assay of virulence, the total cysteine proteinase activity of amebic supernatants was determined and correlated in the same ameba isolates with the cytopathic effect on a fibroblast cell monolayer. The proteinase activity (initial velocity/ 10 μl) of five pathogenic strains containing ACP1 was significantly greater than that of four nonpathogenic strains lacking it ($P < 0.03$) (Fig. 5). The relative CPE (corrected to that of HM-1 as 1.00) by pathogenic strains was also significantly greater ($P < 0.001$) and was completely inhibited, as demonstrated previously (22), by preincubation with 100 μM Z-Phe-Arg-CH$_2$F, an irreversible and specific cysteine proteinase inhibitor. Furthermore, the enhanced proteinase activity correlated with elevated cysteine proteinase mRNA in the pathogenic versus nonpathogenic isolates (Fig. 5). In summary, Fig. 5 shows that the presence of an additional cysteine proteinase gene (ACP1) in pathogenic isolates correlates with increased proteinase mRNA, increased proteinase activity, and

increased CPE that is inhibited by Z-Phe-Arg-CH$_2$F, a cysteine proteinase inhibitor.

The coding sequence of ACP1 predicted an amino acid sequence with only 35–45% identity to the other two genes. In contrast, ACP2 and ACP3 were 85% identical to each other in
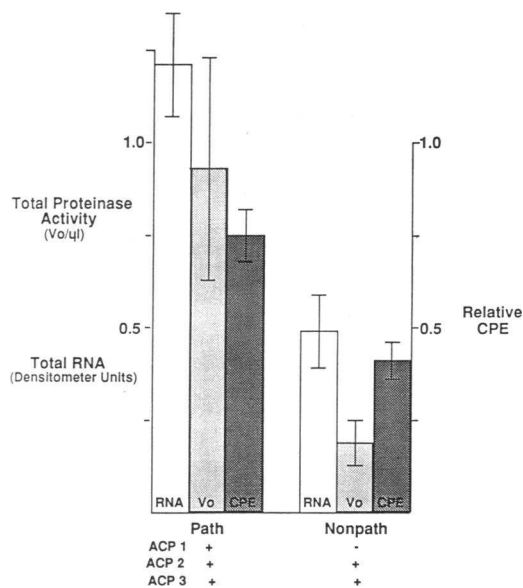
Figure 5. Proteinase activity (measured as initial velocity, Vo/10 ml) of supernatants of clinical isolates (10^7 amebae/ml) compared with the relative CPE (measured as percent of monolayer destruction/ percent destruction by axenic strain HM-1) and total cysteine proteinase mRNA. Proteinase activity is the mean±SEM for 13 measurements of five pathogenic strains and 18 measurements of four nonpathogenic strains. CPE is mean±SEM for five pathogenic and four nonpathogenic isolates. Sum of mRNA for ACP1 and ACP3 assayed by densitometer scan of signals from RNase protection assays of three pathogenic and four nonpathogenic isolates. + or − indicates presence or absence of each gene by Southern blot analysis (see Fig. 2).

predicted amino acid sequence and 90 and 100% identical at the nucleic acid level to the two corresponding genes (cEh-CPnp and cEh-CPp) isolated previously by Tannich et al. (10). All three genes were also homologous to the major excreted protein of transformed fibroblasts (26) as well as other members of the cysteine proteinase family of enzymes (Fig. 6). Specifically, all three amino acid residues of the catalytic triad of cysteine proteinases (cysteine, histidine, and asparagine) were present, and structural motifs flanking these residues were highly conserved (Fig. 6).

Discussion

The potential multifactorial roles of cysteine proteinases in invasion of pathogenic amebae are well documented (16–22). They include degradation of host extracellular matrix and mucoproteins, dislodgment of epithelial cells and degradation of epithelial basement membrane, and possibly recruitment of inflammatory cells to sites of ameba invasion by activation of complement.

We have now shown that the enhanced expression and release of cysteine proteinase activity by virulent laboratory strains (17, 20) and pathogenic clinical isolates (16) correlates with the presence of a unique cysteine proteinase gene (ACP1), which was not detected in Southern blots of DNA of nonpathogenic amebae. Tannich et al. (10) also identified restriction fragment–length polymorphisms between pathogenic and nonpathogenic strains by hybridization with two other cysteine proteinase genes. One of these genes (cEh-CPnp) may correspond to ACP2 and the other (cEh-CPp) is identical to ACP3. However, our examination of a larger set of clinical isolates suggests that neither of these latter two closely related genes is unique to nonpathogenic or pathogenic amebae (Fig. 2). We speculate that ACP2 and ACP3 arose from copies of the same gene and remain closely linked on one chromosome. Their presence in both pathogenic and nonpathogenic isolates of E.

histolytica suggests they may play a role in basic metabolism of amebae, most likely intracellular protein degradation.

ACP1 is quite divergent in sequence from ACP2 and ACP3, and copies of the ACP1 gene are present on different chromosome-sized DNA fragments than ACP2 and ACP3 (Fig. 2 B). Nevertheless, the data from both studies presented here together with those reported by Tannich et al. (10) are consistent with the presence of an additional cysteine proteinase gene in pathogenic strains, resulting in increased cysteine proteinase mRNA and consequently higher levels of extracellular proteolytic activity. Our observation that amino-terminal sequencing of the E. histolytica cysteine proteinase yields more than one signal (APE[S]K[A] . .) also supports the hypothesis that the enhanced production and release of cysteine proteinases by virulent E. histolytica is due to the presence and expression of multiple genes. Although our Southern blot analysis (Fig. 3 A) suggested that the ACP1 gene was unique to pathogenic isolates, Orozco (38) has offered an alternative explanation for these results. She speculates that gene amplification may lead to enhanced expression of virulence factors in pathogenic amebae. We cannot exclude the possibility that our hybridization conditions did not identify a single copy of ACP1 in nonpathogenic isolates and that the gene is highly amplified in pathogenic DNA, leading to enhanced cysteine proteinase expression. In fact, the results of varying EcoRI digestion conditions (0.1–8 U of enzyme) suggested that multiple tandem copies of ACP1 were present. FIGE also suggested that there were at least two distinct chromosome locations for ACP1 gene copies (Fig. 2 B). In either case, we speculate that in pathogenic isolates the expression of the ACP1 gene(s) in addition to expression of ACP2 and ACP3 genes results in extracellular proteinase activity exceeding a critical threshold, which, coupled with other virulence factors (11–18), may lead to tissue lysis, invasion of the bowel wall by trophozoites, and disseminated infection.

There are striking parallels in the release of cathepsin L and B by invasive cancer cells or activated macrophages and cys-
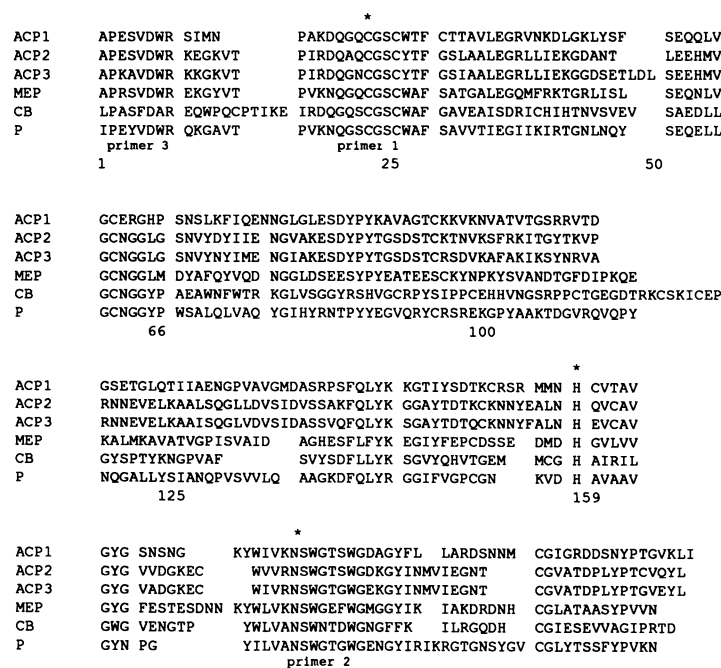
```
                                  *
ACP1   APESVDWR SIMN      PAKDQGQCGSCWTF CTTAVLEGRVNKDLGKLYSF    SEQQLVDCDASDN
ACP2   APESVDWR KEGKVT    PIRDQAQCGSCYTF GSLAALEGRLLIEKGDANT     LEEHMVQCTRDNGNN
ACP3   APKAVDWR KKGKVT    PIRDQGNCGSCYTF GSIAALEGRLLIEKGGDSETLDL SEEHMVQCTREDGNN
MEP    APRSVDWR EKGYVT    PVKNQGQCGSCWAF SATGALEGQMFRKTGRLISL    SEQNLVDCSGPQGNE
CB     LPASFDAR EQWPQCPTIKE IRDQGQSCGSCWAF GAVEAISDRICHIHTNVSVEV SAEDLLLDCCGIQCGD
P      IPEYVDWR QKGAVT    PVKNQGSCGSCWAF SAVVTIEGIIKIRTGNLNQY    SEQELLDCDRRSY
       primer 3           primer 1
       1                  25                     50


ACP1   GCERGHP SNSLKFIQENNGLGLESDYPYKAVAGTCKKVKNVATVTGSRRVTD
ACP2   GCNGGLG SNVYDYIIE NGVAKESDYPYTGSDSTCKTNVKSFRKITGYTKVP
ACP3   GCNGGLG SNVYNYIME NGIAKESDYPYTGSDSTCRSDVKAFAKIKSYNRVA
MEP    GCNGGLM DYAFQYVQD NGGLDSEESYPYEATEESCKYNPKYSVANDTGFDIPKQE
CB     GCNGGYP AEAWNFWTR KGLVSGGYRSHVGCRPYSIPPCEHHVNGSRPPCTGEGDTRKCSKICEP
P      GCNGGYP WSALQLVAQ YGIHYRNTPYYEGVQRYCRSREKGPYAAKTDGVRQVQPY
       66                           100


                                                      *
ACP1   GSETGLQTIIAENGPVAVGMDASRPSFQLYK KGTIYSDTKCRSR MMN H CVTAV
ACP2   RNNEVELKAALSQGLLDVSIDVSSAKFQLYK GGAYTDTKCKNNYEALN H QVCAV
ACP3   RNNEVELKAAISQGLVDVSIDASSVQFQLYK SGAYTDTQCKNNYFALN H EVCAV
MEP    KALMKAVATVGPISVAID  AGHESFLFYK EGIYFEPCDSSE DMD H GVLVV
CB     GYSPTYKNGPVAF       SVYSDFLLYK SGVYQHVTGEM  MCG H AIRIL
P      NQGALLYSIANQPVSVVLQ AAGKDFQLYR GGIFVGPCGN   KVD H AVAAV
       125                                        159


                *
ACP1   GYG SNSNG     KYWIVKNSWGTSWGDAGYFL LARDSNNM CGIGRDDSNYPTGVKLI
ACP2   GYG VVDGKEC   WVVRNSWGTSWGDKGYINMVIEGNT       CGVATDPLYPTCVQYL
ACP3   GYG VADGKEC   WIVRNSWGTGWGEKGYINMVIEGNT       CGVATDPLYPTGVEYL
MEP    GYG FESTESDNN KYWLVKNSWGEFWGMGGYIK IAKDRDNH  CGLATAASYPVVN
CB     GWG VENGTP    YWLVANSWNTDWGNGFFK   ILRGQDH   CGIESEVVAGIPRTD
P      GYN PG        YILVANSWGTGWGENGYIRIKRGTGNSYGV CGLYTSSFYPVKN
                 primer 2
```

Figure 6. Alignment of predicted sequences of the cysteine proteinases of E. histolytica (ACP1, ACP2, and ACP3) to the secreted cysteine proteinase (cathepsin L) of transformed fibroblasts (*MEP*), human cathepsin B (*CB*), and papain (*P*). Active site residues marked by asterisk. Conserved regions used for PCR primers are indicated. Sequence of ACP3 determined from this study is from primer[1] to primer[2] site; the remaining sequence shown is that published for the identical gene cEh-CPp (10).

teine proteinase release by pathogenic amebae. In each case, cysteine proteinases are found both in intracellular digestive organelles as well as released extracellularly in multiple molecular forms, including active or activatable proenzymes (17, 24–26). Like the targeting of cysteine proteinases to lysosomes or endosomes in mammalian cells, the cysteine proteinases of both pathogenic and nonpathogenic strains of *E. histolytica* appear to be targeted at least in part to endosome-like cytoplasmic vacuoles where endocytosed bacteria and cells are degraded (39). Default secretion due to alterations in mannose-6-phosphate containing signal moieties or transcriptional overexpression of specific proteinase genes are two mechanisms by which transformed mammalian cells are thought to release lysosomal cysteine proteinases into the extracellular milieu (26). In amebae there is a correlation between elevated rates of phagocytosis and virulence (40), and phagocytosis-deficient mutants release less extracellular cysteine proteinase (22). The adaptive advantage for the pathogenic amebae of an additional cysteine proteinase gene and, consequently, higher levels of extracellular cysteine proteinase may therefore be related to enhanced phagocytosis and/or extracellular digestion of bacteria and cells.

At this time we cannot exclude the possibility that the differences in the coding sequence of the three cysteine proteinase genes of *E. histolytica* might reflect distinct transport or biochemical properties. However, at least the specificity for synthetic peptide substrates among the three gene products appears to be identical (17–20) and, unlike their mammalian counterparts, none of the ameba cysteine proteinases contain asparagine-linked carbohydrate addition sites for mannose-6-phosphate lysosome targeting signals (Fig. 6) (10).

All previous studies of virulence have used axenized pathogenic strains that have become attenuated to varying degrees during cultivation. A direct comparison to nonpathogenic strains has not been possible because these strains have never been axenized. On the basis of the observations reported here, we can now confirm that there are quantitative differences in CPE between authentic pathogenic and nonpathogenic clinical isolates, which in turn correlate with the presence of the ACP1 gene in pathogenic strains and enhanced proteinase release (Fig. 5). Noteworthy, however, is the observation that CPE is not totally absent in nonpathogenic isolates and therefore cannot be used alone as an assay of pathogenicity, as was suggested by its historical use with axenic strains. We also confirm previous studies (22) that CPE is inhibited by a specific inhibitor of cysteine proteinases.

It is likely that host immune status, nutritional status, and associated intestinal bacteria influence susceptibility to colonization and invasion by *E. histolytica*. However, our identification of a cysteine proteinase gene that is specific to pathogenic strains and correlates with increased extracellular proteinase activity argues strongly that the potential for pathogenicity is an intrinsic property of the organism.

Acknowledgments

References

1. Walsh, J. A. 1988. Prevalence of *Entamoeba histolytica* infection. *In* Amebiasis. Human Infection by *Entamoeba histolytica*. J. I. Ravdin, editor. John Wiley and Sons, New York. 93–105.

2. Brumpt, E. 1925. Etude sommaire de l'*Entamoeba dispar* n. sp. amibe a kystes quadrinuclées, parasite de l'homme. *Bull. Acad. Natl. Med. (Paris)*. 94:942–952.

3. Mirelman, D., R. Bracha, A. Wexler, and A. Chayen. 1986. Changes in isoenzyme patterns of a cloned culture of nonpathogenic *Entamoeba histolytica* during axenization. *Infect. Immun.* 54:827–832.

4. Sargeaunt, P. G. 1987. The reliability of *Entamoeba histolytica* zymodemes in clinical diagnosis. *Parasitol. Today*. 3:40–43.

5. Sargeaunt, P. G., J. E. Williams, and J. D. Greene. 1978. The differentiation of invasive and noninvasive *Entamoeba histolytica* by isoenzyme electrophoresis. *Trans. R. Soc. Trop. Med. Hyg.* 72:519–521.

6. Tannich, E., R. D. Horstmann, J. Knobloch, and H. H. Arnold. 1989. Genomic DNA differences between pathogenic and nonpathogenic *Entamoeba histolytica*. *Proc. Natl. Acad. Sci. USA.* 86:5118–5122.

7. Garfinkel, L. I., M. Giladi, M. Huber, C. Gitler, D. Mirelman, M. Revel, and S. Rozenblatt. 1989. DNA probes specific for *Entamoeba histolytica* possessing pathogenic and nonpathogenic zymodemes. *Infect. Immun.* 57:926–931.

8. Que, X., and S. L. Reed. 1991. Nucleotide sequence of a small subunit ribosomal RNA (16s-like rRNA) gene from *Entamoeba histolytica*: differentiation of pathogenic from nonpathogenic isolates. *Nucl. Acids Res.* 19:5438.

9. Edman, U., M. A. Meraz, S. Rausser, N. Agabian, and I. Meza. 1990. Characterization of an immuno-dominant variable surface antigen from pathogenic and nonpathogenic *Entamoeba histolytica*. *J. Exp. Med.* 172:879–888.

10. Tannich, E., H. Scholze, R. Nickel, and R. D. Horstmann. 1991. Homologous cysteine proteinases of pathogenic and nonpathogenic *Entamoeba histolytica*. *J. Biol. Chem.* 266:4798–4803.

11. Ravdin, J. I. 1986. Pathogenesis of disease caused by *Entamoeba histolytica*: studies of adherence, secreted toxins, and contact-dependent cytolysis. *Rev. Infect. Dis.* 8:247–260.

12. Petri, W. A., R. D. Smith, P. H. Schlesinger, C. F. Murphy, and J. I. Ravdin. 1987. Isolation of the galactose-binding lectin that mediates the in vitro adherence of *Entamoeba histolytica*. *J. Clin. Invest.* 80:1238–1244.

13. Gitler, C., E. Calef, and I. Rosenberg. 1984. Cytopathogenicity of *Entamoeba histolytica*. *Philos. Trans. R. Soc. Lond. Biol. Sci.* B307:73–85.

14. Young, J. D., T. M. Young, L. P. Lu, J. C. Unkeless, and Z. A. Cohn. 1982. Characterization of a membrane pore-forming protein from *Entamoeba histolytica*. *J. Exp. Med.* 156:1677–1690.

15. Leippe, M., E. Sebastian, O. L. Schoenberger, R. D. Horstmann, and H. J. Müller-Eberhard. 1991. Pore-forming peptide of pathogenic *Entamoeba histolytica*. *Proc. Natl. Acad. Sci. USA.* 88:7659–7663.

16. Reed, S. L., W. E. Keene, and J. H. McKerrow. 1989. Thiol proteinase expression correlates with pathogenicity of *Entamoeba histolytica*. *J. Clin. Microbiol.* 27:2772–2777.

17. Keene, W. E., M. G. Pettit, S. Allen, and J. H. McKerrow. 1986. The major neutral proteinase of *Entamoeba histolytica*. *J. Exp. Med.* 163:536–549.

18. Luaces, A. L., and A. J. Barrett. 1988. Affinity purification and biochemical characterization of Histolysin, the major cysteine proteinase of *Entamoeba histolytica*. *Biochem. J.* 250:903–909.

19. Scholze, H., and E. Werries. 1986. Cysteine proteinase of *Entamoeba histolytica*. I. Partial purification and action on different enzymes. *Mol. Biochem. Parasitol.* 18:103–112.

20. Lushbaugh, W. B. 1988. Proteinases of *Entamoeba histolytica*. *In* Amebiasis: Human Infection by *Entamoeba histolytica*. J. I. Ravdin, editor. John Wiley and Sons, New York. 219–231.

21. Reed, S. L., W. E. Keene, J. H. McKerrow, and I. Gigli. 1989. Cleavage of C3 by a neutral cysteine proteinase of *Entamoeba histolytica*. *J. Immunol.* 143:189–195.

22. Keene, W. E., M. E. Hidalgo, E. Orozco, and J. H. McKerrow. 1990. *Entamoeba histolytica*: correlation of the cytopathic effect of virulent trophozoites with secretion of a cysteine proteinase. *Exp. Parasitol.* 71:199–206.

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

24. Sloane, B. F., J. R. Dunn, and K. V. Honn. 1981. Lysosomal Cathepsin B: correlation with metastatic potential. *Science (Wash. DC)*. 212:1151–1153.

25. Mort, J. S., M. Leduc, and A. D. Recklies. 1981. A latent thiol proteinase from ascitic fluid of patients with neoplasia. *Biochim. Biophys. Acta.* 662:173–180.

26. Troen, B. R., D. Ascherman, D. Atlas, and M. M. Gottesman. 1988. Cloning and expression of the gene for the major excreted protein of transformed mouse fibroblasts: a secreted lysosomal protease regulated by transformation. *J. Biol. Chem.* 263:254–261.

27. Eakin, A. E., J. Bouvier, J. A. Sakanari, C. S. Craik, and J. H. McKerrow. 1990. Amplification and sequencing of genomic DNA fragments encoding cysteine proteases from protozoan parasites. *Mol. Biochem. Parasitol.* 39:1–8.

28. Diamond, L. S., D. R. Harlow, and C. C. Cunnick. 1978. A new medium

for the axenic cultivation of *Entamoeba histolytica* and other *Entamoeba. Trans. R. Soc. Trop. Med. Hyg.* 72:431–432.

29. Robinson, G. L. 1968. The laboratory diagnosis of human parasitic amoebae. *Trans. R. Soc. Hyg. Trop. Med. Hyg.* 62:285–294.

30. Huber, M., L. Garfinkel, C. Gitler, D. Mirelman, M. Revel, and S. Rozenblatt. 1987. *Entamoeba histolytica:* cloning and characterization of actin cDNA. *Mol. Biochem. Parasitol.* 24:227–235.

31. Chomczynski, P., and N. Sacchi. 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal. Biochem.* 162:156–159.

32. Torian, B. E., B. M. Flores, V. L. Stroeher, F. S. Hagen, and W. E. Stamm. 1990. cDNA sequence analysis of a 29-kDa cysteine-rich surface antigen of pathogenic *Entamoeba histolytica. Pro. Natl. Acad. Sci. USA.* 87:6358–6362.

33. Church, G. M., and W. Gilbert. 1984. Genomic sequencing. *Proc. Natl. Acad. Sci. USA.* 81:1991–1995.

34. Edman, U., I. Meza, and N. Agabian. 1987. Genomic and cDNA actin sequences from a virulent strain of *Entamoeba histolytica. Proc. Natl. Acad. Sci. USA.* 84:3024–3028.

35. Sambrook, J. C., E. F. Fritsch, and T. Maniatis. 1989. Molecular Cloning: A Laboratory Manual. 2nd ed. C. Nolan, editor. Cold Spring Harbor Press, Cold Spring Harbor, New York.

36. Matsudaira, P. 1987. Sequence from picomole quantities of proteins electroblotted onto polyvinylidene difluoride membranes. *J. Biol. Chem.* 262:10035–10038.

37. Eakin, A. E., A. A. Mills, G. Harth, J. H. McKerrow, and C. S. Craik. 1992. The sequence, organization, and expression of the major cysteine protease (cruzain) from *Trypanosoma cruzi. J. Biol. Chem.* 267:7411–7420.

38. Orozco, E. 1992. Pathogenesis in amebiasis. *Infectious Agents and Disease.* 1:19–21.

39. De Meester, F., E. Shaw, H. Scholze, T. Stolarsky, and D. Mirelman. 1990. Specific labeling of cysteine proteinases in pathogenic and nonpathogenic *Entamoeba histolytica. Infect. Immun.* 58:1396–1401.

40. Orozco, E., M. A. Rodriguez, and F. Hernandez. 1988. The role of phagocytosis in the pathogenic mechanism of *Entamoeba histolytica. In* Amebiasis: Human Infection by *Entamoeba histolytica.* J. I. Ravdin, editor. John Wiley and Sons, New York. 326–338.