UNIVERSITY OF CALIFORNIA, SAN DIEGO

EhCP1 and EhCP5: key released cysteine proteinases of *Entamoeba histolytica* as drug targets

A Thesis submitted in partial satisfaction of the
requirements for the degree of Master of Science

in

Biology

by

Jasmine Ta

Committee in Charge:

Professor Sharon Reed, Chair
Professor Raffi Aroian, Co-Chair
Professor Stephanie Mel

2012

The Thesis of Jasmine Ta is approved, and is acceptable in quality and form for publication on microfilm and electronically

_____

_____

Co-Chair

_____

Chair

University of California, San Diego

2012

DEDICATION

I would like to dedicate this thesis to my mom and dad, who have guided and supported me throughout my undergraduate and graduate education at UCSD.

I would also like to thank my friends Cindy and Lisa for their wonderful, home-cooked food and late-night company throughout my Masters education.

Table of Contents

List of Figures

List of Tables

Acknowledgments

I would first like to express my appreciation to Dr. Sharon Reed for allowing me the opportunity to work in her lab and supporting me throughout my numerous projects. This past year has provided me invaluable knowledge about the skills and critical thinking needed to succeed in the biological sciences.

I would also like to thank Dr. Raffi Aroian and Dr. Stephanie Mel for spending the time to serve on my thesis committee.

Next, I would like to thank Uyen Tran, who served as my mentor for the past year and taught me various experimental techniques. I would also like to express my gratitude to Ken Hirata, who helped me with experimental design, interpreting results, and troubleshooting. I would not have been able to graduate without both of your guidance and support.

I would also like to thank all of the members of the Reed lab: Rosa for her kindness; Grace for helping me build character and feeding me food. I will try to BLG and think about WWKS. And, of course, I can't forget WUSU.

Lastly, I would like to thank my family and friends who have encouraged and supported me while I worked towards this graduate degree.

ABSTRACT OF THE THESIS

EhCP1 and EhCP5: key released cysteine proteinases of *Entamoeba histolytica* as drug targets

by

Jasmine Ta

Master of Science in Biology

University of Califorina, San Diego, 2012

Professor Sharon Reed, Chair

Professor Raffi Aroian, Co-Chair

The protozoan parasite *Entamoeba histolytica* is the causative agent of amebiasis, which affects more than 50 million people worldwide resulting in 70,000 deaths annually. Current treatment relies on metronidazole; however, metronidazole-resistant *E. histolytica* have been induced in the laboratory. Therefore, new drugs based on defined targets are needed to treat amebiasis. Recently, new drugs designed to inhibit *E. histolytica* cysteine proteinases, which are major virulence factors in amebic invasion and evasion of the host immune system, have been developed. To gain a better understanding of how these inhibitors irreversibly bind their target proteinases, recombinant EhCP1 has been resynthesized and expressed in bacteria to cocrystallize with cysteine proteinase

inhibitors. Although soluble rEhCP1 has been expressed, obtaining the necessary large quantity for crystallization is still in progress. Biologically active rEhCP5, the other cysteine proteinase unique to invasive *E. histolytica* was assayed against vinyl sulfone derivatives of K11777, a cysteine proteinase inhibitor of cruzain that is currently in Phase I clinical trials for treatment of Chagas disease. Through these inhibition assays, both rEhCP5 and rEhCP1 have a substrate preference for positively charged amino acids at the P_2 position, more specifically an arginine, despite possessing different amino acid residues at the base of the active pockets. While the readily cell-permeable inhibitors WRR-666 and WRR-668 were effective against rEhCP1 and rEhCP5 *in vitro*, minimal inhibition of infection by *E. histolytica* trophozoites was observed in mouse models, possibly due to drug instability. rEhCP1 and rEhCP5 expression and corresponding inhibition assays provide insight into the development of new therapeutics to treat amebiasis.

Introduction

Amebiasis is an infectious disease affecting more than 50 million people worldwide. (Mortimer et al, 2009). Amebiasis is most common in developing countries such as Mexico, India, South Africa, and South America due to poor sanitation regulations, but travel and tourism have led to an increase in amebiasis cases within developed countries as well (Hughes et al, 2000). Early diagnosis and treatment is important as fatal complications cause approximately 100,000 deaths annually, second only to malaria as a cause of death from protozoal diseases (Mortimer et al, 2009).

Entamoeba infection start with ingestion of contaminated food or water containing cysts. These amebic cysts may asymptomatically colonize the colon. The majority of patients harbor non-invasive *Entamoeba dispar*, whose cysts are identical to *E. histolytica* and make diagnosis of patients at risk of invasive disease difficult. Following excystation, however, trophozoites — the active, mobile stage — are released (Huston et al, 2004). Only *E. histolytica* trophozoites are capable of invading the mucus layer of the colon to cause diarrhea or enter the bloodstream to cause liver abscesses. These released trophozoites can ingest and kill host cells in a contact-dependent manner (Ralston et al, 2011). The combination of host cell death and the resulting host inflammatory response can lead to destruction of host tissue with ulcers in the colon and large abscesses in the liver. Treatment of amebiasis worldwide relies primarily on one class of drugs, imidazoles, particularly metronitrodazole. Because resistance has been produced in the laboratory (Wassman et al, 1999), the identification of new drug targets is important.

Several virulence factors have been implicated in the pathogenesis of *E. histolytica.* One such virulence factor is gal-lectin, an *E. histolytica* surface protein. Gal-

lectin allows for trophozoite adherence to the colonic mucus layer, the first step in invasion (Chadee et al, 1987). Within mice, immunization to gal-lectin prevented infection with an efficacy of 80% (Houpt et al, 2004). Another membrane-bound protein that serves as a virulence factor is *E. histolytica* lipophosphopeptidoglycan (EhLPPG), which activates a host inflammatory response. This inflammation leads to microerosions within epithelia, allowing for entry and dissemination of trophozoites. Within *in vivo* xenograph models using severe combined immunodeficient mice (SCID mice) with fetal human intestine transplanted on their backs, inhibition of NFK-B leads to prevention of trophozoite invasion (Seydel et al, 1997). Once trophozoites adhere to epithelia, they can release amoebapores. These pore-forming proteins present in both *E. histolytica and E. dispar* — can be inserted into host cell membranes and cause necrosis. Knockdown of amoebapore expression within SCID mice reduces the size of amebic liver abscesses (Zhang et al, 2004).

One important virulence factor of *E. histolytica* and potential drug target are cysteine proteinases. Cysteine proteinases degrade proteins for food digestion and protein turnover; they also serve as signaling molecules for important biological functions (Turk, 2006). Disregulation or mutations of cysteine proteinases within humans has been shown to cause multiple diseases, such as cancer and hypertension. Cysteine proteinases are also crucial for protozoan development, host cell destruction, and host immune system evasion. Inhibition of cysteine proteinases blocks intracellular development of the protozoa *Trypanosoma cruzi* and *Leishmania* (Que et al, 2000). Cysteine proteinases are also vital for the excystation of *Giardia* (Que et al, 2000). This important role of cysteine proteinases has been utilized in the development of antiprotozoal drugs; a cysteine

proteinase inhibitor of *Trypanosoma cruzi* is currently in clinical trials for the treatment of Chagas disease (McKerrow et al, 2009).

Cysteine proteinases also play a key role in the virulence of *E. histolytica.* From *Entamoeba* genome study, there are 84 proteinase genes, 50 of which encode for cysteine proteinases (Tillack et al, 2007). These cysteine proteinases degrade the extracellular matrix of the colon, allowing for trophozoite invasion into the bloodstream (Hou et al, 2010). Cysteine proteinases have also been shown to degrade IgA and IgG, antibodies needed to activate the host immune response (Que et al, 2003).This attachment and penetration of host tissue and evasion of host immune system allows for potential dissemination of metastatic infections (Que et al, 2000).

Although there are 50 cysteine proteinases genes in *E. histolytica*, 90% of the total proteinase activity *in vitro* is derived from only three cysteine proteinases: EhCP1, EhCP2 and EhCP5 (Melendez-Lopez et al, 2007). Of these three highly expressed and secreted proteinases, only EhCP1 and EhCP5 are present in the invasive strain *E. histolytica* versus the noninvasive strain *Entamoeba dispar,* implicating their potential importance in colonic invasion (Hou et al, 2010). EhCP1 mRNA expression is increased twofold during invasion in the mouse cecal model, while EhCP5 is not (Gilchrist et al, 2005). In addition, recombinant EhCP1 has been shown to circumvent the host immune system by cleaving the α-chain of human complement component C3 (Reed, 1990), pro-IL-18 (Que et al, 2003), and IgG (Tran et al, 1998). Specific inhibitors developed against EhCP1 blocked invasion in human intestinal xenografts on SCID mice, validating EhCP1 as a drug target (Melendez-Lopez et al, 2007).

EhCP5 also plays an important role in the invasion process. EhCP5 is likely involved in the processing of other proteinases as overexpression of EhCP5 by episomal transfection leads to the increased activity of EhCP1, EhCP2, and EhCP5 (Tillack et al, 2006). Down-regulation of EhCP5 through antisense RNA down-regulates all EhCPs (Tillack et al, 2006). Previous studies have shown that EhCP5 can degrade MUC2 — the major glycosylated protein comprising 90% of colonic mucus — *in vitro*, allowing for attachment to host cells (Moncada et al, 2005). Inhibition of EhCP5 within *E. histolytica* through antisense RNA causes a significant decrease in the amoeba's ability to degrade MUC2 within Chinese hamster cells (Moncada et al, 2005) and limits the formation of liver abscesses in hamsters (Ankri et al, 1999). This EhCP5 knock-down and the correlating decrease in liver abscess formation provide compelling evidence for EhCP5 as an *E. histolytica* virulence factor and potential drug target.

Based on both *in vitro* and *in vivo* data detailed above, EhCP1 and EhCP5 may be effective targets for novel drugs treating amebiasis. However, due to the difficulty in isolating large quantities of native EhCPs, recombinant amebic cysteine proteinases expressed in *Escherichia coli* and *Pichia pastoris* have been generated and used in inhibition assays. Two drawbacks of the bacterial expression system are that the recombinant proteins tends to be toxic and insoluble due to different codon usage between *Entamoeba* and *E. coli* GC-content and nucleotide bias. *E. coli* cells usually package the over-expressed proteinases within inclusion bodies (He et al, 2010), which may cause the proteinases to become misfolded and inactivated.

In order to achieve soluble rEhCP1, the first part of my project focuses on optimizing the expression of recombinant his-tagged rEhCP1 for co-crystallization with

cysteine proteinase inhibitors. To improve soluble expression within bacteria systems, rEhCP1 was expressed within three different bacterial cell lines: Origami, ArcticExpress, and BL21Codon-Plus. The Origami strain contains mutations in the thioredoxin reductase and glutathione reductase genes, allowing for greater disulfide bond formation within the cytoplasm. In contrast, the ArcticExpress strain can be grown at the low temperature of 16°C, providing slower growth conditions for protein folding. The ArcticExpress strain also expresses recombinant chaperonins Cpn10 and Cpn60, which retain protein refolding activity at low temperatures. In addition, rEhCP1 was also expressed within BL21Codon-Plus, a strain containing rare bacterial tRNAs for protein translation of the AT-rich *E. histolytica*.

The second part of my project focuses on creating soluble, histidine-tagged recombinant EhCP5 in the yeast *P. pastoris* expression vector. As a eukaryotic expression system, the *P. pastoris* vector provides several advantages, including the addition of posttranslational modifications to proteins. These posttranslational modifications, such as glycosylation and disulfide bond formation allow for proper refolding of EhCP5 (Que et al, 2003). In addition, *P. pastoris* is a strain of methyltropic yeast that can use a single-carbon compound, such as methanol, as its energy source (Cereghin et al, 1999). By placing the EhCP5 gene downstream of the alcohol oxidase promoter — the promoter normally used for the expression of protein alcohol oxidase, which catalyzes the first step in methanol metabolism — EhCP5 expression can be induced by the addition of methanol (Cereghin et al, 1999). Lastly, the *P. pastoris* expression vector contains a yeast alpha secretion factor, which signals for the export of recombinant protein from the cell (Invitrogen). This secretion of protein from the cell into

the medium provides a simple, first round of purification. The addition of the 6X-histidine tag to EhCP5 and ion-exchange chromatography through a monoQ column also provides further purification of the soluble protein.

Recombinant EhCP5 can then be used in inhibition assays to test the effectiveness of new inhibitors. A number of irreversible vinyl sulfone cysteine proteinase inhibitors were optimized based on derivatives of K11777, a cysteine proteinase which has undergone extensive toxicity and pharmacokinetic testing and is now in clinical trials for Chagas disease. These inhibitors were optimized specifically for EhCP1 because of its marked preference for arginine in the P2 position (Melendez-Lopez et al, 2007). The new inhibitors will be tested against EhCPs, and if active, can be tested in an *in vivo* model of cecal amebiasis using CBA/J mice (He et al, 2010).

Materials and Methods

Recombinant Expression of EhCP1 in Bacteria

The coding region of the pro-EhCP1 gene was PCR amplified from *E. histolytica* genomic DNA using a combination of three different primers: NdeI forward primer (5' GGCAATTCATATGATTGATTTCAACACTTGGGTTG 3'), EcoRI forward primer (5' CCAGAATTC ATGATTGATTTCAACACTTGGGTTG 3'), and XhoI reverse primer (5'CGGCTCGAGTAAATATTCAACACCAGTTG 3'). The PCR product containing NdeI and XhoI restriction sites were used for cloning of the EhCP1 DNA into the plasmid pET-22b(+) (Novagen, Billerica, MA) without the peIB leader sequence; whereas, the PCR product containing the EcoRI and XhoI restriction sites were used to clone EhCP1 DNA into pET-22b(+) with an intact peIB leader sequence. This plasmid encodes a C-terminal His tag sequence for protein purification. PCR of the genomic DNA was performed in a Perkin Elmer GeneAmp 9600 Thermal Cycler, according to the following program: 95°C for 3 minutes; 40 cycles of 95°C for 30 seconds, 57°C for 30 seconds, 72°C for 2 minutes; 72°C for 10 minutes. The pro-EhCP1 PCR product was purified with QIAquick PCR Purification Kit (Qiagen, Valencia, CA), in which 5 volumes of Buffer PB was added to the PCR sample, passed over a QIAquick spin column at 13,000 rpms for 1 minute, washed with 750 µls Buffer PE, and eluted with 35 µls of Buffer EB. The pro-EhCP1 PCR product was confirmed by running the purified product on a 1% agarose gel. The pro-EhCP1 gene and pET-22b(+) plasmid were incubated with NdeI and XhoI restriction enzymes for 2 hours at 37°C and ligated together in a 1:3 vector to insert molar ratio with T4 DNA ligase (New England Biolabs, Ipswich, MA). The recombinant pET-22b(+)-proEhCP1 plasmid was transformed into *E.*

coli TOP10 (Invitrogen, Carlsbad, CA) competent cells by placing 1 µl of the ligation mixture into thawed TOP10 competent cells and incubating on ice for 30 minutes. The transformed cells were heat shocked for 30 seconds at 42°C, and cooled on ice for 2 minutes, before adding Super Optimal Broth with Catabolite Repression (2% peptone, 0.5% Yeast extract, 10mM NaCl, 2.5mM KCl, 10mM $MgCl_2$, 10mM $MgSO_4$, 20 mM glucose) medium to a final volume of 250 µls followed by shaking the cells for 1 hour at 37°C. The cells were plated onto Luria Broth (LB) plates supplemented with 100 µgs/ml ampicillin. Recombinant pET-22b(+)-proEhCP1 plasmids were isolated from TOP10 cells using a Miniprep Purification Kit (Qiagen, Valencia, CA) by centrifuging 5 mls of cells at 4000 rpm for 20 minutes, resuspending the pellet in 250 µls Buffer P1, lysing cells with 250 µls Buffer P2, followed by neutralizing the solution with 350 µls Buffer N3 prior to centrifuging the mixture at 13,000 rpms for 10 minutes. The supernatant was applied to a QIAPrep spin column and centrifuged at 13,000 rpms for 1 minute, washed with 500 µls Buffer PB (centrifuged) and 250 µl Buffer PE (centrifuged), and eluted with 50 µls Buffer EB. The isolated EhCP1/pET22b(+) plasmid DNA sequence was confirmed by Sanger dideoxynucleotide methodology (Genewiz, La Jolla, CA).

Once the identity of EhCP1 gene was confirmed, the isolated, recombinant pET-22b(+)-proEhCP1 plasmid was transformed into *E. coli* ArcticExpress competent cells (Stratagene, La Jolla, CA) as described above and plated onto LB plates (1% peptone, 0.5% yeast extract, 1% NaCl, 2% agar) supplemented with 25 µg/ml ampicillin and 20 µg/ml gentamycin. Single colonies were inoculated into 1 ml LB medium (1% peptone, 0.5% yeast extract, 1% NaCl, with ampicillin and gentamycin) and grown at 37°C overnight, then subcultured in 1.5 mls LB medium containing ampicillin and gentamycin

and grown at 30°C for 3 hours. Colonies were induced with addition of isopropyl-β-D-thiogalactoside (IPTG) to a final concentration of 1 mM and cultured at 16°C for 1, 3, and 24 hours. Induced cultures were pelleted by centrifugation at 5000 rpm for 10 minutes, resuspended in 100 μls B-PER reagent (Thermo Scientific, Rockford, IL), and vortexed for 1 minute to form a homogenous suspension. The soluble and insoluble protein mixture was separated by centrifugation at 15,000 rpms for 5 minutes. The soluble and insoluble protein fractions were analyzed via SDS-PAGE on 12% acrylamide gels and stained with Coomassie Blue.

The recombinant pET-22b(+)-proEhCP1 plasmid was also transformed into *E. coli* BL21 Codon-Plus (Agilent Technologies, Santa Clara, CA) and Origami competent cells (EMD Chemicals, San Diego, CA). Transformed BL21 Codon-Plus cells were plated onto LB plates supplemented with 25 μg/ml ampicillin, 20 μg/ml gentamycin, and 34 μg/ml chloramphenicol. Transformed Origami cells were plated onto LB plates supplemented with 100 μg/ml ampicillin, 50 μg/ml kanamycin, and 100 μg/ml tetracycline. Induction, expression, purification, and characterization of BL21 Codon-Plus and Origami EhCP1 transformed clones were carried out in a similar manner as ArticExpress, except for incubation at 37˚C (described above).

EhCP1 gene was resynthesized by McKerrow's group using GenScript Optimization Analysis. The rProEhCP1 gene, re-engineered for bacterial expression, was cloned into pET28b and transformed into ArcticExpress. After a 2 hour induction at room temperature with 1 mM IPTG, soluble, nickel column-purified rEhCP1 was analyzed by Coomassie stain and Western blot with EhCP1 polyclonal antibody.

Expression of Recombinant His-tagged EhCP5 in *Pichia* (pPICZα)

The proEhCP5 gene containing a 6x-histidine tag was cloned into the pPICZα vector for transformation into *Pichia pastoris* cells. The pro-EhCP5-his tag gene was PCR amplified from *E. histolytica* genomic DNA using a forward primer (5'CCAGAATTCCAACAAATTTCAATACTTGGGTTGC 3'), incorporating an EcoRI restriction site, and a reverse primer (5'CCTGGTACCCTATTAATGATGATGATGATGATGAGCATCATGAACCCCAAC 3') encoding a 3'-His tag and incorporating a KpnI restriction site. PCR of the genomic DNA was performed in a Perkin Elmer GeneAmp 9600 Thermal Cycler, according to the following program: 95°C for 5 minutes; 32 cycles of 95°C for 2 minutes, 62°C for 15 seconds, 70°C for 2 minutes; 70°C for 10 minutes. The PCR product was purified with the Qiagen QIAquick PCR Purification Kit as described above (Qiagen, Valencia, CA) and confirmed as pro-EhCP5-his tag. Once the correct pro-EhCP5 gene size was confirmed, both the pro-EhCP5 gene and pPICZαB plasmid were incubated with EcoRI and KpnI restriction enzymes for 2 hours at 37°C and ligated together in a 1:3 vector to insert molar ratio with T4 DNA ligase.

The recombinant pPICZαB-proEhCP5 plasmid was transformed into *E. coli* NEB Turbo competent cells (New England BioLabs, Ipswich, MA) and plated onto low-salt LB plates (1% peptone, 0.5% yeast extract, 0.5% NaCl) supplemented with 25 μg/ml zeocin. Recombinant pPICZα-proEhCP5 plasmids were isolated from NEB Turbo (New England BioLabs, Ipswich, MA) cells using a Miniprep Purification Kit (Qiagen, Valencia, CA). The plasmid DNA was concentrated through ethanol precipitation by incubating with 1:10 volume of 3M sodium acetate and 2 volumes of 100% ethanol for

30 minutes at 4°C, centrifuging the mixture at 13,000 rpms for 30 minutes at 4°C, washing the pellet with 70% ethanol and centrifuging, air drying the pellet for 30 minutes, and resuspending the DNA in Buffer EB to a concentration of 1 μg/μl. The purified pPICZα-EhCP5 plasmid DNA sequence was confirmed by Sanger dideoxynucleotide methodology (Genewiz, La Jolla, CA). Five micrograms of the recombinant pPICZα-proEhCP5 vector was linearized with SacI and transformed into *Pichia pastoris X-33* by electroporation. To prepare the *Pichia* cells for electroporation, 5 mls of cells were grown in YPD (1% yeast extract, 2% peptone, 2% dextrose) at 30°C overnight, subcultured into 400 mls YPD overnight, centrifuged at 1,500 rpm for 5 minutes at 4°C, washed with 250 mls water, and resuspended in 1 ml of 1 M sorbitol. These cells were incubated with linearized pPICZα-proEhCP5 DNA for 5 minutes at 4°C and electroporated using Electro Cell Manipulator (BTX Harvard Apparatus, Holliston, MA) at 2.5 kV/resistance and 186 ohms for 9 ms in a 1 mm gap disposable cuvette (BTX, Holliston, MA). The electroporation of *Pichia* cells was followed by adding 1 ml of 1M sorbitol and 1 ml of YPD and shaking the cells at 200 rpm for 2 hours at 30°C. Transformed *Pichia* cells were selected by plating onto YPD plates supplemented with 200 μg/ml zeocin.

To induce transformed *Pichia*, single colonies were inoculated into 5 ml YPD medium and grown at 30°C overnight, then subcultured into 300 mls YPD medium and grown at 25°C for 3 days. *Pichia* cells were pelleted at 2000 rpm for 5 minutes and resuspended in 1L buffered minimal medium (100 mM potassium phosphate, 1.34% yeast nitrogen base, $4x10^{-5}$% biotin, pH 7) containing a final concentration of 1.5%

methanol for 72 hours (methanol was replenished at 24 and 48 hours). At 72 hours, the induced *Pichia* were pelleted at 4,000 rpms for 5 minutes. The culture supernatant was filtered through a 0.45 μM-pore size filter and concentrated through Centricon-10 centrifugation units. The concentrated supernatant was dialyzed against 3L Tris-EDTA buffer (25 mM Tris, 2mM EDTA, pH 8), and assayed for proteolytic activity against fluorogenic peptidyl substrate Z-arginine-arginine-AMC. The induced culture cell pellets were resuspended in breaking buffer (50 mM sodium phosphate, 1 mM EDTA, 1 mM pefabloc, 5% glycerol) and lysed by vortexing with glass beads. Soluble and insoluble proteins from the cell pellet were isolated by centrifugation at 12,000 rpm for 5 minutes. Both the culture supernatant (*Pichia* culture spent media) and soluble protein fractions (from lysis of *Pichia* cell pellet) were analyzed on 12% SDS-PAGE and Coomassie Blue stain.

Expression of Recombinant EhCP5 in *Pichia* without a Histidine-tag (pPICZα)

The cloning of recombinant EhCP5 without a histidine-tag into *P. pastoris* vector pPICZα was previously performed in our lab (Que et al, 2003). The cloning, induction, and purification of recombinant EhCP5 were identical to the histidine-tagged rEhCP5 strain above. However, the coding region of the pro-EhCP5 gene was PCR amplified from *E. histolytica* genomic DNA through forward XhoI primer (5' GTGCTCGAGAAAAGAACAAATTTCAATACTTGGGTTG 3') and NotI reverse primer (5'ATAGCGGCCGCTTAACGATCACGAACCCCAACT3'). Recombinant *Pichia* colonies containing EhCP5 gene from glycerol stocks were streaked onto YPD plates. Single colonies were induced and purified in the same manner as above.

Lyophilized culture supernatant containing rEhCP5 was used in inhibition assays. The lyophilized rEhCP5 was resuspended in water to a volume one-tenth of the original volume and dialyzed against 4L Tris-EDTA buffer before purification on a MonoQ column.

MonoQ Purification of Recombinant EhCP5

Recombinant EhCP5 was purified by ion-exchange chromatography. Two milliliters of dialyzed EhCP5 supernatant was loaded onto a Mono-Q column (FPLC System, Amersham Pharmacia Biotech) and eluted with a linear gradient of 0 to 500 mM NaCl in 25 mM Tris- 2 mM dithiothreitol (DTT), pH 7.4. Column fractions were tested for proteolytic activity against Z-arginine-arginine-AMC. Active fractions were pooled and used in inhibition assays with cysteine proteinase inhibitors and stored at 4˚C.

Recombinant EhCP5 Inhibition Assays with Cysteine Proteinase Inhibitors

To test the inhibition of EhCP5 proteolytic activity, recombinant EhCP5 was incubated with various concentrations of cysteine proteinase inhibitors. The active, MonoQ-purified recombinant EhCP5 from lyophilized culture supernatant was activated by adding 1M dithiothreitol (DTT) for a final concentration of 10 mM DTT for 1 minute. Activated rEhCP5 was incubated with varying concentrations of cysteine proteinase inhibitors for 10 minutes at room temperature. After addition of substrate buffer (50 mM Tris, 2 mM EDTA, 2 mM DTT, pH 7.5) containing the fluorogenic peptidyl substrate carboxybenzyl-arginine-arginine-7-amino-4-methylcoumarin (Z-arginine-arginine-AMC, Bachem Americas, Inc., Torrance, CA) to a final volume of 200 µl, fluorescence released

by the cleaved leaving group (AMC) was measured for 10 minutes using a Labsystems Fluorskan Plate II reader. For each inhibitor, the normalized percent inhibition of proteolytic activity compared to the uninhibited control rEhCP5 was plotted against the log of the concentration of inhibitor using GraphPad Prism Software (GraphPad, La Jolla, CA). These graphs were used to calculate the 50% inhibitory concentration (IC_{50}).

For these inhibition assays, cysteine proteinase inhibitors WRR-401, WRR-483, WRR-500, WRR-518, WRR-554, WRR-605, WRR-666, WRR-667, WRR-668, WRR-669, and WRR-670 were used (Chen et al, 2010). These irreversible inhibitors are vinyl sulfone derivatives of K11777, a cysteine proteinase inhibitor developed to treat Chagas disease and is currently in Phase I clinical trials (Doyle et al, 2007). All inhibitors were prepared with dimethyl sulfoxides 10, 20, or 40 mM stocks and stored at -20°C.

Western Blot Analysis of Recombinant EhCP5

Mono Q fractions containing active rEhCP5 were precipitated by adding one-fourth volume of trichloroacetic acid, incubating on ice for 1 hour, centrifuging sample at 13,000 rpm for 30 minutes at 4°C, air drying pellet for 30 minutes, resuspending in 30 µl Thorner sample buffer (40 mM Tris-HCL, 8 M urea, 5% SDS, 143 mM β-mercaptoethanol, 0.1 mM EDTA, and 0.4 mg/ml bromophenol blue), and neutralizing residual TCA by addition of 4 µls of 1 M Tris-HCl. Precipitated recombinant EhCP5 was boiled for 5 minutes, ran on a 12% SDS-PAGE, and transferred to a polyvinylidene fluoride membrane (Immobilon P, Millipore, Billerica, MA). The polyvinylidene fluoride membrane was blocked in a 5% milk-Tween-TBS (20mM Tris, 0.15 M NaCl, 0.1% Tween) solution for 2 hours at room temperature, then incubated with primary rabbit anti-

EhCP5 polyclonal antibody for 1 hour. After three 5-minute washes with Tween-TBS, the membrane was incubated with secondary goat-anti-rabbit-alkaline phosphatase antibody (Zymed, Carlsbad, CA) for 1 hour. Following three 5-minute washes with Tween-TBS, the membrane was developed with 1 ml nitro-blue tetrazolium chloride and 5-bromo-4-chloro-3-indolyphosphate p-toluidine salt (NBT/BCIP, Thermo Scientific, Rockford, IL) for 5 minutes before stopping the reaction with water.

Western Blot Analysis of Recombinant EhCP5 with Biotinylated Inhibitors

MonoQ-purified, active rEhCP5 (5μg) was incubated with biotinylated E-64 derivative DCG04 (Greenbaum et al, 2000) for 30 minutes at room temperature. Recombinant EhCP5 incubated with biotinylated inhibitor was then resuspended in 1x sample buffer containing 5% β-mercaptoethanol, run on a 12% SDS-PAGE, and transferred to a polyvinylidene fluoride membrane (Immobilon P, Millipore, Billerica, MA). The polyvinylidene fluoride membrane was blocked in a 5% milk-Tween-TBS solution for 2 hours at room temperature followed by incubation with ExtrAvidin (Sigma Aldrich, St. Louis, MO) for 1 hour in TBS/Tween at 1:10,000. Following three 5-minute washes with Tween-TBS, the membrane was developed with NBT/BCIP.

Mouse Model of Amebiasis

The mouse model of cecal amebiasis was adapted from the model developed by Houpt's group (Houpt et al, 2002) and adapted for drug testing by Reed's group (He et al, 2010). Six week old male CBA/J mice (Jackson Laboratory, Bar Harbor, ME) were administered 0.2 mg/100 μls of dexamethasone (Sigma Aldrich, St Louis, MO)

intraperitoneally for four days before surgery. Prior to surgery, the mice were gavaged with 0.6 mg/100 µl lopermide-HCl to inhibit bowel motility. Mice were anesthetized through IP injections of 7ul/gram of mice body weight from a stock of 12 mg/ml ketamine and 0.5 mg/ml xylazine. 0.25% bupivicane, a local analgesic, was also administered to the wound site. The cecum of each mouse was injected with 2 million cells/200 µls of *E. histolytica* trophozoites. Test mice were administered 50 mg/kg/dose given twice a day intraperitoneally with WRR-666 cysteine proteinase inhibitor, while control mice were administered water for 7 days after surgery. Seven days after surgery, the cecum of each mouse was collected and assayed for the presence of myeloperoxidase (MPO), a measure of neutrophil inflammation) and *E. histolytica* DNA to determine the number of trophozoites present.

To assay for the presence of MPO, the cecum was resuspended in 500 µls PBS and homogenized as described in Chen et al, 2010. Homogenized cecum (100 mg) was resuspended in 500 µls 80 mM sodium phosphate buffer with 1% hexadecyltrimethylammonium bromide, 5 mM EDTA, and Complete One protease inhibitor mixture (Roche Applied Sciences, Indianapolis, IN), pH 5.4. This mixture was freeze-thawed 3 times by incubation for 5 minutes in dry ice, followed by incubation in water for 10 minutes and vortexing. The sample was centrifuged at 2000 rpm for 10 minutes. Five microliters of the sample supernatant was incubated with 15 µls buffer and 100 µls SureBlue Reserve TMB Microwell Peroxidase Substrate for 10 minutes at room temperature. After the 10-minute incubation, the reaction was stopped with 120 µls of 50% H_3PO_4 solution and read at an optical density of 450 nm. As a positive control, a

sample containing 5 units MPO were also assayed. The negative controls contained a sample without MPO, as well as a sample containing *E. histolytica* alone.

To determine the number of *E. histolytica* trophozoites in the cecal tissue, a quantitative PCR assay for *E. histolytica* DNA was performed. DNA from 100 mg of homogenized cecum was purified using a DNeasy Mini Stool Kit (Qiagen, Valencia, CA). The homogenized cecum was resuspended in 200 µls ASL buffer and vortexed for 10 seconds, followed by addition of 200 µls ATL buffer and incubation at 56°C overnight. After incubation, 200 µls of ASL buffer and half of an InhibitEx tablet were added to the digested cecum mixture, vortexed for 30 seconds, incubated for 1 minute, and centrifuged twice at 14,000 rpm for 5 minutes. Additional 200 µls of AL buffer was added to the supernatant and vortexed for 5 seconds, followed by washing with 200 µls 100% ethanol. This mixture was passed through a DNeasy spin column and washed with 500 µls buffer AW1 and AW2 at 8,000 rpm for 2 minutes. After drying the membrane through an additional 2 minute spin at 12,000 rpm, the DNA was eluted with 200 µls AE buffer. Each DNA sample was then quantitated using qPCR using SYBR Green Quantitative PCR in a Step One Plus real-time PCR machine (Applied Biosystems). The number of amebic trophozoites in cecal tissues was determined by comparison with a standard curve generated with the DNA extracted from trophozoites added to an uninfected control cecum (Melendez-Lopez et al, 2007).

Results

Expression of Histidine-Tagged, Recombinant EhCP1 in Bacterial Expression Systems

The ability to design more efficient inhibitors is reliant upon obtaining the crystal structure of the target molecule. The *E. histolytica* genome is very AT-rich (Loftus, 2005), resulting in lower expression in bacteria and yeast systems. In order to obtain soluble, active rEhCP1 for protein crystallization, a synthetic EhCP1 gene was first re-engineered for a yeast codon bias (Eton Biosciences) and cloned into the pET22b(+) plasmid to take advantage of the peIB leader sequence (Figure 1). This leader sequence potentially localizes recombinant protein to the periplasmic space, providing better conditions for protein folding and purification. Two rEhCP1 constructs, one with the leader sequence, one without the leader sequence, were expressed in three different bacterial cell lines: Origami, ArcticExpress, and BL21-CodonPlus. The Origami cell line provides mutations within the thioredoxin reductase and glutathione reductase genes, allowing for greater disulfide bond formation within the cytoplasm. ArcticExpress is able to grow at lower temperatures in order to slow down translation of the protein, allowing more protein to fold into the correct conformation. The ArcticExpress cells also contain recombinant cpn10/cpn60 chaperonin proteins that are active at low temperatures. Bl21-CodonPlus strain provides rare tRNAs for expression of heterologous recombinant proteins.

Clones from each of the two constructs were induced for 20 hours with 1 mM IPTG at 16°C (ArcticExpress) and 37°C (Origami, Bl21-CodonPlus). Induced and corresponding non-induced clones were resuspended in B-PER lysis reagent. Equal volumes of soluble and insoluble fractions from each transformant were analyzed on 12%

SDS-PAGE gels and stained with Coomassie blue. The induced rEhCP1 gene containing the peIB leader sequence in the Origami cell line presents an over-expressed band of approximately 30 kDa (indicated by arrows in Figure 2). This protein band may represent an incompletely processed form of rEhCP1, which has a predicted 37 kDa pro-rEhCP1-His and 27 kDa mature rEhCP1-His molecular weight. Recombinant protein expression in the other bacterial strains showed a similar expression profile (Table 1). Regardless of the presence of the peIB leader sequence, there was only expression of rEhCP1 in the insoluble pellet. Within the insoluble pellet fractions containing rEhCP1, the ability to obtain sufficiently high levels of pure and active rEhCP1 for protein crystallization with specific inhibitors was not achievable.

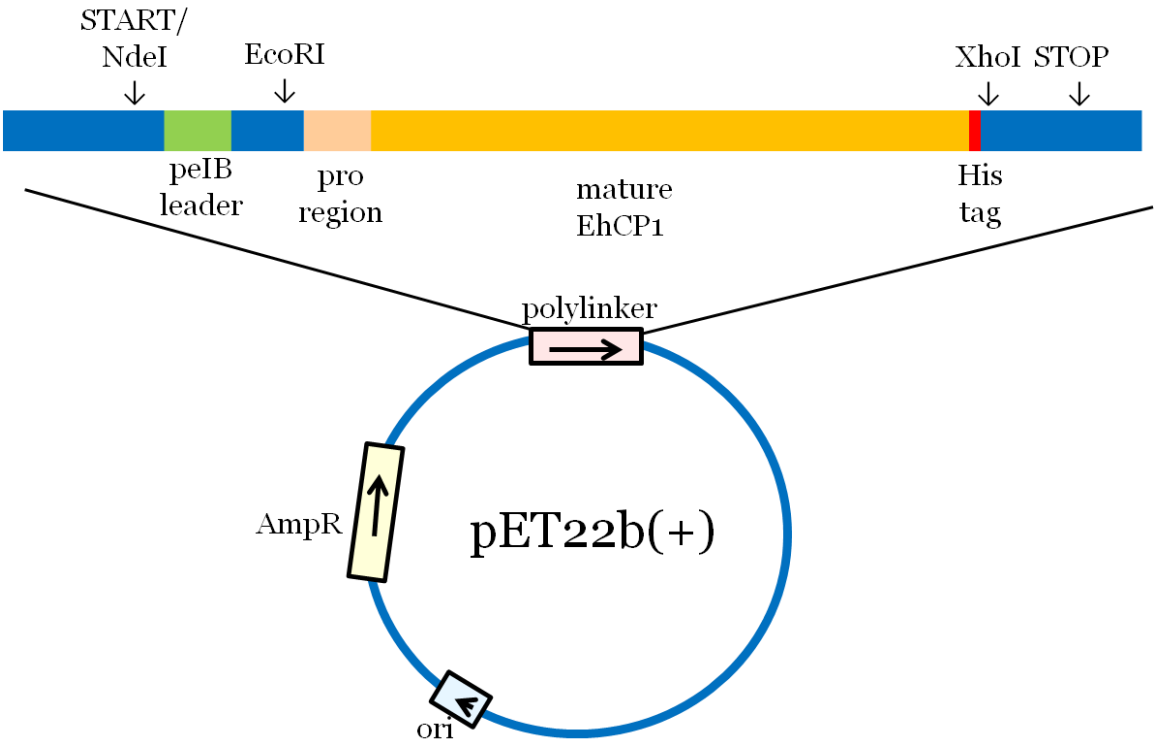

Figure 1: rEhCP1 Expression Construct
The proEhCP1 gene with a histidine tag was cloned into the pET22b(+) plasmid with or without the peIB leader sequence using a combination of NdeI, EcoRI and XhoI restriction sites.

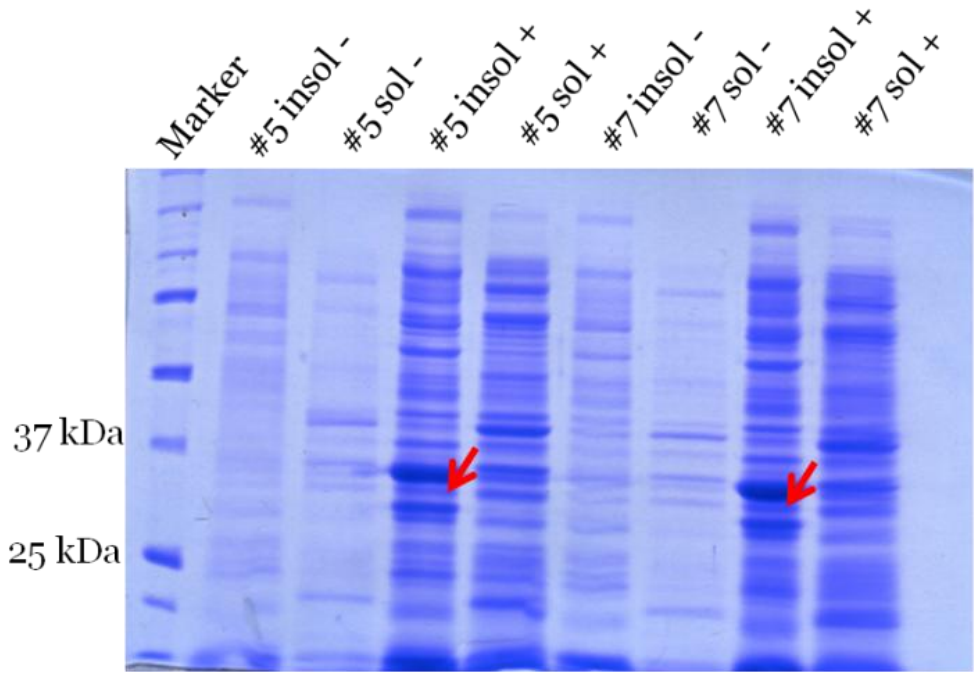

Figure 2: rEhCP1 Expression in Origami
Two Origami clones (#5, 7) containing recombinant peIB leader-rEhCP1-pET22b(+) plasmids were induced with 1 mM IPTG for 20 hours. Both soluble (sol) and insoluble (insol) protein fractions from induced (+) and uninduced (-) clones were analyzed by 12% SDS-PAGE gel analysis. The arrows indicate an approximate 30 kDa overexpressed band in the induced insoluble fraction from clones 5 and 7.

Table 1: Expression of rEhCP1 in Bacteria
Two rEhCP1 constructs, proEhCP1 gene incorporating the peIB leader sequence or proEhCP1 gene without peIB leader sequence, were cloned into the pET22b(+) plasmid and transformed into either Origami, ArcticExpress, or BL21Codon-Plus. After 20-hour induction, both soluble and insoluble protein fractions were analyzed on a 12% SDS-PAGE gel. + represents expressed rEhCP1 protein; - represents no expressed protein.

Strain	Plasmid Construct	Insoluble Protein	Soluble Protein
Origami	rEhCP1	+	-
	leader-rEhPC1	+	-
ArcticExpress	rEhCP1	+	-
	leader-rEhPC1	+	-
BL21Codon-Plus	rEhCP1	+	-
	leader-rEhPC1	+	-

To overcome previous expression problems, EhCP1 was then resynthesized for bacterial codon bias (GenOptics). The synthetic rEhCP1 was ligated into pET28b and expressed in ArcticExpress. After a 2 hour induction with 1 mM IPTG at room temperature, the induced cells were resuspended in B-PER lysis buffer. The soluble supernatant, potentially containing rEhCP1, was further purified through NiNTA chromatography. This purified rEhCP1 was analyzed in a Western blot against EhCP1 antibody (Figure 3), showing a prominent band around 24 kDa. This protein band reflects the approximate expected rEhCP1 size, indicating successful expression of rEhCP1. However, the soluble rEhCP1 did not show much proteolytic activity. Therefore, the soluble rEhCP1 will be tested under various refolding conditions to optimize proper folding and proteolytic activity. Once soluble rEhCP1 is properly refolded to increase proteolytic activity, expression can be scaled up to obtain larger quantities for protein crystallization.

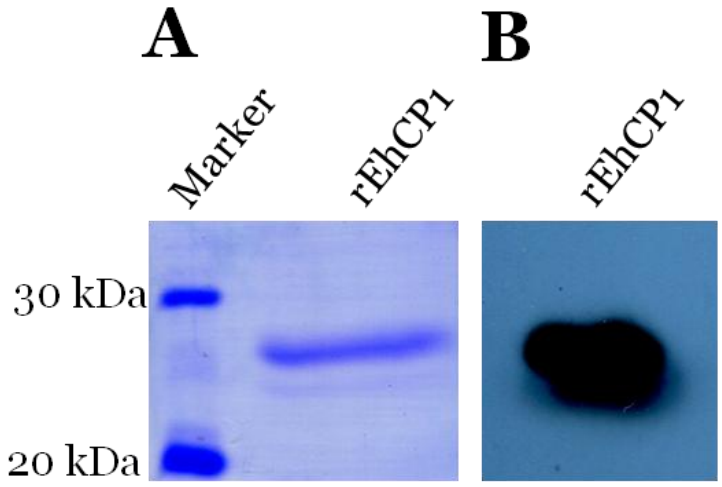

Figure 3: rEhCP1 Expression in ArcticExpress
The rProEhCP1 gene, re-engineered for bacterial expression, was cloned into pET28b and transformed into ArcticExpress. After a 2 hour induction at room temperature with 1 mM IPTG, soluble, nickel column-purified rEhCP1 was analyzed by Coomassie stain (A) and Western blot with EhCP1 polyclonal antibody (B). A band of the expected size for the mature EhCP1 is seen in the rEhCP1 lane.

Expression of Histidine-Tagged, Recombinant EhCP5 in Yeast

To ensure that any new inhibitors would also inhibit the other major released cysteine proteinase, EhCP5, a histidine-tagged EhCP5 for expression in *Pichia* was constructed. Histidine-tagged rEhCP5 was cloned into pPICZαB and expressed in *P. pastoris* strain X-33. The pPICZα vector encodes an α-secretion factor, enabling expressed recombinant proteins to be secreted into the culture supernatant (Figure 4). Seven clones were induced with methanol to a final concentration of 1.5%, replenished every 24 hours. After 72 hours, secreted proteins were collected and analyzed on a 12% SDS-PAGE gel. One clone (# 3) showed an overexpressed band of approximately 34 kDa, reflecting the pro-mature form of rEhCP5 (Figure 5). This 34 kDa protein band is absent from the induced vector control. MonoQ-purified secreted protein from clones 3-8

were then assayed for proteolytic activity against Z-arginine-arginine-AMC, a substrate for EhCP5. However, no proteolytic activity was observed in any of the clones.

Because the secreted recombinant protein did not possess proteolytic activity, the possibility that the recombinant protein was not properly processed for secretion was analyzed. The *Pichia* transformants were lysed and purified over a nickel column in case rEhCP5 was unable to be secreted. The purified intracellular proteins were analyzed on a 12% SDS-PAGE gel for the presence of recombinant EhCP5. Coomassie and western blot analysis did not indicate the presence of intracellular rEhCP5 (data not shown).

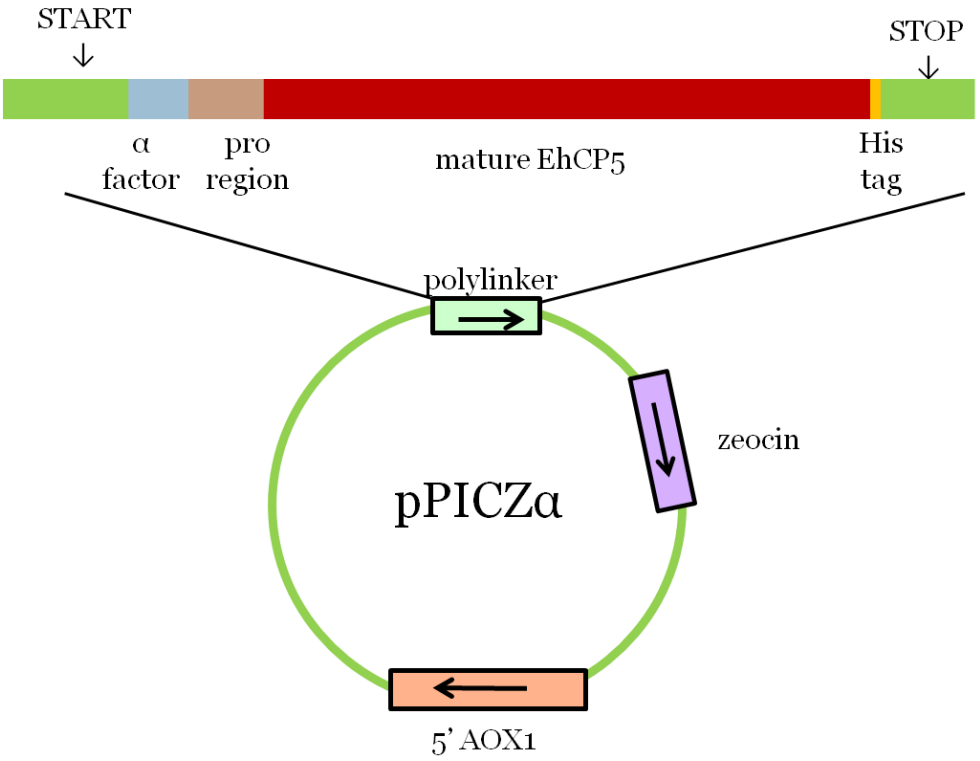

Figure 4: rEhCP5 Expression Construct
The proEhCP5 gene with a histidine tag was cloned into the pPICZα plasmid, utilizing zeocin as the selectable marker, 5' AOX1 promoter for methanol induction, and α factor for secretion of recombinant protein into culture media.

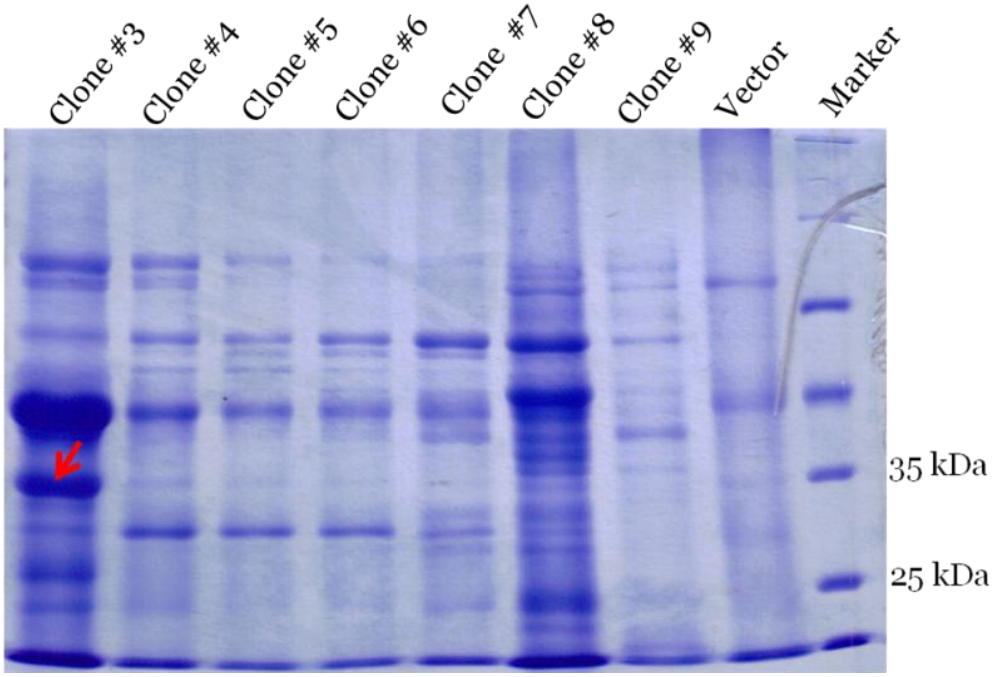

Figure 5: rEhCP5 Expression in pPICZα
Seven *Pichia* clones (#3-9) containing recombinant pPICZα-rEhCP5his and a *Pichia* control vector were induced with methanol to a final concentration of 1.5% for 72 hours. Secreted protein from each clone was collected, concentrated, dialyzed, purified through MonoQ chromatography, TCA precipitated, and analyzed on a 12% SDS-PAGE gel. The arrows indicate the over-expressed rEhCP5.

Expression of Recombinant EhCP5 without a Histidine Tag in Yeast

To avoid the possibility that the 6x-histidine tag may have interfered with rEhCP5 proteolytic activity, the EhCP5 was cloned into pPICZα without the histidine-tag and transformed into *Pichia.* Following induction and subsequent purification processes described above, purified rEhCP5 was shown to proteolytically cleave the synthetic substrate Z-RR-AMC, and the activity was inhibited with 10 μM cysteine protease inhibitor E-64 (Table 2). To further characterize the active rEhCP5, purified recombinant protein was incubated with DCG-04, a biotinylated cysteine protease inhibitor, and confirmed by Western blot (Figure 6).

Table 2: rEhCP5 Proteolytic Activity in *P. pastoris*
rEhCP5 was activated with 10 mM DTT, incubated with 10 μM E-64 for 10 minutes, and measured for proteolytic cleavage of the synthetic peptide Z-RR-AMC. Percent activity was measured against uninhibited rEhCP5 control.

Contents	Percent Activity
rEhCP5	100
rEhCP5 + E-64	1.21

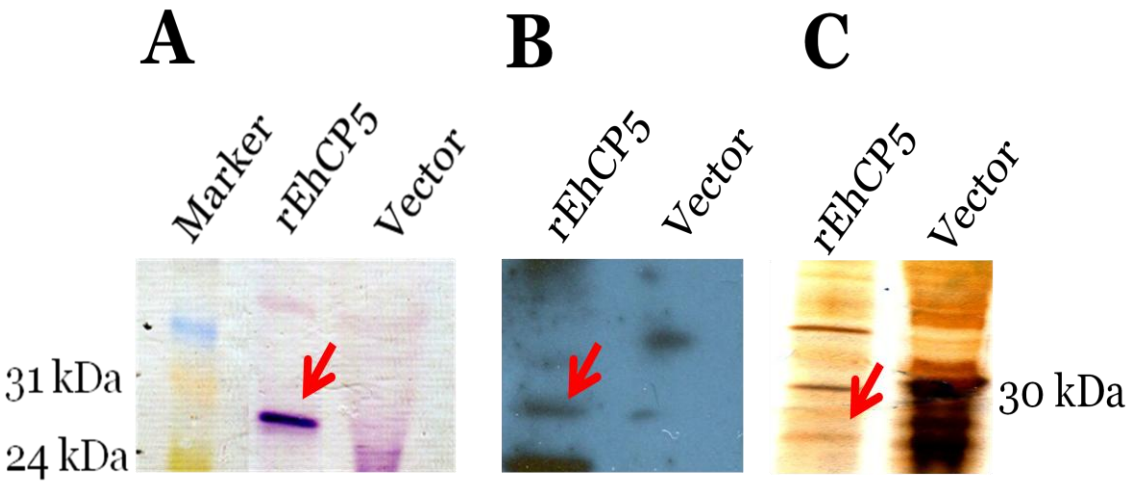

Figure 6: rEhCP5 Expression in *P. pastoris*
The rProEhCp5 gene was cloned into pPICZαA vector, transformed into *P. pastoris*, and induced with 1.5 % methanol. FPLC-purified rEhCP5 protein was analyzed by silver stain (C), or Western blot after a 30-minute incubation with 5 μM of DCG04, a biotin-labeled cysteine proteinase inhibitor (A), or with EhCP5 antibody (B). rEhCP5 lane: secreted protein from induced *P. pastoris* containing rEhCP5 gene; vector lane: secreted protein from induced *P. pastoris* with the pPICZαA vector alone. A band of the expected size is seen in the rEhCP5 lane, but not the vector control in both Western blots.

Inhibition of rEhCP5 with Vinyl Sulfone Inhibitors

We next tested the active, recombinant EhCP5 against vinyl sulfone inhibitors designed to inhibit EhCP1. Utilizing the scaffold of the cysteine proteinase inhibitor

K11777 (Figure 7B-C), which has shown efficacy in treating *T. cruzi* (Chen et al, 2010), derivatives were designed to optimize effectiveness.

To assay these synthesized derivatives, DTT-activated rEhCP5 was tested against twelve cysteine protease inhibitors by incubating for ten minutes with varying concentrations of inhibitor. Proteolytic activity was determined through the cleavage of the peptide substrate Z-RR-AMC in buffer (50 mM Tris, 2 mM EDTA, 2 mM DTT, pH 7.5). The IC50 values, the half maximal inhibition concentration used as a measure of drug effectiveness, were calculated by taking the average of the triplicates and dividing by the positive control (rEhCP5). An example of the IC50 curve for WRR-666 is shown in Figure 8. Similar curves were produced with all of the 12 tested inhibitors. Among the tested inhibitors, WRR-483 was the best inhibitor with the WRR-666 series, synthesized for optimal cell permeability, inhibiting at similar values. The IC50 for rEhCP5 was then compared against the IC50 for rEhCP1 (Unpublished data, Hwang and Reed, 2012) (Table 3). While the best inhibitor for rEhCP5 was WRR-483 (Figure 7D), the best inhibitor for rEhCP1 was WRR-554 with an IC50 of less than 6 nM. Comparing the individual IC50 values obtained for rEhCP1 and rEhCP5 indicates that they are very similar.

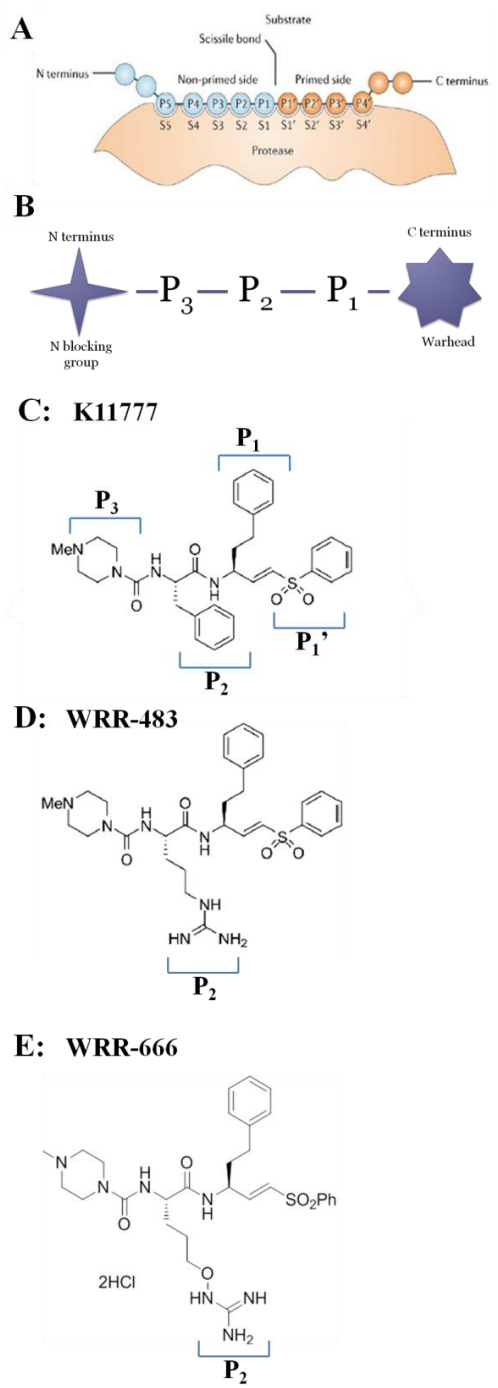

Figure 7: Structure of Specific Cysteine Proteinase Inhibitors
Cysteine proteinase inhibitors bind at specific sites within the active site pocket (A). The general structure of cysteine proteinase inhibitors include an N-blocking group, P_1-P_3 sites, and a C-terminus warhead group (B). All tested inhibitors were derived from the scaffold of the cysteine proteinase inhibitor K11777, whose P_1' - P_3 subsites are labeled (C). WRR-483 has an arginine in place of a phenylalanine at P_2 (D). WRR-666 introduces an oxyguanidine group within the arginine at P_2 (E).

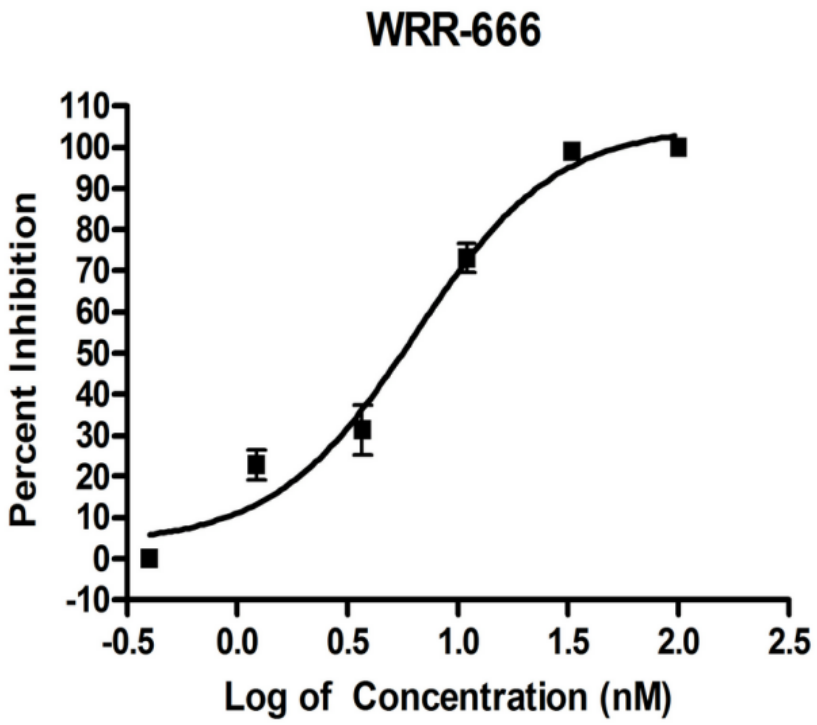

Figure 8: WRR-666 Inhibition of rEhCP5
rEhCP5 was activated with 10 mM DTT, incubated with different concentrations of WRR-666 for 10 minutes, and measured for proteolytic cleavage of the synthetic peptide Z-RR-AMC. The IC50 was calculated by taking the average of the triplicates and dividing by the positive control (rEhCP5).

Table 3: Inhibition of rEhCP5 with Irreversible, Cysteine Proteinase Inhibitors
rEhCP5 was activated with 10 mM DTT, incubated with increasing concentrations of cysteine proteinase inhibitors for 10 minutes, and measured for proteolytic cleavage of the synthetic peptide Z-RR-AMC.

Inhibitor	P_2 Residue	IC50 (nM)	
		EhCP5	EhCP1
WRR-290	Leucine	8.6×10^3	2×10^3
WRR-401	Leucine	$> 1 \times 10^6$	$> 1 \times 10^6$
WRR-483	Arginine	5.0	15
WRR-500	Leucine	3.4×10^3	3.8×10^3
WRR-518	Lysine	152.5	180
WRR-554	Arginine	16.8	< 6
WRR-605	Valine	$> 1 \times 10^6$	$> 1 \times 10^6$
WRR-666	Arginine mimic	6.4	26.0
WRR-667	Arginine mimic	11.9	42.6
WRR-668	Arginine mimic	9.7	7.7
WRR-669	Arginine mimic	78.5	261.7
WRR-670	Arginine mimic	129.3	157.7

In Vivo Inhibition of *E. histolytica* Trophozoites

Since cysteine proteinase inhibitor WRR-666, an inhibitor specifically designed to be cell permeable, was effective against both rEhCP1 and rEhCP5, WRR-666 was tested *in vivo* using our mouse cecal model of amebiasis. Six-week old male CBA/J mice were pretreated daily for four days prior to surgery with 0.2 milligrams of dexamethasone intraperitoneally. Surgically exposed cecums were inoculated with 1×10^6 *E. histolytica* trophozoites and treated with either 50 mg/kg/dose WRR-666 or water twice daily for 7 days intraperitoneally. Following treatment, the cecum was harvested and immediately

homogenized for assaying the levels of trophozoite burden and inflammation through qPCR and myeloperoxidase (MPO) respectively. The amount of trophozoite burden, as determined as the percent of the mean of the control, was found to be similar in both the WRR-666 treated and control mice (Figure 9A). The levels of cecal inflammation were also similar between the treated and control mice (Figure 9B). Similar results were also observed with WRR-668 treatment (data not shown).

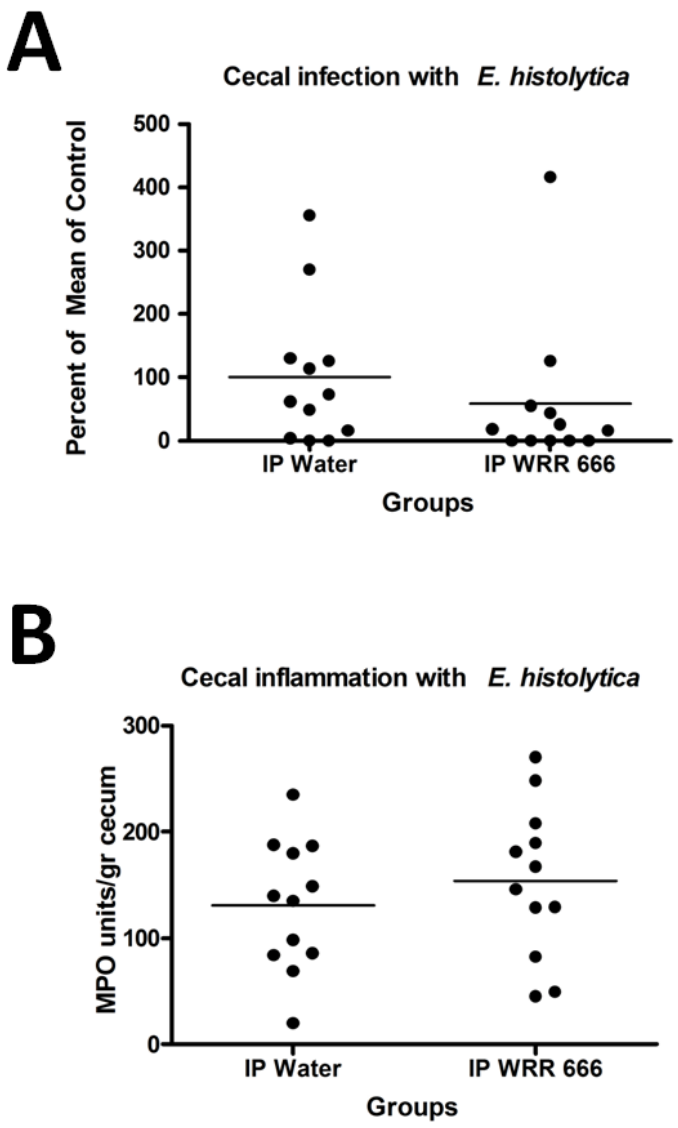

Figure 9: Cecal Infection and Inflammation with *E. histolytica*
CBA/J mice were infected with *E. histolytica* trophozoites and treated with either water or cysteine proteinase inhibitor WRR-666 or water. (A) *E. histolytica* trophozoite levels for the WRR-666 and control groups were measured by qPCR. (B) Cecal inflammation for the WRR-666-treated groups were measured by myeloperoxidase per gram of cecum tissue.

Discussion

Amebiasis is a major cause of water-borne infections worldwide. Currently, treatment relies on a 7-10 day course of an imidazole drug. While metronidazole is effective at treating amebiasis, drug resistance is a rising concern. Metronidazole-resistant *E. histolytica* have been induced in the laboratory (Bansal et al, 2006; Wassmann, 1999). For two other anaerobic protozoa, *Trichomoniasis vaginalis* and *Giardia lamblia,* also treated with metronidazole, clinical cases of drug resistance have been documented (Bansal et al, 2006). Therefore, new drugs based on defined targets are needed to treat amebiasis. The *E. histolytica* genome project has revealed 50 cysteine proteinase genes (Tillack et al, 2007) of which, EhCP1 and EhCP5 are unique to the invasive strain *E. histolytica* versus the noninvasive strain *E. dispar*. There is significant data supporting EhCP1 and EhCP5 as virulence factors (Hou et al, 2010; Melendez-Lopez et al, 2007; Moncada et al, 2005; Que et al, 2000), making these cysteine proteinases ideal targets for treating amebiasis. As a result, expression of active recombinant EhCP1 and EhCP5 is needed for *in vitro* assays evaluating the efficacy of new therapeutics.

Other cysteine proteinases have been targeted for treatment of parasitic diseases, including *Trypanosoma cruzi, Leishmania,* and *Schistosoma*. A new treatment for Chagas disease centers around inhibition of *T. cruzi* cysteine proteinase cruzain, which is vital for the survival of *T. cruzi* within the host (Choe et al, 2004). Currently, the vinyl sulfone inhibitor K11777 is in phase I clinical trials for the treatment of Chagas disease. Similarly, K11777 was also used as an effective inhibitor of a L*eishmania tropica* cysteine proteinase to arrest development (Mahmoudzadeh-Niknam et al, 2004).

K11777 has also been shown to inhibit *Schistosomiasis mansoni* cysteine proteinase cathepsin B1, achieving a parasitologic cure in a murine model (Abdulla et al, 2007).

As one of the three major secreted proteinases *in vitro*, EhCP1 is an important virulence factor for amebic invasion. When *E. histolytica* trophozoites are incubated with stomach porcine mucin, EhCP1 mRNA is upregulated 3-fold, indicating its potential importance in degrading mucin to aid for amebic invasion (Debnath et al, 2006). To circumvent the host immune response, EhCP1 cleaves pro-IL-18, IgG, and the third component of complement (Melendez-Lopez et al, 2007). EhCP1 has a unique specificity for arginine in the P_2 position, and a derivative of K11777, WRR-483, was designed to specifically target EhCP1. Not only was it 2500-fold more active against EhCP1 than K11777 *in vitro,* but it also significantly decreased amebic invasion in colonic xenografts using SCID mice (Melendez-Lopez et al, 2007).

In order to design optimal drugs, co-crystals of EhCP1 and potential inhibitors are needed. However, the crystallization of EhCP1 has been difficult, requiring a minimum of 1 to 5 milligrams of pure, active enzyme to obtain a crystal structure. Previous attempts in our lab to purify such large quantities of active rEhCP1 expressed in bacteria were unsuccessful. Because the *E. histolytica* genome is AT-rich (75%), the EhCP1 gene was re-synthesized to reflect the codon bias of *Saccharomyces* (62% AT) to express in *Pichia pastoris* and bacteria. To optimize expression of soluble proteinase in bacteria, this engineered EhCP1 gene was ligated into the pET-22b(+) plasmid either retaining or removing the peIB leader sequence, which targets the recombinant protein to the periplasmic space. This localization of the protein provides better conditions for soluble

protein purification. Three different strains of *E. coli* (Origami, ArcticExpress, BL21-CodonPlus) were utilized to find the best soluble expression.

Regardless of the presence or absence of the peIB leader sequence, or the *E. coli* cell line utilized, no expression of soluble rEhCP1 was observed with this construct (Figure 2, Table 1). There was limited expression of rEhCP1 within the insoluble fractions, as an over-expressed band of approximately 30 kDa was present. Although this 30 kDa band does not represent the predicted 37 kDa pro-rEhCP1-His or 27 kDa mature rEhCP1-His molecular weight, the over-expressed protein may represent an incompletely processed form of rEhCP1. The exact reasons for the lack of success in expressing soluble and active recombinant cysteine proteinase could be multifactorial, including, but not limited to, host cell toxicity and recombinant gene codon bias. This poor translation efficiency may be due to differences in codon bias between bacteria and yeast. These preferred bacterial codons are important because they specifically use codons matching the most abundant tRNA present for faster, more accurate translation (Hershberg et al, 2008). Therefore, this engineered EhCP1 specific for yeast codon bias may express EhCP1, but with slight errors that might cause problems with protein processing such as cleavage of the pro-sequence or folding. In order to optimize protein recovery, periplasmic solubilization, NiNTA column purification, and protein re-folding may be performed to recover active rEhCP1. However, current efforts to obtain active soluble protein for crystallization were unsuccessful.

To overcome poor rEhCP1 expression within bacteria, EhCP1 was again resynthesized to encode for a bacterial codon bias by McKerrow's group using GenScript Optimization Analysis (Sharp et al, 2005). Since the EhCP1 gene contains tandem rare

codons that can reduce translation efficiency and possibly interrupt the translational machinery, the GC content was adjusted from approximately 35 to 45% in order to prolong the half-life of mRNA. mRNA secondary structures were also altered to stabilize ribosomal binding. In addition, the codon adaptation index value (CAI), a measure of codon usage bias in reference to a highly expressed gene, was increased from 0.78 to 0.89. Unlike the previous bacterial construct containing the EhCP1 gene resynthesized for *Saccharomyces*, this new synthetic EhCP1 gene with a C-terminal histidine tag expressed in ArcticExpress produced soluble protein that was purified through NiNTA affinity chromatography in a much larger quantity with fewer bacterial background proteins (Figure 3). Despite obtaining soluble protein, the proteolytic activity did not appear to match the level of rEhCP1 protein seen on SDS-PAGE or Western blot. This inefficient level of active rEhCP1 could be due to incomplete cleavage of the pro-region, which may possibly block the active site pocket from binding potential substrates. Considering the ultimate goal of determining the crystal structure of EhCP1 and the requirement for active protein, the possibility that the recombinant protein was not correctly folded is a valid concern. In order to improve the proteolytic activity of the purified rEhCP1 and to facilitate the formation of active, mature proteinase, different activation and refolding conditions to optimize the yield will be tested. Upon determining the correct conditions, the rEhCP1 will undergo preliminary inhibition screening with different inhibitors and mass spectrophotometry to confirm correct expression, followed by crystallization experiments.

Along with EhCP1, EhCP5 is one of two specific cysteine proteinases which is actively expressed, translated, and secreted in the invasive species *E. histolytica* but not

in the noninvasive species *E. dispar*. Like EhCP1, EhCP5 is another important virulence factor. *In vitro*, EhCP5 degrades MUC2, a protein found within the colonic mucus layer, allowing the attachment of trophozoites to the epithelial cells. EhCP5 may also be involved in the processing of other proteinases; overexpression of EhCP5 by episomal transfection leads to the increased proteolytic activity of EhCP1, EhCP2, and EhCP5 (Tillack et al, 2006). In addition, EhCP5 plays a role in host immune system evasion by degrading pro-IL-18, which activates macrophages. The development of specific inhibitors which target EhCP5 could prove valuable in the treatment of amebic colitis or liver abscesses as well. Since EhCP5 comprises 90% of proteolytic activity along with EhCP1 and EhCP2, optimal inhibitors should be designed to inhibit all three of these proteinases.

Previously in our lab, rEhCP5 was successfully expressed within a yeast expression system. However, this construct did not incorporate an affinity purification tag. As a eukaryote, yeast have many advantages over bacteria, including protein processing, protein folding, and posttranslational modifications. Therefore, the *P. pastoris* expression system was used to express a histidine-tagged rEhCP5, instead of pursuing a bacterial expression system. Additionally, the inclusion of the α-secretion signal sequence allows the expressed, recombinant protein to be secreted into the culture supernatant. Because *P. pastoris* secretes very low levels of native proteins when grown under minimal growth medium, recombinant protein secretion provides the first-step of purification of protein (Invitrogen). The incorporation of the histidine tag provides a second purification step. Unfortunately, the expression and proteolytic activity of rEhCP5 was relatively low (Figure 5). There may be several possible reasons for the lack of

protein expression. Even though rEhCP5 should contain an α-secretion factor, allowing it to be secreted into the supernatant, the majority of rEhCP5 could have remained within the *Pichia* cells. To determine if rEhCP5 was expressed, but not secreted, the *Pichia* cells were lysed, and the intracellular proteins were purified over a NiNTA column. However, no rEhCP5 was detected by Western blot. It is possible that the histidine tag was not properly exposed, so that the rEhCP5 would not be purified over the nickel column. Although there seemed to be some promature rEhCP5 expression within the culture supernatant, there was no detectable proteolytic activity. Given the presence of an expressed secreted protein, but lacking in proteolytic activity, the assumption is the rEhCP5's functionality may be inhibited by improper protein folding, possibly the result of a 6x-histidine affinity tag (personal communication, Dr. James McKerrow). Although this occurrence is relatively rare, cases of histidine tags inducing conformational changes within the active sites of certain enzymes have been documented (Freydank et al, 2008).

Since the histidine-tagged rEhCP5 expression was not optimal, rEhCP5 without a histidine tag (previously expressed in our lab) was induced for inhibition assays. Utilizing this active, recombinant EhCP5, a panel of vinyl sulfone-based cysteine proteinase inhibitors were developed for efficacy testing. These vinyl sulfone inhibitors contain a warhead moiety that binds to the cysteine of the catalytic triad, forming a stable covalent bond. This Michael addition of the warhead at S_1' protein subsite is stabilized by other covalent interactions between the P_1, P_2, and P_3 inhibitor subsites and the rEhCP5 active pocket, creating an irreversible binding of the substrate. Inhibitor specificity derives from the P_2 residue, which binds to proteinase's S_2 subsite. Since the S_2 site for papain family proteinases is buried deeply within the proteinase's active pocket, the inhibitor's P_2

residue must be highly specific to fit and bind within the active pocket. In contrast, the inhibitor's P_1' and P_1 residues do not have to be as specific, since the proteinase's S_1' and S_1 are highly solvent exposed and serve as an anchor for stable bond formation. The addition of a blocking group at the inhibitor's P_3 subsite prevents digestion of these small molecules from other proteinases.

One such successful vinyl sulfone is K11777, an inhibitor of cruzain, a *T. cruzi* cysteine proteinase (Chen et al, 2010). This inhibitor is currently in phase I clinical trials and was used as the scaffold for the synthesis of new inhibitors targeting the *E. histolytica* cysteine proteinases. In order to make the K11777 inhibitor more specific to EhCP1, EhCP5, and other *E. histolytica* cysteine proteinases, the P_2 phenylalanine of K11777 was replaced with an arginine group to create WRR-483 (Figure 7).

The best inhibitor for rEhCP5 is also WRR-483 (Table 3), which mimics rEhCP1's preferred arginine substrate (Helberg et al, 2001). Because WRR-483 contains a positively charged arginine at P_2 subsite, it is attracted by the negatively charged S_2 aspartic group within the active pocket of EhCP1. Although WRR-483 is the best rEhCP5 inhibitor *in vitro* with an IC50 value of 5 nM, it has poor oral bioavailability because of the charged arginine group (Sun et al, 2011). In consideration of eventual use *in vivo*, attempts to synthesize a more readily cell permeable version of WRR-483 began by adding an oxyguanidine to the P_2 arginine group to form another series of inhibitors (WRR-666 – WRR-670) (personal communication, William R. Roush). Since oxyguanidine is more basic than guanidine, intestinal absorption is increased (Sun et al, 2011). This arginine mimetic is also a bioisostere of oxyguanidine, so it has the same physical properties as the original guanidine group. Therefore, as expected, WRR-666

and WRR-668 inhibit rEhCP5 well with an IC50 of 6.4 nM and 9.7 nM respectively, making both inhibitors potential *in vivo* candidates.

Because rEhCP5 has a glycine in the P_2 subsite, inhibition of rEhCP5 may be slightly different from rEhCP1, which contains an aspartic acid in the P_2 subsite (Brinen et al, 2000). Due to this difference in P_2 site, we hypothesized that EhCP5 has a broader substrate preference, including hydrophobic residues. Therefore, WRR-290, 401, and 500, which incorporate a leucine at P_2 were assayed against rEhCP5. However, all the vinyl sulfone inhibitors had similar inhibition of EhCP1 and EhCP5 with a preference for positively charged amino acids at the P_2 position. Therefore, the smaller glycine residue at the rEhCP5 P_2 site, appears to allow bulky residues, including positively charged arginine to fit within the active pocket.

To further test rEhCP5's substrate preference, WRR-605, another hydrophobic inhibitor containing valine at the P_2 subsite, was tested against both proteinases. WRR-605 was not an effective inhibitor with an IC50 value of greater than 1 mM for both enzymes. This lack of inhibition is most likely due to the valine present at the P_2 subsite. This is not surprising as WRR-605 was designed inhibit rEhCP4, which has a substrate preference for nonpolar phenylalanine over arginine at the P_2 site (He et al, 2010).

Since humans and apes are the only naturally infected hosts of *E. histolytica*, a simple *in vivo* rodent model is difficult. However, Houpt's group found that CBA/J mice deficient in toll-like receptor 4 can be infected with amoeba if the cecum is directly injected with trophozoites (Houpt et al, 2002). The use of steroids, which suppresses the immune system, and loperamide, which slows digestive motility, also enhances the rate of infection. As markers of infection, trophozoite burden and inflammation are measured.

The number of trophozoites is quantified by PCR using the peroxiredoxin gene as a marker, using previously extracted trophozoites added to mouse cecum as a standard curve. Inflammation is measured by quantification of myeloperoxidase, a protein released by neutrophils.

Since WRR-666 and WRR-668 are more readily cell-permeable than WRR-483, both inhibitors were tested *in vivo*. Even though WRR-666 and WRR-668 are efficient *in vitro*, they were not effective *in vivo* with no significant difference between control and infected mice (Figure 9). This lack of activity of both inhibitors *in vivo* could result from the cleavage of the weak N-O bond in the oxyguanidine group. This potential event was substantiated *in vitro* by synthesizing WRR-671, an inactive metabolite which was not a biologically active inhibitor of EhCP1 or EhCP5. When WRR-668 was injected intravenously in mice, it was cleaved into the inactive metabolite WRR-671 (unpublished data, Dr. William R. Roush).

Both EhCP1 and EhCP5 are important virulence factors for amebic invasion and are good drug targets for several reasons: 1) the mechanism of action of cysteine proteinase has been well-defined, 2) a number of inhibitors, including the scaffold compound, K11777 have undergone extensive safety and pharmacokinetic studies, 3) we and others have shown that the inhibition of EhCP1 and EhCP5 blocks invasive disease in animal models, and 4) cysteine proteinases with similar conserved structures are critical for survival of other parasites (Sajid, 2002). Although the resynthesized EhCP1 gene for yeast was not optimal for bacteria expression, a second resynthesized EhCP1 gene favoring a bacterial codon bias produced soluble protein that is currently being optimized under different refolding conditions to purify large quantities of active protein

for crystallization. This crystallization will provide insight into the mechanism behind proteinase-substrate binding. Previously, a crystallized K11777-cruzain complex revealed the interaction with K11777 P_3-P_1' subsites with the cruzain active pocket, allowing structural details for design of other inhibitors (Sajid et al, 2011). We have successfully expressed active rEhCP5 in *Pichia* and shown that key inhibitors can effectively target EhcP1 and EhCP5. We would anticipate that EhCP2 and EhCP3, which have an identical P_2 pocket to EhCP1, would be inhibited as well, explaining the efficacy of these inhibitors against all released native amebic cysteine proteinases (Grace Hwang, unpublished data). These inhibition studies with two cysteine proteinases implicated in the virulence of *E. histolytica* provide a better understanding of the challenge of obtaining inhibitors with good activity, stability, and bioavailability. The design of novel inhibitors may provide new effective therapeutics for treating amebiasis, as well as other protozoal infections in the future.

References

Abdulla, M. Abdulla, K., McKerrow, J.H., Caffrey, C.R. 2007. Schistosomiasis Mansoni: Novel Chemotherapy Using a Cysteine Protease Inhibitor. *PLOS Medicine*.4(1):e14. activity-dependent cysteine protease profiling and discovery tools. Chem. Biol. 7:569-581.

Ankri, S., Stolarsky, T., Bracha, R., Padilla-Vaca, F., and Mirelman, D. 1999. Antisense inhibition of expression of cysteine proteinases affects Entamoeba histolytica- induced formation of liver abscess in hamsters. *Infection and Immunity* 67:421-322.

Bansal, D., Malla, N., Majajan, R.C. 2006. Drug resistance in amoebiasis. *Indian J. Med Res*. 123(2): 115-118.

Cereghin, J.L, and J. Cregg. 1999. F Heterologous protein expression in methyltropic yeast *Pichia pastoris*. *EMS Microbiology Review*. 24.1: 45-66.

Chadee K, Petri WA Jr, Innes DJ, Ravdin JI. 1987. Rat and human colonic mucins bind to and inhibit adherence lectin of *Entamoeba histolytica*. *J Clin Invest.* 80(5):1245-54.

Chen YT, Brinen LS, Kerr ID, Hansell E, Doyle PS, McKerrow JH, Roush WR. *In vitro* and in vivo studies of thetrypanocidal properties of WRR-483 against Trypanosoma cruzi. PLOS Neglected Tropical Diseases. 20104(9):e825

Choe, Y., Brinen, L.S., Price, M.S., Engel, J.C., Lange, M., Grisostimi, C., Weston, S.G., Pallai, P.V., Cheng, H., Hardy, L.W., Hartsough, D.S., McMakin, M., Tilton, R.F., Baldino, C.M., Craik, C.S. 2005. Development of α-keto-based inhibitors of cruzain, a cysteine protease implicated in Chagas disease. Bioorg. Med. Chem. 13.6: 2141-2156.

Debnath A, Tashker JS, Sajid M, McKerrow JH. 2007. Transcriptional and secretory responses of *Entamoeba histolytica* to mucins, epithelial cells and bacteria. *Int J Parasitol.* 37(8-9):897-906.
Freydank, AC, Brandt W, Dräger B. 2008. Protein structure modeling indicates hexahistidine-tag interference with enzyme activity. *Proteins.* 72(1):173-183.

Gilchrist, C.A., Houpt, E., Trapaidze, N., Fei, Z., Crasta, O., Asgharpour, A., Evans, C., Martino-Catt, S., Baba, D.J., Stroup, S., Harrnano, S., Ehrenkaufer, G., Okada, M., Signh, U., Nozaki, T., Mann, B.J., Petri, W.A. 2006. Impact of intestinal colonization and invasion on the *Entamoeba histolytica* transcriptome. *Molecular and Biochemical Parasitology*. 147.2: 163-176,

Greenbaum, D., Medzihradszky, K.F., Burlingame, A., and Bogyo, M. 2000. Epoxide electrophiles as activity-dependent cysteine protease profiling and discovery tools. Chem Biol. 7(8): 569-581.

He, C., Nora, G.P., Schneider, E.L., Kerr, I.D., Hansell, E., Hirata, K., Gonzalez, D., Sajid, M., Boyed, S.E., Hruz, P., Cobo, E.R., Le, C., Liu, W., Eckmann, L., Dorrestein, P.C., Houpt, E.R., Brinen, L.S., Craik, C.S., Roush, W.R., McKerrow, J., and Reed,SL. 2010.A novel *Entamoeba histolytica* cysteine proteinase, EhCP4, is key for invasive amebiasis and a therapeutic target. *Journal of Biological Chemistry*. 285: 18156-185277.

Hellberg A, Nickel R, Lotter H, Tannich E, Bruchhaus I. 2001 Overexpression of cysteine proteinase 2 in *Entamoeba histolytica* or *Entamoeba dispar* increases amoeba-induced monolayer destruction *in vitro* but does not augment amoebic liver abscess formation in gerbils. *Cell Microbiol.* 3(1):13-20.

Hou, Y., Mortimer, L., and Chadee, K. 2010. *Entamoeba histolytica* cysteine proteinase 5 binds integrin on colonic cells and stimulates NFkB-mediated pro-inflammatory responses. *Journal of Biological Chemistry*. 285: 35497-35504.

Houpt ER, Glembocki DJ, Obrig TG, Moskaluk CA, Lockhart LA, Wright RL, Seaner RM, Keepers TR, Wilkins TD, Petri WA Jr.The mouse model of amebic colitis reveals mouse strain susceptibility to infection and exacerbation of disease by CD4+ T cells. *J Immunol.* 2002 Oct 15;169(8):4496-503

Houpt E, Barroso L, Lockhart L, Wright R, Cramer C, Lyerly D, Petri WA. 2004. Prevention of intestinal amebiasis by vaccination with the *Entamoeba histolytica* Gal/GalNac lectin.Vaccine. 22(5-6):611-7.

Hughes, M.A., and W.A. Petri. Amebic liver abscess. 2000. Infectious disease clinics of North America. 14.3: 565-582.

Huston, C.D. Parasite and host contributions to the pathogenesis of amebic colitis. 2004. *Trends in Parasitology.* 20.1: 23-26.

Loftus, B., Anderson, I., Davies, R., Alsmark, C.M., Samuleson, J. et al. (there are at least 50 authors!). 2005. The genome of the protest parasite *Entamoeba histolytica. Nature*. 433:865-8.

Mahmoudzadeh-Niknam, H., and J.H. McKerrow. 2004. *Leishmania tropica*: cysteine proteases are essential for growth and pathogenecity. 106.3: 158-163.

McKerrow, J.H., Doyle, P.S., Engel, J.C., Podust, L.M., Robertson, S.A., Ferreira, R., Saxton, T., Arkin, M., Kerr, I.D., Brinen, L.S., and Craik, C.S. 2009. Two approaches to discovering and developing new drugs for Chagas disease. *Mem Inst Osswaldo Cruz*. 104: 263-269.

Melendez-Lopez, S.G., Herdman, S., Hirata, K., Choi, M., Choe, Y., Craik, C., Caffrey, C.R., Hansell, E., Chavez-Muguia, B., Chen, Y.T., Roush, W.R., McKerrow, J., Eckmann, L., Guo, J., Stanley, S.L., and Reed, S.L. 2007. Use of recombinant

Entamoeba histolytica cysteine proteinase 1 to identify a potent inhibitor of amebic invasion in a human colonic model. Eukaryotic Cell. 6.7: 1130-1136.

Moncada, D., Keller, K., Ankri, S., Mirelman, D., and Chadee, K. 2006. Antisense inhibition of *Entamoeba histolytica* cysteine proteases inhibits colonic mucus degradation. Gastroenterology. 130.3: 721-730.

Mortimer, L., and K. Chadee. 2010. The immunopathogenesis of *Entamoeba histolytica*. Experimental Parasitology. *126*: 366-380.

Que, X., Kim, S., Sajid, M., Eckmann, L., Dinarello, C.A., McKerrow, J.H., and Reed, S.L. 2003. A surface amebic cysteine proteinase inactivates interleukin-18. Infection and Immunity. *71.3*: 1274-1280.

Que. X. and S.L. Reed. 2000. Cysteine proteinases and the pathogenesis of amebiasis. Clin Microbiol Rev. 13(2): 196-206.

Ralston, K.S., and W.A. Petri. 2011. Tissue destruction and invasion by *Entamoeba histolytica*. Trends in Parasitology. *27.6*: 254-263.

Reed, S.L., and Gigli, I. 1990. Lysis of complement-sensitive *Entamoeba histolytica* by activated terminal complement components. Initiation of complement activation by an extracellular neutral cysteine proteinase. *J.Clin.Invest.* 86:1815-1822.

Seydel, K.B., Li, E., Swanson, P.E., and Stanley, S.L. 1997. Human intestinal epithelial cells produce proinflammatory cytokines in response to infection in a SCID mouse-human intestinal xenograft model of amebiasis. *Infect.Immun.* 65:1631-1639.

Sun, J., Miller, J. M., Beig, A., Rozen, L., Amidon, G, and Gahan, L.P. 2011. A. Mechanistic enhancement of the intestinal absorption of drubs containing the polar guanidine functionality. *Expert Opinion Drug Metabol. Toxicol*. 7: 313-323.

Tillack, M., Biller, L., Irmer, H. Freitas, M., Gomes, MA, Tannich, E., and Bruchhaus, I. 2007. The *Entameoba histolytica* genome: primary structure and expression of proteolytic enzymes. *BMC Genomics*. 8: 170.

Tillack, M., Nowack, N., Lotter, H., Bracha, R., Mirelman, D., Tannich, E., Bracha, R. 2006. Increased expression of the major cysteine proteinases by stable episomal transfection underlines the important role of EhCP5 for the pathogenicity of Entamoeba histolytica. *Molecular & Biochemical Parasitology*. 149: 58-64.

Tran, V.Q., Herdman, D.S., Torian, B.E., and Reed, S.L. 1998. The neutral cysteine proteinase of *Entamoeba histolytica* degrades IgG and prevents its binding. *J.Infect.Dis.* 177:508-5

Turk, Boris. 2006. Targeting proteases: successes, failures and future prospects. *Nature Reviews*. 5: 785-797.

Wassmann, C., Hellberg, A., Tannich, E., and Bruchhaus, I. 1999. Metronidazole resistance in the protozoan parasite *Entamoeba histolytica* is associated with increased expression of iron-containing superoxide dismutase and peroxiredoxin and decreased expression of ferredoxin 1 and flavin reductase. *J. Biol. Chem.* 274:26051-26056.

Zhang, X., Zhang, Z., Alexander, D., Bracha, R., Mirelman, D., Stanley, S. L. 2004. Expression of amoebapores is required for full expression of Entamoeba histolytica virulence in amebic liver abscess but is not necessary for the induction of inflammation or tissue damage in amebic colitis. Infect. Immun. 72.2: 678-683.