Title
Trafficking and Biological Functions of Giardia Cysteine Proteases

Author
DuBois, Kelly N.

Publication Date
2007-09-13

Peer reviewed|Thesis/dissertation

TRAFFICKING AND BIOLOGICAL FUNCTIONS OF GIARDIA CYSTEINE PROTEASES

by

Kelly Nicole DuBois

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Biomedical Sciences

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

UMI Number: 3274658

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

Dedications

This work is dedicated to those individuals without whom it would not have been possible:

My parents, who always told me I could do whatever I put my mind to

My husband, who gave me unlimited love and support

Marla, a wonderful colleague and friend

Saj, whose teaching and mentoring any person is lucky to have

Jim, who got me interested in tropical diseases and made sure I always had the resources available to do the best work possible

Declarations and Acknowledgements

The text of this dissertation is in part (Chapter 1) based on material from a manuscript submitted to *Traffic* under the title "A Pluripotent Compartment Functions as ER in Endosome/Lysosome in *Giardia lamblia*."

The text of this dissertation is in part (Chapter 2) based on material from a manuscript submitted to the *Journal of Biological Chemistry* under the title "Identification of the Major Cysteine Protease of *Giardia lamblia* and its Role in Encystation."

The following individuals contributed to the data presented in this dissertation:

Marla Abodeely: contributed technical help to many of the experiments presented in chapter one, created the *Gl*CP2-GFP expression construct, contributed experimental ideas to chapter one, co-authored *Traffic* paper based on chapter one.

Mohammed Sajid: contributed experimental ideas throughout this dissertation

Adrian Hehl: contributed the data for Figures 1.7B-1.7D.

Ivy Hsieh: contributed the data for Figures 1.9A and 1.9B.

Richard Fetter: contributed the data for Figures 1.9C, 1.10B, and 3.7.

Juan C. Engel: contributed the data for Figure 1.10A.

Katalin Medzihradsky: contributed the data for Table 2.1.

Christopher Franklin: created parts of Figures I.2 and 1.11.

TRAFFICKING AND BIOLOGICAL FUNCTIONS OF *GIARDIA* CYSTEINE PROTEASES

Kelly N. DuBois

Abstract

Giardia lamblia is a eukaryotic protozoan parasite and the causative agent of *giardiasis*, a debilitating enteric disease resulting in much morbidity and mortality worldwide. It is of interest not only as a target for the development of improved *giardiasis* therapies, but also as a model eukaryotic system. *Giardia* represents the earliest branching clade of eukaryotic cells. It is thus an ideal system for investigating the evolution of cell processes, organelle compartmentalization, and critical protein families. Analysis of the structure and function of the *Giardia* endomembrane system, cysteine proteases, and clathrin orthologues are the focus of this dissertation project.

The ER has been studied recently as a putative endocytic organelle. Examination of *Giardia* endocytosis using fluorophore-labeled proteins revealed that proteins were rapidly trafficked to a tubulovesicular network with ER-like properties. Using reporter constructs, cysteine proteases that are orthologues of lysosomal hydrolases were localized to the same tubulovesicular network. Functional protease assays helped define the role that cysteine proteases play in the degradation of exogenous proteins. Organelle-specific markers were used to describe the tubulovesicular compartment in which endocytosis and subsequent proteolysis takes place.

Cysteine proteases have been implicated in life cycle transitions (encystation

and excystation) of *Giardia*. The completion of the *Giardia* genome indicated that there are twenty-seven cysteine protease genes in *Giardia*. Cysteine protease 2 (*Gl*CP2) was identified as the major expressed cysteine protease gene in *Giardia*. Biochemical analysis of heterologously expressed *Gl*CP2 suggested that this gene indeed plays an important role in both encystation and excystation.

Giardia clathrin is also key to the processes of endocytosis and encystation. Clathrin is associated with the peripheral vacuoles of vegetative *Giardia* and may facilitate the inital endocytic uptake of proteins. During encystation, clathrin localizes to encystation specific vesicles and may function in cyst formation. A dominant negative clathrin heavy chain disrupted cyst formation but did not affect endocytosis.

Table of Contents

Introduction **1**

I.1. *Giardia lamblia* as a pathogen and simple eukaryotic cell 1

I.2. *Giardiasis* 4

I.3. Cysteine proteases in protozoa 6

I.4. Figures and tables for introduction 9

Chapter One: Cysteine proteases process endocytosed proteins in the ER of *Giardia*, which acts as an endocytic compartment. **12**

1.1. Introduction 12

1.1.1. Endocytosis 12

1.1.2. The ER as a pluripotent compartment 13

1.1.3. Endocytosis in *Giardia* 13

1.1.4. Protein hydrolases as markers of an endocytic compartment 15

1.2. Results 16

1.2.1. Proteins and Fluospheres are endocytosed by *Giardia* trophozoites and rapidly traverse a TVN that extends to the perinuclear region. 16

1.2.2. Cathepsin B-like proteases co-localize with and degrade endocytosed proteins in the TVN. 17

1.2.3. *Giardia* cysteine proteases are not secreted upon contact with intestinal epithelial cells. 18

1.2.4. *Giardia* cysteine proteases are not required for trophozoite survival. 18

1.2.5. The *Giardia* TVN contains active cysteine proteases and ER markers. 19

1.2.6. Ultrastructural analysis confirms that the TVN has properties of ER and

identifies foci of ER fusion with PVs. 20

1.3. Discussion 21

1.4. Figures and tables for chapter one 28

Chapter 2: Identification of the major cysteine protease of *Giardia lamblia* and its role in encystation. 45

2.1. Introduction 45

2.1.1. Clan CA cysteine proteases 45

2.1.2. Activation and catalytic mechanism 46

2.1.3. Protease substrate specificity 47

2.1.4. Cysteine proteases in *Giardia* encystation and excystation 48

2.2. Results 49

2.2.1. There are twenty-seven Clan CA cysteine proteases in the genome of *Giardia lamblia*. 49

2.2.2. *Gl*CP2 is the most highly expressed cysteine protease of the twenty-seven in the *Giardia* genome. 50

2.2.3. *Gl*CP2 is identified in lysate fractions enriched for cysteine protease activity. 51

2.2.4. *Giardia* cysteine proteases exhibit difficulty in recombinant expression. 52

2.2.5. Expression and characterization of recombinant *Gl*CP2. 52

2.2.6. The activity profile of recombinant r*Gl*CP2 and the dominant cysteine protease activity found in total *Giardia* lysates are identical. 53

2.2.7. *Gl*CP2 is found in *Giardia* ESVs and can proteolytically process CWP2 to the predicted size found in the cyst wall. 54

2.2.8. *Gl*CP2 is localized to the periphery of cysts during in vitro excystation. 55

2.3. Discussion 55

2.4. Figures and tables for chapter two 59

Chapter 3: Clathrin may facilitate the initial uptake of endocytosed proteins and play a role in ESV maturation in *Giardia*. 74

3.1. Introduction 74

3.1.1. Clathrin and cargo transport 74

3.1.2. Selectivity of cargo recruitment 76

3.2. Results 77

3.2.1. Clathrin localizes to the PVs in vegetative trophozoites and this localization is distinct from that of other *Giardia* coat proteins. 77

3.2.2. Clathrin remains PV associated during endocytosis. 78

3.2.3. Clathrin hub expression did not affect clathrin localization or endocytosis in vegetative trophozoites. 79

3.2.4. Clathrin hub expression did not significantly affect clathrin localization or cyst formation in encysting *Giardia*. 79

3.2.5. Electron microscopy shows invaginations of the plasma membrane, but the characteristic clathrin coat is not seen on the forming vesicles. 80

3.3. Discussion 80

3.4. Figures and tables for chapter three 85

Chapter 4: General discussion and future directions 97

4.1. Cysteine proteases and the endocytic system in vegetative *Giardia* 97

4.2. Cysteine proteases and encystation 99

4.3. Cysteine proteases and excystation 101

4.4. *Giardia* clathrin 103

4.5. General Conclusions 104

Chapter 5: Materials and methods **105**

5.18. Figures and tables for chapter five 117

Chapter 6: Abbreviations and References **119**

6.1. Abbreviations used 119

6.2. References 119

List of Figures and Tables

Figure I.1: Structure of a *Giardia* trophozoite. 10

Figure I.2: The life cycle of *Giardia lamblia*. 11

Figure 1.1: The endocytic pathway. 29

Figure 1.2: Proteins were rapidly endocytosed by trophozoites into the TVN and the perinuclear region. 30

Figure 1.3: *Giardia* orthologues of mammalian lysosomal proteases are present in the endocytic TVN where they co-localized with endocytosed proteins. 31

Figure 1.4: *Giardia* cysteine proteases degraded endocytosed proteins and were optimally active at a higher pH than their mammalian orthologues. 33

Figure 1.5: Cysteine proteases are not secreted by *Giardia* upon contact with intestinal epithelial cells in culture. 35

Figure 1.6: Cysteine protease activity is not essential for trophozoite survival. 36

Figure 1.7: *Gl*CP2-GFP and *Gl*CP protease activity co-localized with the ER marker PDI2 in the perinuclear region and TVN. 37

Figure 1.8: Cathepsin-like cysteine proteases are not significantly concentrated in the acidified clathrin-rich PVs of *Giardia* trophozoites. 39

Figure 1.9: Ultrastructural analysis revealed that gold particles were endocytosed into the TVN. 41

Figure 1.10: Identification of sites of fusion between the TVN and PVs. 43

Figure 1.11: Model of endocytic network of *Giardia* trophozoites. 44

Figure 2.1: The *Giardia* genome contains twenty-seven genes encoding clan CA cysteine

proteases, of which *Gl*CP2 is the most highly expressed as measured by quantitative RT-PCR. 60

Figure 2.2: Two distinct cysteine protease activities were resolved by anion exchange chromatography of *Giardia* lysates. 61

Figure 2.3: *Giardia* CP1 (*Gl*CP1) was heterologously expressed in *P. pastoris*. 63

Figure 2.4: Resynthesized *Gl*CP2 (r*Gl*CP2) was heterologously expressed, purified, and biochemically characterized. 64

Figure 2.5: r*Gl*CP2 was fractionated by anion exchange chromatography and its activity was found to be identical to cysteine protease activities from *Giardia* lysates. 66

Figure 2.6: *Gl*CP2 and other clan CA cysteine proteases co-localized with CWPs in ESVs during *Giardia* encystation. 68

Figure 2.7: r*Gl*CP2 can accomplish the proteolytic processing of recombinant CWP2 (rCWP2) to the predicted 26kDa size shown to be necessary for incorporation into the cyst wall. 69

Figure 2.8: *Gl*CP2-GFP was found around the periphery of cysts following *in vitro* excystation. 70

Table 2.1: Amino acid sequences of peptide fragments identified by LC-MS/MS of cysteine protease activity Peak A and Peak B eluted from anion exchange chromatography of *Giardia* lysates. 71

Table 2.2: *Giardia* cysteine proteases presented resistance to recombinant expression. 72

Table 2.3: Inhibition of r*Gl*CP2 activity against the N-terminally blocked fluorogenic peptide substrate Z-FR-AMC. 73

Figure 3.1: Structure of the clathrin triskelion. 86

Figure 3.2: *Giardia* clathrin heavy chain (*Gl*CLH) is localized to PVs and this localization is distinct from other coat proteins. 87

Figure 3.3: *Gl*CLH remains localized to the PVs during endocytosis, but does not resemble the expression pattern of *H. sapiens* CHC22. 88

Figure 3.4: Both *Gl*CLH-GFP and *Gl*CLHhub-GFP co-localize with an antibody against the consensus sequence of clathrin light chain (from many model organisms). 90

Figure 3.5: *Gl*CLHhub-GFP expression did not disrupt endocytosis of proteins or lipids in *Giardia* trophozoites. 91

Figure 3.6: *Gl*CLHhub-GFP expression did not disrupt ESV formation or clathrin localization during early encystation, but caused irregularities in cyst formation. 93

Figure 3.7: Ultrastructural analysis showed membrane pits forming very near PVs and clathrin localized in and around PVs. 95

Table 5.1: GenBank accession numbers and primer sets used for quantitative RT-PCR of the clan CA cysteine proteases of *Giardia lamblia*. 118

TRAFFICKING AND BIOLOGICAL FUNCTIONS OF *GIARDIA* CYSTEINE PROTEASES

Introduction

I.1. *Giardia lamblia* as a pathogen and simple eukaryotic cell

Giardia was first described in 1681 by Anton van Leeuwenhoek, a Dutch dry-goods dealer who ground the first microscope lens (1). He examined his own diarrheal stool under a microscope saying that "his watery excrements do contain much more little animals (and different species) than a normal solid stool" (2). Lambl described the organism in greater detail in 1859 and therefore the species of *Giardia* infectious to humans is named for him (*G. lamblia*) (3). The organism that both of these scientists described is a protozoan parasite that inhabits the upper small intestine of many vertebrate hosts and is the most commonly isolated intestinal parasite world wide (3). The binucleate trophozoite is a flattened pear shaped cell roughly 12 μm long with a convex dorsal surface and a concave ventral surface. The bi-lobed adhesive disc by which the cell adheres to surfaces comprises much of the ventral side of the cell and is formed by a single layer of microtubules. One nucleus is located above each lobe of the disc (4). Four pairs of flagella emerge from the cell body from anterior, posterior lateral, caudal, and ventral positions (5) (Figure I.1). With these flagella the trophozoite produces locomotion, often in a twisting or coiled path. The median body, another structural

element of *Giardia* trophozoites, is crescent-shaped and is made up of microtubules.

Giardia has a simple two-stage life cycle that includes a replicating trophozoite stage and, under certain environmental conditions, an infectious environmentally resistant cyst (Figure I.2). The cyst form is taken up by a host and following passage through the acidic stomach into the alkaline duodenum, the trophozoite form excysts and resides in the upper small intestine of its vertebrate host, causing the disease *giardiasis*. Trophozoites asexually reproduce by binary fission and opportunistically scavenge nutrients by uncharacterized endocytic pathways. In response to environmental cues, the trophozoite begins the process of encystation and commences encystation-specific gene expression (6,7). Numerous media compositions are sufficient to trigger encystation *in vitro,* though the efficiency of producing viable cysts varies greatly among the medias (8-12). These stimuli include alkaline pH, increased bile salts, and cholesterol starvation. It has been suggested that the combination of alkaline pH and increased bile salts actually limits cholesterol uptake and therefore acts indirectly to cause encystation by way of cholesterol starvation (8). Encystation-induced gene expression mainly consists of genes encoding for components of the cyst wall, largely the cyst wall proteins (CWPs). Following translation, these proteins are concentrated into encystation-specific vesicles (ESVs) that traffic the protein to the plasma membrane, where it is assembled into the cyst wall. The cyst wall is a fibrillar extracellular matrix composed of forty three percent carbohydrates (13,14) and three known CWPs (7,15,16). Upon completion of the cyst wall, the trophozoite divides once and the quadrinucleated cyst is excreted from the host with the host stool. The cyst wall allows the replicative trophozoite to persist under conditions of desiccation outside the host and in harsh chemical environments like the

stomach after uptake by a host.

G. *lamblia* is of interest to study as a model system because it represents the most early branching clade of eukaryotic cells (17,18). Historically, the term "model organism" described organisms of small size and short generation times that could be studied in a laboratory to gain insight into general principles that underlying specific disciplines. However, with the proliferation of genome-sequencing projects, many organisms are called "model organisms" and are studied for specific traits that may answer a specific research question (17). As one of the earliest eukaryotic cells, analysis of the cell biology and biochemistry of *Giardia* may illuminate the mechanisms of early evolutionary development of the clade eukaryota and specifically the development of organelle complexity and function seen in higher order eukaryotic cells. In addition to the DNA, rRNA, and protein sequences that have been used to determine the phylogenic position of *Giardia*, this organism has many traits that support its early evolutionary position. First, *Giardia* is a facultative anaerobe, lacking true mitochondria or any of the components of oxidative phosphorylation (3). *Giardia* metabolism is fermentative, but it can live in a low-oxygen environment for limited amounts of time and does have the ability to reduce oxygen (19). This is by the action of an NADH oxidase of mixed function that can reduce oxygen to water. The only previously identified NADH oxidases with this capacity were isolated from bacteria species (19). Finally, *Giardia* also lacks the typical eukaryotic pyruvate dehydrogenase. Instead, the bacterial enzyme pyruvate: ferredoxin oxidoreductase reduces pyruvate (3,19). The identification of "mitosomes" in *Giardia*, double membrane-bound organelles that function in iron-sulfur protein maturation and possibly ATP production, has led some to question the postulation that

Giardia is primitively amitochondrial, suggesting that *Giardia* is instead reductively evolved to its current primitive position (20). This is a matter of ongoing debate.

The primitive endomembrane structure of *Giardia* also suggests that it is an evolutionarily ancient organism. *Giardia* lacks peroxisomes, hydrogenosomes, true mitochondria, or a true Golgi apparatus, all of which are characteristic of eukaryotic cells (3,5). *Giardia* does contain two nuclei, a glycogen-rich cytoplasm, acidified peripheral vacuoles (PVs) and a tubulovesicular network (TVN), segments of which are decorated with ribosomes reminiscent of rough endoplasmic reticulum (ER) (Figure I.1) (16,21,22). Previous studies have revealed that the two diploid nuclei of *Giardia* trophozoites are equal in size, contain equal amounts of DNA, and are both transcriptionally active (3,23). The two nuclei are partitioned equally during mitosis and the left and right nuclei are inherited with mirror image symmetry in the two daughter cells (21,24). The cytoplasm of *Giardia* is densely packed with two types of granules, ribosomes and glycogen deposits. Ribosomes are likely associated with the extensive ER elements found throughout the cytoplasm. The PVs are vacuoles encompassing the periphery of the cell. They are acidified as evidenced by the fact that they contain of acid phosphatase activity and take up of acridine orange. PVs may have some function in endocytosis (25-27). The TVN, which is the tubular structure recognized to house functions of the ER, has extensions that protrude into the cytoplasm of the cell and even contact the PVs at the cell periphery (16). The structure and function of the endomembrane system of *Giardia* remains poorly characterized and its analysis is one of the goals of this thesis.

I.2. *Giardiasis*

Giardia lamblia causes the enteric disease known as *giardiasis*. As previously stated, the infection is initiated by ingestion of the cyst form of the parasite. The infectious dose is as low as ten cysts making it a highly successful and widespread parasite, found in every climate and country (28). There are three major modes of transmission; waterborne, fecal-oral, and foodborne. Waterborne transmission is responsible for the majority of cases of *giardiasis* worldwide, being especially prevalent in developing countries lacking access to clean water. The incubation period for people with symptomatic infection is usually one to two weeks but can vary from one to forty-five days depending on the individual. In about sixty percent of the population the infection remains asymptomatic and asymptomatic infection is more common in children and people with prior infections (29). Symptoms of *giardiasis* are highly varied, which often compounds the diagnosis of the disease. Common symptoms include diarrhea, fatty stool, abdominal cramps, bloating, anorexia, malaise, and weight loss (29,30). The epidemiology of the disease causing these symptoms has not been completely elucidated. Theories range from direct damage by trophozoites to the intestinal epithelium, inflammation of the intestinal epithelium, or passive interference of nutrient and fluid uptake due to large numbers of trophozoites coating the villi of endothelial cells in the upper small intestine (31). No *Giardia* toxin similar to that found in many pathogenic enteric bacteria has been identified (30). Regardless of the exact cause of the disease pathogenesis, *giardiasis* is a debilitating disease that causes much mortality and morbidity, particularly in developing countries. The drug of choice for treating *Giardia* infections is quinacrine hydrochloride. Side effects of this drug include dizziness, headache, and vomiting. Metronidazole and furazolidone also may

be used. Metronidizole is highly effective, but the development of nausea is a frequent barrier to use (29). Tinidazole, another nitroimidazole is widely used throughout the world but it has not been officially approved in the United States for treating *giardiasis* due to uncertainties about its carcinogenic and mutagenic properties. Furazolidone is less effective against *Giardia* than the other two drugs and its availability in liquid form makes it useful for treating young children. None of these drugs can cure all *Giardia* infections, and none is particularly well tolerated (32). Additionally, the emergence of drug-resistant strains is a topic of concern.

I.3. Cysteine proteases in protozoa

Cysteine proteases are one of the four major types of eukaryotic proteases (serine, cysteine, aspartyl, metallo). In higher eukaryotes cysteine proteases function predominantly, although not exclusively, as lysosomal hydrolases. As such, they are useful markers of the final components in the endosome-lysosome pathway. Serine proteases dominate in higher eukaryotes as the most abundant protease gene products. However, in invertebrates cysteine proteases are dominant and likely to play unique roles in these organisms. The function of the primitive cathepsin B-like proteases of *Giardia* is a major focus of this dissertation. For a discussion of the classification of cysteine proteases, see Chapter Two.

Clan CA (papain) cysteine proteases in *Giardia lamblia* remain largely uncharacterized. This class of enzymes has been found to be important in the processes of encystation and excystation and therefore may be indispensable for the life cycle of *Giardia* (37,38). While these proteases are of interest as putative drug targets, they are

also of biochemical and evolutionary interest as the most early diverging eukaryotic clan CA cysteine proteases. Examination of *Giardia* cysteine proteases may provide insight into the molecular evolution of this protein family that is of prime importance especially in invertebrate eukaryotic cells. This thesis reports an analysis of the cysteine proteases in *Giardia lamblia*, specifically focusing on the trafficking pattern, location, and the activity of this enzyme class. The ultimate goal of this study is to describe the putative biological functions of cysteine proteases in *Giardia*.

The emergence of drug resistant parasitic diseases and the global burden of animal and human parasitic diseases raise the need for new chemotherapies to target these infections. The development of well-tolerated and less costly drugs for the treatment of parasitic diseases has been the subject of many investigations. From both *in vitro* and *in vivo* studies, cysteine proteases have emerged as a protein family of highly promising drug targets in parasites. Aiding in the efficacy of the drug development against this protein family is the fact that these enzymes are highly redundant in the mammalian system, while the cysteine proteases of parasites generally have more specific and important roles. In contrast to the redundant catabolic roles of most of these enzymes in vertebrates, cysteine proteases perform essential roles in parasite biology such as growth, differentiation, cell-signalling, host invasion, and immune evasion (33-35). Therefore, blocking the function of these proteases in many instances selectively incapacitates the parasite. For example, the hemoglobin-degrading cysteine proteases in *P. falciparum* have been successfully targeted *in vitro* and in a mouse model (34,36). Success has also been seen with cysteine protease inhibitors against the trypanosomatids *T. cruzi, T. brucei, and Leishmania*. Inhibitors bound to cruzipain (cruzain), the major cysteine

protease of *T. cruzi*, cause an accumulation of the proform of the enzyme in the Golgi apparatus, resulting in parasite cell death (34). Cysteine protease inhibitors kill *T. brucei* in culture and inhibit the growth and replication of *L. major* and *L. mexicana* (36). This class of enzymes is especially of interest because homologues of cysteine proteases are present in many pathogenic parasitic organisms and therefore drugs developed to treat one parasite may also be effective against other parasites (33).

I.4. Figures and tables for Introduction

Figure I.1: Structure of a *Giardia* trophozoite. 8

Figure I.2: The life cycle of *Giardia lamblia*. 9

Figure I.1: Structure of a *Giardia* trophozoite.

The binucleate trophozoite is a flattened pear shaped cell with a convex dorsal surface and a concave ventral surface. The bi-lobed adhesive disc by which the cell adheres to surfaces comprises much of the ventral side of the cel. One nucleus is located above each lobe of the disc. Four pairs of flagella emerge from the cell body from anterior, posterior lateral, caudal, and ventral positions. The median body is a crescent shaped structural element composed of microtubules. Peripheral vacuoles (PVs) are large acidified compartments that are thought to play a role in the endocytic system. Variant surface proteins (VSPs) are constitutively produced and are continuously trafficked to the surface of the cell and recycled. These may play a role in immune evasion. Encystation-specfic vesicles (ESVs) carry cyst wall proteins (CWPs) and cyst wall components to the cell surface to form the cyst wall during encystation. The ER network/ tubulovesicular network (TVN) is an extensive branched network, studded with ribosomes in some areas, that appears to house functions of anabolism and catabolism.

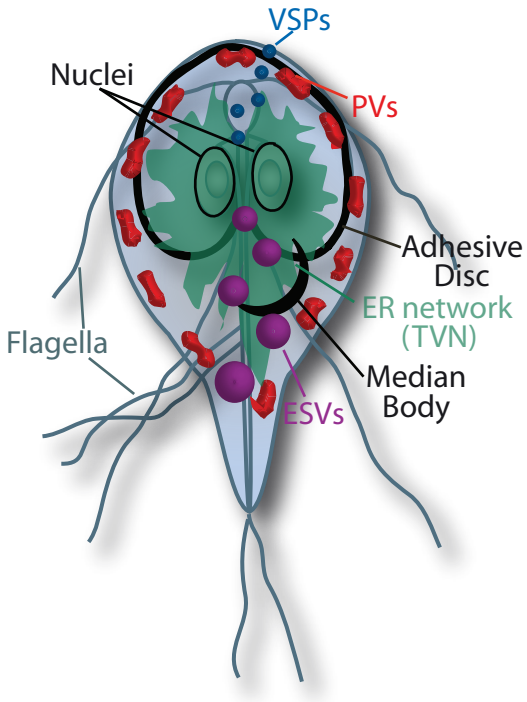
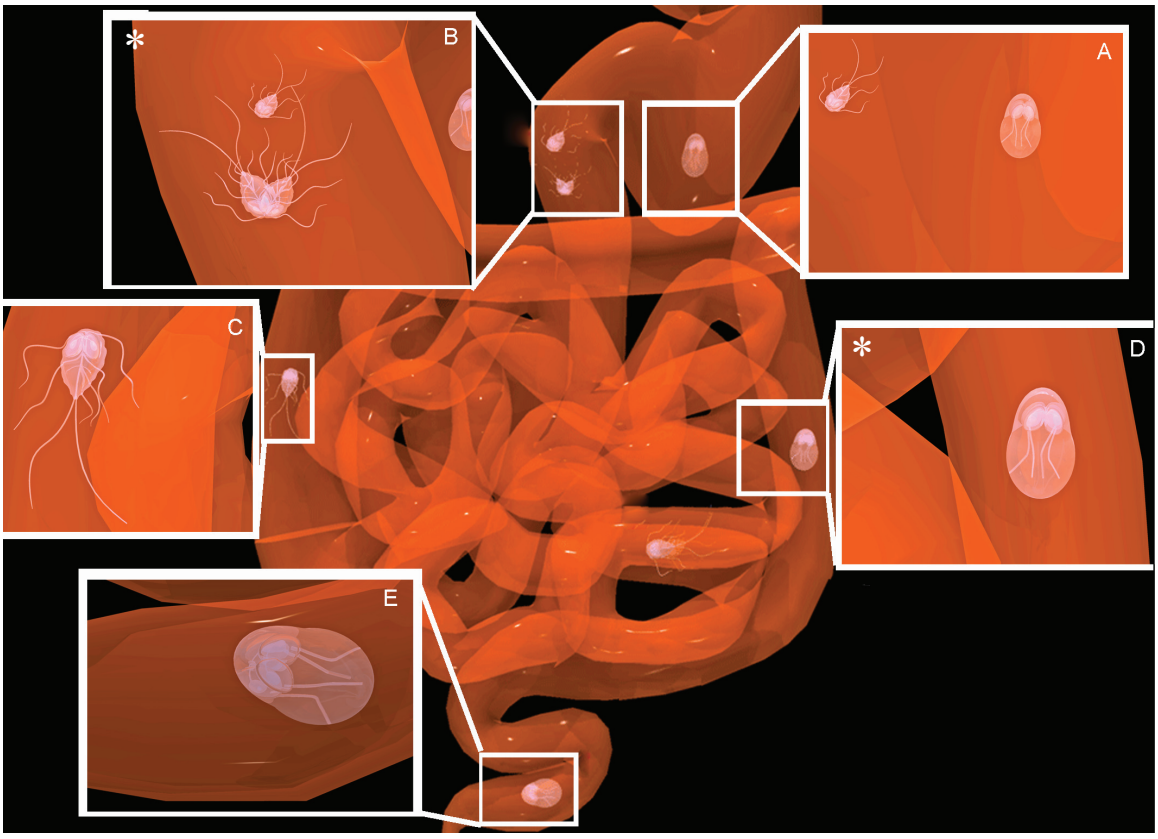


Figure I.2: The life cycle of *Giardia lamblia*.

(A) A host ingests a *Giardia* cyst with contaminated food or water. (B) In the host duodenum the cyst undergoes excystation and two vegetative trophozoites emerge and begin to asexually reproduce by binary fission. (C) Trophozoites attach to intestinal epithelia by a ventral adhesive disc. Flagella produce locomotion in unattached cells to remain in the small intestine. (D) As cells move toward the large intestine, they begin to encyst. These quadrinucleate cysts travel into the large intestine. (E) Cysts are passed out of the host with the feces and can be ingested by another host to perpetuate the cycle of infection. (*) Cysteine proteases have been shown to be necessary for both excystation (B) and encystation (D).



Chapter One: Cysteine proteases process endocytosed proteins in the endocytic ER of *Giardia*.

1.1. Introduction

1.1.1. Endocytosis

Endocytosis is a time-, temperature-, and energy-dependent process by which eukaryotic cells internalize extracellular fluids and particles as well as plasma membrane proteins by invagination of the plasma membrane. In fluid-phase endocytosis, the internalization occurs without a binding step. In contrast, receptor-mediated endocytosis occurs when a ligand is internalize with the membrane receptor it is bound to (39-42). Endocytosis begins with the formation of an endocytic vesicle, often clathrin-associated, that is internalized and fuses with the early endosome (43,44). Also called the "sorting endosomes", these compartments begin to separate cargos bound for recycling from those bound for degradation (45-47). Receptors to be recycled to the plasma membrane are rapidly removed from the early endosome (48). Cargo is then transported along microtubules either by vesicle transport or early endosome maturation into a lower density, more acidified compartment called the late endosome (48). Late endosomes are more spherical in shape than early endosomes and often have the appearance of multivesicular bodies (49). The delivery of endosomal contents from late endosomes to lysosomes is thought to occur by fusion of these endosomes with pre-existing lysosomes. In lysosomes, the cell accomplishes one of the most well-characterized functions of the endocytic system—endocytosed material is catabolyzed by hydrolases to provide the

nutrients needed for cell sustenance and growth (41). These hydrolases, including the cathepsin class of proteases, are delivered to the lysosomes by a mannose-6-phosphate modification to allow for packaging into mannose-6-phosphate containing vesicles that fuse with lysosomes. Early endosomes, late endosomes, and lysosomes can all exchange cargo with the trans-Golgi network and some protein degradation can occur in each of these compartments (Figure 1.1) (41,50,51).

1.1.2. The ER as a pluripotent compartment

The ER has been studied recently as an additional endosomal/phagosomal cell compartment. Gagnon et al. (2002) recently proposed that the ER is involved in direct uptake of material from the extracellular environment via fusion with the plasma membrane (52). This hypothesis was based on the presence of ER markers at the initial stage of phagosome formation in mammalian macrophages. Touret et al. (2005), however, found no evidence for direct ER to plasma membrane communication in either macrophages or dendritic cells (53). The concept of pluripotent functions for the ER was left unresolved, but this activity has been demonstrated to be important for antigen processing for major histocompatability complex (MHC) class I presentation. In addition, there are also intriguing examples of exogenous toxins and viruses entering mammalian cells via the ER (54,55). These studies underscore the potential for an ER function in phagocytosis or endocytosis.

1.1.3. Endocytosis in *Giardia*

The process by which *Giardia lamblia* endocytoses proteins and scavenges

nutrients remains poorly understood. *Giardia* is found in the upper region of the small intestine, including the jejunum and duodenum. The preference of *Giardia* for this area suggests that trophozoites require a high concentration of nutrients, especially those that the parasite is unable to synthesize. *Giardia*, like most eukaryotic cells, requires exogenous lipids for proliferation. Trophozoites are unable to synthesize cholesterol de novo, though trophozoites can generate isoprenoid compounds from mevalonate, a precursor of cholesterol in eukaryotes (56). To multiply and colonize the midjejunum of mammals, trophozoites might depend on pre-formed biliary lipids present in the upper small intestine. These nutrients and others are endocytosed by trophozoites into their unique endomembrane system. The vegetative trophozoite form of *Giardia* lacks aspects of the compartmentalization and diversity seen in the endocytic pathways of higher eukaryotic cells. Previous work has elucidated parts of the endocytic sytem in *Giardia*. For example, when examining the trafficking of variant surface proteins (VSPs), Mccaffery et al. found that these proteins localized to ER and cytoplasmic membrane cisternae and also to PVs during vegetative growth, suggesting a role for the PVs as endocytic compartments (57). This role for PVs was further supported upon the discovery of soluble PV-resident enzymes such as acid phosphatase and encystation-specific cysteine protease (ESCP) (16,38). The group of Adrian Hehl also found that proteins from the surface of *Giardia* are internalized into the PVs (University of Zurich, personal communication). However, the *Giardia* PVs were never formally shown to act as a lysosomal compartment and the continuous nature of the *Giardia* endomembrane system has suggested an unconventional mechanism for uptake and degradation of protein (5,16,58).

1.1.4. Protein hydrolases as markers of an endocytic compartment

Protein hydrolases are an important component of any endocytic degradative system. Acid hydrolases are enzymes that are concentrated in lysosomes and are active under acidic conditions. The lysosome is maintained at an acidic pH by a proton ATPase in the membrane that pumps protons into the lysosome (1). An important class of hydrolases that process proteins in the lysosome is the cathepsin family of cysteine endopeptidases (59-61). Four of them, cathepsins B, C, H and L, are ubiquitous in lysosomes of animals (62). Cathepsins L and B are, together with the aspartic protease cathepsin D, the most abundant lysosomal proteases, their lysosomal concentration being as high as 1 mM (60). These enzymes are therefore an important component of the endocytic system of eukaryotic cells.

Giardia contains a multigene family of cysteine endopeptidases that are orthologous to cathepsin L and cathepsin B of higher organisms and are therefore useful markers of cell compartments where protein degradation takes place (37,63) (*Giardia* genome project, www.mbl.edu/*Giardia*). Due to the limited availability of established ER/endocytic subcompartment markers and the lack of classical genetic techniques, the discrete endocytic pathway of *Giardia* has not been fully characterized. Despite reports of preliminary efforts, RNAi targeted gene disruption has not been established as a reliable and consistent genetic approach (64).

To circumvent these limitations I used reporter gene constructs and fluorophore-labeled proteins to examine the localization of cysteine proteases and endocytosed proteins in *Giardia*. Functional protease cytochemistry was used to define the role that cysteine proteases play in the degradation of exogenous

proteins. Coupling these techniques with organelle-specific markers, we defined the compartment in which endocytosis and subsequent proteolysis takes place.

1.2. Results

1.2.1. Proteins and Fluospheres are endocytosed by *Giardia* trophozoites and rapidly traverse the TVN that extends to the perinuclear region.

By confocal microscopy, the intracellular localization of fluorescein-labeled albumin and casein was first ascertained following a 30 minute incubation of labeled proteins with *Giardia* in culture (Figure 1.2.A, left and middle panels). Live cell imaging verified that uptake is extremely rapid; cells were saturated with fluorescent biotin-conjugated Fluospheres (small inert fluorophore-conjugated particles of a specific size)within 40 sec following initial uptake (Figure 1.2.B). During live cell imaging, the initial Fluosphere signal was observed near the site of flagellar attachment in the region of the ventral groove (Figure 1.2.B) (65). This was followed by rapid distribution throughout the labyrinthine TVN with accumulation in the perinuclear region. Diffusion of the fluorochrome following protein degradation did not occur, as no difference in fluorescence localization was observed upon pretreatment with protease inhibitors (data not shown). This is in contrast to the diffuse cytosolic, non-compartmental localization seen upon expression of a GFP construct lacking a signal peptide (Figure 1.2.A, right panel). Treatment of trophozoites with sodium azide, cytochalasin D or incubation at 4°C inhibited protein uptake (Figure 1.2.C, data not shown).

1.2.2. Cathepsin B-like proteases co-localize with and degrade endocytosed proteins in the TVN.

Giardia contains a number of genes coding for clan CA cathepsin B-like cysteine proteases of which *Gl*CP1, *Gl*CP2 and *Gl*CP3 are representative members (63). C-terminal fusion GFP reporter constructs indicated that each of *Giardia* cathepsin B-like proteases were present in both the peripheral TVN as well as the perinuclear region and co-localized with endocytosed rhodamine-conjugated albumin and DQRed BSA (Figures 1.3.A, 1.3.B, and 1.3.C). To ensure that the GFP constructs did not mislocalize due to misfolding or due to the presence of the GFP reporter protein, cathepsin activity, identified by *in situ* cleavage of the MNA-derivatized peptide substrate, was localized to the same region of the cell (Figure 1.3.C). *In situ* MNA peptide cleavage, indicated by coupled fluorescent activity, was completely inhibited by treatment with the membrane permeable cysteine protease inhibitors E64d and K11777 (data not shown). This activity correlated with the localization and degradation of endocytosed proteins. An *in vitro* assay demonstrated that *Giardia* lysates could degrade fluorophore-labeled casein and BSA. This degradation was completely inhibited by cysteine protease inhibitors (Figure 1.4.A). A functional *in vivo* assay confirmed that endocytosed casein was degraded and the known clan CA cysteine protease inhibitors E64d, K11777 and WRR477 inhibited this degradation (Figure 1.4.B). Replacing casein-FITC with BSA-FITC in this assay yielded similar results (Figure 1.4.C). Consistent with their location and activity in the ER-like TVN, the pH optimum of the *Giardia* cathepsin activity was 6-7, notably higher than the mammalian cathepsin orthologues, which function in acidified lysosomes and have an acidic pH optimum of 5-6 (Figure 1.4.D) (61).

1.2.3. *Giardia* cysteine proteases are not secreted upon contact with intestinal epithelial cells.

There have been some reports of epithelial damage by infection with *Giardia* (31). To determine if cysteine proteases could also be acting outside of the cells during a *Giardia* infection, *Giardia* trophozoites were added to a monolayer of T84 intestinal epithelial cells. Media was collected from this culture as well as cultures of *Giardia* cells alone and T84 cells alone. Conditioned media was tested for cysteine protease activity against Z-FR-AMC and also probed with an irreversible active site probe, ^{125}I-DCG04, and fractionated by SDS-PAGE. There was no statistically significant increase in cysteine protease activity against the fluorogenic peptide substrate in the media from the co-culture of *Giardia* trophozoites with intestinal epithelial cells. The results were additive (Figure 1.5.A). The active site affinity probe also did not identify any unique *Giardia*-specific extracellular cysteine protease activity in the co-culture of *Giardia* trophozoites and intestinal epithelial cells (Figure 1.5.B).

1.2.4. *Giardia* cysteine proteases are not required for trophozoite survival.

Giardia trophozoites were grown in various concentrations of the cysteine protease inhibitors E64d or K11777. Growth was not affected by inhibitor concentrations up to 50 μM (Figure 1.6.A). Because cysteine proteases degrade endocytosed protein it was hypothesized that the rich vegetative growth media allowed trophozoites to survive in the absence of cysteine protease activity as they are in an abundance of amino acids and small peptides. A minimal growth media was prepared without yeast extract, tryptone, and with 5% FBS, the minimal percentage required for trophozoite survival over the

course of the experiment (K.N. DuBois, unpublished data). The cysteine protease inhibitors E64d and K11777 were added to trophozoite growing in this media at various concentrations. The survival of trophozoites in the presence of the inhibitors did not differ from those without cysteine protease inhibitors (Figure 1.6.B).

1.2.5. The *Giardia* TVN contains both active cysteine proteases and ER markers.

The TVN has ER characteristics as demonstrated by localization of the characterized *Giardia* ER marker, protein disulfide isomerase 2 (PDI2, shown in red) (Figure 1.7) (6). Antibodies against HSC70, an ER-resident protein, and the ER-retention signal (KDEL) also localized to the TVN and particularly to the perinuclear compartment (Figure 1.7.A). PDI2 co-localized with a cathepsin B-like protease, as shown by the *Gl*CP2-GFP reporter construct (green) (Figure 1.7.B, 1.7.C). Confocal images were analyzed and co-localization was quantified (Figure 1.7.D). The majority of GlCP2-GFP co-localized with PDI2. Some GlCP2-GFP-positive compartments lacking PDI2 staining were visible and may represent subcompartments of the TVN. PDI2 also co-localized with other cathepsin B-like proteases expressed episomally in *Giardia*, such as the protease represented by GenBank accession number EAA37074 (data not shown). Similarly, cysteine protease activity, highlighted by cleavage of the MNA peptide substrate, was found in a PDI2-positive compartment (Figure 1.7.E).

GFP-labeled cathepsins and cysteine protease activity, while predominantly in the TVN, did appear peripheral to ER markers in the cell in some optical cuts and we therefore examined if the peripheral cathepsins localized to the PVs. The acidified PVs were previously thought to represent a lysosome-like compartment where endocytosed

proteins are degraded (16). The PVs were identified by accumulation of Lucifer yellow (Figure 1.8.A) (16), and unexpectedly contained abundant clathrin, which co-localized with Lucifer yellow (Figure 1.8.B). A concentration of PVs was observed near the site of flagellar attachment in the region of the ventral groove corresponding to the site of initial Fluosphere uptake observed in Figure 3B (Figure 1.8.A, 1.8.C). However, PVs were clearly distinct from the labyrinthine TVN and the *Gl*CP2-GFP (Figure 1.8.C). To ensure that *Gl*CP2 was not present in the PVs and just lacking fluorescence due to pH effect, *Gl*CP2-GFP expressing cells were probed with an anti-GFP antibody (Figure 1.8.D). No signal from the anti-GFP antibody was seen in the PVs and the anti-GFP antibody co-localized with the GFP signal.

1.2.6. Ultrastructural analysis confirms that the TVN has properties of ER and identifies foci of ER fusion with PVs.

Sequential confocal images of the TVN along the Z-axis and ultrastructural analyses were both consistent with a complex branching tubular network (Figures 1.8.C, 1.9.A, 1.9.B, and 1.10.B). This morphology was consistent with previous reports of the *Giardia* ER and was confirmed by co-localization of ER markers. The ER network of *Giardia* appears slit-like in ultrastructural planes as previously reported and is highlighted by the glucose-6-phosphatase (G6P) reaction (16,66)(Figure 1.9.A). While the PVs are differentially labeled with acid phosphatase, G6P and acid phosphatase activities overlap, especially at focal points of ER-PV fusion (16,66). Ultrasmall (≤5 nm) albumin-conjugated gold particles were endocytosed by *Giardia* and distributed throughout the TVN, including the perinuclear region (Figure 1.9.B), consistent with results of confocal

microscopy. Larger (10 nm) gold particles entered the PVs but progressed no further (Figure 1.9.C). While at first surprising, this result may reflect both the increased charge and relative mass of the larger gold particle. Three-dimensional reconstruction of G6P distribution provides a detailed view of the tubular network of ER and confirms focal fusion between ER and PVs (Figure 1.10.A) (16). Figure 1.10.B also provides evidence for a slit-like TVN extending to and contacting the PVs at the periphery of the cell (elongated arrow head). These fusiform clefts were first recognized by Friend (1966) and were noted to have ribosomes at their margins (4).

1.3. Discussion

The ER was first identified and characterized in mammalian cells by Keith Porter (67). Accordingly, our current concepts about the function and compartmentalization of the ER derive primarily from studies with mammalian cells and yeast. The ER is a labyrinthine tubular endomembrane system that co-evolved with the nucleus and serves as a conduit for newly translated proteins, a subset of which will be sorted by the Golgi apparatus to other membrane-bounded organelles such as lysosomes and secretory granules (68-70). *Giardia lamblia*, one of the earliest branches of the eukaryotic tree (18,71), lacks several functioning organelles of higher eukaryotes and, as an early evolutionary form, can provide fundamental insights into the evolution of cell compartmentalization. The ER of *Giardia* was first identified through localization of BIP, an HSP70 homolog that functions as a protein-folding chaperone (70). This helped to distinguish the ER from the PVs of *Giardia*, which were thought to be secretory granules, endosomes or lysosomes (3). Based on the observations presented in this study,

it appears that the *Giardia* ER is a tubulovesicular endomembranous system (the TVN) in which the disparate functions of ER, endosomes and lysosomes of higher eukaryotes have not fully diverged.

By video microscopy, endocytosed proteins and Fluospheres primarily appear first in a region near the flagellar attachment site (65). This region appears to be abundant in clathrin-containing vacuoles (Figures 1.8.A, 1.8.C, white arrowheads). In other protozoan parasites, particularly the kinetoplastids, secretion and endocytosis are polarized, occurring only at a region known as the flagellar pocket (72). Therefore, it is possible that *Giardia* most efficiently (though not exclusively) takes up extracellular material at a similar region where the action of flagellar movement may optimize sampling of the environment. The observation by electron microscopy that 10 nm gold particles only enter PVs (Figure 1.9.C) suggests that the clathrin-rich peripheral vesicle network may be the initial site of uptake followed by rapid dynamic fusion with the TVN leading to the transfer of endocytosed material. This model is consistent with both previous ultrastructural observations that showed fusion of PVs with the plasma membrane (16) and the tomographic electron micrographs that show focal fusion of TVN with PVs in vegetative trophozoites (Figure 1.10). Both studies suggest that direct delivery of endocytosed material from PVs to the TVN could occur. A recently documented pathway of ceramide uptake also supports this model of *Giardia* endocytosis (73). While we cannot rule out the possibility of a clathrin coated vesicle population trafficking endocytosed material between the PVs and the TVN, we saw no direct evidence of this.

It has been proposed that the ER is an extremely dynamic organelle (69,74); its

size and shape can undergo drastic changes to meet the needs of ER-related functions. The focal fusions noted between TVN and PVs in vegetative trophozoites and the appearance of the PVs as either spherical organelles or a fused network, might both be reflections of dynamic changes occurring within the endomembranous system of *Giardia* (summarized in Figure 1.11). Furthermore, rather than serving as lysosomes at the end of a typical mammalian cell endocytic pathway, the acidified clathrin-rich PVs may either denature recently endocytosed proteins, facilitating degradation by proteases in the "downstream" TVN; or neutralize the alkaline secretions of the pancreas present in the mammalian duodenum where *Giardia* trophozoites reside and replicate. A previous report that cathepsin B-like cysteine protease activity resides in PVs (37) was made before the availability of ER and PV markers. Based on current observations, the "vesicles" observed by Ward et al., (1997) (37) are likely TVN in cross section or compartmentalized TVN with active protease restricted to discrete locations in the organelle.

Circumstantial support for this proposed pathway of protein uptake by *Giardia* comes from a study of the uptake of *Giardia lamblia* virus (GLV), a double stranded RNA virus that is first translocated to the PVs after binding to the plasma membrane. GLV entry is arrested in the PVs of sodium azide-treated cells. Like some mammalian cell viruses, GLV appears to exploit the endocytic pathway of *Giardia* (75). More direct support for the proposed endocytic pathway in *Giardia* comes from a recent publication in which the acquisition of ceramide is initiated in peripheral clathrin-coated vesicles that in turn deliver ceramide to the ER and the same perinuclear location seen for protein uptake in Figures 1.2.A and 1.3.A (73).

Proteins that first enter *Giardia*'s endocytic system through the PVs concentrate in the TVN and co-localize with *Gl*CP-GFP fusion proteins. Although GFP reporter constructs have not been routinely used in anaerobic organisms, our studies with the pGFP.pac expression vector yielded consistent results. We attribute this to the fact that the *Giardia* is microaerophilic and cells were pulsed with sufficient oxygen levels prior to GFP visualization. Cathepsin proteolytic activity, confined to lysosomal organelles in other eukaryotes, is distributed throughout the TVN of *Giardia*, extending from the peripheral TVN to the perinuclear region (Figures 1.3.C, 1.8.C). In keeping with the localization of *Giardia* cathepsins in a pH neutral compartment, the pH optimum of the *Giardia* cathepsin enzymes are shifted towards neutral pH (6 –7). The orthologous proteases of mammalian lysosomes have a strictly acidic 5 - 6 pH optimum (61). While some cathepsin activity or GFP reporter construct was seen peripheral to the TVN, this does not appear to be in the PVs. More likely it is either peripheral TVN devoid of ER markers or the vesicles like those proposed to carry ceramide from the PVs to the ER (73).

The observation that the *Giardia* ER functions in endocytosis and degradation of proteins has bearing on the recent debate concerning the role of the ER in mammalian macrophages. Gagnon et al. (2002) proposed that the ER plays a pluripotent role, including direct involvement in formation of the phagosome at the early stages of macrophage phagocytosis (52). Specifically, they proposed a direct fusion of the ER to the plasma membrane to initiate this process. In support of this model, they demonstrated that ER proteins are enriched in phagosomes. In contrast, however, Touret et al. (2005) could not verify a physical continuity between the ER and the plasma membrane using

a combination of biochemical fluorescent imaging and electron microscopy (53). Nevertheless, the latter authors pointed out the appeal of ER-mediated phagocytosis as an explanation of how phagocytes are able to internalize multiple large particles without large decreases in surface area. In light of this debate, two observations made in this study are noteworthy. First, the TVN of the protist *Giardia lamblia* serves as both a site of protein synthesis, endocytosis and degradation of material from the extracellular milieu. Second, a direct fusion of the ER in *Giardia* with the plasma membrane was not seen, rather the clathrin-rich PV network appears to be the initial site of endocytosis (Figure 1.11). Gagnon et al. (2002) in fact suggested that the ER fuses with plasma membrane at phagocytic cups in an area of plasma membrane known to contain endocytic membranes, and perhaps even clathrin. Therefore, the observation that *Giardia* endocytoses proteins into clathrin-coated PVs that then fuse with the ER may be more analogous to the proposed model of phagocytosis than first appears.

This study raises the question of how a single compartment could function as a site of protein catabolism as well as a pathway for import and sorting of newly translated protein. There is evidence of some polarized compartmentalization within the *Giardia* ER in that ribosomes are attached to the ER predominantly, though not exclusively, in the perinuclear region where most KDEL (ER retention signal) localization is also seen (Figure 1.7.A) (76). Nevertheless, cathepsin GFP reporter constructs and functional protease activity are present both in the perinuclear region as well as in more peripheral regions of the TVN (Figure 1.2, 1.3). In fine confocal sections, it appears that though *Gl*CP2-GFP and PDI2 are dispersed throughout the TVN, single-marker compartments exist for both proteins (Figure 1.7.B and 1.7.C). The nature of these subcompartments,

including how and if they are stably maintained, remains unclear. However, though the entire compartment may be a pathway for protein import, degradation may only occur in limited sectors of the TVN.

The chaperones PDI2, which also contains an ER retention signal (Knodler et al., 1999), and the BIP homologue HSC70 are dispersed throughout the TVN including near the periphery of the cell and may represent another possible mechanism to protect newly translated proteins from the action of proteases functioning in the ER is sequestration by chaperones (77-79). Newly translated peptide chains are rapidly enveloped and protected by chaperones as they thread into the ER lumen (69). This may allow functional separation of protein degradation and folding/transport of newly translated proteins in the same membrane-bound compartment in *Giardia*. Separating anabolic and catabolic ER functions might also be aided by the fact that newly folded globular proteins should be less efficiently cleaved in contrast to endocytosed, denatured proteins that have passed sequentially through the host's acidic upper digestive tract and acidified *Giardia* PVs and have been acted on by host digestive proteases.

The function of the TVN seen in *Giardia* may have parallels in higher eukaryotes, for example in the delivery of exogenous oligopeptides to the ER of MHC class I-presenting cells (80). Day et al. found that after a brief incubation of antigen-presenting cells with fluorescein-labeled peptides in the medium, the peptides are detected within the ER bound to MHC class I molecules. Such ER internalization of exogenous peptides is blocked by pinocytosis inhibitors and occurs without traversing the Golgi or the cytosol. Most importantly, in cells lacking MHC class I, peptides are delivered to the ER but then are quickly removed. The authors speculated that this was due to the action of resident

ER proteases (80). There are other examples of direct delivery of extracellular material to the ER in higher eukaryotic cells including endocytosis of AB5 toxins (55) and SV40 variants (54,81). We propose that both MHC class I retrieval, and thus protection of exogenous peptides, and the proposed chaperone-mediated protection of newly translated peptides in *Giardia*, represent related examples of a functional ER compartmentalization that has its origin in primitive ancestors of the modern eukaryotic cell.

1.4. Figures and tables for Chapter One

Figure 1.1: The endocytic pathway. 26

Figure 1.2: Proteins were rapidly endocytosed by trophozoites into the TVN and the perinuclear region. 27

Figure 1.3: *Giardia* orthologues of mammalian lysosomal proteases are present in the endocytic TVN where they co-localized with endocytosed proteins. 28

Figure 1.4: *Giardia* cysteine proteases degraded endocytosed proteins and were optimally active at a higher pH than their mammalian orthologues. 29

Figure 1.5: Cysteine proteases are not secreted by *Giardia* upon contact with intestinal epithelial cells in culture. 31

Figure 1.6: Cysteine protease activity is not essential for trophozoite survival. 32

Figure 1.7: *Gl*CP2-GFP and *Gl*CP protease activity co-localized with the ER marker PDI2 in the perinuclear region and TVN. 33

Figure 1.8: Cathepsin-like cysteine proteases are not significantly concentrated in the acidified clathrin-rich PVs of *Giardia* trophozoites. 35

Figure 1.9: Ultrastructural analysis revealed that gold particles were endocytosed into the TVN. 37

Figure 1.10: Identification of sites of fusion between the TVN and PVs. 39

Figure 1.11: Model of endocytic network of *Giardia* trophozoites. 40

Figure 1.1: The endocytic pathway.

(Figure adapted from (1)). Endocytosis/phagocytosis begins with the formation of a phagocytic cup on the plasma membrane, to which the ER may contribute membrane. Maturation from early to late endosomes occurs through the formation of multivesicular bodies. These bodies move inward along microtubules, and recycling of components to the plasma membrane continues as the bodies move. The multivesicular bodies gradually turn into late endosomes, either by fusing with each other or by fusing with preexisting late endosomes. The late endosomes no longer send vesicles to the plasma membrane but communicate with the trans Golgi network via transport vesicles, which deliver the proteins that will convert the late endosome into a lysosome.

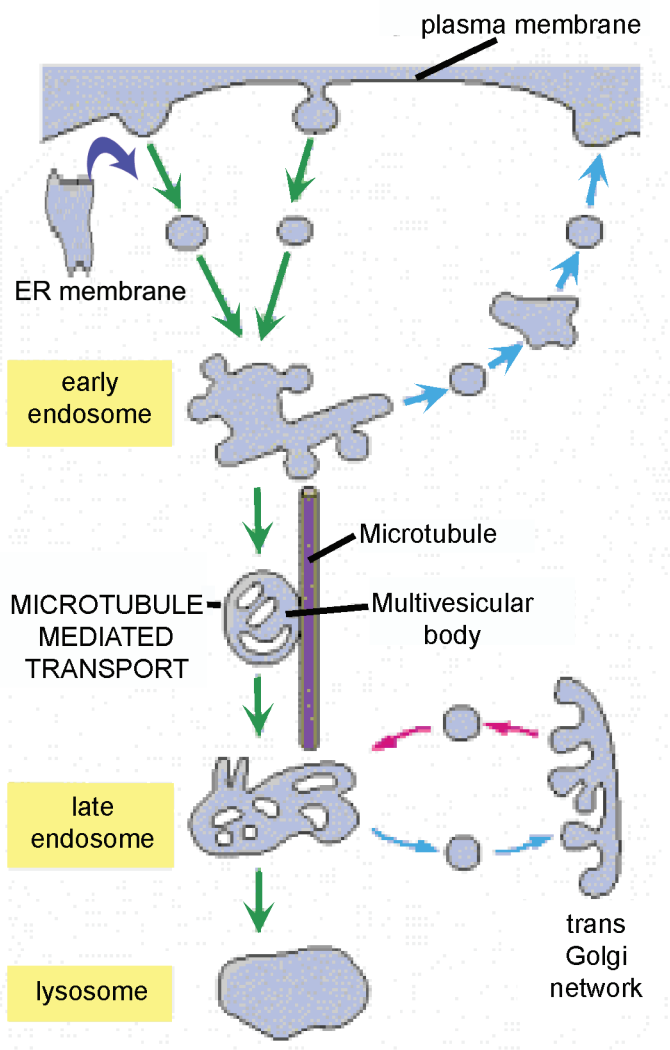


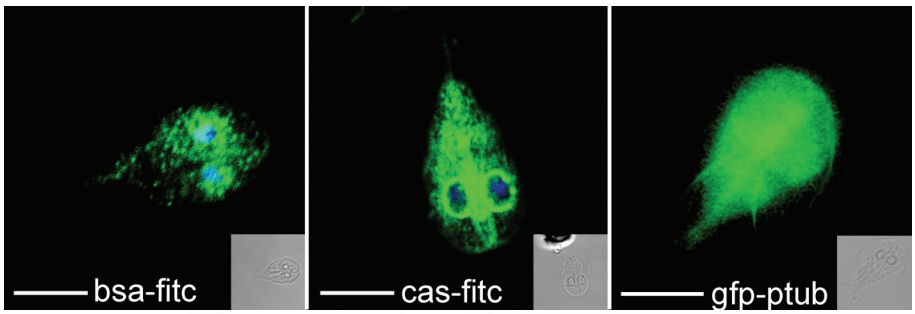
Figure 1.2: Proteins were rapidly endocytosed by trophozoites into the TVN and the perinuclear region.

A. Endocytosis of albumin (BSA-FITC, left panel), and casein (casein-FITC, center panel) resulted in both a fine "vesicular" pattern in the cytoplasm and intense perinuclear localization. DAPI (blue) highlighted the nuclei of *Giardia*. The perinuclear localization is more apparent with casein due to the specific optical cut that included the nuclei. Transfection with the empty GFP vector (right panel) resulted in cytoplasmic localization that was outside of any membrane-bound cellular compartment. Bars, 5 μm.

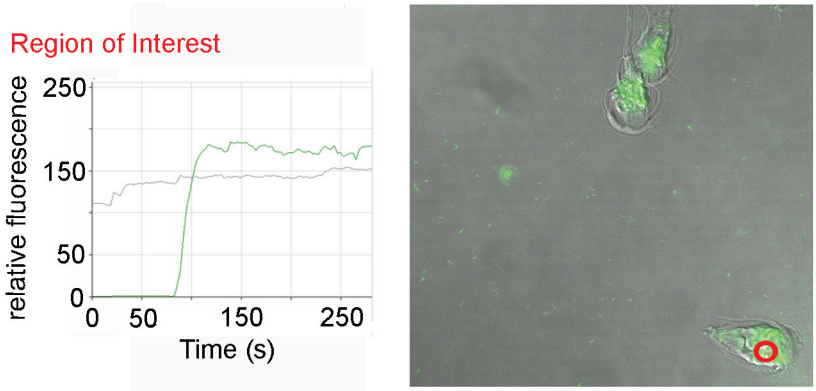
B. Uptake of biotin-conjugated Fluospheres. The kinetics of uptake were tracked by monitoring the increase in relative fluorescent units (RFU) in the perinuclear region of a cell (red circle) compared to the background level of fluorescence. The green line in the graph corresponds to fluorescence intensity over time. (The gray line represents background intensity).

C. Pretreatment of cells with either (A) 20 μM cytochalasin D or (B) 10 μM sodium azide inhibited albumin (green) uptake. Some accumulation of BSA was visible on the outside of cells (white arrowhead), but the pattern distinctly differed from that seen without inhibitors.

A



B



C

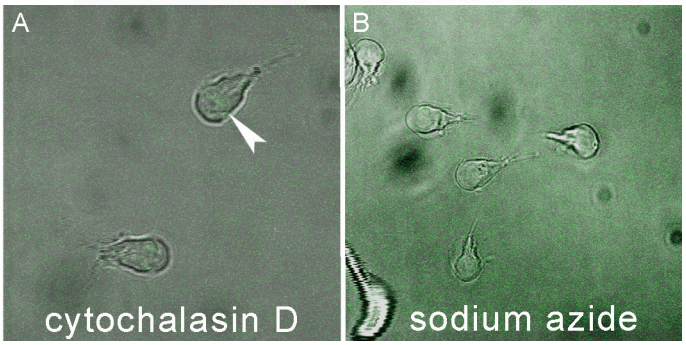
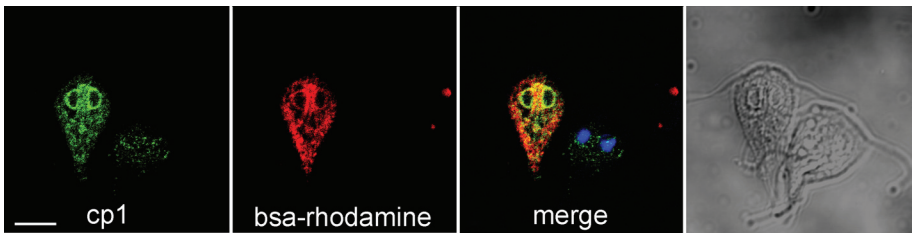


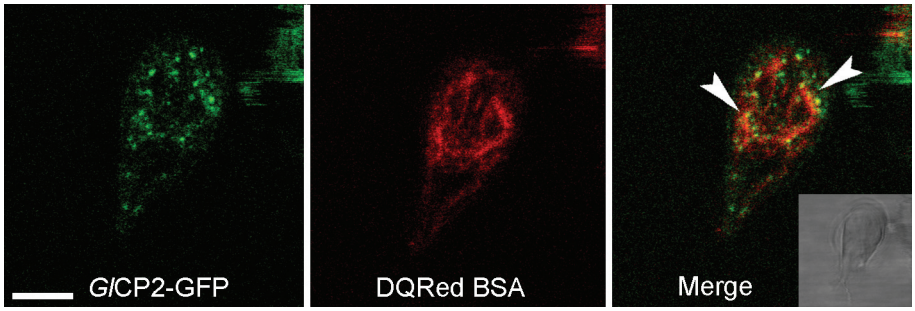
Figure 1.3: *Giardia* orthologues of mammalian lysosomal proteases are present in the endocytic TVN where they co-localized with endocytosed proteins.

A. The cathepsin B-like protease *Gl*CP1-GFP (green) also co-localized with endocytosed albumin (red) in the TVN. DAPI (blue) marks the nuclei. Bar, 5 μm
B. The cathepsin B-like protease *Gl*CP2-GFP (green, *Giardia* orthologue of mammalian lysosomal proteases) co-localized with endocytosed albumin (red, DQRed BSA) during live cell imaging. The merged image shows overlap (yellow) in both the perinuclear region of the trophozoite as well as the finer network of the TVN (left cell). (Cell in upper right is in motion). Bar, 5 μm.
C. Quantification of co-localization shown by plotting pixel fluorescence intensity from channels for DQRed BSA and GlCP2-GFP. The majority of pixels are positive for both markers.
D. Cathepsin endopeptidase activity detected by fluorescent substrate cleavage (Z-FR-MNA). This activity is seen in the outer TVN but is particularly intense in the perinuclear region (short arrows). Bar, 5 μm.

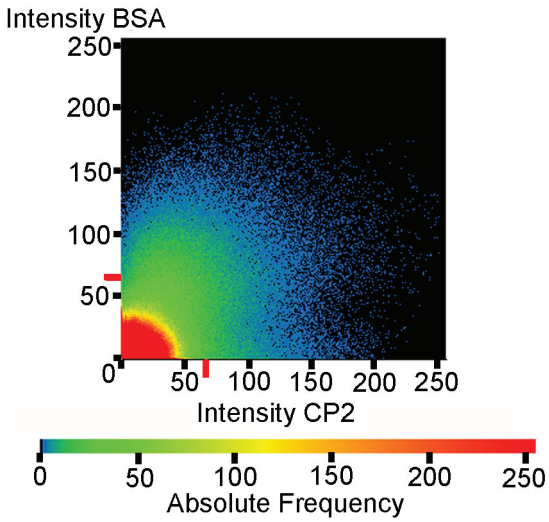
A


cp1
bsa-rhodamine
merge

B


GlCP2-GFP
DQRed BSA
Merge

C


Intensity BSA
250
200
150
100
50
0
50
100
150
200
250
Intensity CP2
0
50
100
150
200
250
Absolute Frequency

D

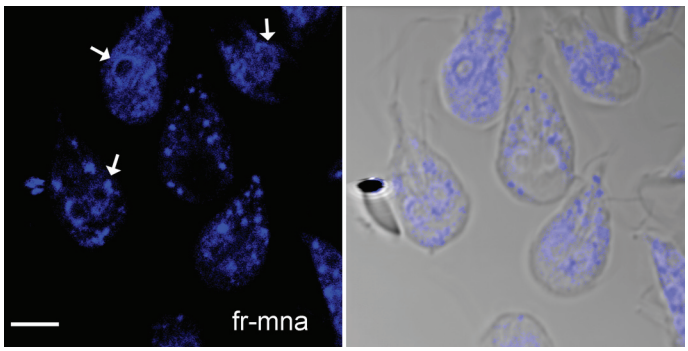

fr-mna

Figure 1.4: *Giardia* cysteine proteases degraded endocytosed proteins and were optimally active at a higher pH than their mammalian orthologues.

A. *Giardia* lysates were incubated with BSA-FITC (short arrowhead) and Casein-FITC (elongated arrowhead). Lane A: BSA 24 hr. no lysate control. Lane B: Casein 24 hr. no lysate control. Lane 1: BSA 5 hr. Lane 2: BSA 16 hr. Lane 3: BSA 16 hr + 20 μM E64. Lane 4: Casein 5 hr. Lane 5: Casein 16 hr. Lane 6: Casein 16 hr. + 20 μM E64. BSA-FITC is clearly protected by the addition of E64. The bottom band of Casein-FITC is also highly sensitive to cysteine protease degradation.

B. Cysteine proteases degraded endocytosed casein-FITC. Cells were pulsed for 30 min with substrate and chased with fresh media for the times indicated. Cell lysates were examined. Lanes 1, 2 and 3 shows loss of substrate in the cells following 1, 5 and 16 h of incubation. In the presence of three known cysteine protease inhibitors (10 μM) (Lanes 4, 5 and 6) casein degradation was blocked (arrows).

C. Cysteine proteases degraded endocytosed albumin-FITC. Experiment was done as in Figure 5B. Lane 1 shows undegraded albumin Lanes 2 and 3 show a decrease in substrate following 24 h of incubation. In the presence of E64d (Lane 4), cells contained much more intact albumin

D. pH profile of *Giardia lamblia* cathepsin B-like cysteine protease activity in total *Giardia* lysates against the fluorogenic substrates Z-FR-AMC and Z-RR-AMC and against the macromolecular substrate casein-resorufin. pH optimum was 6-7 for activity against all substrates.

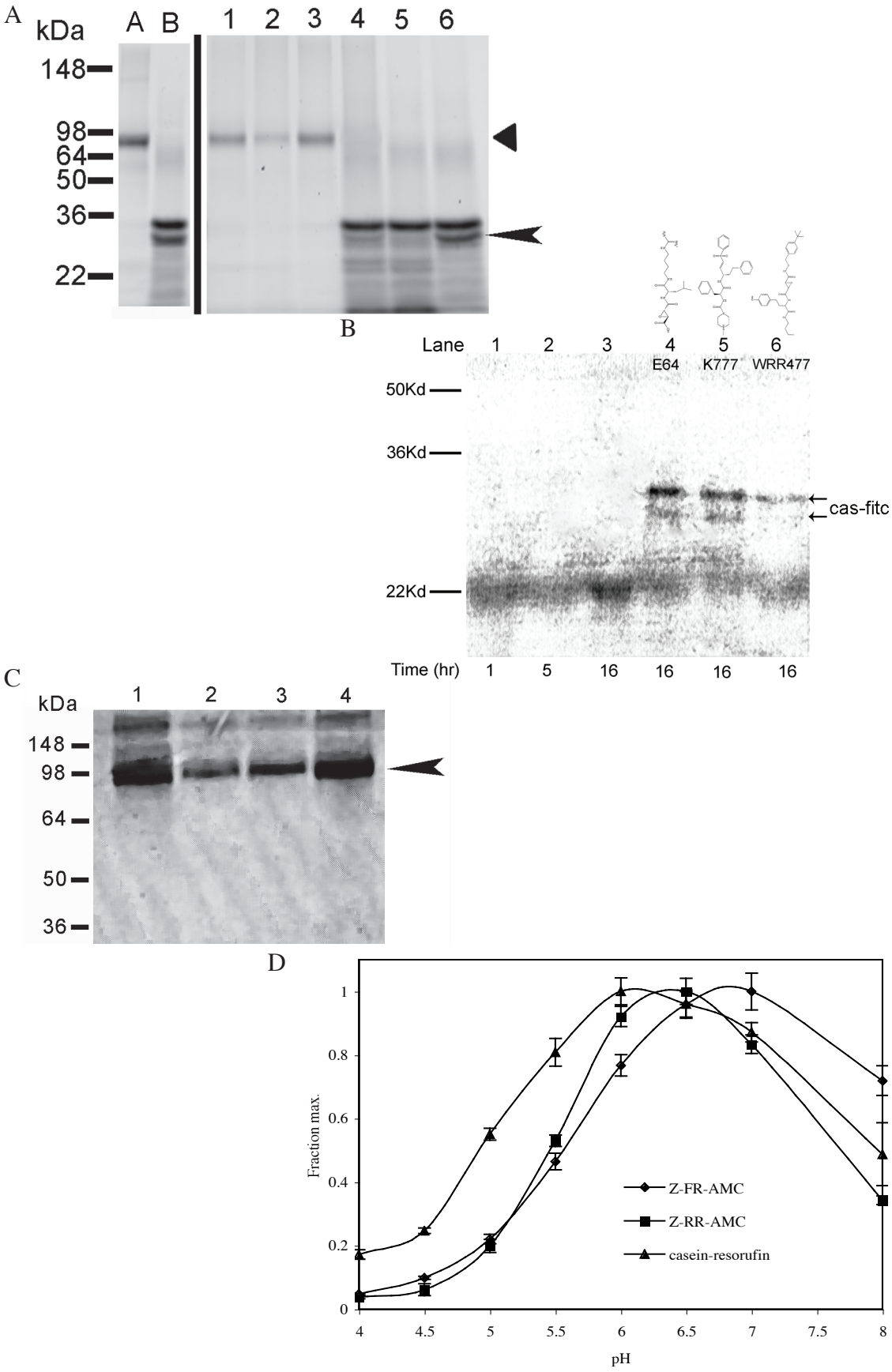
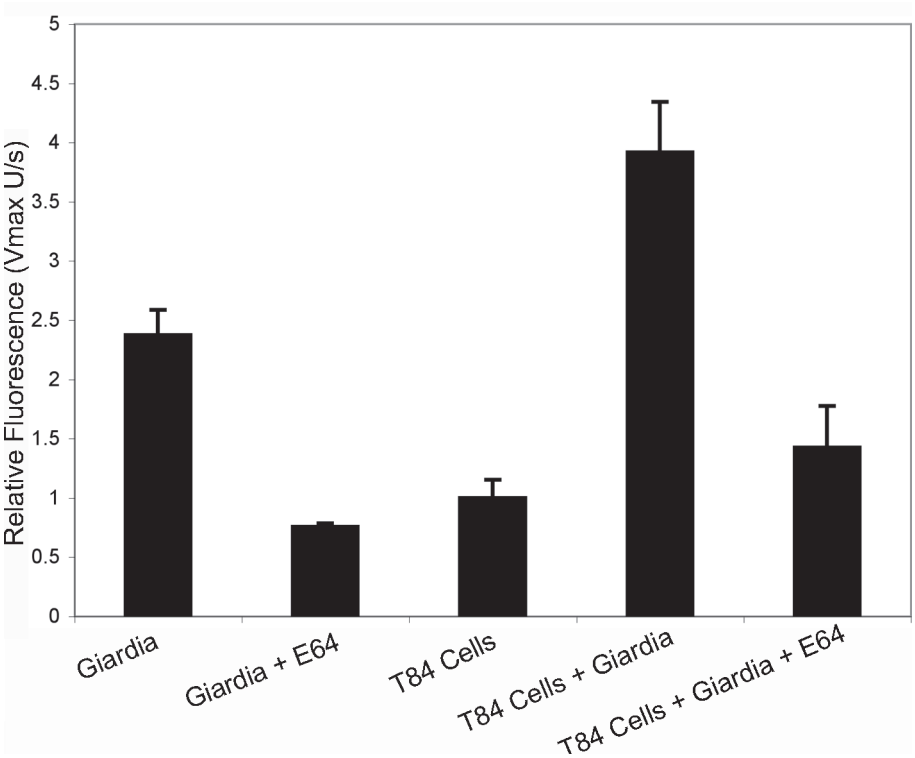

A
kDa
A B
1 2 3 4 5 6
148
98
64
50
36
22
B
Lane 1 2 3 4 5 6
E64 K777 WRR477
50Kd
36Kd
22Kd
cas-fitc
Time (hr) 1 5 16 16 16 16
C
kDa
1 2 3 4
148
98
64
50
36
D
Fraction max.
Z-FR-AMC
Z-RR-AMC
casein-resorufin
pH

Figure 1.5: Cysteine proteases are not secreted by *Giardia* upon contact with intestinal epithelial cells in culture.

A. *Giardia* trophozoites were preincubated in media or media containing E64 and added to a monolayer of T84 intestinal epithelial cells in culture. Supernatants were tested for activity against the fluorogenic peptide substrate Z-FR-AMC. No statistically significant increase was seen in *Giardia* incubated with T84 cells over the additive amount of the *Giardia* and the cells alone.

B. The cysteine protease inhibitor ^{125}I-DCG04 was incubated with media from the following: (1) *Giardia* lysates, (2) *Giardia* cells, (3) *Giardia* cells preincubated with E64d, (4) T84 cells, (5) *Giardia* + T84 cells, and (6) *Giardia* preincubated with E64d + T84 cells. No specific cysteine protease activity was seen in *Giardia* + T84 cells. No activity was detected at the size of the known endogenous cysteine protease activity in *Giardia* (black arrowhead).

A



B

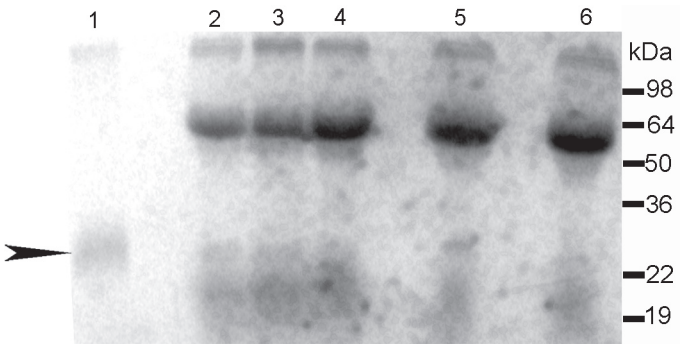


Figure 1.6: Cysteine protease activity is not essential for trophozoite survival.

A. Addition of the general cysteine protease inhibitors E64d and K11777 (K777) at varying concentrations did not affect the survival curve of *Giardia* trophozoite cultures.

B. As cysteine proteases were shown to degrade endocytosed proteins, *Giardia* in minimal media and 5% FBS (compared to 10% in normal growth medium) were grown in the presence of E64d and K11777. Cysteine protease inhibitors did not affect the survival of trophozoites in this minimal media.

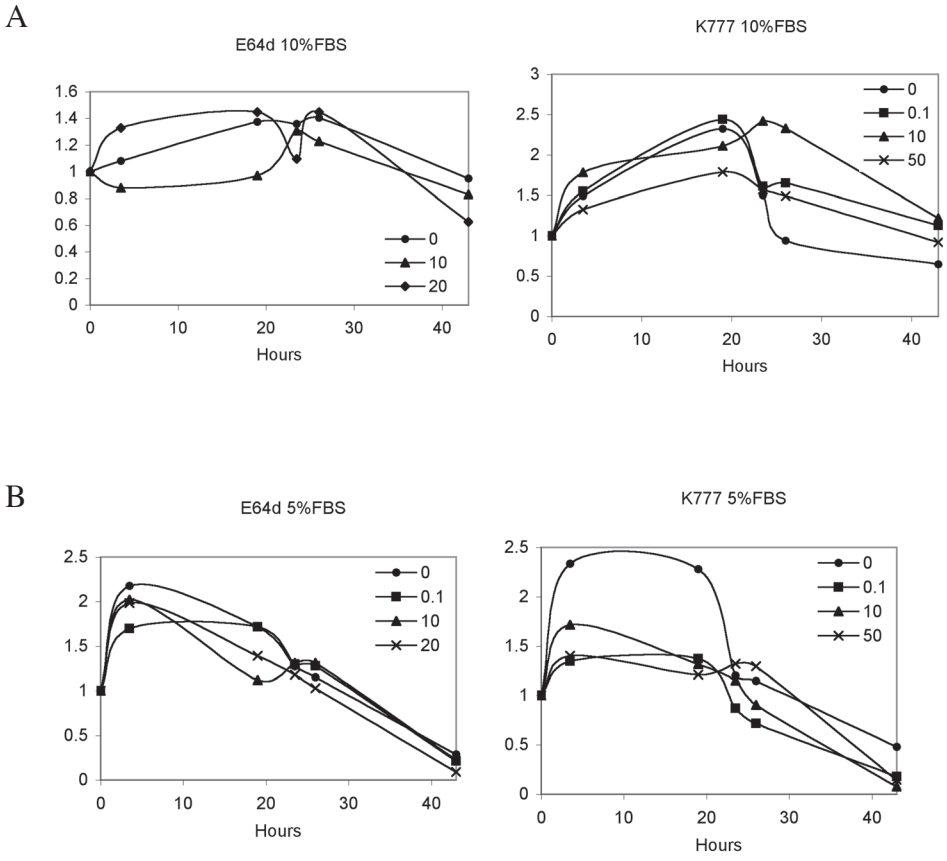
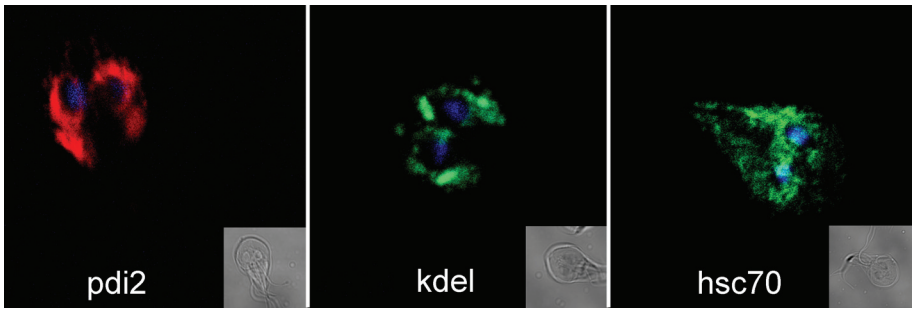


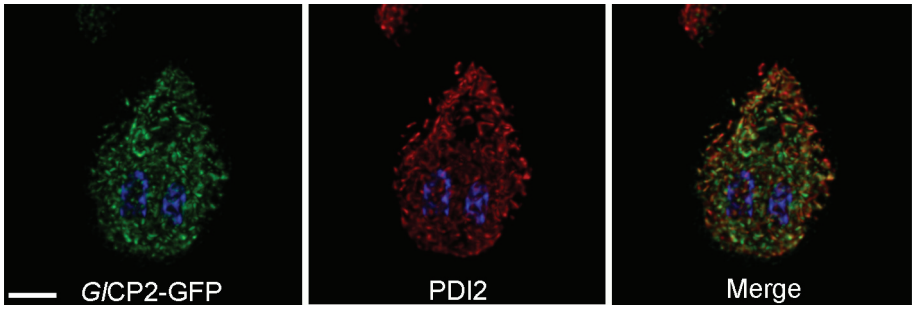
Figure 1.7: *Gl*CP2-GFP and *Gl*CP protease activity co-localized with the ER marker PDI2 in the perinuclear region and TVN.

A. PDI2, KDEL and HSC70 identified the ER of the trophozoites in a perinuclear and tubulovesicular pattern. KDEL-containing proteins immunolocalized discretely to the perinuclear region. DAPI (blue) marks the nuclei.
B. Reconstructed three dimensional images from confocal microscopy fine sections of trohozoites showed cathepsin B-like cysteine protease *Gl*CP2-GFP (green) co-localized with PDI2 (red), extending throughout the TVN. Tubulovesicular nature of the TVN is apparent. DAPI (blue) marks the nuclei. Bar, 3 μm.
C. Software analysis enabled visualization of co-localization. Left panel, co-localization colored in pale yellow. Right panel, volumetric reconstruction of confocal sections. Inset shows that PDI2 is membrane-localized while *Gl*CP2-GFP is luminal (elongated arrowhead). *Gl*CP2-GFP compartments lacking PDI2 staining (short arrowhead) may be subcompartments of the ER.
D. Quantification of co-localization shown by plotting pixel fluorescence intensity from two channels. Left panel, co-localization of PDI2 and *Gl*CP2-GFP. Right panel, co-localization of DAPI and *Gl*CP2-GFP as a negative control for co-localization.
E. Three-dimensional projection of cysteine protease activity against the substrate Z-RR-MNA (green) co-localized with PDI2 (red) in the perinuclear region and in portions of the TVN. The projection highlights the continuity of the TVN. Protease activity (green) is less prevalent than in Figure 4A due to loss of product during PDI2 immunolocalization. n=nucleus. Bar, 5 μm.

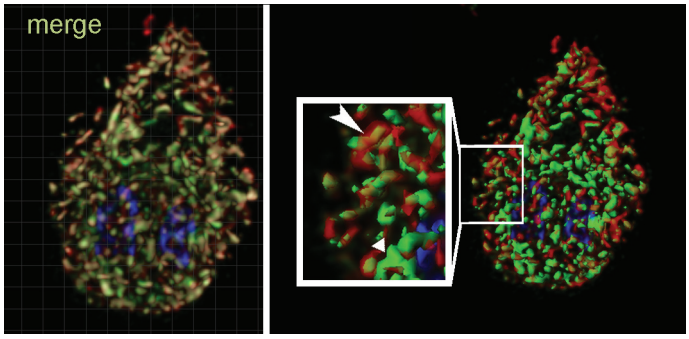
A


pdi2
kdel
hsc70

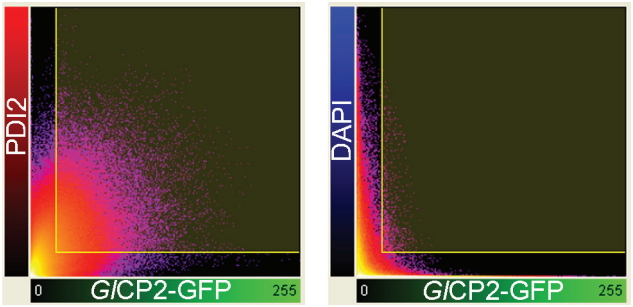
B


GlCP2-GFP
PDI2
Merge

C


merge

D


PDI2
0
GlCP2-GFP
255
DAPI
0
GlCP2-GFP
255

E

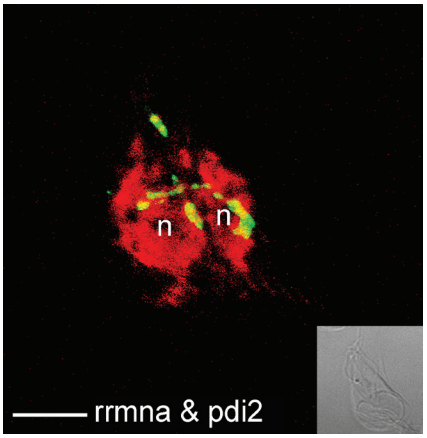
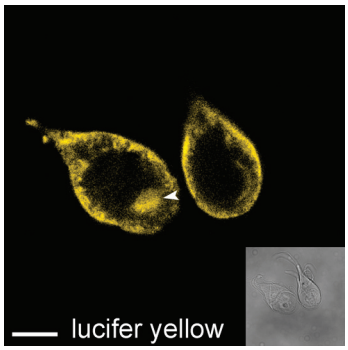

n
n
rrmna & pdi2

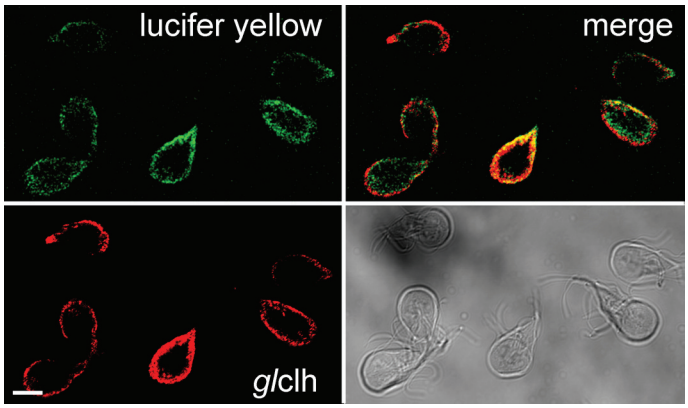
Figure 1.8: Cathepsin-like cysteine proteases are not localized to the acidified clathrin-rich PVs of *Giardia* trophozoites.

A. Lucifer yellow staining (yellow) highlights the acidic nature of the PV network aside from the cell periphery. PVs are also abundant in an area of the cell near the flagella origin (white arrow head). Bar, 5 μm.
B. Clathrin (red) and Lucifer yellow (green) overlap at the PV network. Merge is yellow. Bar, 5 μm.
C. Cathepsin B-like cysteine protease *Gl*CP2-GFP (green) marks the endocytic TVN of *Giardia*. This network is distinct from the clathrin-associated PVs (red). Occasional co-localization of *Gl*CP2-GFP with clathrin was seen at the interface between the two staining patterns, but was the exception. Clathrin is also abundant near the site of flagellar attachment (white arrow head). Bar, 5 μm.
D. *Gl*CP2-GFP (green) co-localized with an anti-GFP antibody (red), demonstrating that the GFP fluorescence is an inclusive marker of the localization of GFP regardless of the cellular environment in which it is found. DAPI (blue) marks the nuclei. Bar, 5 μm.

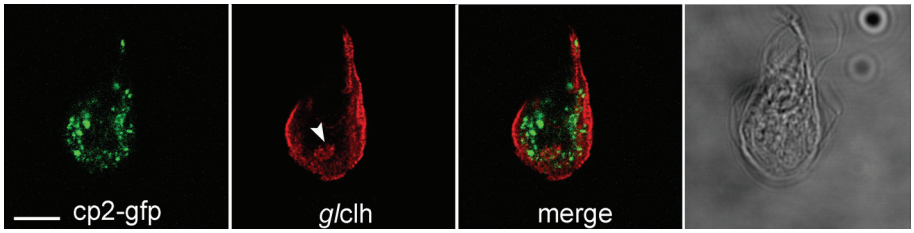
A


lucifer yellow

B


lucifer yellow
merge
glclh

C


cp2-gfp
glclh
merge

D

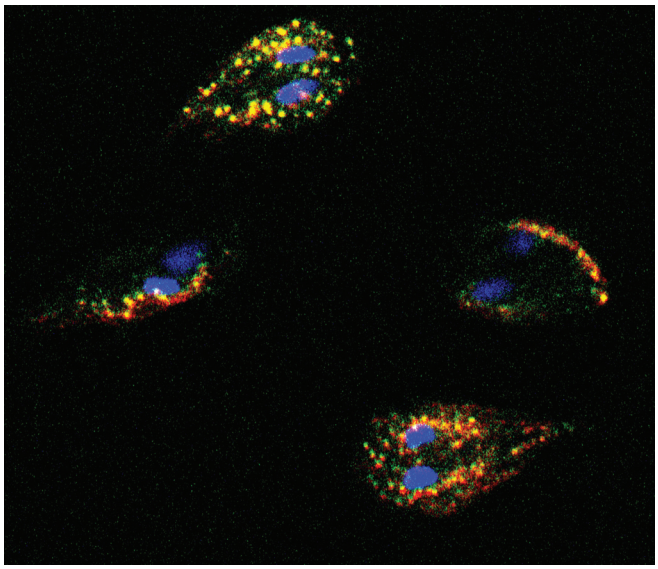


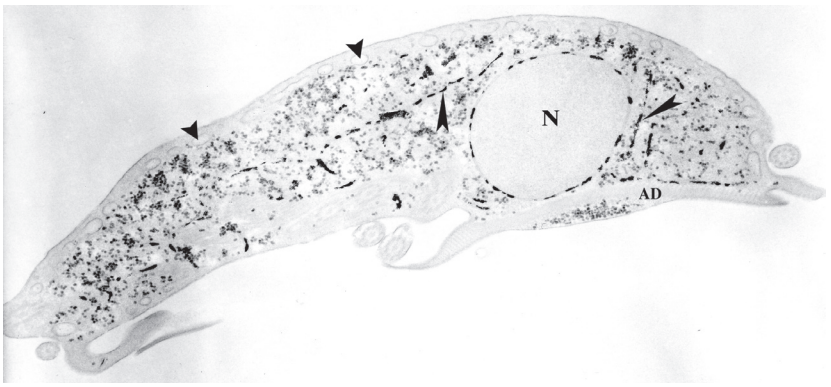
Figure 1.9: Ultrastructural analysis revealed that gold particles were endocytosed into the TVN.

A. Glucose-6-phosphatase activity highlighted the ER-like TVN in electron micorographs of *Giardia* trophozoites. The perinuclear region of the TVN was easily seen by the electron-dense G-6-P reaction product against the nucleus (N). Portions of the labyrinthine cytoplasmic region of the TVN (elongated arrow heads) were seen as short, slit-like G-6-P reaction product on the background of less electron dense glycogen particles as previously described (21). Focally fused, PVs (short arrow heads) can be seen at the edge of the cell at the top of the micrograph.

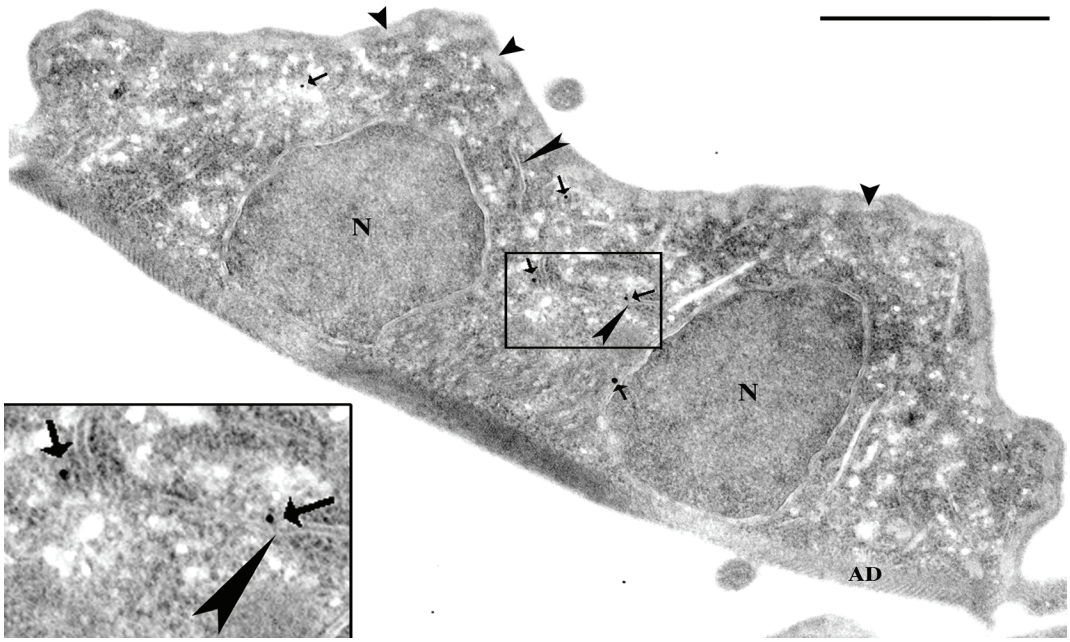
B. Endocytosed albumin-coated ultrasmall (≤5 nm) gold particles (short arrows) localized to the TVN (elongated arrow heads) including the perinuclear region (see inset). This correlated with the endocytic pattern seen in the TVN and perinuclear region by confocal analysis in Figures 1-4 (short arrow heads denote PVs; N = nucleus, AD = adhesive disc). Bar, 1 μm.

C. Endocytosis of larger albumin-coated gold particles (10 nm) (elongated arrow heads) was arrested in the PVs. Bar, 200 nm.

A



B



C

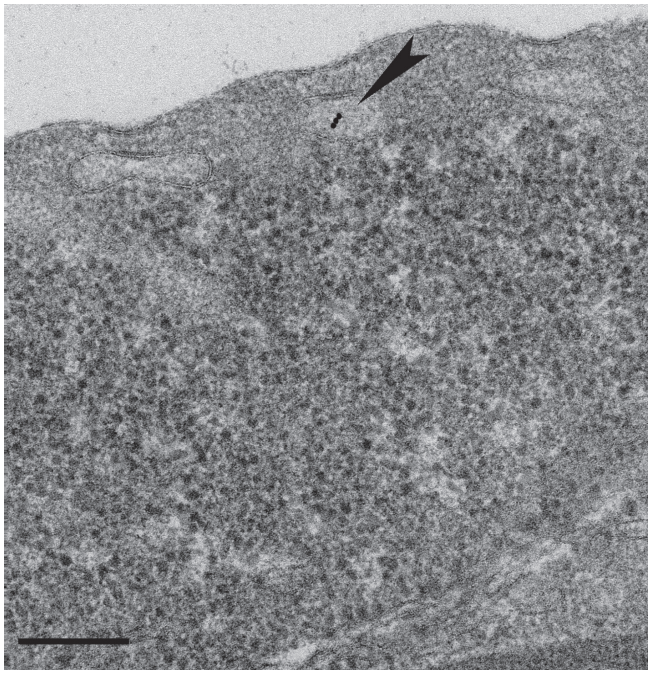
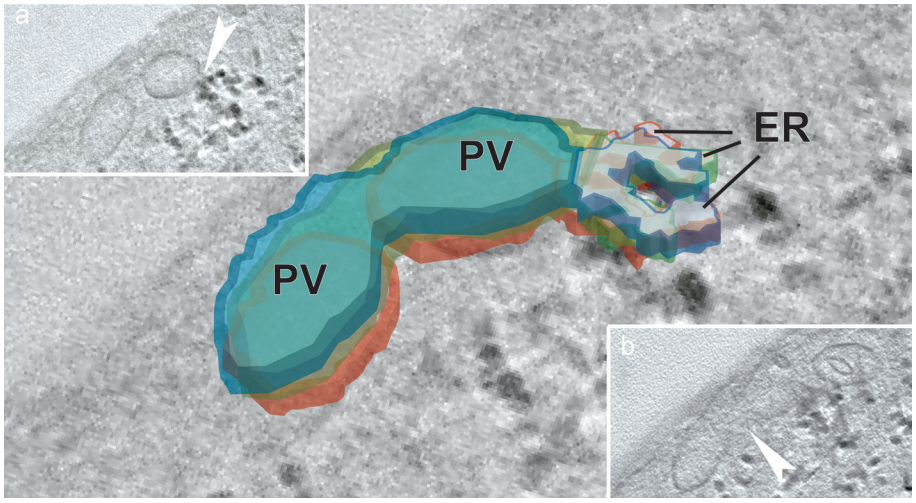


Figure 1.10: Identification of sites of fusion between the TVN and PVs.

A. In a stack of three consecutive EM tomography images, fusion between a large PV and the branching ER-like TVN (identified by the glucose-6-phosphatase reaction product, black) was observed (inset a, white arrowhead, inferior z-stack slice). Fusion between two adjacent PVs is also seen (inset b, white arrowhead, superior z-stack slice). Each subsequent section is represented by a different color (red, green, and blue). The TVN is outlined with color, while PVs are filled.

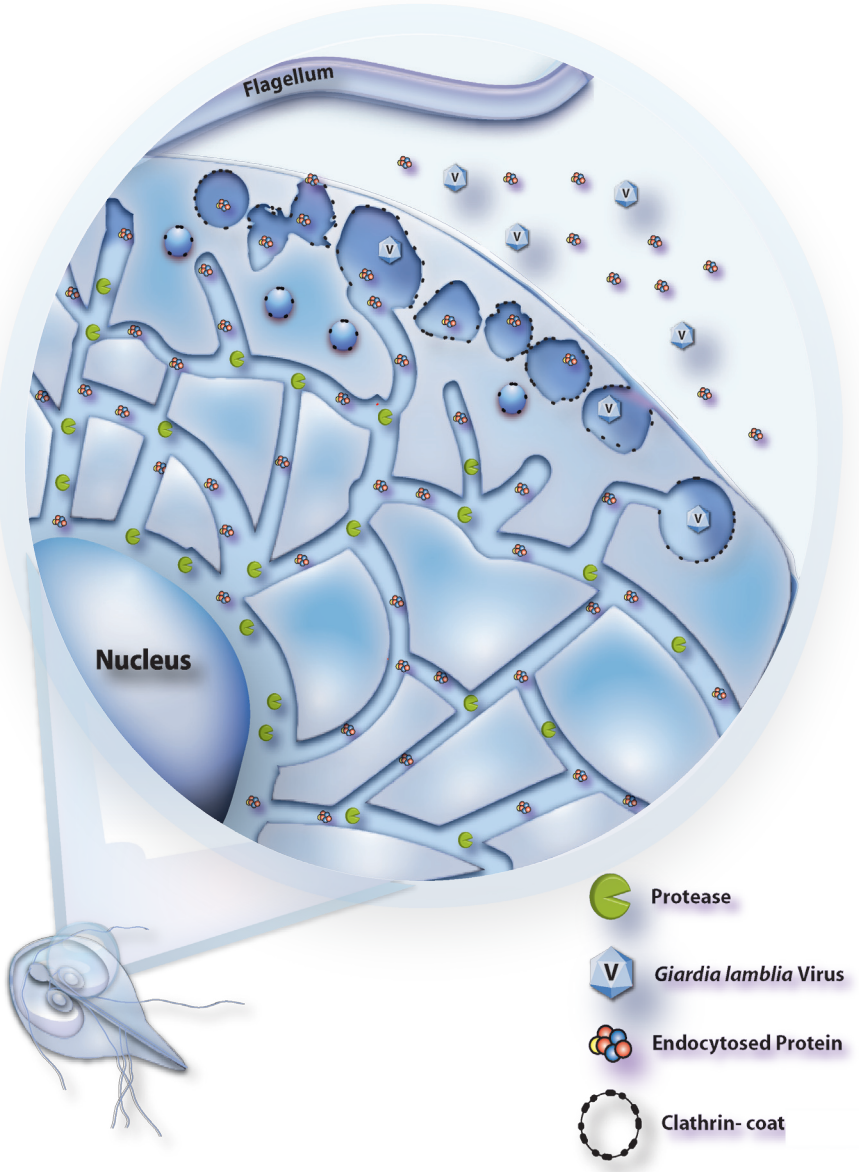
B. Fusions between tubular extensions of the TVN and PVs at the periphery of the cell were seen on EM sections (elongated arrow head).

A



B

Figure 1.11: Model of endocytic network of *Giardia* trophozoites.
Active proteases reside primarily in the TVN where endocytosed proteins are degraded. PVs contain clathrin and are the site of initial uptake. Membrane fusions between PVs as well as between PVs and the TVN are dynamic. Endocytosed proteins are passed from PVs to the TVN by these dynamic fusions and possibly by a clathrin-coated vesicle population trafficking between the two compartments as evidenced by Hernandez et al. (2007) (73). *Giardia* virus subverts this network for entry into the cell (75).



Chapter 2: Identification of the major cysteine protease of *Giardia lamblia* and its role in encystation.

2.1. Introduction

2.1.1. Clan CA cysteine proteases

Peptidases (proteases) in which the active site nucleophile is the sulfhydryl group of a cysteine are known as cysteine proteases (59). Cysteine proteases catalyze the cleavage of peptide, amide, ester, and thiol ester linkages in macromolecular proteins and oligomeric peptides (82). Cysteine proteases are classified into ten different clans. Clans CB and CC are composed of viral cysteine endopeptidases. The caspase family of proteins is included in clan CD. Clan CE includes the adenovirus endopeptidase and related proteins. Clans PA(C) and PB(C) comprise peptidases that are chymotrypsin-like, but the active site serine has been replaced by a cysteine residue. A miscellaneous clan CX contains families of proteases for which too little is known to classify with confidence. Clans CF, CH, CJ, and CK also have distinct tertiary scaffolds (59). The most abundant, clan CA, are those proteases resembling papain, the first clearly recognized cysteine protease (59). Within clan CA, proteases are further divided into families, of which family C1 is the most abundant (34).

The papain-like cysteine proteases in clan CA, family C1, can be found in the animal and plant kingdoms and also in some viruses and bacteria (34,83). In eukaryotic cells, the proteases that fall in this family are the lysosomal cathepsins. The name "cathepsin" was coined by Willstätter and Bamann in 1929 to describe a protease that was active at an acidic pH but could be differentiated from pepsin (84-86). The term

evolved through the years to mainly refer to intracellular proteases, often localized in the lysosomes, that are active at a weakly acidic pH (84). However, the list of exceptions to these requisites continues to grow as more proteases classified in this family by amino acid sequence similarity exhibit varying traits (33,83). Papain-like cathepsins are involved in such diverse processes in mammalian cells as bone resorption, hormone maturation, antigen processing, and epidermal homeostasis, and cancer metastasis (34). The destination and location of activity for many cathepsins is the lysosome. These cathepsins, such as C (also called dipeptidyl peptidase I), B, H, and L function in an acidic environment and degrade proteins in the cell. In parasites, the function of cathepsins diverges even more, with enzymes from this class involved in growth and development, evasion of host immune response, tissue/skin penetration, destruction of host proteins, and a multitude of other diverse roles (33,34). It is clear in recent years that the structure and not the function of these enzymes is what groups them together into this class.

2.1.2. Activation and catalytic mechanism

The tertiary structure of all family C1 cysteine proteases resembles that of papain and this three dimensional fold is highly conserved (34). The cathepsins are synthesized as zymogens, and are activated by proteolytic removal of the N-terminal propeptide. This activation can be accomplished intramolecularly (cis), where the molecule itself catalyzes proteolysis, or a second molecule can activate the protein in an intermolecular method (trans) (83). The pro-region of the protease has a number of functions; it can act as an intramolecular chaperone to assist in protein folding, an inhibitor regulating

protease activity, and a trafficking signal that targets the protease to its destination (87). Following removal of the inhibitory pro-peptide (and in some enzymes the removal of a C-terminal peptide as well), the active site, located at the interface between two domains of the enzyme, is exposed and it can catalyze the cleavage of substrate peptides (59,86). The mechanism of activity of family C1 cysteine proteases involves three key residues; Cys25, His159, and Asn175 (papain numbering). During peptide hydrolysis, the nucleophilic cysteine attacks the carbonyl carbon of the substrate peptide and forms a tetrahedral intermediate, which has a negative charge that is stabilized by the "oxyanion hole." The oxyanion hole is formed by the backbone amine of Cys25 and the amino group of Gln19 (84,88). The leaving amine is then protonated and the acyl enzyme is produced. Reaction of the acyl enzyme with water produces the second tetrahedral intermediate followed by the release of the second product and the regeneration of the free enzyme (59,84).

2.1.3. Protease substrate specificity

The specificity of each protease in the C1 family of cysteine proteases is determined by a series of peptide-binding subsites in the enzyme catalytic site cleft. Peptide amino acids are denoted by "P" and interact with binding subsites denoted by "S". The amino acids and binding subsites surrounding the scissile bond are numbered. The carboxyl side of the substrate peptide and binding subsites are referred to as the prime side and those on the amino side of the substrate peptide are called the non-prime side according to Schechter and Berger nomenclature (33,89). The S2 subsite of papain-like cysteine proteases is a hydrophobic pocket and generally prefers hydrophobic side

chains such as Phe and Leu (84). In most clan CA proteases, the S2 subsite has a more defined specificity than the S1 subsite. The amino acid preference of a protease can be determined by assaying the enzyme against a non-prime side positional scanning synthetic combinatorial library (PS-SCL). This type of library, which is made up of peptide substrates that at any given position the amino acids are fixed and known, has been used to characterize many of the cathepsin family of proteases (90-92). Additional structural elements of cathepsins can influence their activity and specificity. For example, cathepsin B contains an "occluding loop" that can stabilize the C-terminal carbonyl of a protein substrate in the active site cleft and endow cathepsin B with exopeptidase activity (59,93). Interestingly, the major cysteine proteases in *Giardia,* although cathepsin B-like in sequence homology, lack the aforementioned occluding loop and therefore are expected to lack exopeptidase activity.

2.1.4. Cysteine proteases in *Giardia* encystation and excystation

In *G. lamblia*, indispensable roles for cysteine proteases of the C1 family have been documented in the processes of encystation and excystation. The process of encystation is a coordinated secretion of cyst wall materials to the periphery of a cell to form the cyst wall (8,94). In response to environmental cues, trophozoites produce abundant CWPs that are packaged into ESVs (Figure I.1). These vesicles grow, mature, and eventually traffic to the plasma membrane of the trophozoite, where the cyst wall material is secreted to form the environmentally stable cyst wall (8,15,95). The expression of many proteins is upregulated during the encystation process (8). Ward et al. validated a role for cysteine endopeptidase activity in the excystation process by

demonstrating that excystation was inhibited by the addition of small molecule cysteine protease inhibitors to the excystation media (37). Touz et al. implicated a cysteine exopeptidase in the process of encystation. Processing of CWP2, one of the main CWPs that form the structure of the cyst wall, was blocked by cysteine protease inhibitors (38).

While chemical knockout experiments focused attention on clan CA cysteine proteases in *Giardia* encystation (38), the recent completion of the *Giardia* genome (www.mbl.edu/giardia) indicated that there are twenty-seven candidate clan CA cysteine protease genes in *Giardia*. To address the question of which gene product(s) was responsible for key events in the life cycle such as cyst wall processing, we analyzed the transcription levels of all twenty-seven genes and found that cysteine protease 2 (*Gl*CP2) was in fact the major expressed cysteine protease gene in *Giardia*. We therefore cloned, heterologously expressed, and biochemically characterized this protease, and specifically evaluated its role in encystation.

2.2. Results

2.2.1. There are twenty-seven Clan CA cysteine proteases in the genome of *Giardia lamblia*.

Prior to the completion of the *Giardia* genome, only four cysteine proteases from *Giardia* had been identified (37,38). Three of these genes encode cathepsin B-like cysteine proteases, the fourth a cathepsin C-like protease. Using these papain family enzymes as a query, the *Giardia* genome was mined for additional genes coding for cysteine proteases. In total, twenty-seven clan CA cysteine protease genes were located in the *Giardia* genome and these could be classified by sequence homology into cathepsin

B-like, cathepsin C-like, or cathepsin K/L like categories (Figure 2.1). Three of these (GenBank accession number EAA38990) are greater than 95% identical to each other and yet are found in three discrete regions of the genome. There is also a set of four genes greater than 95% identical and assigned to the same GenBank accession number (AAK92150).

2.2.2. *Gl*CP2 is the most highly expressed cysteine protease of the twenty-seven in the *Giardia* genome.

RT-PCR was used to determine if each of these genes was expressed in the vegetative and encysting stages of the *Giardia* life cycle. It was found that twenty-five of these twenty-seven genes are expressed while no expression could be seen for the genes with GenBank accession numbers EAA37191 and EAA39030 (data not shown). Quantitative RT-PCR was performed to determine the relative levels of gene expression among the expressed *Giardia* cysteine proteases in the vegetative and encysting life stages. Expression was normalized to the expression of the housekeeping gene glutaraldehyde phosphate dehydrogenase, (GAPDH) which has been found to have stable expression during the *Giardia* life cycle (7,94). It is notable that the cathepsin B-like cysteine proteases are more highly expressed than the cathepsin C-like or K/L-like proteases (Figure 2.1). Expression of all of the cysteine protease genes was increased marginally during encystation. *Giardia* cysteine protease 2 (*Gl*CP2) was the most highly expressed transcript in both vegetative and encysting life stages by 1.6-fold and 3.5-fold respectively over the next most highly expressed transcript (*Giardia* cysteine protease 3, with 89% homology to *Gl*CP2). This is consistent with *Gl*CP2 being the only cysteine

protease Ward et al. biochemically identified from the parasite by affinity purification and N-terminal sequencing (37). The expression of this gene was also increased by 7-fold during encystation (Figure 2.1).

2.2.3. *Gl*CP2 is identified in lysate fractions enriched for cysteine protease activity.

Biochemical characterization of total cysteine protease activity found in *Giardia* lysates was concurrently undertaken to complement the gene expression analysis. *Giardia* lysates were fractionated by ion exchange chromatography and each fraction assayed against an array of N-terminally blocked fluorescent peptide substrates (data not shown). Two main peaks of cysteine protease activity were resolved against the substrates Z-FR-AMC and Z-RR-AMC (Figure 2.2.A). The first peak eluted (Peak A) exhibited activity against both Z-RR-AMC and Z-FR-AMC, while the second peak (Peak B) had activity against Z-FR-AMC but far less against Z-RR-AMC. The activity-containing fractions from each peak were subsequently enriched with two additional rounds of ion exchange chromatography, concentrated, probed with a labeled irreversible cysteine protease active site inhibitor and resolved by 1D SDS-PAGE or 2D gel electrophoresis. The active site probe labeled two discrete protein bands in Peak A and only one protein band in Peak B (Figure 2.2.B). Protein bands from 1D SDS-PAGE or spots from 2D gel electrophoresis were subjected to tryptic digest and analyzed using LC-MS/MS (Figures 2.2.C and 2.2.D). The only cysteine protease identified from these methods was *Gl*CP2, of which peptides were identified in both Peak A and Peak B (Table 2.1). This was consistent with the observation that *Gl*CP2 was the major cysteine protease transcript expressed by *Giardia lamblia*.

2.2.4. *Giardia* cysteine proteases exhibit difficulty in recombinant expression.

Recombinant expression was attempted for *Gl*CP2 (EAA37433), *Gl*CP1 (CAC18646), and EAA37074 in several different heterologous systems. Table 2.2 summarizes the recombinant expression attempts and products. Protein was expressed in some of the systems used, but largely in an inactive form. *Gl*CP1 protein was produced in the *Pichia pastoris* expression sytem but the amount produced was small and the expression was not repeatable. However, this protein did bind ^{125}I-DCG04 (Figure 2.3), demonstrating that active protease was produced. *Gl*CP2 was functionally expressed in the *P. pastoris* system.

2.2.5. Expression and characterization of recombinant *Gl*CP2.

To further analyze the activity and biological role of *Gl*CP2, a resynthesized *Gl*CP2 (r*Gl*CP2) gene (resynthesized to optimize for yeast codon bias) of 903 bp was expressed heterologously in *Pichia pastoris* as a 34 kDa protein. The polyhistidine tagged r*Gl*CP2 was purified by affinity and anion exchange chromatography and was found to autoactivate during the purification process to the mature form of 28 kDa (Figure 2.4.A). The pro-mature and mature forms of r*Gl*CP2 had activity on a native 10% gelatin zymogram gel (Figure 2.4.B). An array of protease inhibitors was tested for their ability to inhibit r*Gl*CP2 activity against Z-FR-AMC. Leupeptin and E64 were the most effective inhibitors of r*Gl*CP2 (Table 2.3). The pH optimum of r*Gl*CP2 was elucidated using the peptide substrates Z-FR-AMC and Z-RR-AMC and the macromolecular substrate casein-resorufin. The pH optimum for r*Gl*CP2 was found to be in the neutral

range for each of these substrates (Figure 2.4.C). This is consistent with the localization of *Gl*CP2 in non-acidified compartments and its absence in the acidified PVs. The K_m of Z-FR-AMC and Z-RR-AMC for *Gl*CP2 was found to be 40 μM and 8 μM, respectively.

To characterize the substrate specificity of r*Gl*CP2, a positional scanning synthetic combinatorial library (PS-SCL) was used to determine the substrate preference of the substrate binding sites for P1-P4 (90) (Figure 2.4.D). r*Gl*CP2 displays an amino acid preference at subsites P1 and P2 (P1: K>>R, Q, P; P2: L, M, V, F) while subsites P3 and P4 have relaxed specificity. This PS-SCL was tested both at the optimal pH for the enzyme (7.2) and at pH 5.5, the conventional pH for this class of enzymes. The substrate specificity did not change over this pH range, though the level of enzyme activity was decreased by approximately fifty percent at the lower pH (data not shown). Based on the substrate specificity, an ideal substrate (Z-VLK-AMC) was used to measure the K_m, which was found to be 19 μM.

2.2.6. The activity profile of recombinant r*Gl*CP2 and the dominant cysteine protease activity found in total *Giardia* lysates are identical.

To compare r*Gl*CP2 activity to that predominantly seen in *Giardia* lysates, the activity profile of r*Gl*CP2 by ion exchange chromatography was examined against Z-FR-AMC and Z-RR-AMC. The two activity peaks seen in *Giardia* lysates were reproduced with purified recombinant protein; the peaks of activity represent the pro and mature forms of the protease in Peak A and the mature protease alone in Peak B (Figure 2.5.A). This correlates with the two protein bands labeled with the cysteine protease active site probe in Peak A, and the single band labeled in Peak B (Figure 2.5.B). To determine

if full-length r*Gl*CP2 has activity against a peptide substrate, protein from Peak A and Peak B was fractionated by SDS-PAGE in non-reducing conditions. In-gel activity was tested against the fluorogenic substrate Z-FR-MNA. Two bands of activity could be visualized in Peak A, while only one band of endoprotease activity was resolved in Peak B (Figure 2.5.B). A western blot of these fractions using an antibody against *Gl*CP2 also demonstrates that two bands in Peak A and one in Peak B are identified as *Gl*CP2 (Figure 2.5.C). This data is consistent with the biochemical and mass spectrometry evidence that *Gl*CP2 is responsible for the activity found in both Peaks A and B.

2.2.7. *Gl*CP2 is found in *Giardia* ESVs and can proteolytically process CWP2 to the predicted size found in the cyst wall.

The localization of *Gl*CP2 during encystation was determined by episomal expression of a *Gl*CP2-GFP fusion in *Giardia*. Encysting cells were probed with an antibody against CWPs to highlight the ESVs and *Gl*CP2-GFP was found to localize to ESVs (Figure 2.6.A). Interestingly, other clan CA cysteine proteases were also localized to ESVs by this method, such as EAA37074 (Figure 2.6.B). Ward et al. previously implicated *Gl*CP2 in *Giardia* excystation (37). However, whether this protease could also play a role in the encystation process was not addressed directly.

Total *Giardia* lysates or purified r*Gl*CP2 was incubated with recombinant CWP2 (rCWP2). In the presence of either *Giardia* lysates or r*Gl*CP2, rCWP2 was processed from its original 39 kDa size to a 26 kDa fragment, the same size of the protein found in the cyst wall and shown to be produced by incubation with a purified fraction of cysteine protease activity containing ESCP (Figures 2.7.A and 2.7.B) (8,38). rCWP2 was also

processed to this 26 kDa fragment in the presence of the endopeptidases trypsin and chymotrypsin, suggesting that the processing of CWP2 is not dependent on protease specificity but instead dependent on the structure of CWP2 and the protein segments accessible to an endopeptidase (Figure 2.7.B). *R. norvegicus* cathepsin C was also tested and no processing of rCWP2 was seen in the presence of this enzyme (Figure 2.7.B). At high concentrations (tested at five times higher than what was used in CWP2 degradation experiments) of r*Gl*CP2, trypsin, or chymotrypsin the 26 kDa fragment of rCWP2 is further degraded to small peptides (data not shown). Interestingly, only r*Gl*CP2 Peak B and not Peak A from the anion exchange column exhibited proteolytic activity against rCWP2 in the time frame of this assay (Figure 2.7.B).

2.2.8. *Gl*CP2 is localized to the periphery of cysts during *in vitro* excystation.

Cells expressing *Gl*CP2-GFP were encysted as previously described (96). Cysts were induced to excyst through a two-stage process as described by Lauwaet et al. ((97). By confocal microscopy, the majority of cysts appeared slightly degraded after the excystation process, which is not surprising as the efficiency of *in vitro* excystation is very low. In these cysts, *Gl*CP2-GFP appeared localized to the trophozoites visible within the cyst wall (Figure 2.8.A). However, in some cases, the cyst still appeared smooth and undegraded. In these cysts, *Gl*CP2-GFP could be seen under the cyst wall around the periphery of the cyst (Figure 2.8.B).

2.3. Discussion

Because *Giardia lamblia* is an early diverging branch of the eukaryotic evolutionary

tree, as defined by 16S ribosomal RNA sequence and protein coding sequences, it is an intriguing model system to investigate the evolution of protein families and their functions (17,98,99). The clan CA cysteine protease family has essential functions in numerous organisms including well- characterized lysosomal protein degradation and a wide array of other indispensable cellular tasks (60). There are twenty-seven gene sequences for the *Giardia* clan CA cysteine protease family. Twenty-five of these genes are expressed in vegetative and encysting life cycle stages. Of these expressed genes, *Gl*CP2 emerges as the most highly expressed cysteine protease (3.5 fold higher expression than the next most highly expressed gene) and exhibits developmental regulation, with expression increasing dramatically during encystation (Figure 2.1). It is also the only cysteine protease that was identified in protein fractions enriched for cysteine protease activity.

*Gl*CP2 is a cathepsin B-like cysteine protease (37). The protein is produced as a zymogen and is activated by proteolytic removal of an N-terminal pro peptide of fifty-one amino acids. *Giardia* cathepsin B-like cysteine proteases lack the "occluding loop" that is characteristic of cathepsin B-like enzymes of higher eukaryotes (33). This loop endows the protease with exopeptidase activity by stabilizing the free carboxyl at the C terminus of a peptide substrate (100). Therefore, the *Giardia* cathepsin B-like proteins are expected to exhibit only endopeptidase activity. Though the mammalian orthologues of this enzyme are lysosomal enzymes and are optimally active at an acidic pH, *Giardia* cysteine protease activity, and in particular the activity of *Gl*CP2, exhibits optimal substrate degradation at neutral pH (Figure 2.4.C) (59-61).

During purification of enzyme activity, in both native lysates and with r*Gl*CP2, a

unique specificity was seen in fractions containing both the pro enzyme and the mature enzyme not observed in those fractions containing only mature *Gl*CP2. As seen in Figure 2.5.B, Peak A contains predominantly the proform of the protease and exhibits much higher activity against Z-RR-AMC than against Z-FR-AMC. The biochemical basis for this disparity is not currently understood but is likely related to the ability of the proform of the protease to cleave small peptide substrates but not proteins like CWP2. Peak B, which has more mature protease, efficiently cleaves CWP2 (Figure 2.7).

Previously it was reported that a cathepsin C-like, encystation specific cysteine protease (ESCP), was responsible for the essential proteolytic processing of CWP2 from a 39 kDa protein to 26 kDa fragment. This processing step removes a highly basic carboxy terminal domain, allowing polymerization and formation of the cyst wall (8,10,38). However, ESCP has all of the conserved domains of cathepsin C-like proteins including the N-terminal exclusion domain that limits cathepsin C to dipeptidyl exopeptidase activity (59-61,101,102). Therefore, it would not be predicted to possess any endopeptidase activity, such as would be necessary to accomplish CWP2 processing. Recombinant *R. norvegicus* cathepsin C did not process CWP2 (Figure 2.7.B). Because Touz et al. used ESCP purified from *Giardia* lysates and not recombinant protein, the possibility exists that the enzyme preparation was contaminated with one of the other much more abundant *Giardia* clan CA cysteine endopeptidases. In this study we show that *Gl*CP2 is capable of processing CWP2 to the expected 26 kDa fragment and is found in ESVs with CWPs. *Gl*CP2 is expressed at levels 20-fold higher than ESCP during encystation (Figure 2.1) so a small amount of the active *Gl*CP2 could easily contaminate the "purified" preparation of ESCP (95). Indeed, the protease inhibitors demonstrated

by Touz et al. to interfere with cyst production (E64, ALLN, ALLM) (38) also efficiently inhibit *Gl*CP2 activity (Table 2.3). The ability of trypsin and chymotrypsin, but not cathepsin C, to accurately process CWP2 suggests that it is the structure of CWP2 that presents specific endoprotease processing sites. The processing of CWP2 may therefore be redundant, especially as other cysteine endopeptidases in *Giardia* were localized to the ESVs during encystation, such as EAA37074 (Figure 2.6.B). However, the high level of *Gl*CP2 expression strongly suggests that *Gl*CP2 is the key proteolytic constituent of the CWP2 processing machinery.

The fact that *Gl*CP2 can, at high concentrations or long incubation periods, degrade CWP2 to small peptides suggests that there must be a mechanism in place for regulating the activity of *Gl*CP2 against CWP2 in the ESVs. This may be accomplished by delayed activation of *Gl*CP2, since it was found that the zymogen-containing fractions of r*Gl*CP2 did not exhibit any proteolytic activity against CWP2 in the degradation assay (Figure 2.7.B). It could also be regulated by acidification of the ESVs, as it has been suggested that ESVs fuse with PVs prior to formation of the cyst wall and the activity of *Gl*CP2 is greatly reduced in an acidic compartment (38) (Figure 2.4.C). The ability of *Gl*CP2 to completely degrade CWP2 under other specific conditions also supports a second role for this enzyme in the process of excystation, as was previously postulated by Ward et al. (37). Indeed, the presence of *Gl*CP2-GFP around the periphery of cysts following the excystation procedure suggests that *Gl*CP2 may be secreted at the cyst wall to process CWPs and allow trophozoite emergence from the cyst (Figure 2.8).

2.4. Figures and tables for Chapter Two

Figure 2.1: The *Giardia* genome contains twenty-seven genes encoding clan CA cysteine proteases, of which *Gl*CP2 is the most highly expressed as measured by qRT-PCR. 55

Figure 2.2: Two distinct cysteine protease activities were resolved by anion exchange chromatography of *Giardia* lysates. 56

Figure 2.3: *Giardia* CP1 (*Gl*CP1) was heterologously expressed in *P. pastoris*. 58

Figure 2.4: Resynthesized *Gl*CP2 (r*Gl*CP2) was heterologously expressed, purified, and biochemically characterized. 59

Figure 2.5: r*Gl*CP2 was fractionated by anion exchange chromatography and its activity was found to be identical to cysteine protease activities from *Giardia* lysates. 61

Figure 2.6: *Gl*CP2 and other clan CA cysteine proteases co-localized with CWPs in ESVs during *Giardia* encystation. 63

Figure 2.7: r*Gl*CP2 can accomplish the proteolytic processing of recombinant CWP2 (rCWP2) to the predicted 26kDa size shown to be necessary for incorporation into the cyst wall. 64

Figure 2.8: *Gl*CP2-GFP was found around the periphery of cysts following *in vitro* excystation. 65

Table 2.1: Amino acid sequences of peptide fragments identified by LC-MS/MS of cysteine protease activity Peak A and Peak B eluted from anion exchange chromatography of *Giardia* lysates. 66

Table 2.2: *Giardia* cysteine proteases presented resistance to recombinant expression. 67

Table 2.3: Inhibition of r*Gl*CP2 activity against the N-terminally blocked fluorogenic peptide substrate Z-FR-AMC. 68

Figure 2.1: The *Giardia* genome contains twenty-seven genes encoding clan CA cysteine proteases, of which *Gl*CP2 is the most highly expressed as measured by quantitative RT-PCR.

Genes were classified by sequence homology into cathepsin B-like (8), cathepsin C-like (4), and cathepsin K/L-like (10) categories. GenBank accession number EAA38990 represents three greater than 95% identical genes and GenBank accession number AAK92150 represents four genes with greater than 95% homology. Gene expression was measured during vegetative growth and twenty-four h after induction of encystation. Transcript level is expressed relative to the housekeeping gene GAPDH. Two genes, both cathepsins K/L-like, were not expressed (GenBank accession numbers EAA37191 and EAA39030). *Gl*CP2 was the most highly expressed cysteine protease gene in both vegetative and encysting life cycle stages.

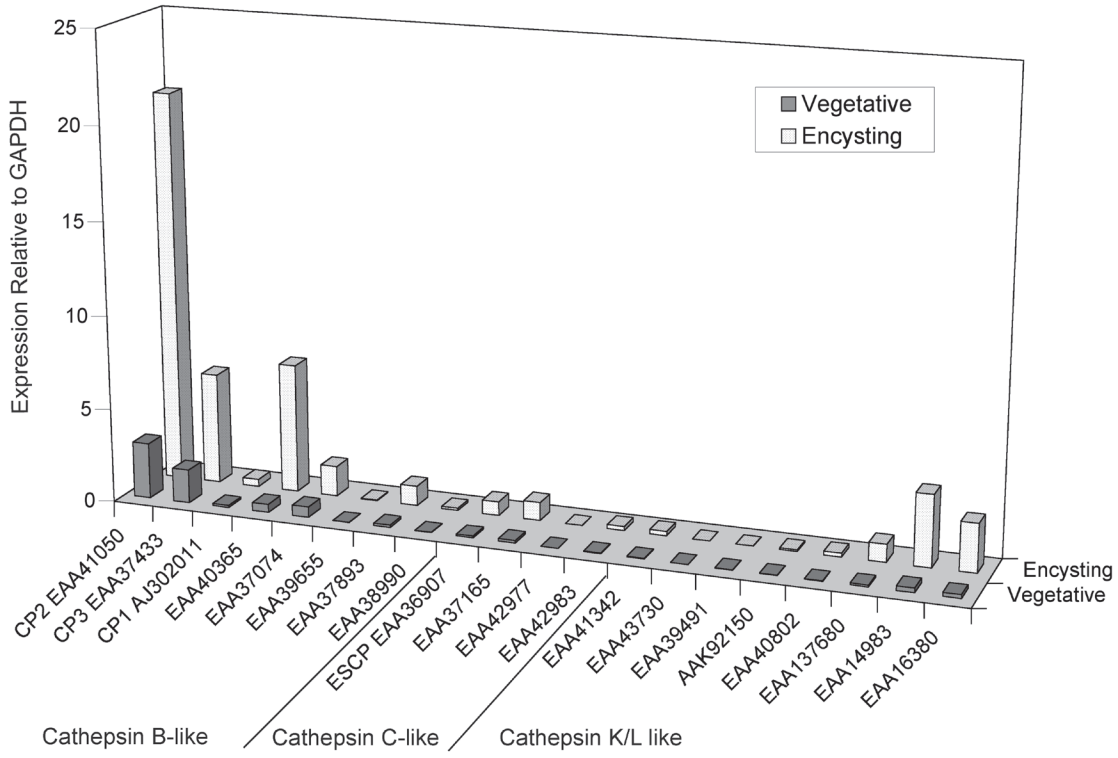


Figure 2.2: Two distinct cysteine protease activities were resolved by anion exchange chromatography of *Giardia* lysates.

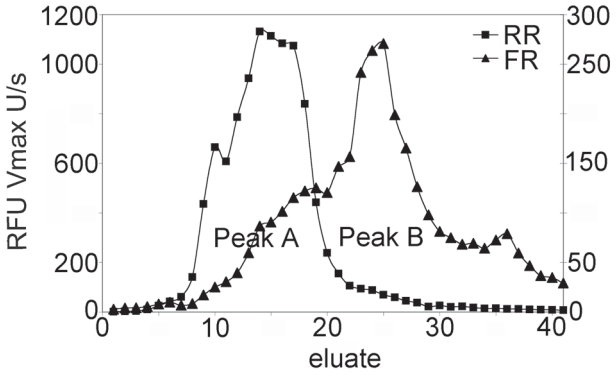
A. Protease activity from eluates (collected from a 0-500 mM NaCl gradient) was tested against the N-terminally blocked fluorogenic peptide substrates Z-FR-AMC (FR ▲) and Z-RR-AMC (RR ■). The two cysteine protease activity peaks exhibited distinct substrate preferences. Peak A had higher activity against Z-RR-AMC, while Peak B had higher activity against Z-FR-AMC and far less activity against Z-RR-AMC.

B. Eluates from the cysteine protease activity Peaks A and B were incubated with the ^{125}I-labeled active site probe DCG04. The probe labeled two protein bands in Peak A while only one band could be seen in Peak B.

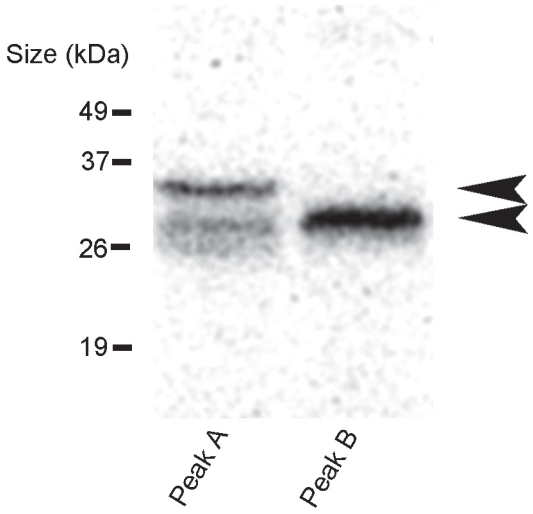
C. A 1D gel of purified and concentrated Peak A and Peak B was silver stained. Numbered bands are those excised and digested with trypsin for mass spectrometric analysis. Black arrowheads indicate gel segments in which *Gl*CP2 was identified.

D. 2D gel electrophoresis of purified and concentrated Peaks A and B. Spots were excised and analyzed by mass spectrometry.

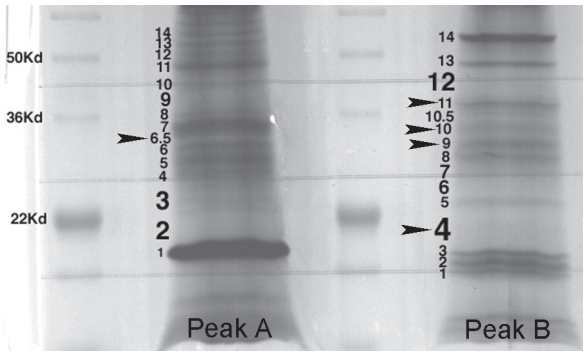
A


1200
1000
600
200
0
RFU Vmax U/s
300
250
150
50
0
RR
FR
Peak A
Peak B
0
10
20
30
40
eluate

B


Size (kDa)
49
37
26
19
Peak A
Peak B

C


50Kd
36Kd
22Kd
Peak A
Peak B

D

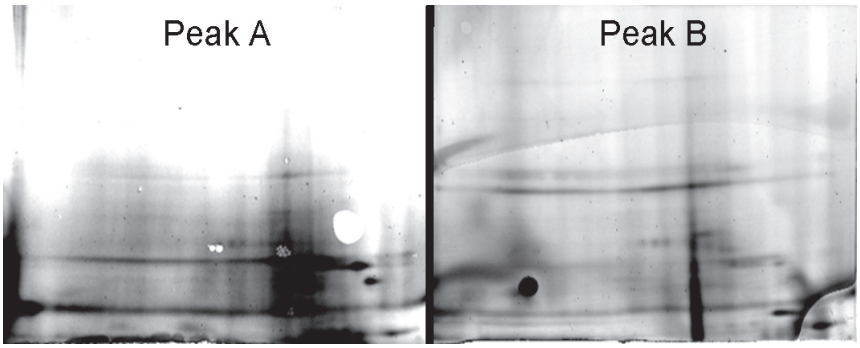

Peak A
Peak B

Figure 2.3: *Giardia* CP1 (*Gl*CP1) was heterologously expressed in *P. pastoris*. Recombinant *Giardia* CP1, labeled with the cysteine protease active site probe ^{125}I-DCG04, was expressed as a zymogen in *P. pastoris* and following SDS-PAGE fractionation ran at the mature size (elongated arrowhead), similar to the cysteine protease activity seen in wild type *Giardia* lysates.

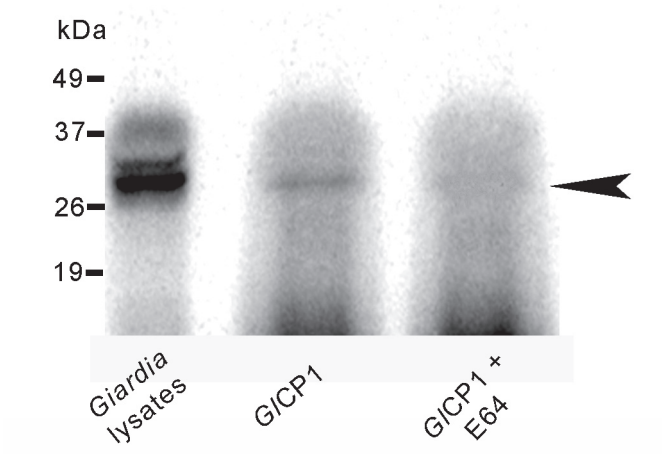


Figure 2.4: Resynthesized *Gl*CP2 (r*Gl*CP2) was heterologously expressed, purified, and biochemically characterized.

A. A western blot probed with anti-*Gl*CP2 peptide antibody showed that r*Gl*CP2 was expressed as a zymogen of 34 kDa. During the process of purification r*Gl*CP2 was autoactivated to the mature form of 28 kDa.
B. Activity of r*Gl*CP2 was observed on a native 10% zymogram (gelatin) gel. Activity is seen in the pro-mature and mature forms of r*Gl*CP2.
C. A pH profile of r*Gl*CP2 against the fluorogenic substrates Z-FR-AMC (▲) and Z-RR-AMC (■) and against the macromolecular substrate casein-resorufin (◆) demonstrated that the pH optimum was 6-7 for activity against all substrates.
D. A positional scanning synthetic combinatorial library (PS-SCL) was used to determine the amino acid preference of the substrate binding sites P1-P4. Sites P1 and P2 exhibited clear selectivity for certain amino acids, while sites P3 and P4 had relaxed specificity. Y-axis is fluorescence, correlating to protease activity (Vmax U/s), normalized by setting the highest activity to 100.

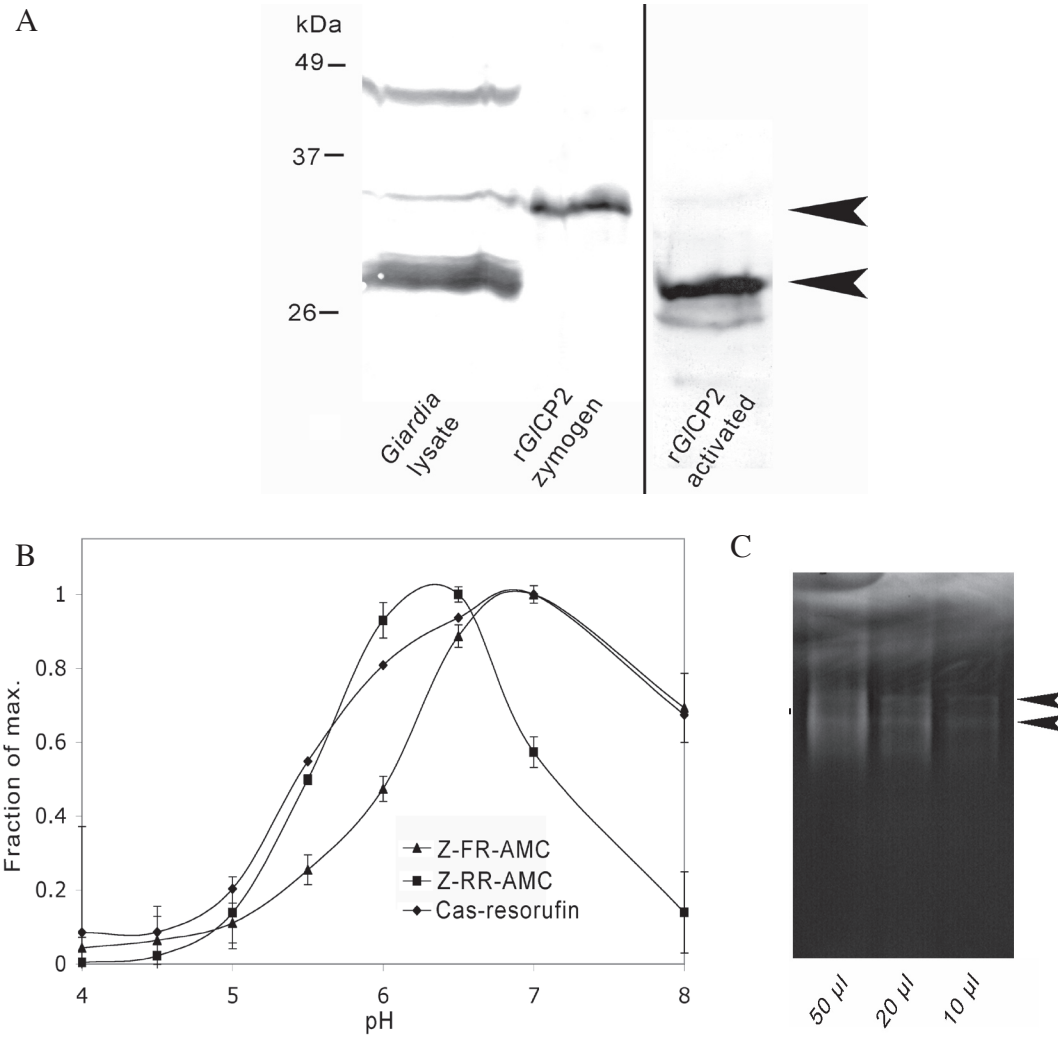

A
kDa
49
37
26
Giardia lysate
rGICP2 zymogen
rGICP2 activated
B
Fraction of max.
pH
Z-FR-AMC
Z-RR-AMC
Cas-resorufin
C
50 µl
20 µl
10 µl
D
P4
P3
P2
P1
KRHDENQSTGPFWYAV I LnM

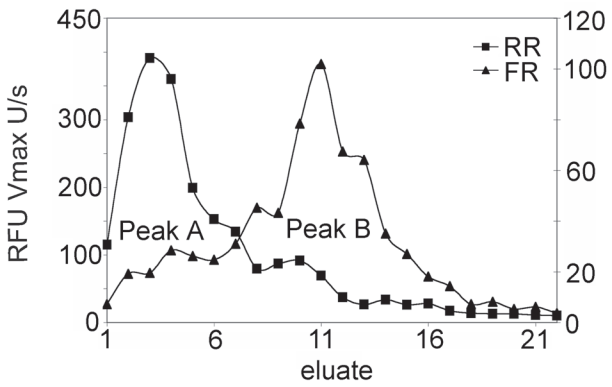
Figure 2.5: r*Gl*CP2 was fractionated by anion exchange chromatography and its activity was found to be identical to cysteine protease activities from *Giardia* lysates.

A. r*Gl*CP2 eluates from anion exchange chromatography (collect from a 0-500 mM NaCl gradient) were tested against the N-terminally blocked fluorescent peptide substrates Z-FR-AMC (FR ▲) and Z-RR-AMC (RR ■). As found with protease activity from lysates, two cysteine protease activity peaks were resolved and exhibited distinct substrate preferences. Peak A had higher activity against Z-RR-AMC, while Peak B had higher activity against Z-FR-AMC and far less activity against Z-RR-AMC.

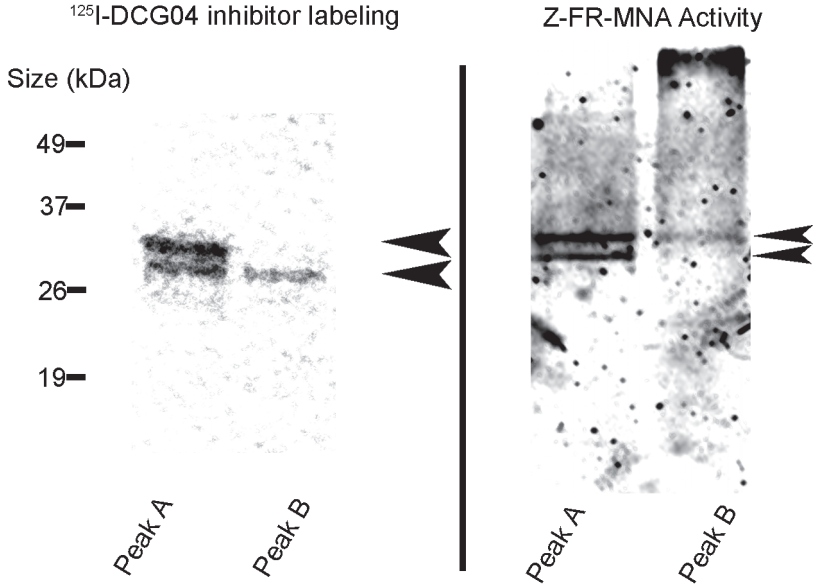
B. Left panel: fractions of cysteine protease activity Peaks A and B were incubated with ^{125}I-DCG04 and fractionated by SDS-PAGE. The probe labeled two protein bands in Peak A while only one labeled protein band could be seen in Peak B (arrowheads). Right panel: Peaks A and B proteins were fractionated by SDS-PAGE under native conditions and the resulting gel was incubated with the fluorogenic peptide substrate Z-FR-MNA. Two bands of fluorescence, indicative of endoprotease activity, were observed in Peak A and one band in Peak B (arrowheads).

C. A western blot probed with anti-*Gl*CP2 peptide antibody demonstrated that two protein bands in Peak A and one protein band in Peak B were identified as *Gl*CP2.

A


RFU Vmax U/s
450
300
200
100
0
120
100
60
20
0
RR
FR
Peak A
Peak B
1
6
11
16
21
eluate

B


125I-DCG04 inhibitor labeling
Z-FR-MNA Activity
Size (kDa)
49
37
26
19
Peak A
Peak B
Peak A
Peak B

C

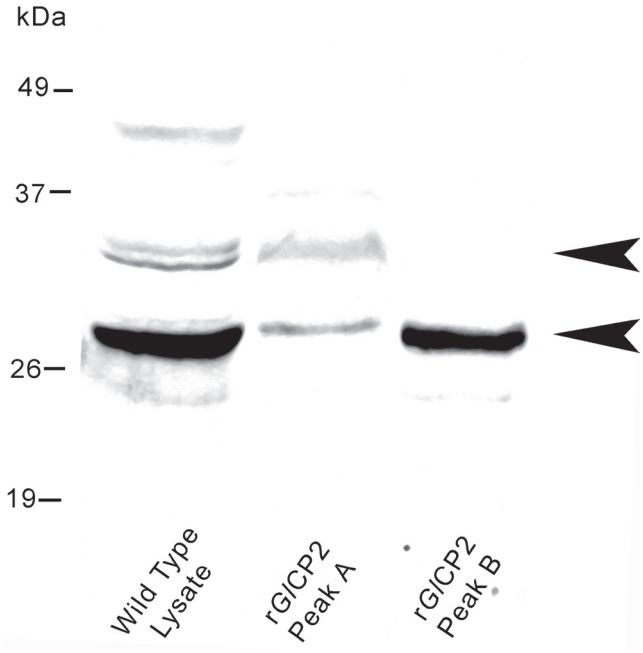

kDa
49
37
26
19
Wild Type Lysate
rGICP2 Peak A
rGICP2 Peak B

Figure 2.6: *Gl*CP2 and other clan CA cysteine proteases co-localized with CWPs in ESVs during *Giardia* encystation.

A. A C-terminal GFP fusion of *Gl*CP2 (green) was expressed episomally in *Giardia*. Cells were fixed twenty-four h after induction of encystation and probed with an antibody against CWPs (red). Image represents a 3D projection of confocal images taken along the Z axis. *Gl*CP2 clearly co-localized with CWP in ESVs. DAPI (blue) marks the nuclei. Bar, 5μm.

B. A C-terminal GFP fusion of *Giardia* EAA37074 was expressed episomally in *Giardia*. Cells were fixed twenty-four h after induction of encystation and probed with an antibody against CWPs. DAPI (blue) marks the nuclei. Bar, 5μm.

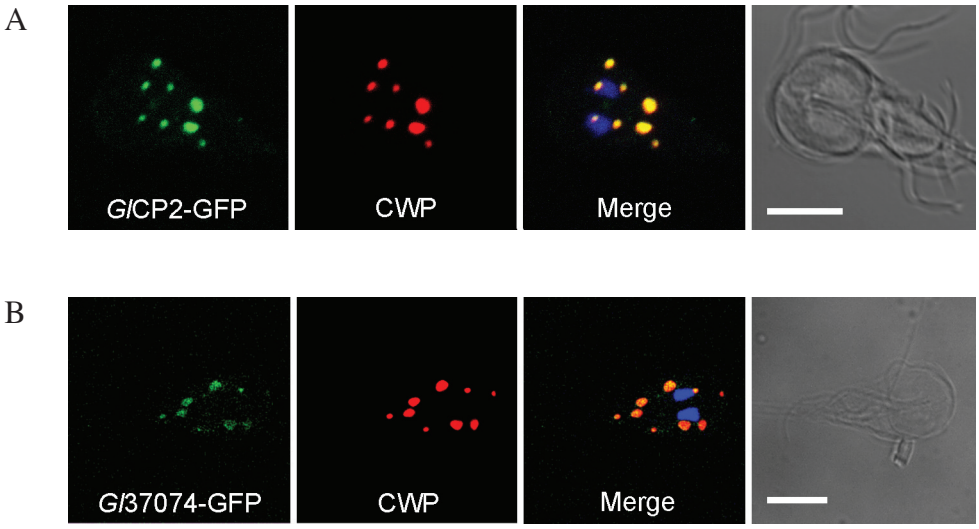
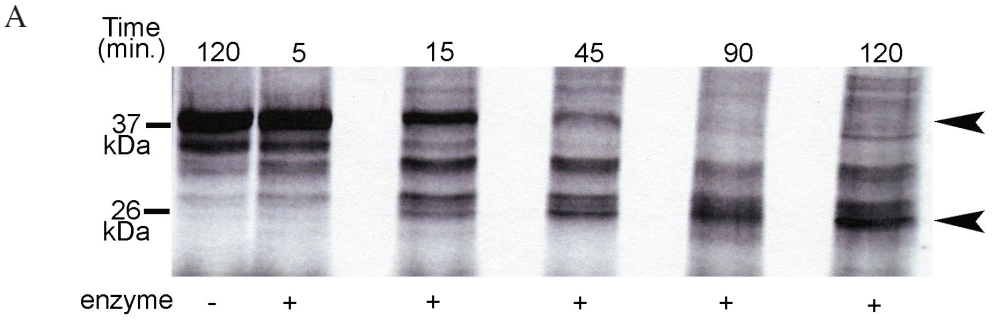


Figure 2.7: r*Gl*CP2 can accomplish the proteolytic processing of recombinant CWP2 (rCWP2) to the predicted 26kDa size shown to be necessary for incorporation into the cyst wall.

A. In solution cleavage of ^{35}S-labeled rCWP2 by purified r*Gl*CP2. rCWP2 was exposed to r*Gl*CP2 in Tris buffer, pH 7.6, and incubated at 25 °C for 0, 5, 15, 45, 90, and 120 min r*Gl*CP2 can cleave the 39kDa rCWP2 to a 26kDa fragment (arrowheads) in a time dependent manner.

B. rCWP2 was incubated for 35 min with a variety of proteases. Peak A from anion exchange purification of r*Gl*CP2 (Lane 1) and *R. norvegicus* cathepsin C (Lane 6) did not process rCWP2 from its full-length size of 39 kDa (arrowhead). r*Gl*CP2 (Lane 2) and *Giardia* lysates (Lane 3) exhibited identical processing patterns of rCWP2. Other endopeptidases, such as trypsin (Lane 4) and chymotrypsin (Lane 5), can also process rCWP2 to 26kDa (arrowhead). Lanes 5 and 6 are set apart as these samples were run on separate gels.



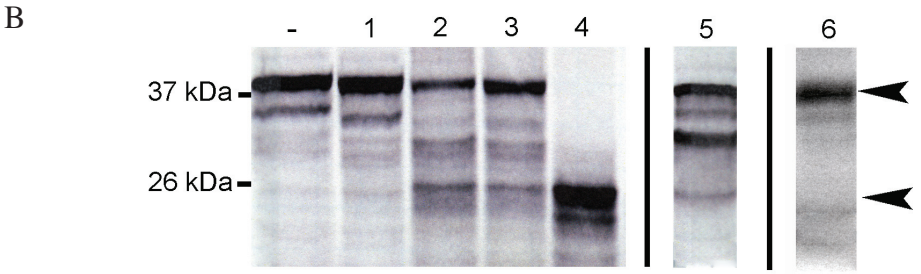
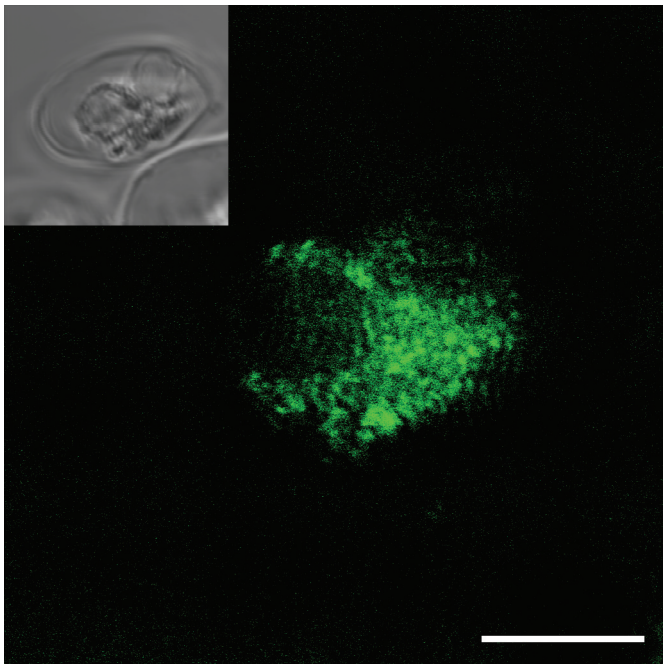


Figure 2.8: *Gl*CP2-GFP was found around the periphery of cysts following *in vitro* excystation.

A. Many cysts appeared to have degraded contents and *Gl*CP2-GFP was still observed within trophozoites inside of the cysts. Bar, 5 μm.

B. In some cysts, *Gl*CP2-GFP was observed just inside the periphery of the cyst, apparently excluded completely from the trophozoites. Bar, 5 μm.

A



B

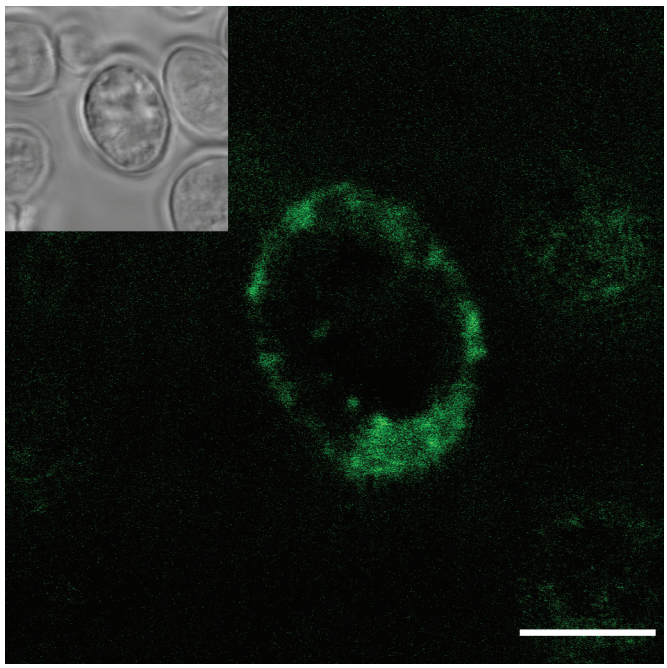


Table 2.1: Amino acid sequences of peptide fragments identified by LC-MS/MS of cysteine protease activity Peak A and Peak B eluted from anion exchange chromatography of *Giardia* lysates.

Peak A	Peak B
CVAGLDK	CVAGLDK
TGTTTDECVPYK	TGTTTDECVPYK
	VHLATATSYK
	DYGLDIPAMMK
	GINDCSIEEQAYAGFFDE
	NSWGPDWGEDGYFR
	SGSTTLR
	GTCPTK
	CADGSSK

Expression system	Organism	Protein (GenBank acc.)	Protein produced	Active protease
pET-28a(+) (Novagen)	*E. coli*	CAC18646	X	X
pET-28a(+) (Novagen)	*E. coli*	EAA37433	X	X
pET102/DTopo® (Invitrogen)	*E. coli*	EAA37433	X	X
pMal-p2X (NEB)	*E. coli*	EAA37433	X	X
pMal-c2X (NEB)	*E. coli*	EAA37433	*	X
pMal-c2X (NEB)	*E. coli*	EAA37074	*	X
TNT® Rabbit reticulocyte (Promega)	Cell free	EAA37433	*	X
TNT® Rabbit reticulocyte (Promega)	Cell free	EAA37074	*	X
TNT® Wheat germ (Promega)	Cell free	EAA37433	*	X
TNT® Wheat germ (Promega)	Cell free	EAA37074	*	X
pTub.pac VSP fusion	*G. lamblia*	EAA37433	X	X
pGSTag	*E. coli*	CAC18646	X	X
pXG (†)	*L. donovani*	EAA37433	*	X
pIEX-2 (EMD biosciences)	High FiveTM (Invitrogen)	EAA37433	X	X
pPicZaB (Invitrogen)	*P. pastoris*	EAA37433	X	X
pPicZaB (Invitrogen)	*P. pastoris*	resynthesized EAA37433	*	*
pPicZaB (Invitrogen)	*P. pastoris*	CAC18646	*	*
pPicZaB (Invitrogen)	*P. pastoris*	EAA37074	X	X
pET102/DTopo® + Chaperones (Takara Mirus Bio)	*E. coli*	EAA37433	X	X
pTub.pac	*G. lamblia*	EAA37433	X	X

† Ha, DS, Schwarz, JK, Turco SJ, and Beverley SM (1996), "Use of the green fluorescent protein as a marker in transfected *Leishmania*", *Mol. Biochem. Parasitol*. 77:57-64.

Table 2.2: ***Giardia*** **cysteine proteases presented resistance to recombinant expression. Many systems were used for the attempted expression of several** ***Giardia*** **cysteine protease. Though protein was produced by some methods, soluble and active protease was only obtained by one of the many methods utilized.**

Table 2.3: Inhibition of r*Gl*CP2 activity against the N-terminally blocked fluorogenic peptide substrate Z-FR-AMC. Activity is expressed as percent activity relative to a control reaction. r*Gl*CP2 was preincubated with inhibitor 10 min before activity was assayed.

Inhibitor	% Activity	Inhibitor	% Activity
PMSF 1 mM	113	E64 1 μM	0
EDTA 10 mM	47	E64 10 μM	0
Aprotinin 10 μg/ml	93	CA074 1 μM	77
Pepstatin A 1 μM	100	CA074 10 μM	34
Pepstatin A 10 μM	118	Lactacystin 1 μM	100
Leupeptin 1 μM	1	Lactacystin 10 μM	119
Leupeptin 10 μM	0	α1antitrypsin 1 μM	90
TLCK 1 μM	38	α1antitrypsin 10 μM	74
TLCK 10 μM	17	ALLN 1 μM	5
TPCK 1 μM	50	ALLN 10 μM	5
TPCK 10 μM	25	ALLM 1 μM	6
		ALLM 10 μM	5

Chapter 3: Clathrin may facilitate the initial uptake of endocytosed proteins and play a role in ESV maturation in *Giardia*.

3.1. Introduction

3.1.1. Clathrin and cargo transport

The transfer of protein cargo in the endocytic and secretory systems of eukaryotic cells often involves vesicle transport between donor and acceptor membranes. The process of vesicle formation begins by the nucleation of a budding vesicle on a membrane. Cargo to be transported becomes concentrated at a location on the donor membrane. When a protein coat has assembled on this patch, the membrane becomes deformed and the fully formed vesicle leaves the donor membrane whereupon the protein coat is removed and the vesicle traffics to the acceptor membrane (103). Several different coat proteins, such as clathrin, copI, and copII, and their corresponding adaptor proteins can form vesicles and facilitate vesicle mediated protein trafficking (104,105). In most eukaryotic cells studied to date, copI mediates intra-Golgi transport and also possibly retrograde transport from the Golgi apparatus to the ER(106). CopII shuttles proteins from ER exit sites towards the Golgi (107). Clathrin is the most widely studied and well characterized of these coat proteins and is involved in numerous cell processed including endocytosis, secretion, receptor recycling, and even mitosis (103,108,109). In many cells, clathrin-mediated endocytosis represents the major pathway of entry into the cell (110).

Clathrin knockouts have further emphasized the importance of this coat protein to the normal functioning of eukaryotic cell (108,111). Yeast cells deficient in clathrin

heavy or light chains display a delay in endocytosis and mislocalize Golgi proteins, though some plasticity in trafficking does occur, with TGN to endosome traffic restored via an alternate route (112,113). It is notable, however, that yeast cells deleted for clathrin can continue to endocytose proteins by a receptor-mediated mechanism (114). In *D. discoidium,* amoebae lacking clathrin were demonstrated to lack coated pits, coated vesicles and large vacuoles that serve as endosomes and contractile vacuoles (115). They were also demonstrated to be defective in the endocytosis of fluid phase markers, osmoregulation, morphogenesis, and regulation of cell polarity (115,116). Another mechanism for disrupting the function of the clathrin heavy chain (CLH) is the expression of the C-terminal third of the gene, called "hub" (117,118). Hub expression acts as a dominant negative clathrin inhibitor by sequestering the available clathrin light chain. This has been used in mammalian cells to show the involvement of clathrin in trafficking from the TGN and endocytosis (117,119,120). Hub expression has also defined a role for clathrin in maintaining the cellular distribution of early endosomes (121).

The clathrin coated vesicle (CCV) is generally a tri-level structure, with a cage of clathrin triskelions forming the outermost stratum. Each clathrin triskelion is composed of three kinked CLHs that radiate from a central vertex and three clathrin light chains bound to the heavy chains (Figure 3.1) (1,108). In addition to the conventional CLH, a CLH homologue was recently discovered in humans that is enriched in muscle tissue and exhibits distinct intracellular function and distribution (122). The N-terminal domains of the CLHs are directed into the vesicle and are bound to a variety of adaptor proteins. These adaptor proteins bind both clathrin triskelions and membrane-bound cargo

receptors, which form the innermost layer of the CCV, thereby mediating the recruitment of both membrane and cargo molecules into the vesicle (1).

3.1.2. Selectivity of cargo recruitment

The selective recruitment of cargo is determined by many factors. Probably the most widely studied selectivity is at the level of the binding specificity and localization of the clathrin adaptor proteins. At least twenty clathrin adaptor proteins have been explored, the first-described class being the heterotetrameric adaptor protein complexes (APs) (originally named assembly polypeptides) (103,123). The four known APs exhibit a similar organization, but differ in their limited membrane localization. AP-1, AP-3, and AP-4 function at the TGN and/or the endosomal/lysosomal sorting pathway while AP-2 functions primarily at the plasma membrane (124-126). Solving the structure of these APs elucidated several folded domains and unstructured regions that contain multiple peptide-binding motifs (109,123,127). The affinity of these motifs for domains on cargo proteins can enrich vesicles for specific cargo. These domains can also bind phosphoinositides, which were proven to be important in recruiting AP complexes to membranes (111). The restricted localization of various phosphoinositides provides another layer of specificity in the determination of cargo trafficking and localization. Other clathrin adaptor proteins have been identified, such as the Golgi-localized gamma ear containing ARF-binding proteins (GGAs) and stonins that differ greatly in structure from the APs and yet have similar folds and can therefore catalyze the mechanisms of cargo recruitment and clathrin binding (111,128,129).

Orthologues of the coat proteins copI, copII, and the CLH have been

identified in the *Giardia* genome and their localization has been determined using specific antibodies. During encystation, an antibody against the copI subunit β cop localized to early ESVs and additional endomembrane structures. An antibody against the copII-specific GTPase sar1p labeled the ER/nuclear envelope structure. A specific antibody against the *Giardia* CLH (*Gl*CLH) localized to the PVs during early encystation and to the ESVs during mid- and late-stages of encystation (130). It was recently suggested that clathrin is involved in lipid uptake by *Giardia* (73). Beyond these studies, very little has been done to characterize *Giardia* clathrin and define the role that it plays in the life of the cell. In this study antibody labeling, reporter constructs, and dominant negative hub expression were used to describe the role of clathrin in the processes of endocytosis and encystation.

3.2. Results

3.2.1. Clathrin localizes to the PVs in vegetative trophozoites and this localization is distinct from that of other *Giardia* coat proteins.

The acidified PVs were previously thought to represent a lysosome-like compartment where endocytosed proteins are degraded (16). The PVs unexpectedly contained abundant *Gl*CLH that co-localized with Lucifer yellow, the marker initially used to identify the PVs (Figure 1.8.B). A concentration of PVs was observed near the site of flagellar attachment in the region of the ventral groove corresponding to the site of initial Fluosphere uptake observed in Chapter 1 (Figure 1.3.B). A C-terminal fusion GFP reporter construct expressed in trophozoites (*Gl*CLH-GFP) further indicated that *Gl*CLH is localized in the area of the PVs (Figure 3.2.A). The previously mentioned

concentration of clathrin near the site of flagellar attachment can also be seen in *Gl*CLH-GFP transgenic trophozoites. This localization differs somewhat from the localization pattern seen with antibodies against *Giardia* orthologues of a copI subunit or the copII-specific GTPase sar1p. CopI appears to remain near the periphery of the cell as well, but the staining pattern is less broad and continuous than that seen with *Gl*CLH and copI is not seen near the flagellar exit site (Figure 3.2.B). *Giardia* sar1p localizes to the ER/nuclear envelope region of the cell, clearly distinct from the pattern of clathrin localization (Figure 3.2.C).

3.2.2. Clathrin remains PV associated during endocytosis.

To determine if the PV associated localization of clathrin was altered during endocytosis, *Giardia* cells were incubated with DQRed BSA and live cell imaging was performed during endocytosis. The localization of clathrin appeared to remain quite stable and little colocalization was observed between clathrin and DQRed BSA (Figure 3.3.A). Clathrin was previously implicated in lipid endocytosis, namely ceramide (73). *Gl*CLH-GFP transgenic cells indeed readily endocytosed FITC-labeled ceramide. Again, *Gl*CLH-GFP remained localized to the periphery of the cell throughout the endocytic process (Figure 3.3.B). In humans, a CLH isoform was identified (CHC22) with an expression pattern and function distinct from well-characterized CCVs. To determine if the structure and function of *Giardia* CLH may be similar to CHC22, a C-terminal GFP fusion of *H. sapiens* CHC22 was expressed in *Giardia* trophozoites. The distribution of CHC22-GFP was markedly different than the distribution of *Gl*CLH-GFP. Much appeared localized to the cytoplasm, while more intense staining was seen at the base of the flagella (Figure

3.3.C).

3.2.3. Clathrin hub expression did not affect clathrin localization or endocytosis in vegetative trophozoites.

*Gl*CLHhub-GFP, comprising the C-terminal 780 (of 1870) amino acids of the CLH fused to GFP, was expressed in *Giardia* trophozoites. The localization of *Gl*CLHhub-GFP was indistinguishable from that of *Gl*CLH-GFP. An antibody against the consensus sequence of light chain (131) appeared to co-localize with both *Gl*CLH-GFP and *Gl*CLHhub-GFP (Figures 3.4.A and 3.4.B). As clathrin has been suggested to play a role in uptake of proteins and lipids in *Giardia*, the affect of the hub expression on the process of endocytosis was investigated. *Gl*CLHhub-GFP transgenic trophozoites endocytosed DQRed BSA at the same rate as *Gl*CLH-GFP transgenic cells and the localization patterns of the BSA and clathrin constructs were identical between these two cell types (Figures 3.5.A and 3.5.B). Ceramide uptake was again investigated in *Gl*CLHhub-GFP transgenic cells and was found to be indistinguishable from wild type cells (Figure 3.5.C).

3.2.4. Clathrin hub expression did not significantly affect clathrin localization but may affect cyst formation in encysting *Giardia*.

*Gl*CLH-GFP transgenic trophozoites were induced to encyst *in vitro* and cells were probed with an antibody against CWPs. *Gl*CLH was distinctly localized to ESVs in twenty four h encysting cells and to the cyst wall in newly formed cysts at 24 h encystation, suggesting a role for clathrin in the encystation process (Figure 3.6.A).

The presence of *Giardia* clathrin in ESVs suggests that it has an essential function in trafficking of CWPs to the cell surface. To ascertain if hub expression affected the process of encystation, *Gl*CLHhub-GFP transgenic trophozoites were induced to encyst *in vitro* and the localization and effect of hub expression were investigated. *Gl*CLHhub-GFP clearly localized to ESVs in 24hr encysting cells and to the cyst wall in newly formed cysts as was seen with *Gl*CLH-GFP (Figure 3.6.B). Some of the cysts did show signs of being poorly formed, such as abnormal shapes and incompletely formed cyst walls (Figure 3.6.C).

3.2.5. Electron microscopy shows invaginations of the plasma membrane, but the characteristic clathrin coat is not seen on the forming vesicles.

As has been reported previously, membrane invaginations occur around the plasma membrane near the PVs. (Figure 3.7.A). Clathrin was demonstrated by immunofluorescence microscopy to be localized around the periphery of *Giardia* trophozoites, co-localized with the PVs (Figures 3.2.A and 1.8.B). Clathrin also appeared to localize in and around the PVs by immuno-EM (Figure 3.7.B). However, electron microscopy of PVs also did not clearly reveal any characteristic clathrin coat on these large vacuoles (Figure 3.7.C).

3.3. Discussion

In most eukaryotic cells, CLH is found on vesicles trafficking between the plasma membrane and organelles involved in the endocytic and secretory systems. In *Giardia*, the CLH is found associated with the large, acidified PVs. This pattern of

clathrin localization in *Giardia* differs from what is seen in many eukaryotic cells. Immunocytochemistry studies have shown that clathrin in mammalian cells is found in numerous puncta at the plasma membrane, in the cytoplasm and in an accumulation at the Golgi apparatus (108). Total internal reflection fluorescence microscopy (TIR-FM) has been used to see the dynamics of clathrin coated pit formation at the plasma membrane. In one minute of imaging, roughly fifteen percent of visible clathrin coated pits at the plasma membrane disappeared, presumably internalized (132). Though TIR-FM has not been used to visualize dynamics of *Giardia* clathrin, the punctate pattern characteristic of vesicular localization was not observed. In contrast, *Giardia* clathrin appears to remain in a stable confirmation even during endocytosis, with little co-localization apparent between clathrin and protein endocytosed by vegetative cells (Figure 3.3.A).

The stable presence of clathrin on PVs may at first seem surprising given their relatively large size (50-200 nm). However, CHC22 also exhibits an expression pattern that is distinct from well-characterized clathrin-coated vesicles in mammalian myocytes (133). Therefore, there is precedent for unconventional and yet functional clathrin localization. As an ancient eukaryote, the organization of the endocytic system may in fact predate what is thought to be conventional. Furthermore, recent analysis of ceramide uptake highlighted a clathrin-mediated pathway from peripheral vesicles to the perinuclear space so clathrin may be mainly involved in the initial steps of endocytosis in *Giardia* (73). As shown in Chapter 1, dynamic fusion of ER tubules may be the chief mechanism for the movement of proteins after they are endocytosed by CCVs into the PVs. The presence of clathrin inside of the PV lumen (Figure 3.7B) suggests that the PVs may contain membranes and possibly smaller vesicles. This may be analagous to the

formation of multivesicular bodies.

Expression of clathrin hub in eukaryotic cells has been shown to decrease the rate of receptor mediated endocytosis at the plasma membrane (121). In the cells expressing hub, the localization of the clathrin heavy and light chains was also disrupted, being diffuse and nonvesicular, implying its cytosolic distribution (117). In *Giardia*, no effect was observed of *Gl*CLHhub-GFP expression on endocytosis of protein (DQRed BSA) or lipid (Ceramide-FITC) substrates (Figure 3.5). There are several possible explanations for this. First, the level of episomal expression of the *Gl*CLHhub-GFP construct cannot be regulated. Though it is under the control of the *Gl*tubulin promotor, the level of expression is highly variable even within a cell population (unpublished observation) and therefore may not be high enough to disable clathrin function. Additionally, light chain co-localization with *Gl*CLHhub-GFP shows that the light chain is not being sequestered in the cytoplasm as has been shown for hub expression in mammalian cells (Figure 3.4.B). Since the localization of *Gl*CLHhub-GFP and the clathrin light chain appear unaffected, even the *Gl*CLHhub-GFP construct may be able to functionally trimerize with wild type CLH and also bind to clathrin light chain and join an operative clathrin lattice.

*Gl*CLHhub-GFP expression did not appear to affect early stages of encystation, as *Gl*CLHhub-GFP co-localized with CWPs in ESVs and was indistinguishable from wild type *Gl*CLH. The only possible encystation effect arose very late in encystation at the point of cyst formation. Some of the cysts formed by *Gl*CLHhub-GFP transgenic cells were small, unevenly shaped, and incompletely closed compared to the *Gl*CLH-GFP cells (Figure 3.6.C). In mammalian cells, evidence suggests that clathrin plays a role in trafficking mature proteases away from forming secretory granules, thus maturing

them and protecting the contents from aberrant proteolysis (134). It may be the case, in *Giardia* cells, that clathrin mediates the removal of unwanted material from the ESVs (the *Giardia* equivalent of secretory vesicles) and enables their maturation. If this process was inhibited, the result would be an increase in protolytic activity in the ESV, which may be causative of the phenotypes seen in *Gl*CLHhub-GFP cysts. However, in the experiments described here, *Gl*CLH-GFP vesicles lacking CWP (as would be expected for vesicles trafficking unwanted material out of ESVs) were not observed. Additional evidence comes from the laboratory of Dr. Adrian Hehl at the University of Zurich, Switzerland. Their laboratory also expressed *Gl*CLHhub from a genome-integrated construct and observed normal localization of *Gl*CLHhub in vegetative cells and did not detect any abnormalities cysts. However, their hub construct failed to localize to ESVs in encysting cells (Adrian Hehl, personal communication). The isolation of hub from ESVs may have prevented the deleterious effects on cyst formation that were observed in this thesis. Another interpretation consistent with the results presented here is that clathrin may play a role in organizing the CWP to assist in its polymerization.

Although by immunofluorescent microscopy *Gl*CLH is localized near the PVs, EM studies have been unable to detect a classic clathrin coat on the PVs or the invaginations of the plasma membrane near the PVs (Figure 3.7.C). This is not entirely suprising, as the attempts to visualize clathrin-coated pits and clathrin-coated vesicles in ultrastructural studies of wild-type yeast have generally been unsuccessful (114). Furthermore, the stable localization of *Gl*CLH suggests that its structure and function may differ from the classical mammalian CLH and it may therefore appear different in EM studies. In this study, immunoEM was used to localize clathrin in and around

the PVs, though the level of staining was low (Figure 3.7.B). Further studies should be undertaken to more thoroughly describe the function of *Giardia* clathrin and better understand the evolution of the system of clathrin-mediated trafficking widely used by higher eukaryotic organisms.

3.4. Figures and tables for Chapter Three

Figure 3.1: Structure of the clathrin triskelion. 79

Figure 3.2: *Giardia* clathrin heavy chain (*Gl*CLH) is localized to PVs and this localization is distinct from other coat proteins. 80

Figure 3.3: *Gl*CLH remains localized to the PVs during endocytosis, but does not resemble the expression pattern of *H. sapiens* CHC22. 81

Figure 3.4: Both *Gl*CLH-GFP and *Gl*CLHhub-GFP co-localize with an antibody against the consensus sequence of clathrin light chain (from many model organisms). 83

Figure 3.5: *Gl*CLHhub-GFP expression did not disrupt endocytosis of proteins or lipids in *Giardia* trophozoites. 84

Figure 3.6: *Gl*CLHhub-GFP expression did not disrupt ESV formation or clathrin localization during early encystation, but caused irregularities in cyst formation. 86

Figure 3.7: Ultrastructural analysis showed membrane pits forming very near PVs and clathrin localized in and around PVs. 88

Figure 3.1: Structure of the clathrin triskelion.

Each triskelion is composed of 3 clathrin heavy chains and 3 clathrin light chains. The light chains bind to the C-terminus and proximal leg of the heavy chains.

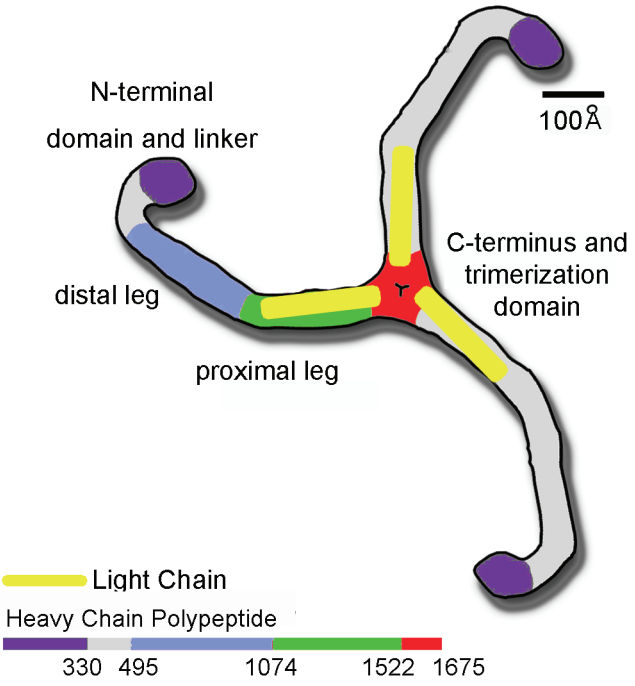


Figure 3.2: *Giardia* clathrin heavy chain (*Gl*CLH) is localized to PVs and this localization is distinct from other coat proteins.

A. *Gl*CLH-GFP (green) was observed around the periphery of *Giardia* trophozoites, corresponding to the localization of the PVs. DAPI (blue) labels the nuclei. Bar, 5 μm.

B. The *Giardia* copI orthologue staining pattern (red) was less broad and continuous than that of *Gl*CLH. Additionally, *Giardia* copI was not seen at the flagellar exit site as seen for *Gl*CLH (Figure 9). DAPI (blue) labels the nuclei. Bar, 5 μm.

C. The *Giardia* Sar1p orthologue (copII subunit, red) localized around the nuclei (labeled blue by DAPI staining), differing greatly from the localization of *Gl*CLH. Bar, 5 μm.

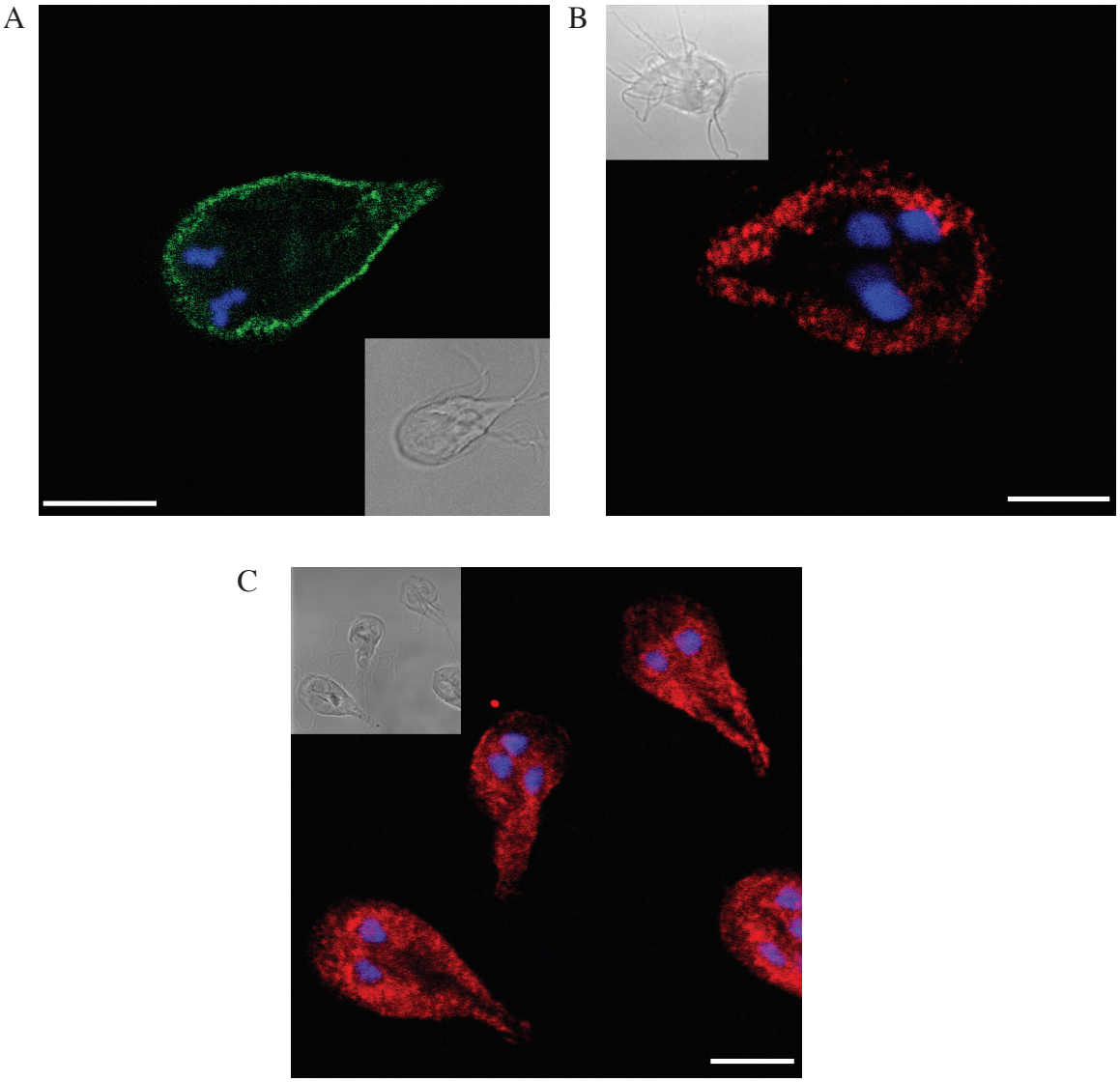


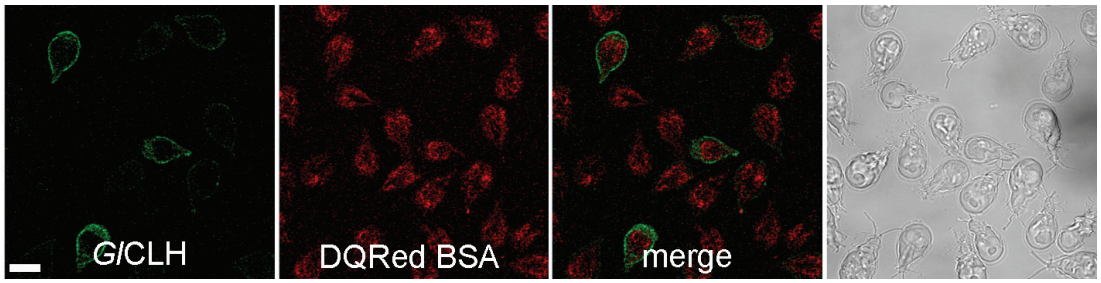
Figure 3.3: *Gl*CLH remains localized to the PVs during endocytosis, but does not resemble the expression pattern of *H. sapiens* CHC22.

A. *Gl*CLH-GFP remained stably localized to the periphery of trophozoites during live cell imaging of DQRed BSA uptake. No apparent co-localization was observed between *Gl*CLH-GFP and DQRed BSA. Bar, 5 μm.

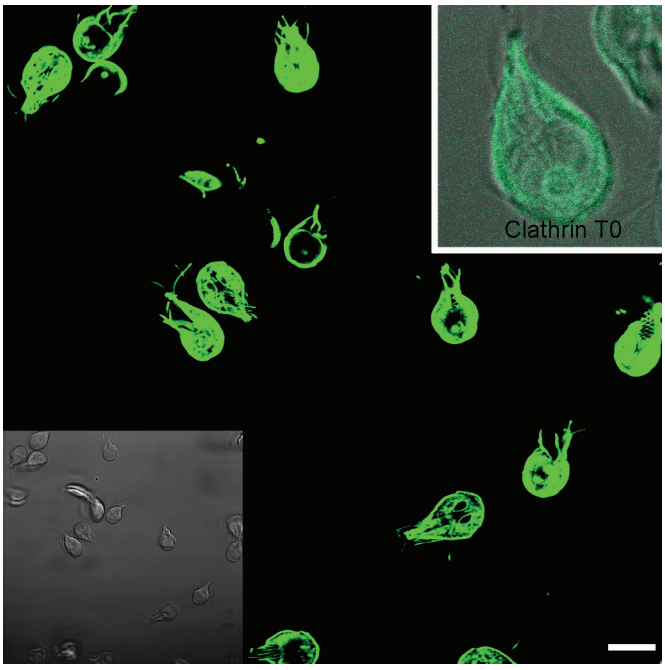
B. *Gl*CLH-GFP (green) also localized to PVs during live cell imaging of ceramide-FITC (green) endocytosis. Transfectant cell at time zero shown in upper right inset. Bar, 5 μm.

C. CHC22-GFP was expressed in trophozoites and was found to largely localize to the cytoplasm while more intense staining was seen at the base of the flagella (white arrowhead). The localization pattern of *Gl*CLH-GFP bears no resemblance to this. DAPI (blue) labels the nuclei. Bar, 5 μm.

A



B



C

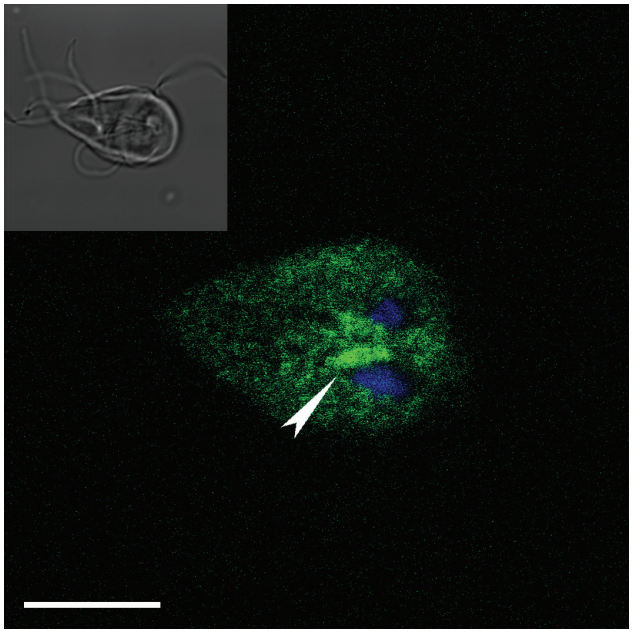


Figure 3.4: Both *Gl*CLH-GFP and *Gl*CLHhub-GFP co-localize with an antibody against the consensus sequence of vertebrate clathrin light chains (131).

A. *Gl*CLH-GFP (green) was observed co-localizing with clathrin light chain (red) along the periphery of the cell in three dimensional reconstructions of image stacks taken along the Z-axis. DAPI (blue) labels the nuclei. Bar, 5 μm.

B. *Gl*CLHhub-GFP expression did not disrupt the light chain localization. *Gl*CLHhub-GFP (green) co-localized with the clathrin light chain (red) along periphery of the cell in three dimensional reconstructions of image stacks taken along the Z-axis. DAPI (blue) labels the nuclei. Bar, 5 μm.

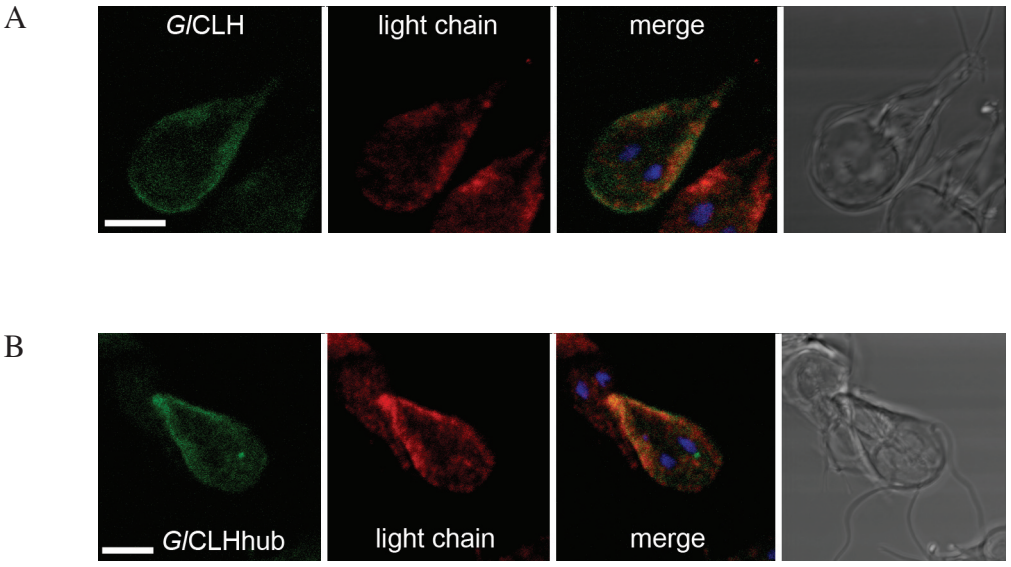


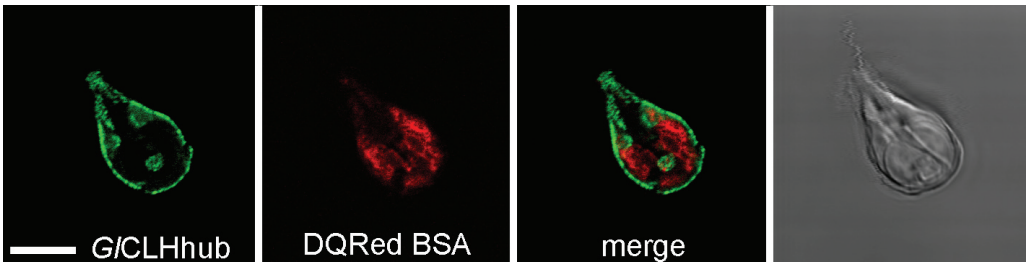
Figure 3.5: *Gl*CLHhub-GFP expression did not disrupt endocytosis of proteins or lipids in *Giardia* trophozoites.

A. *Gl*CLHhub-GFP (green) was localized near the PVs, identical to the localization of *Gl*CLH. Following uptake of DQRed BSA by trophozoites, no co-localization was observed between the BSA (red) and *Gl*CLHhub-GFP (green). Bar, 5 μm.

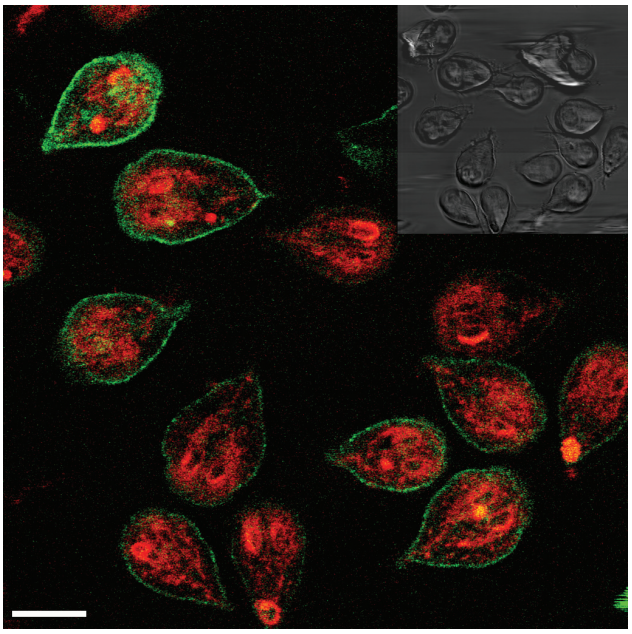
B. Live cell imaging further exemplified that the *Gl*CLHhub-GFP (green) remained stably localized to the PVs during endocytosis and did not co-localize with endocytosed DQRed BSA (red). Bar, 5 μm.

C. Ceramide-FITC (green) uptake by trophozoites expressing *Gl*CLHhub-GFP (green) was indistinguishable, both in the rate of uptake and the localization pattern, from wild type cells. Transfectant cell at time zero shown in upper left inset. Bar, 5 μm.

A



B



C

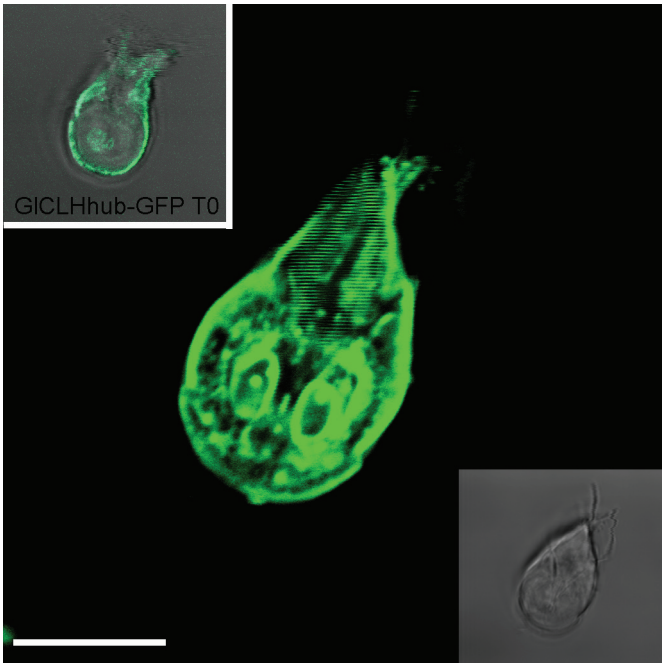


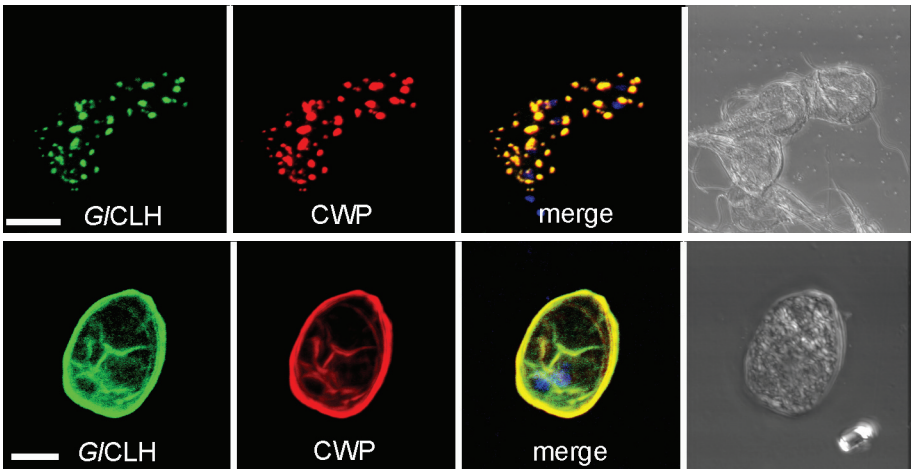
Figure 3.6: *Gl*CLHhub-GFP expression did not disrupt ESV formation or clathrin localization during early encystation, but may cause some irregularities in cyst formation.

A. *Gl*CLH-GFP (green) co-localized with CWPs (red) in ESVs during early encystation (top) and at the cyst wall during late encystation/cyst formation (bottom). Figures are projections of images taken along the Z-axis. DAPI (blue) labels the nuclei. Bar, 5 μm.

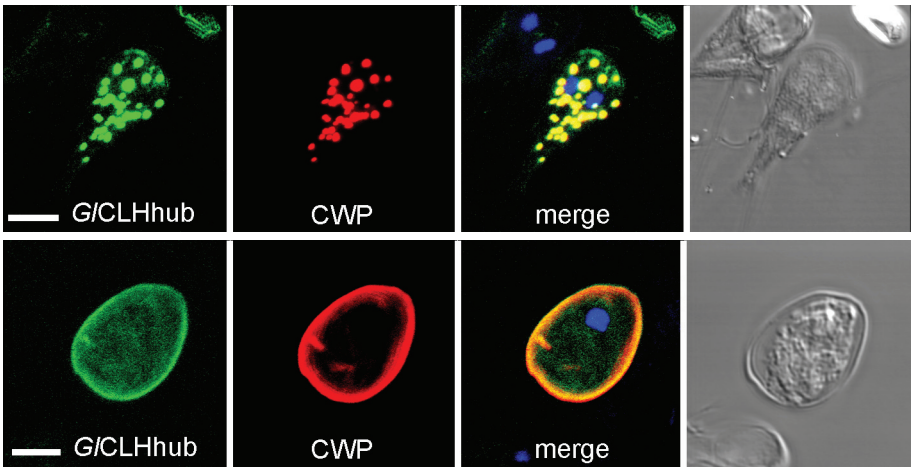
B. *Gl*CLHhub-GFP (green) co-localized with CWPs (red) in ESVs during early encystation (top) and at the cyst wall during late encystation/cyst formation (bottom). Figures are projections of images taken along the Z-axis. DAPI (blue) labels the nuclei. Bar, 5 μm.

C. Cysts in *Gl*CLHhub-GFP (green) expressing trophozoites often appeared up to 50% smaller than cysts in *Gl*CLH-GFP transgenic cells. Additionally, some cysts were incompletely formed, with unretracted flagella and membrane blebbing visible at points along the cyst wall (white arrowheads). DAPI (blue) labels the nuclei. Bar, 5 μm.

A


GlCLH
CWP
merge
GlCLH
CWP
merge

B


GlCLHhub
CWP
merge
GlCLHhub
CWP
merge

C

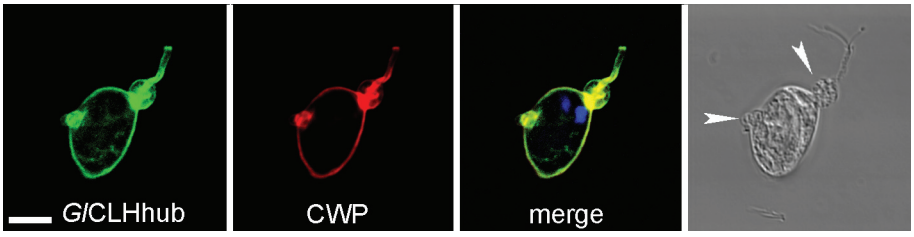
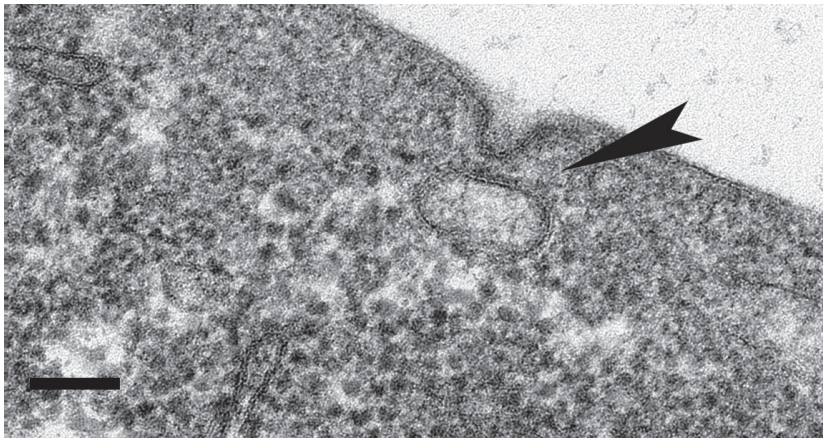

GlCLHhub
CWP
merge

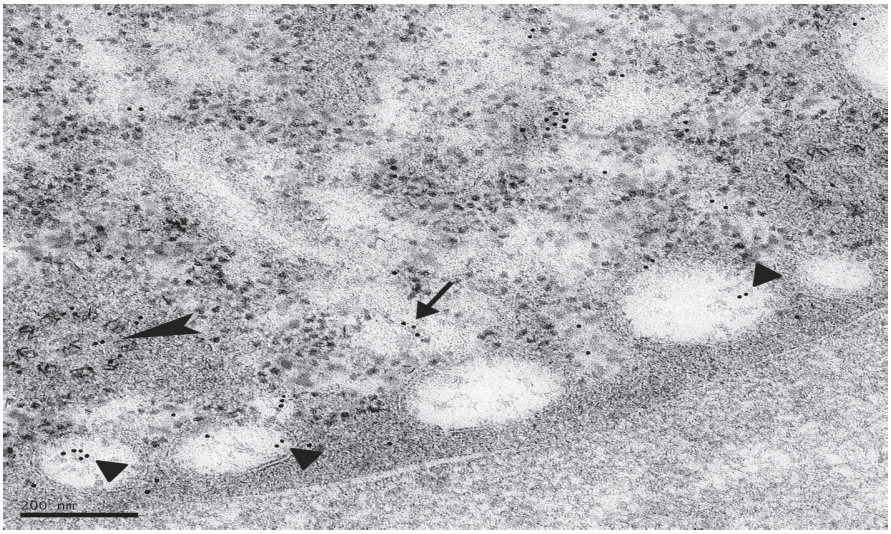
Figure 3.7: Ultrastructural analysis showed membrane pits forming very near PVs and clathrin localized in and around PVs.

A. Invaginations of the plasma membrane were visible in close proximity to PVs (black arrowhead), suggestive of the possibility of fusion between the plasma membrane and the PVs. Bar, 200 nm.
B. Immuno EM was performed with a gold-conjugated antibody specific for *Gl*CLH. Though the level of staining was low, gold particles were visible in and around PVs (short arrowheads). Some gold was apparent in membrane-bound tubules near the PVs (arrow). Gold was also observed in some flagella cross-sections (elongated arrowhead), which is in agreement with fluorescence microscopy studies showing abundant *Gl*CLH near the point of flagella exit from the cell body (Figure 9C). Bar, 200 nm.
C. Though plasma membrane invaginations (short arrowheads) are readily visible by EM, the presence of a classical clathrin coat is ambiguous on these forming cups or on PVs (elongated arrowheads), where clathrin has been localized by fluorescence microscopy and immuno EM. Bar, 200 nm.

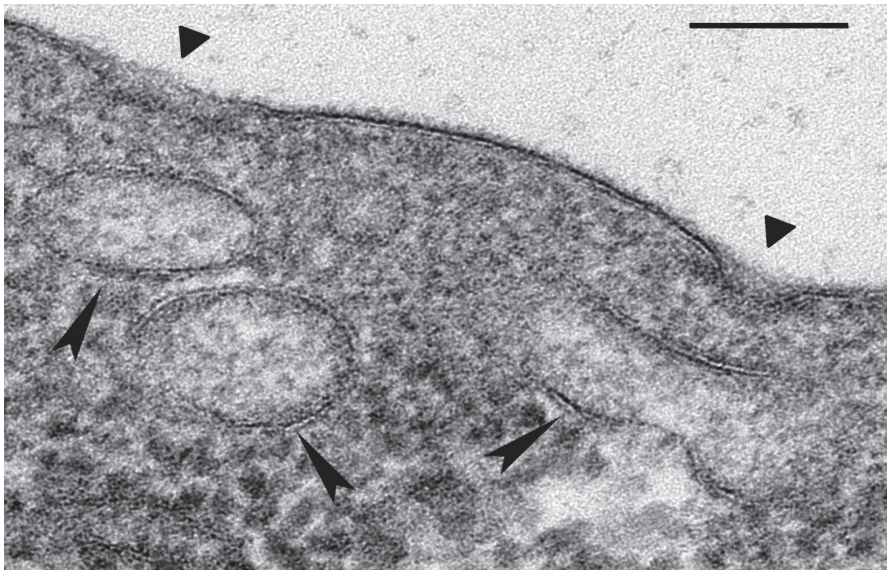
A



B



C



Chapter 4: General discussion and future directions

The aim of this dissertation project was to define the role of the major cysteine proteases in *Giardia lamblia* and utilize cysteine protease activity as a probe for defining the compartmentalization of the endosomal pathway in *Giardia*. As *Giardia* is recognized as an early-branching clade of eukaryotic cells, information obtained from these experiments can provide details regarding the evolution of eukaryotic cell function, compartmentalization, and the evolution of protease families. Prior to this work, very little was known about cysteine proteases in *Giardia*. Before the completion of the *Giardia* genome, only four cysteine proteases from *Giardia* had been identified (37,38). One of these, the endopeptidase *Gl*CP2, was implicated in the process of excystation (37). On the converse side of the life cycle, the exopeptidase ESCP was implicated in the process of encystation (38). In work presented in this dissertation, twenty-seven *Giardia* clan CA cysteine proteases were identified and their expression patterns were determined during vegetative growth and encystation (Figure 2.1, page 55). A role for cysteine proteases was described in vegetative trophozoites. Additionally the endomembrane architecture of vegetative trophozoites, and particularly the endocytic system, was described in greater detail than was previously understood. The importance of cysteine protease activity was also clarified during encystation and excystation. Finally, the localization pattern and function of the *Giardia* clathrin heavy chain (*Gl*CLH) was investigated.

4.1. Cysteine proteases and the endocytic system in vegetative *Giardia*

During vegetative growth, *Giardia* trophozoites do not secrete cysteine proteases,

suggesting that the main function of this protease class is intracellular. Cysteine proteases are also not required for trophozoite survival, at least using the inhibitors presently available. However, cysteine proteases do accomplish the majority of the degradation of endocytosed proteins, suggesting redundancy in activity (Figure 1.11). This study suggests that endocytosed proteins are rapidly taken into the PVs and transferred to the TVN by dynamic fusion of TVN tubules with the PVs. Cysteine proteases, which reside in the TVN and co-localize with ER markers, efficiently degrade the endocytosed protein. This non-canonical mechanism for endocytosis and subsequent proteolysis of endocytosed material in the TVN has parallels in higher eukaryotes. For example, the delivery of exogenous oligopeptides to the ER of MHC class I-presenting cells (80). Gagnon et al. (2002) recently proposed that the ER was involved in direct uptake of material from the extracellular environment via fusion with the plasma membrane (52). Therefore, the mechanism of the ER (TVN) acting as an endocytic compartment may be evolutionarily conserved and utilized by specialized cells of higher eukaryotes.

Future experiments to further characterize the endomembrane system of *Giardia* are dependent on the develpment of additional markers to define membrane compartments. In yeast and mammalian cells, multiple markers are available to define organelles and vesicle populations. As mentioned in Chapter 3, proteins like the adaptor proteins, that exhibit restricted localization and provide selectivity to vesicle cargo, are useful markers of vesicle subsets that appear identical by conventional EM or other currently available methods. Several AP orthologues were identified in *Giardia* recently and may soon be developed into valuable markers (135). Additional putative markers identified were in the categories of Rab-GTPases, SNAREs (SNAP Receptors), and Arfs

(ADP-ribosylation factors). The Ras-like Rab-GTPases have proven useful in defining vesicle populations of the endocytic system in higher eukaryotes. In human cells, almost seventy Rabs and Rab-like proteins are expressed and are often localized to distinct membrane-bound compartments. For example, Rab5 is localized to early endosomes, Rab6 to the Golgi complex, and Rabs 7 and 9 to late endosomes (136). If similar distinct localization patterns could be defined for markers in *Giardia*, the endomembrane system would be understood in much greater detail. Particularly, if subcompartments of the TVN exist, they would be identified and would enable a viable explanation of the paradox of *Giardia* catabolism and anabolism occuring in the same organelle.

4.2. Cysteine proteases and encystation

Following vegetative growth in the duodenum, trophozoites undergo encystation to form environmentally stable cysts. It was previously reported that cysteine protease activity is required for cyst formation (38). CWP2, a primary cyst wall component, requires processing by a cysteine protease to remove a C-terminal peptide and enable multimerization. In this study, it was discovered the *Gl*CP2 co-localized with CWP2 in ESVs and can accomplish CWP2 processing (Figures 2.6 and 2.7). Though ESCP, a cysteine exopeptidase, was previously implicated in this processing, it is more likely that *Gl*CP2 plays this role, as it is a cysteine endopeptidase and is twenty-fold more highly expressed than ESCP during encystation. Indeed, cysteine protease inhibitors that were reported to completely inhibit cyst formation also completely inhibited *Gl*CP2 activity (Table 2.3) (38). Other clan CA cysteine endopeptidases in *Giardia* also co-localized with CWPs in ESVs, so the proteolytic cleavage of CWP2 may be a redundant process.

The relatively large number of distinct cysteine proteases in *Giardia* is unique among protozoan parasitic organisms. In general, parasites do not exhibit the functional redundancy of proteases seen in vertebrates and higher eukaryotes. Twenty-seven cysteine proteases is quite a large number compared to many other parasites. For example, *T. cruzi* only has two cysteine proteases although gene amplification appears to have lead to as many as 100 copies of one depending on the parasite strain (137). Other trypanosome species, toxoplasma, and *T. vaginalis* also have relatively few cysteine proteases. However, there are examples, such as in *P. falciparum* and *Leishmania*, where multiple cysteine proteases are present. *Leishmania* has three distinct cysteine protease gene families (CPa, CPb, and CPc). Gene deletion of any on has a minor or no phenotype, but deletion of two produces a profound effect on macrophage infection or lesion development in the mouse (138, 139). While many distinct cysteine protease genes are found in *Giardia*, the *Gl*CP2 transcript accounts for roughly 40% of total cysteine protease transcript and is therefore clearly the major cysteine protease expressed by the cells.

As ESCP was previously implicated in CWP2 processing, ESCP recombinant expression would enable better clarification of the role, if any, that this exopeptidase has in the encystation process. Though unlikely for the aforementioned reasons, the possibility that the *Giardia* ESCP has the capability for endopeptidase activity remains feasible until the recombinant protein is examined. If recombinant ESCP did not process CWP2 it would more clearly define the requisites for CWP2 processing in encystation. If ESCP did indeed exhibit endopeptidase activity, the evolution of the rest of the cathepsin C-like protein family away from this type of activity would be of great interest. Indeed,

recombinant ESCP would be of evolutionary interest regardless of its defined activity because of the early-branching evolutionary status of *Giardia*.

*Gl*CP2 is also of evolutionary interest to further define the development of the clan CA cysteine proteases. Solving the crystal structure of *Gl*CP2 for the first time will allow a complete analysis of the similarities and differences between the "ancient" cathepsin B-like proteins and those of the more recently evolved higher eukaryotic organisms. It is often the sequence similarities and not functional relationships that define an enzyme family. Crystal structures provide much more detailed information about the connection between two enzymes with respect to protein folds and tertiary structure, active site architecture, and side-chain interactions. The structure may also shed light on the structural basis for the high pH optimum seen experimentally.

4.3. Cysteine proteases and excystation

Cysts that are ingested by a host must pass through the stomach acid and excyst in the small intestine. Ward et al. (1997) demonstrated that the process of excystation was dependent on cysteine protease activity in *G. muris* (37). It was suggested that cysteine proteases are secreted near the cyst wall to break down the cyst wall and allow trophozoites to emerge. *In vitro* excystation of *G. lamblia* remains quite inefficient, but when cysts formed by cells expressing *Gl*CP2-GFP were induced to excyst, the cysteine protease reporter construct was seen adjacent to the inside of the cyst wall (Figure 2.8). It was also demonstrated that while *Gl*CP2 processed CWP2 at lower concentrations, with more activity *Gl*CP2 completely degraded CWP2. These data suggest that *Gl*CP2 has an important role in excystation. Cysteine protease inhibitors, though not effective

as a treatment for *giardiasis* against vegetative trophozoite growth, may limit the transmission of *giardiasis* by inhibiting the processes of encystation and excystation. An appropriately safe cysteine protease inhibitor could prove quite useful as a *Giardia* prophylaxis for groups of people (hikers, travelers, military personnel) that may need to use a contaminated water source as drinking water for a short period of time.

The expression profiling presented in Chapter Two provides useful information to determine the relative importance of specific cysteine proteases or groups of proteases. Heterologous expression followed by crystallography-based structural studies would aid in the development of a specific inhibitor. *In silico* docking could also prove useful for the selection of candidate compounds for inhibitor develpment. Inhibitors as drugs would only need to be taken for a short treatment course and would not need to be effectively absorbed to reach the target organ (the small intestine). Therefore, many drug develpment obstacles might not interefere with the selection of candidate compounds for this indication. Potential inhibitors could then be tested in an *in vivo* mouse model to determine if cysteine protease inhibitors can indeed act as effective prophylactic drugs to prevent *Giardia* excystation and therefore prevent *giardiasis*.

To gain a greater understanding of the cell biology of excystation, two things are necessary; better microscopy techniques and more efficient *in vitro* excystation. Currently, the process of *in vitro* excystation is inefficient. One cyst from an entire culture (5×10^6 cells) might exhibit movement after completion of the procedure. The remaining cysts appear degraded and probably are harmed during the harsh excystment protocol which involves a highly acidic solution and large amounts of trypsin. Therefore, to do any microscopy-based investigations, the reliability of the sample and the biological

relevence fall into question. A higher percentage of cells excysting would allow for more confidence in data. Additionally, higher resolution microscopy and real time live cell imaging would allow for a detailed description of the mechanism of excystation. Presently, hypotheses have been proffered to explain how trophozoites emerge from a cyst, but little compelling data has been presented in this area of study.

4.4. *Giardia* clathrin

The localization and function of *Gl*CLH was investigated and found to differ greatly from canonical CLH found in higher eukaryotes. Clathrin in higher eukaryotes is found to localize to the cytoplasm and transiently localize to forming vesicles. *Gl*CLH was stably localized to the periphery of the cell, in and around the PVs (Figures 3.2.A and 3.3). This localization and the data presented in Chapter One suggest a role for clathrin in the initial uptake of proteins from the cell environment into the PVs. A dominant negative CLH (*Gl*CLHhub) did not disrupt the processes of protein endocytosis, lipid endocytosis, or early encystation (Figures 3.5 and 3.6). *Gl*CLHhub expression may have slightly disrupted late stage cyst formation, but these results were not corroborated by another laboratory investigating *Gl*CLHhub expression. However, the construct from that laboratory did not localize to ESVs and therefore may have been isolated from the area in which it would have affected cyst formation.

The structure and role of *Gl*CLH and its evolutionary relationship to the highly clathrin-dependent endocytic system of higher eukaryotic cells will require further experimentation to clearly define. First, the incorportation of clathrin into the cyst wall suggests that it may be an integral part of the cyst wall structure. Platinum-replica EM

methods may be able to determine if a clathrin coat is associated with the cyst wall. Ernst Ungewickell (Max-Planck-Institute für biochemie, Germany) has successfully used this technique to identify and characterize clathrin coats. The clathrin light chain of *Giardia* has not been identified to date, so its identification and characterization are of interest. Its identification may also lead to better antibodies that could be used for immunoprecipitation experiments to identify binding partners throughout the *Giardia* life cycle. In the work presented in this dissertation, using antibodies against the consensus light chain sequence, it appears that the light chain does co-localize with wild type *Gl*CLH and also *Gl*CLHhub during vegetative growth. Localization of the light chain during encystation and in the cyst may provide more detail about the structure of the clathrin lattice in *Giardia*.

4.5. General Conclusion

Giardia lamblia is a useful model system for studying the evolution of the eukaryotic cell. It can serve as a model for conventional function, such as degradation of exogenous protein by proteases, or a model of functions maintained by specialized cells, like the uptake of protein from the environment into the ER. Further analysis of the endomembrane system of *Giardia* may provide additional insights into the compartmentalization and development of the eukaryotic cell and may lead to the development of valuable therapies to combat *giardiasis*.

Chapter 5: Materials and methods

5.1. Cell culture and transfection

WB isolate *Giardia lamblia* trophozoites from the American Type Culture Collection (ATCC) were maintained in a modified TYI-S-33 medium supplemented with 10% FBS (Omega Scientific, Inc.), penicillin-streptomycin (UCSF CCF), vitamins (Gibco/ Invitrogen) and Fungizone (UCSF). The pGFP.pac vector (Gift from Theodore Nash, NIH; modified by Lei Li, CC Wang laboratory, UCSF by substituting the *Giardia* tubulin gene promotor for the *Giardia* Giardin gene promoter) was used to transiently and episomally express C-terminal GFP fusion proteins in *Giardia* trophozoites. The transfection protocol used by Singer et al. (99) was followed with modifications as follows: 1-2x10^6 trophozoites were incubated on ice for 20 min with 50 μg of circular plasmid DNA and then electroporated (GenePulser XCell, Bio-Rad) at 0.45 kv, 950 μF. Transfectants were selected in a dose-dependent manner using puromycin dihydrochloride (Sigma Inc.) increased in 5-20 μg/ml increments to a final concentration of 80 to 120 μg/ml. Trophozoites were induced to encyst as indicated by Abel et al (96). Excystation was performed as described by Lauwaet et al., 2007 (97).

5.2. T84 cell assay

1X10^6 *Giardia* trophozoites were preincubated in media containing 20 μM E64d (Sigma) or a DMSO control. Trophozoites were added to a monolayer of T84 intestinal epithelial cells and incubated for 30 min Supernatants were collected, 0.2 μm filtered, and tested for activity against Z-FR-AMC as described above.

5.3. Minimal media/starvation

Minimal media was prepared by omitting tryptone and yeast extract from the normal growth medium and using 5% FBS instead of the normal 10%. Cells were grown for 42 h under starvation conditions or in normal growth medium and in varying concentrations of the cysteine protease inhibitors E64d (Sigma) and K11777. Cells were counted in three demarcated areas of the culture tube regularly through the experiment and the number of cells was recorded.

5.4. Microscopy

A confocal microscope (LSM510 META; Carl Zeiss MicroImaging, Inc.) equipped with multiline (458, 477, 488, and 514 nm) Ar, 543 nm HeNe, and 633 nm HeNe visible lasers and a Chameleon two-photon laser module (Coherent, Inc.) with a "Plan-Apochromat" 63x/1.40 Oil DIC oil immersion lens (Carl Zeiss MicroImaging, Inc.) was used for fluorescence and live cell imaging experiments. Microscopy was done at room temperature. The cells were pulsed with sufficient oxygen at 37°C for 1 to 3 h and fixed in 3% paraformaldehyde (Electron Microscopy Sciences) for 40 min at room temperature and mounted with ProLong Gold mounting media (+ or – DAPI) (Molecular Probes). Live imaging was done in PBS. In the time frame of these experiments, *Giardia* are viable and active in PBS. No difference was noted in comparisons to control imaging in *Giardia* media, but PBS was preferred as background fluorescence was minimal. LSM Image Browser software (Carl Zeiss MicroImaging, Inc.) was used for confocal image acquisition and analysis. Adobe Photoshop CS (Adobe Systems, Inc.) was used for

subsequent processing.

5.5. Antibodies and reagents

Anti-*Giardia* CWP polyclonal was used at 1:100 (Waterborne, Inc.) Anti-*Gl*CP2 peptide polyclonal (raised against the peptide SSKVHLATATSYKDYGLDI) was used at 1:500. Anti-*Gl*CLH polyclonal and anti-*Giardia* PDI2 polyclonal (134) were used at 1:500 and 1:3500, respectively for immunofluorescence microscopy. Anti-*Gl*CLH polyclonal was used at 1:250 for electron microscopy. Anti light chain consensus sequence polyclonal was used at 1:2000. Anti-KDEL monoclonal antibody (Stressgen Bioreagents) and anti-spinach HSC70 monoclonal (Stressgen Bioreagents) were used at the recommended dilutions. Although the HSC70 monoclonal antibody was directed against the spinach HSC70, the spinach HSC70 protein sequence is identical to the *Giardia* HSC70 protein sequence (www.mbl.edu/giardia). Rhodamine-labeled anti-mouse IgG (γ) (KPL Inc.) and fluorescein-labeled anti-mouse IgG (γ) (KPL Inc.) were used at 1:10. Inhibitors: PMSF (Sigma), EDTA (Sigma), Aprotinin (Sigma), Pepstatin A (Calbiochem), Leupeptin (Sigma), TLCK (1-chloro-3-tosylamido-7-amino-2-heptanone HCl) (Sigma), TPCK (1-chloro-3-tosylamido-4-phenyl-2-butanone) (Sigma), E64 (Sigma), CA074 (Sigma), lactacystin (Sigma), α-1 antitrypsin (Sigma), calpain inhibitor I (ALLN, Sigma), calpain inhibitor II (ALLM, Sigma). *R. norvegicus* cathepsin C was a gift from John Pederson (Unizyme, Denmark). For endocytosis studies, fluorescent proteins (3.75 μg/ml casein; 5.0 μg/ml albumin) or 0.04 μm biotin-conjugated Fluospheres (2.8 x 1014 particles/ml; 1.0 μl/ml final) (excitation/emission: 505 nm/515 nm) (Molecular Probes) were incubated with trophozoites for 30 min at 37°C in media or PBS. Sodium azide and cytochalasin D

were used at 20 mM and 10 μM, respectively. Cells were fixed and visualized. Lucifer yellow lithium salt (excitation/emission: 428 nm/536 nm; Molecular Probes) was used to visualize PVs. Trophozoites were incubated in 1 ml of PBS and Lucifer yellow to a final concentration of 1 mg/ml in PBS for 15 min at 37°C. Following incubation, cells were then fixed and visualized.

5.6. Endocytosis visualized by live cell imaging

Trophozoites were allowed to adhere to chambered coverglass (Lab-Tec) in vegetative growth media. Media was replaced with 400 μl PBS at 37°C. Filming began as 5 × 109 to 5 × 1010 biotin-conjugated Fluospheres (excitation/emission: 505 nm/515 nm) 0.04 μm (Molecular Probes) or 250 ng of DQRed BSA (excitation/emission: 590 nm/620 nm) (Molecular Probes) were added to the chamber of adhered trophozoites.

5.7. Immunofluorescence localization of organelle markers

Trophozoites were stained as previously described (135) with modifications. All experiments were carried out at room temperature, unless otherwise stated. Cells were harvested following a 15 min incubation on ice and a 15-20 min spin at 411xg to pellet the cells. Cells were washed with cold PBS (Ca2+ and Mg2+ free) before fixing for 40 min with fresh 3% paraformaldehyde (EMS). Following a 5 min incubation in 0.1 M glycine in PBS, the cells were permeabilized in 0.1% Triton X-100 in PBS for 30 min and blocked with 2% BSA in PBS. Trophozoites were incubated with primary and secondary antibodies (diluted in 2% BSA/0.1% Triton X-100 in PBS) for 1 h each.

5.8. Protease activity assays

5.8.1. pH optima with peptide substrates.

The fluorogenic substrates Z-FR-AMC (N-carbobenzoxy-phenylalanyl-arginyl-7-amido-4-methylcoumarin) and Z-RR-AMC (N-carbobenzoxy-arginyl-arginyl-7-amido-4-methylcoumarin; excitation/emission of AMC: 360 nm/470 nm) (Bachem) were incubated with *Giardia* lysates or recombinant enzyme in Tris-HCl buffer (pH 7.2) or citrate/dibasic sodium phosphate buffers (pH 4.0-8.0) containing 4 mM DTT, 1 mM Pefabloc, and 10 mM EDTA. Subsequent protease activity was measured by monitoring the increase in relative fluorescence units (RFU) over time.

5.8.2. pH optima with casein-resorufin.

r*Gl*CP2 was incubated for 30 min with 50 μg casein-resorufin (Molecular Probes) in 200 μl citrate/dibasic sodium phosphate buffers. 960 μl 5% (w/v) TCA was added, samples were incubated 10 min, and centrifuged. 400 μl supernatant was added to 600 μl 0.5M Tris, pH 8.8. Hydrolysis was quantified by measuring fluorescence (excitation/emission: 574 nm/584 nm).

5.8.3. Processing of rCWP2.

Purified ^{35}S-rCWP2 was incubated with enzyme in Tris-HCl buffer, pH 7.2, 4 mM DTT at 25°C. Sample was fractionated by SDS-PAGE, dried, and visualized by phosphorimaging (Typhoon Trio, GE Healthcare).

5.8.4. Identification of protease activity using an active site probe.

DCG04 (^{125}I- or BODIPY-labeled), the clan CA cysteine protease inhibitor (140), was incubated with enzyme and 4 mM DTT for 30 min Proteins were fractionated by SDS-PAGE, dried, and visualized by phosphorimaging (Typhoon Trio,GE Healthcare).

5.8.5. K_m determination.

The fluorogenic peptide substrates Z-FR-AMC, Z-RR-AMC, and Z-VLK-AMC (Bachem) were incubated with *Gl*CP2 at a range of concentrations and the Vmax U/s was recorded. The non-linear regression and K_m calculations were determined using Prism 4 software (Graphpad).

5.8.6. Zymogram detection of protease activity.

r*Gl*CP2 was fractionated on a Novex® 10% zymogram (gelatin) gel (Invitrogen) under native conditions as recommended by manufacturer. Gel was stained with SimplyBlue™ Safestain (Invitrogen) and destained in water to visualize bands of protease activity.

5.8.7. *In vitro* identification of protease activity.

r*Gl*CP2 was fractionated by SDS-PAGE under non-reducing conditions on a 15% Tris-HCl gel. Gel was washed 2X in 20 mM Tris-HCl, 0.2% Triton X-100. Gel was incubated in 20 mM Tris-HCl, 0.2% Triton X-100, 5 mM DTT, and 10 μM Z-FR-MNA (N-carbobenzoxy-phenylalanyl-arginyl-4-methoxy-β-naphthylamide) for 2 hr. at room temperature. Two volumes (compared to substrate) of 2 M coupling reagent (5-nitro-2-salicylaldehyde) (excitation/emission: 395 nm/575 nm) was added to the reaction, which was incubated for an additional 4 h at room temperature. Fluorescence was visualized on a Typhoon Trio (GE Healthcare).

5.8.8. *In vivo* identification of protease activity.

The fluorescent substrates Z-FR-MNA (N-carbobenzoxy-phenylalanyl-arginyl-4-methoxy-β-naphthylamide) and Z-RR-MNA (N-carbobenzoxy-arginyl-arginyl-4-methoxy-β-naphthylamide) (Bachem) were used in an adapted in vivo protease assay (141). Trophozoites were washed once with PBS (Ca2+ and Mg2+ free), then incubated

in PBS with 70 μM Z-FR-MNA and two volumes of 2mM coupling reagent (5-nitro-2-salicylaldehyde; 2 M cacodylate buffer (excitation/emission: 395 nm/575 nm) for 30 min at 37°C. For Z-RR-MNA, trophozoites were washed in 100 mM cacodylate buffer pH 6.8 and 5% sucrose and incubated with Z-RR-MNA at a final concentration of 500 μM in 100 mM cacodylate buffer pH 6.8 for 1 h at 37°C followed by a 20 min incubation at 37°C with coupling reagent. The cells were fixed and visualized.

5.8.9. *In vitro* substrate degradation.

1 μg BSA-FITC or Casein-FITC (Molecular Probes) was incubated with 1 mM PMSF, 4 mM DTT, 10 mM EDTA and *Giardia* lysates. Proteins were precipitated by 10% TCA to stop the reaction and precipitated proteins were fractionated by SDS-PAGE. Fluorescence was detected using a 8600 Variable Mode Imager (Amersham).

5.8.10. *In vivo* substrate degradation. 1.0 X 10^6 *Giardia* trophozoites were incubated with 20 μg casein-FITC for 30 min in vegetative media and chased with fresh media. Cells were incubated at 37°C for 1 h, 5 h, and 16 h. Cells were also incubated for 16 h in the presence of three known cell permeable cysteine protease inhibitors (10 μM, DMSO<1%): E64d, K11777 (K777), and WRR477. Cells were lysed and proteins fractionated by SDS-PAGE. In-gel casein-FITC was detected using a Typhoon 8600 Variable Mode Imager (Amersham).

5.9. Electron microscopy

5.9.1. Localization of glucose-6-phosphatase activity by electron microscopy.

A trophozoite monolayer was fixed for 30 min Glucose-6-phosphatase reaction was performed as previously described (142). Cells were washed with 0.1 M Tris/maleate, 3 mM lead nitrate, 4 mM disodium glucose-6-phosphate and 5% sucrose, pH 6.5 at

37°C for 60 min Cells were post-fixed with the aforementioned fixative and osmium tetroxide and embedded in epon. Blocks were sectioned with a Leica ultracut UCT ultramicrotome. Sections were viewed with FEI Tecnai 10 electron microscope (FEI Company).

5.9.2. Ultrastructural tomography studies.

A Tecnai T20 electron microscope (FEI Company) equipped with a bottom mounted four-quadrant 4K x 4K Gatan UltraScann CCD (Gatan Inc.) was run at 200 kV and images were collected every 2 degrees from –60 to +60 degrees. The cameral was set at binning 2 (2K x 2K) resolution and images were collected at 21,500X magnification. Samples were 250 nm thick. Images were processed, analyzed and visualized using PRIISM software (143). Data were analyzed further using Openlab software (Improvision Ltd., England). To obtain a pseudo 3D view, the transparent background of 3 selected consecutive 0.25 μm images was set to 75, 50, 35, and 0% opacity, respectively, prior to image merging.

5.9.3. Conventional electron microscopy.

Cells were fixed with 2% glutaraldehyde in 0.12 M Na-cacodylate, pH 7.4 for 2 h at room temperature. Cells were post-fixed with 1% OsO_4 for 2 h followed by en bloc staining with 1% aqueous uranyl acetate for 1 h with rotation. Cells were dehydrated with ethanol followed by propylene oxide infiltration and flat embedding in eponate 12.

5.9.4. Immunolocalization of *Gl*CLH.

Cells were fixed by high pressure freezing/freeze substitution. Substitution medium was 1.1 OsUAc (Kent). Substituted samples were rinsed with dry acetone and infiltrated with LR White with heat curing accelerator, and hardened by microwave.

5.10. Transformation and expression of *G. lamblia* CP2 in *P. pastoris*

The *G. lamblia* cysteine protease 2 (*Gl*CP2) gene was re-synthesized to optimize for yeast codon usage (DNA 2.0). The r*Gl*CP2 gene was amplified by PCR from the pJ31:7972 vector into which the full-length cDNA had previously been cloned and modified to include a polyhistidine tag using the forward primer *Gl*CP2pPicF: CTCGAGAAAAGACATCATCATCATCATCATGAGTTGAATCATATTACTC and the reverse primer *Gl*CP2pPicR: TCTAGATTACTCATCGAAAAATCCAGCATAGGCC. The 920 bp amplicon was subcloned in the *Xho*I/*Xba*I site of the *P. pastoris* expression vector pPicZαB (Invitrogen). The plasmid was linearized by digestion with *Sal*I and introduced into *P. pastoris* by electroporation (GenePulser XCell, Bio-Rad) according to manufacturer's specifications. Transformants were screened by growth on YPD + 100 μg/ml Zeocin (Invitrogen).

5.11. Purification of recombinant *Gl*CP2.

P. pastoris was grown under expression induction conditions in a BioFlo 110 Fermentor/ Bioreactor (New Brunswick Scientific) for three days according to manufacturer's specifications. Methanol was maintained at 0.5% (calculated by a methanol sensor) by addition of 100% methanol 2X/day. 0.2 μM-filtered supernatant was lyophilized. 8g lyophilized material was resuspended in 40 ml ddH2O + 1 mM Pefabloc (Sigma). Solution was filtered at 0.2 μm, dialyzed in 10,000 MWC dialysis tubing (Pierce) against 10 mM Tris-HCL, pH 8.0 at 4°C, and fractionated by ion exchange chromatography with Fast Flow Q resin (GE Healthcare) followed by dialysis to desalt and a MonoQ anion

exchange column (GE Healthcare).

5.12. Purification of cysteine protease activity from G. *lamblia* lysates.

Giardia cells were incubated in a 20 mM Tris-HCL (pH 7.2) and 0.2% Triton X-100 (Sigma-Aldrich) buffer at 4°C with stirring for 2 h Debris was pelleted and supernatant was 0.2 μm filtered and subjected to anion exchange chromatography using a MonoQ column.

5.13. RNA methods

5.13.1. RNA preparation.

Total RNA from vegetative or encysting *Giardia* cells was isolated with Trizol reagent (Invitrogen). 2 μg RNA was treated with 1 U of amplification grade DNAse I (Sigma). cDNA was synthesized with Superscript III reverse transcriptase according to manufacturer's specifications (Invitrogen). cDNA samples were stored at -80 °C until use. Control samples were prepared as above using nuclease-free ddH2O in place of RNA.

5.13.2. Real Time PCR.

PCR was performed in an Mx3005P™ QPCR system using MxPro™ QPCR software (Stratagene). Amplification was performed in a final volume of 25 μl, containing cDNA from the reverse transcribed reaction, primer mixture (0.3 μM each of sense and antisense primers), and 12.5 μl of 2x SYBR Green Master Mix (Applied Biosystems). The final mRNA levels of the genes studied were normalized to GAPDH expression using the comparative C_T method (Stratagene).

The sequences of *Giardia* cysteine proteases were obtained from the *Giardia* genome

project (www.mbl.edu/giardia). For GenBank protein accession numbers and primers see Table 5.1.

5.14. Positional scanning synthetic combanitorial library.

Protease activity was assayed at 25°C in a buffer containing 20 mM Tris-HCl (pH 7.2), 5 mM DTT, 0.2% Triton X-100 (Sigma-Aldrich), and 1% Me2SO (from the substrates) or in buffer with NaOAc replacing Tris-HCl (pH 5.5) as referenced in the text. Assays were performed as previously described (90).

5.15. Expression and purification of *Giardia lamblia* CWP2.

The sequence of CWP2 was amplified from genomic DNA and a C-terminal polyhistidine tag was added using the primers CWP2pMalF: TCTAGAATGGCTTGCCCTGCCACCGAGG and CWP2pMalR: GCGGCCGCTTTAATGATGATGATGATGATGCCTTCCCTGGATCCTT CTGCGGACAATAG and inserted into the *Not*I/*Xba*I site of the expression vector pCMVTnT (Promega). 1μg of plasmid DNA was used as a template for *in vitro* transcription using the TnT Quick Coupled Transcription/Translation kit according to manufacturer's specifications (Promega). rCWP2 was further purified on a nickel-nitrilotriacetic acid column (Qiagen).

5.16. Mass spectrometry

Tryptic digest sample was analyzed by LC/MS/MS. Analyses were performed with an LTQ iontrap (Thermo Scientific) and a QSTAR (Applied Biosystems). Database search

was conducted using Mascot (Matrix Science Inc) on the full NCBI protein database. Mass accuracy for the QSTAR data: 100ppm in MS; 0.2 Da in MS/MS. Mass accuracy for the LTQ data: 3 Da in MS; 0.8 Da in MS/MS.

5.17. 2D gel electrophoresis

Protein samples were de-salted with centricon spin columns (Millipore). 2D gel electrophoresis was performed according to manufacturer's specifications using the Zoom IPGRunner system (Invitrogen). Gels were silver stained with the Silver Stain Plus Kit (Biorad) and protein spots were excised and trypsin digested.

5.18. Figures and tables for Chapter Five

Table 5.1: GenBank accession numbers and primer sets used for quantitative RT-PCR of the clan CA cysteine proteases of *Giardia lamblia*. 114

Genbank Accession Number	Forward Primer	Reverse Primer
EAA40112	CGTTATCCCGTCCAG	CGTTGCGTCCTTCTC
EAA41050	GCGGCGTCTACCAGCACACC	CCGTCCTCGCCCCAGTCAG
CAC18646	GATGACGGGACGGACTACTGG	TCATCCTCTATCCTGCACTCATTG
EAA40365	CGAGGCCGGTTGTCAGG	GCCGAGGCAGCTTTGTAAT
EAA37074	CAGCACAGGTGGGGGCTCTGGCTTGGCGGC	CTCTACGTCTGCATAGTGGAAGCCCTCGTG
EAA39655	AACAGAGTTGGTTTTCTTTCTGTTGAGATA	CTTGGTGGCCATGGATGGTTTGTTAGGGCT
EAA37893	TGTTGGCGTGCCCTACTCA	TCCTTATCGGCGTCTGTGT
EAA37433	ATGCGAATGATGCGGAAGTAGC	TACGGAAGAGTCGAGGGTGGTC
EAA38990, Orf:113303, Orf:119224	CCGGCGGTGCAGTTACAAT	GGGGCAGAGCGGGTGAA
EAA39491	CCTCAGACATCTAGCTTCCTCAGGAGCAAT	TGGATGCAGTGACGTATTTTCACTTTTCCG
EAA42730	ACTATCAAAAGCAAGAGGTGGAGGGTTAGC	CCTAAATAGTCCCCGCTCTCCTCTAATGTT
EAA41342	ATCGTTCGTAGTAGGCAGAGGAAGCTTCAA	CCCAATTATGTTTTGTAACTGCGCTCCAGG
EAA40802	CTACCCGCATCGCAGAACT	ACACCGAAGCCAACCAGAA
AAK92150, Orf:114773, Orf:114915, Orf:112831	GCACGTCGACGAGGAAGAAGC	GAACTGGGGCCGGACCTGAAG
EAA38408	ACCCCTCTGGGAGTGCTT	GACCCAGTCCGCAACTCT
EAA39488	GCCTGTCTCCATCGCTGT	CCCAGCCGACGAGTAGAA
EAA39487	GCCGTTCTTGCTGTTGGT	CATCCATCCCCCAAAGTG
EAA37165	GTCCACCCGCTCTTCCTCTC	TAGTTCCACGCGCTGAACCTC
EAA36907	GTGGCGTTCGGGGTCGTCT	CACGGCAGGCCAGCAGTTG
EAA42977	ACGCATCCAAACAAATTGGTCATTCAGCAA	TAATGCATGGTCAGACATCATTAGACACGA
EAA42983	TTAGACGCGCCTCAGAAGGTTATGCCTTTA	TGGATTTGAAAGCATGCTTAGCTCCTCACT

Table 5.1: GenBank accession numbers and primer sets used for quantitative RT-PCR of the clan CA cysteine proteases of *Giardia lamblia*

Chapter 6: Abbreviations and References

6.1. Abbreviations used

CWP: cyst wall protein

PV: peripheral vacuole

TVN: tubulovesicular network

ER: endoplasmic reticulum

MHC: major histocompatability complex

VSP: variant surface protein

ESCP: encystation specific cysteine protease

PDI2: protein disulfide isomerase 2

GLV: *Giardia lamblia* virus

G6P: glucose-6-phosphatase

*Gl*CP: *Giardia lamblia* cysteine protease

CCV: clathrin-coated vesicle

AP: adaptor protein complex

GGA: Golgi-localized gamma ear containing ARF-binding protein

CLH: clathrin heavy chain

TIR-FM: Total internal reflection fluorescence microscopy

6.2. References

1. Alberts B (2002) Molecular biology of the cell. New York: Garland Science. xxxiv,

1463, [1486] p.

2. Dobell C (1920) The discovery of intestinal protozoa in man. Proc R Soc Med: 1-15.

3. Adam RD (2001) Biology of Giardia lamblia. Clin Microbiol Rev 14: 447-475.

4. Friend DS (1966) The fine structure of Giardia muris. J Cell Biol 29: 317-332.

5. Gillin FD, Reiner DS, McCaffery JM (1996) Cell biology of the primitive eukaryote Giardia lamblia. Annu Rev Microbiol 50: 679-705.

6. Knodler LA, Noiva R, Mehta K, McCaffery JM, Aley SB, et al. (1999) Novel protein-disulfide isomerases from the early-diverging protist Giardia lamblia. J Biol Chem 274: 29805-29811.

7. Mowatt MR, Lujan HD, Cotten DB, Bowers B, Yee J, et al. (1995) Developmentally regulated expression of a Giardia lamblia cyst wall protein gene. Mol Microbiol 15: 955-963.

8. Lujan HD, Mowatt MR, Nash TE (1997) Mechanisms of Giardia lamblia differentiation into cysts. Microbiol Mol Biol Rev 61: 294-304.

9. Gillin FD, Boucher SE, Rossi SS, Reiner DS (1989) Giardia lamblia: the roles of bile, lactic acid, and pH in the completion of the life cycle in vitro. Exp Parasitol 69: 164-174.

10. Gillin FD, Reiner DS, Gault MJ, Douglas H, Das S, et al. (1987) Encystation and expression of cyst antigens by Giardia lamblia in vitro. Science 235: 1040-1043.

11. Kane AV, Ward HD, Keusch GT, Pereira ME (1991) In vitro encystation of Giardia lamblia: large-scale production of in vitro cysts and strain and clone differences in encystation efficiency. J Parasitol 77: 974-981.

12. Reiner DS, Hetsko ML, Das S, Ward HD, McCaffery M, et al. (1993) Giardia

lamblia: absence of cyst antigens and reduced secretory vesicle formation and bile salt uptake in an encystation-deficient subline. Exp Parasitol 77: 461-472.

13. Manning P, Erlandsen SL, Jarroll EL (1992) Carbohydrate and amino acid analyses of Giardia muris cysts. J Protozool 39: 290-296.

14. Gerwig GJ, van Kuik JA, Leeflang BR, Kamerling JP, Vliegenthart JF, et al. (2002) The Giardia intestinalis filamentous cyst wall contains a novel beta(1-3)-N-acetyl-D-galactosamine polymer: a structural and conformational study. Glycobiology 12: 499-505.

15. Reiner DS, McCaffery M, Gillin FD (1990) Sorting of cyst wall proteins to a regulated secretory pathway during differentiation of the primitive eukaryote, Giardia lamblia. Eur J Cell Biol 53: 142-153.

16. Lanfredi-Rangel A, Attias M, de Carvalho TM, Kattenbach WM, De Souza W (1998) The peripheral vesicles of trophozoites of the primitive protozoan Giardia lamblia may correspond to early and late endosomes and to lysosomes. J Struct Biol 123: 225-235.

17. Hedges SB (2002) The origin and evolution of model organisms. Nat Rev Genet 3: 838-849.

18. Sogin ML, Gunderson JH, Elwood HJ, Alonso RA, Peattie DA (1989) Phylogenetic meaning of the kingdom concept: an unusual ribosomal RNA from Giardia lamblia. Science 243: 75-77.

19. Upcroft J, Upcroft P (1998) My favorite cell: Giardia. Bioessays 20: 256-263.

20. Tovar J, Leon-Avila G, Sanchez LB, Sutak R, Tachezy J, et al. (2003) Mitochondrial remnant organelles of Giardia function in iron-sulphur protein maturation. Nature

426: 172-176.

21. Benchimol M (2004) Giardia lamblia: behavior of the nuclear envelope. Parasitol Res 94: 254-264.

22. Kulda J, Nohynkova, E (1996) *Giardia* in humans and animals. In: Kreier J, editor. Parasitic Protozoa. 2nd ed. San Diego: Academic Press.

23. Kabnick KS, Peattie DA (1990) In situ analyses reveal that the two nuclei of Giardia lamblia are equivalent. J Cell Sci 95 (Pt 3): 353-360.

24. Sagolla MS, Dawson SC, Mancuso JJ, Cande WZ (2006) Three-dimensional analysis of mitosis and cytokinesis in the binucleate parasite Giardia intestinalis. J Cell Sci 119: 4889-4900.

25. de Souza W, Lanfredi-Rangel A, Campanati L (2004) Contribution of microscopy to a better knowledge of the biology of Giardia lamblia. Microsc Microanal 10: 513-527.

26. Kattenbach WM, Pimenta PF, de Souza W, Pinto da Silva P (1991) Giardia duodenalis: a freeze-fracture, fracture-flip and cytochemistry study. Parasitol Res 77: 651-658.

27. Feely DE, Dyer JK (1987) Localization of acid phosphatase activity in Giardia lamblia and Giardia muris trophozoites. J Protozool 34: 80-83.

28. Guerrant RL, Walker DH, Weller PF (1999) Tropical infectious diseases : principles, pathogens & practice. Philadelphia: Churchill Livingstone. 2 v. (xxiii, 1644, lx), [1648] of col. plates p.

29. Ortega YR, Adam RD (1997) Giardia: overview and update. Clin Infect Dis 25: 545-549; quiz 550.

30. Ali SA, Hill DR (2003) Giardia intestinalis. Curr Opin Infect Dis 16: 453-460.

31. Eckmann L (2003) Mucosal defences against Giardia. Parasite Immunol 25: 259-270.

32. Baron S (1998) Medical microbiology. Galveston, Tex.: University of Texas Medical Branch at Galveston.

33. Sajid M, McKerrow JH (2002) Cysteine proteases of parasitic organisms. Mol Biochem Parasitol 120: 1-21.

34. Lecaille F, Kaleta J, Bromme D (2002) Human and parasitic papain-like cysteine proteases: their role in physiology and pathology and recent developments in inhibitor design. Chem Rev 102: 4459-4488.

35. McKerrow JH (1999) Development of cysteine protease inhibitors as chemotherapy for parasitic diseases: insights on safety, target validation, and mechanism of action. Int J Parasitol 29: 833-837.

36. McKerrow JH, Engel JC, Caffrey CR (1999) Cysteine protease inhibitors as chemotherapy for parasitic infections. Bioorg Med Chem 7: 639-644.

37. Ward W, Alvarado L, Rawlings ND, Engel JC, Franklin C, et al. (1997) A primitive enzyme for a primitive cell: the protease required for excystation of Giardia. Cell 89: 437-444.

38. Touz MC, Nores MJ, Slavin I, Carmona C, Conrad JT, et al. (2002) The activity of a developmentally regulated cysteine proteinase is required for cyst wall formation in the primitive eukaryote Giardia lamblia. J Biol Chem 277: 8474-8481.

39. Munn AL (2000) The yeast endocytic membrane transport system. Microsc Res Tech 51: 547-562.

40. Engqvist-Goldstein AE, Drubin DG (2003) Actin assembly and endocytosis: from

yeast to mammals. Annu Rev Cell Dev Biol 19: 287-332.

41. Mukherjee S, Ghosh RN, Maxfield FR (1997) Endocytosis. Physiol Rev 77: 759-803.

42. Conner SD, Schmid SL (2003) Regulated portals of entry into the cell. Nature 422: 37-44.

43. Toret CP, Drubin DG (2006) The budding yeast endocytic pathway. J Cell Sci 119: 4585-4587.

44. Singer-Kruger B, Frank R, Crausaz F, Riezman H (1993) Partial purification and characterization of early and late endosomes from yeast. Identification of four novel proteins. J Biol Chem 268: 14376-14386.

45. Lakadamyali M, Rust MJ, Zhuang X (2006) Ligands for clathrin-mediated endocytosis are differentially sorted into distinct populations of early endosomes. Cell 124: 997-1009.

46. Yamashiro DJ, Maxfield FR (1987) Acidification of morphologically distinct endosomes in mutant and wild-type Chinese hamster ovary cells. J Cell Biol 105: 2723-2733.

47. Mellman I (1996) Endocytosis and molecular sorting. Annu Rev Cell Dev Biol 12: 575-625.

48. Maxfield FR, McGraw TE (2004) Endocytic recycling. Nat Rev Mol Cell Biol 5: 121-132.

49. Piper RC, Luzio JP (2001) Late endosomes: sorting and partitioning in multivesicular bodies. Traffic 2: 612-621.

50. Miaczynska M, Zerial M (2002) Mosaic organization of the endocytic pathway. Exp Cell Res 272: 8-14.

51. Hirota Y, Masuyama N, Kuronita T, Fujita H, Himeno M, et al. (2004) Analysis of post-lysosomal compartments. Biochem Biophys Res Commun 314: 306-312.

52. Gagnon E, Duclos S, Rondeau C, Chevet E, Cameron PH, et al. (2002) Endoplasmic reticulum-mediated phagocytosis is a mechanism of entry into macrophages. Cell 110: 119-131.

53. Touret N, Paroutis P, Terebiznik M, Harrison RE, Trombetta S, et al. (2005) Quantitative and dynamic assessment of the contribution of the ER to phagosome formation. Cell 123: 157-170.

54. Kartenbeck J, Stukenbrok H, Helenius A (1989) Endocytosis of simian virus 40 into the endoplasmic reticulum. J Cell Biol 109: 2721-2729.

55. Sandvig K, van Deurs B (2002) Transport of protein toxins into cells: pathways used by ricin, cholera toxin and Shiga toxin. FEBS Lett 529: 49-53.

56. Das S, Stevens T, Castillo C, Villasenor A, Arredondo H, et al. (2002) Lipid metabolism in mucous-dwelling amitochondriate protozoa. Int J Parasitol 32: 655-675.

57. McCaffery JM, Faubert GM, Gillin FD (1994) Giardia lamblia: traffic of a trophozoite variant surface protein and a major cyst wall epitope during growth, encystation, and antigenic switching. Exp Parasitol 79: 236-249.

58. Lujan HD, Touz MC (2003) Protein trafficking in Giardia lamblia. Cell Microbiol 5: 427-434.

59. Barrett AJ, Rawlings ND, Woessner JF (1998) Handbook of proteolytic enzymes. San Diego: Academic Press. xxix, 1666 p.

60. Turk B, Turk D, Turk V (2000) Lysosomal cysteine proteases: more than scavengers.

Biochim Biophys Acta 1477: 98-111.

61. Turk V, Turk B, Turk D (2001) Lysosomal cysteine proteases: facts and opportunities. Embo J 20: 4629-4633.

62. Turk B, Turk V, Turk D (1997) Structural and functional aspects of papain-like cysteine proteinases and their protein inhibitors. Biol Chem 378: 141-150.

63. Mottram J, North, MJ, and Sajid, M (2004) Trichomonad and Giardia cysteine endopeptidases. In: A.J. Barrett NDR, and J.F. Woessner, editor. Handbook of Proteolytic Enzymes. London: Academic Press. pp. 1170-1173.

64. Touz MC, Kulakova L, Nash TE (2004) Adaptor protein complex 1 mediates the transport of lysosomal proteins from a Golgi-like organelle to peripheral vacuoles in the primitive eukaryote Giardia lamblia. Mol Biol Cell 15: 3053-3060.

65. Benchimol M, Piva B, Campanati L, de Souza W (2004) Visualization of the funis of Giardia lamblia by high-resolution field emission scanning electron microscopy--new insights. J Struct Biol 147: 102-115.

66. Lanfredi-Rangel A, Attias M, Reiner DS, Gillin FD, De Souza W (2003) Fine structure of the biogenesis of Giardia lamblia encystation secretory vesicles. J Struct Biol 143: 153-163.

67. Porter KR (1953) Observations on a submicroscopic basophilic component of cytoplasm. J Exp Med 97: 727-750.

68. Alberts B (2002) Molecular biology of the cell. New York: Garland Science.

69. Federovitch CM, Ron D, Hampton RY (2005) The dynamic ER: experimental approaches and current questions. Curr Opin Cell Biol 17: 409-414.

70. Soltys BJ, Falah M, Gupta RS (1996) Identification of endoplasmic reticulum in the

primitive eukaryote Giardia lamblia using cryoelectron microscopy and antibody to Bip. J Cell Sci 109 (Pt 7): 1909-1917.

71. Gupta RS, Aitken K, Falah M, Singh B (1994) Cloning of Giardia lamblia heat shock protein HSP70 homologs: implications regarding origin of eukaryotic cells and of endoplasmic reticulum. Proc Natl Acad Sci U S A 91: 2895-2899.

72. Morgan GW, Hall BS, Denny PW, Carrington M, Field MC (2002) The kinetoplastida endocytic apparatus. Part I: a dynamic system for nutrition and evasion of host defences. Trends Parasitol 18: 491-496.

73. Hernandez Y, Castillo C, Roychowdhury S, Hehl A, Aley SB, et al. (2007) Clathrin-dependent pathways and the cytoskeleton network are involved in ceramide endocytosis by a parasitic protozoan, Giardia lamblia. Int J Parasitol 37: 21-32.

74. Perret E, Lakkaraju A, Deborde S, Schreiner R, Rodriguez-Boulan E (2005) Evolving endosomes: how many varieties and why? Curr Opin Cell Biol 17: 423-434.

75. Tai JH, Ong SJ, Chang SC, Su HM (1993) Giardiavirus enters Giardia lamblia WB trophozoite via endocytosis. Exp Parasitol 76: 165-174.

76. McCaffery JM, Gillin FD (1994) Giardia lamblia: ultrastructural basis of protein transport during growth and encystation. Exp Parasitol 79: 220-235.

77. Gething MJ, Sambrook J (1992) Protein folding in the cell. Nature 355: 33-45.

78. Nishikawa S, Brodsky JL, Nakatsukasa K (2005) Roles of molecular chaperones in endoplasmic reticulum (ER) quality control and ER-associated degradation (ERAD). J Biochem (Tokyo) 137: 551-555.

79. Paulsson K, Wang P (2003) Chaperones and folding of MHC class I molecules in the endoplasmic reticulum. Biochim Biophys Acta 1641: 1-12.

80. Day PM, Yewdell JW, Porgador A, Germain RN, Bennink JR (1997) Direct delivery of exogenous MHC class I molecule-binding oligopeptides to the endoplasmic reticulum of viable cells. Proc Natl Acad Sci U S A 94: 8064-8069.

81. Lord JM, Roberts LM (1998) Toxin entry: retrograde transport through the secretory pathway. J Cell Biol 140: 733-736.

82. Kirschke H, Barrett AJ, Rawlings ND (1995) Proteinases 1: lysosomal cysteine proteinases. Protein Profile 2: 1581-1643.

83. Wiederanders B (2003) Structure-function relationships in class CA1 cysteine peptidase propeptides. Acta Biochim Pol 50: 691-713.

84. Otto HH, Schirmeister T (1997) Cysteine Proteases and Their Inhibitors. Chem Rev 97: 133-172.

85. Willstatter R, Bamann E (1929) Uber die proteasen der magenschleimhaut. Hoppe-Seyler's Z Physiol Chem: 127-143.

86. McGrath ME (1999) The lysosomal cysteine proteases. Annu Rev Biophys Biomol Struct 28: 181-204.

87. Turk V, Turk B, Guncar G, Turk D, Kos J (2002) Lysosomal cathepsins: structure, role in antigen processing and presentation, and cancer. Adv Enzyme Regul 42: 285-303.

88. Cygler M, Mort JS (1997) Proregion structure of members of the papain superfamily. Mode of inhibition of enzymatic activity. Biochimie 79: 645-652.

89. Schechter I, Berger A (1967) On the size of the active site in proteases. I. Papain. Biochem Biophys Res Commun 27: 157-162.

90. Choe Y, Leonetti F, Greenbaum DC, Lecaille F, Bogyo M, et al. (2006) Substrate

profiling of cysteine proteases using a combinatorial peptide library identifies functionally unique specificities. J Biol Chem 281: 12824-12832.

91. St Hilaire PM, Alves LC, Sanderson SJ, Mottram JC, Juliano MA, et al. (2000) The substrate specificity of a recombinant cysteine protease from Leishmania mexicana: application of a combinatorial peptide library approach. Chembiochem 1: 115-122.

92. Greenbaum DC, Arnold WD, Lu F, Hayrapetian L, Baruch A, et al. (2002) Small molecule affinity fingerprinting. A tool for enzyme family subclassification, target identification, and inhibitor design. Chem Biol 9: 1085-1094.

93. Illy C, Quraishi O, Wang J, Purisima E, Vernet T, et al. (1997) Role of the occluding loop in cathepsin B activity. J Biol Chem 272: 1197-1202.

94. Hehl AB, Marti M, Kohler P (2000) Stage-specific expression and targeting of cyst wall protein-green fluorescent protein chimeras in Giardia. Mol Biol Cell 11: 1789-1800.

95. Lujan HD, Mowatt MR, Conrad JT, Bowers B, Nash TE (1995) Identification of a novel Giardia lamblia cyst wall protein with leucine-rich repeats. Implications for secretory granule formation and protein assembly into the cyst wall. J Biol Chem 270: 29307-29313.

96. Abel ES, Davids BJ, Robles LD, Loflin CE, Gillin FD, et al. (2001) Possible roles of protein kinase A in cell motility and excystation of the early diverging eukaryote Giardia lamblia. J Biol Chem 276: 10320-10329.

97. Lauwaet T, Davids BJ, Torres-Escobar A, Birkeland SR, Cipriano MJ, et al. (2007) Protein phosphatase 2A plays a crucial role in Giardia lamblia differentiation. Mol

Biochem Parasitol 152: 80-89.

98. Hashimoto T, Nakamura Y, Nakamura F, Shirakura T, Adachi J, et al. (1994) Protein phylogeny gives a robust estimation for early divergences of eukaryotes: phylogenetic place of a mitochondria-lacking protozoan, Giardia lamblia. Mol Biol Evol 11: 65-71.

99. Singer SM, Yee J, Nash TE (1998) Episomal and integrated maintenance of foreign DNA in Giardia lamblia. Mol Biochem Parasitol 92: 59-69.

100. Mort JS, Buttle DJ (1997) Cathepsin B. Int J Biochem Cell Biol 29: 715-720.

101. Que X, Engel JC, Ferguson D, Wunderlich A, Tomavo S, et al. (2007) Cathepsin Cs are key for the intracellular survival of the protozoan parasite, Toxoplasma gondii. J Biol Chem 282: 4994-5003.

102. Turk D, Janjic V, Stern I, Podobnik M, Lamba D, et al. (2001) Structure of human dipeptidyl peptidase I (cathepsin C): exclusion domain added to an endopeptidase framework creates the machine for activation of granular serine proteases. Embo J 20: 6570-6582.

103. Edeling MA, Smith C, Owen D (2006) Life of a clathrin coat: insights from clathrin and AP structures. Nat Rev Mol Cell Biol 7: 32-44.

104. Kirchhausen T (2000) Three ways to make a vesicle. Nat Rev Mol Cell Biol 1: 187-198.

105. Pelkmans L, Burli T, Zerial M, Helenius A (2004) Caveolin-stabilized membrane domains as multifunctional transport and sorting devices in endocytic membrane traffic. Cell 118: 767-780.

106. Bethune J, Wieland F, Moelleken J (2006) COPI-mediated transport. J Membr Biol

211: 65-79.

107. Haucke V (2003) Vesicle budding: a coat for the COPs. Trends Cell Biol 13: 59-60.

108. Royle SJ (2006) The cellular functions of clathrin. Cell Mol Life Sci 63: 1823-1832.

109. Maldonado-Baez L, Wendland B (2006) Endocytic adaptors: recruiters, coordinators and regulators. Trends Cell Biol 16: 505-513.

110. Aridor M, Traub LM (2002) Cargo selection in vesicular transport: the making and breaking of a coat. Traffic 3: 537-546.

111. Robinson MS (2004) Adaptable adaptors for coated vesicles. Trends Cell Biol 14: 167-174.

112. Payne GS, Schekman R (1989) Clathrin: a role in the intracellular retention of a Golgi membrane protein. Science 245: 1358-1365.

113. Deloche O, Schekman RW (2002) Vps10p cycles between the TGN and the late endosome via the plasma membrane in clathrin mutants. Mol Biol Cell 13: 4296-4307.

114. Baggett JJ, Wendland B (2001) Clathrin function in yeast endocytosis. Traffic 2: 297-302.

115. O'Halloran TJ, Anderson RG (1992) Clathrin heavy chain is required for pinocytosis, the presence of large vacuoles, and development in Dictyostelium. J Cell Biol 118: 1371-1377.

116. Wessels D, Reynolds J, Johnson O, Voss E, Burns R, et al. (2000) Clathrin plays a novel role in the regulation of cell polarity, pseudopod formation, uropod stability and motility in Dictyostelium. J Cell Sci 113 (Pt 1): 21-36.

117. Liu SH, Marks MS, Brodsky FM (1998) A dominant-negative clathrin mutant

differentially affects trafficking of molecules with distinct sorting motifs in the class II major histocompatibility complex (MHC) pathway. J Cell Biol 140: 1023-1037.

118. Liu SH, Wong ML, Craik CS, Brodsky FM (1995) Regulation of clathrin assembly and trimerization defined using recombinant triskelion hubs. Cell 83: 257-267.

119. Trejo J, Altschuler Y, Fu HW, Mostov KE, Coughlin SR (2000) Protease-activated receptor-1 down-regulation: a mutant HeLa cell line suggests novel requirements for PAR1 phosphorylation and recruitment to clathrin-coated pits. J Biol Chem 275: 31255-31265.

120. Bennett EM, Chen CY, Engqvist-Goldstein AE, Drubin DG, Brodsky FM (2001) Clathrin hub expression dissociates the actin-binding protein Hip1R from coated pits and disrupts their alignment with the actin cytoskeleton. Traffic 2: 851-858.

121. Bennett EM, Lin SX, Towler MC, Maxfield FR, Brodsky FM (2001) Clathrin hub expression affects early endosome distribution with minimal impact on receptor sorting and recycling. Mol Biol Cell 12: 2790-2799.

122. Liu SH, Towler MC, Chen E, Chen CY, Song W, et al. (2001) A novel clathrin homolog that co-distributes with cytoskeletal components functions in the trans-Golgi network. Embo J 20: 272-284.

123. Owen DJ, Collins BM, Evans PR (2004) Adaptors for clathrin coats: structure and function. Annu Rev Cell Dev Biol 20: 153-191.

124. McNiven MA, Thompson HM (2006) Vesicle formation at the plasma membrane and trans-Golgi network: the same but different. Science 313: 1591-1594.

125. Bonifacino JS, Lippincott-Schwartz J (2003) Coat proteins: shaping membrane

transport. Nat Rev Mol Cell Biol 4: 409-414.

126. Nakatsu F, Ohno H (2003) Adaptor protein complexes as the key regulators of protein sorting in the post-Golgi network. Cell Struct Funct 28: 419-429.

127. Bonifacino JS, Traub LM (2003) Signals for sorting of transmembrane proteins to endosomes and lysosomes. Annu Rev Biochem 72: 395-447.

128. van Vliet C, Thomas EC, Merino-Trigo A, Teasdale RD, Gleeson PA (2003) Intracellular sorting and transport of proteins. Prog Biophys Mol Biol 83: 1-45.

129. Boehm M, Bonifacino JS (2001) Adaptins: the final recount. Mol Biol Cell 12: 2907-2920.

130. Marti M, Li Y, Schraner EM, Wild P, Kohler P, et al. (2003) The secretory apparatus of an ancient eukaryote: protein sorting to separate export pathways occurs before formation of transient Golgi-like compartments. Mol Biol Cell 14: 1433-1447.

131. Brodsky FM, Chen CY, Knuehl C, Towler MC, Wakeham DE (2001) Biological basket weaving: formation and function of clathrin-coated vesicles. Annu Rev Cell Dev Biol 17: 517-568.

132. Rappoport JZ, Simon SM (2003) Real-time analysis of clathrin-mediated endocytosis during cell migration. J Cell Sci 116: 847-855.

133. Towler MC, Gleeson PA, Hoshino S, Rahkila P, Manalo V, et al. (2004) Clathrin isoform CHC22, a component of neuromuscular and myotendinous junctions, binds sorting nexin 5 and has increased expression during myogenesis and muscle regeneration. Mol Biol Cell 15: 3181-3195.

134. Molinete M, Dupuis S, Brodsky FM, Halban PA (2001) Role of clathrin in the regulated secretory pathway of pancreatic beta-cells. J Cell Sci 114: 3059-3066.

135. Marti M, Regos A, Li Y, Schraner EM, Wild P, et al. (2003) An ancestral secretory apparatus in the protozoan parasite Giardia intestinalis. J Biol Chem 278: 24837-24848.

136. Pfeffer SR (2005) Structural clues to Rab GTPase functional diversity. J Biol Chem 280: 15485-15488.

137. Lima AP, dos Reis FCG, Serveau C, Lalmanach G, Juliano L, Menard R, Vernet T, Thomas DY, Storer AC, Scharfstein J (2001) Cysteine proease isoforms from Trypanosoma cruzi cruzipain 2 and cruzain, present different substrate preference and susceptibility to inhibitors. Mol Biochem Parasitol 114: 41-52.

138. Alexander J, Coombs GH, Mottram JC (1999) Leishmania mexicana cysteine proteinase-deficient mutants have attenuated virulence or mice and potentiate a TH1 response. J Immunol 161: 6794-6801.

139. Mottram JC, Souza AE, Hutchison JE, Carter R, Frame MJ Coombs GH (1996) Evidence from disruption of the lmcpb gene array of Leishmania mexicana that cysteine proteinases are virulence factors. Proc Natl Acad Sci 93: 6008-6013.

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

138. de Meester F, Shaw E, Scholze H, Stolarsky T, Mirelman D (1990) Specific labeling of cysteine proteinases in pathogenic and nonpathogenic Entamoeba histolytica. Infect Immun 58: 1396-1401.

139. Lewis PR, Knight DP (1982) Staining methods for section material; Glauert A, editor. New York: Elsevier Biomedical Press. 206-207 p.

140. Zheng QS, Braunfeld MB, Sedat JW, Agard DA (2004) An improved strategy for automated electron microscopic tomography. J Struct Biol 147: 91-101.

UCSF Library Release

Publishing Agreement

It is the policy of the University to encourage the distribution of all theses and dissertations. Copies of all UCSF theses and dissertations will be routed to the library via the Graduate Division. The library will make all theses and dissertations accessible to the public and will preserve these to the best of their abilities, in perpetuity.

Please sign the following statement:

I hereby grant permission to the Graduate Division of the University of California, San Francisco to release copies of my thesis or dissertation to the Campus Library to provide access and preservation, in whole or in part, in perpetuity.

Kelly N. DuBois 9-10-07

Author Signature **Date**