INTERACTOME STUDY OF *GIARDIA INTESTINALIS* CYTOCHROMES B_5

TRENT UNIVERSITY

Peterborough, Ontario, Canada

Environmental and Life Sciences Ph.D. Graduate Program

May 2017

ABSTRACT

Interactome study of *Giardia intestinalis* cytochromes b_5

Guillem Dayer

Giardia intestinalis is an anaerobic protozoan that lacks common eukaryotic heme-dependent respiratory complexes and does not encode any proteins involved in heme biosynthesis. Nevertheless, the parasite encodes several hemeproteins, including three members of the Type II cytochrome b_5 sub-group of electron transport proteins found in anaerobic protist and amitochondriate organisms. Unlike the more well-characterized cytochrome *b5*s of animals, no function has been ascribed to any of the Type II proteins. To explore the functions of these Giardia cytochromes (gCYTB5s), I used bioinformatics, immunofluorescence microscopy (IFM) and co-immunoprecipitation assays. The protein-protein interaction *in silico* prediction tool, STRING, failed to identify relevant interacting partners for any of the Type II cytochromes b_5 from Giardia or other organisms. Differential cellular localization of the gCYTB5s was detected by IFM: gCYTB5-I in the perinuclear space; gCYTB5-II in the cytoplasm with a staining pattern similar to peripheral vacuole-associated protein; and gCYTB5-III in the nucleus. Co-immunoprecipitation with the gCYTB5s as bait identified potential interacting proteins for each isotype. The most promising candidate is the uncharacterized protein GL50803_9861, which was identified in the immunoprecipitate of both gCYTB5-I and II, and which co-localizes with both. Structural analysis of GL50803_9861 using Swiss Model, Phyre2, I-TASSER and RaptorX predicts the presence of a nucleotide-binding domain, which is consistent with a potential redox role involving nicotinamide or flavin-containing cofactors. Finally, the protein GL50803_7204 which contains a RNA/DNA binding domain was identified a potential partner of gCYTB5-III. These findings represent the first steps in the discovery of the roles played by these proteins in Giardia.

SPECIFIC GOALS OF THE PROJECT

The goal of the following project is to gain insights into the function of the Giardia cytochrome b_5-I, II and III (gCYTB5-I, II and III). These proteins belong to the type II of cytochrome b_5 for which no functions are yet known. The first objective of the thesis is to identify other organisms encoding type II cytochrome b_5. Each member of this family will be then used as a query in STRING analysis that generates a predicted interactome for each query protein using the known interactions of its orthologues in other species. These interactomes could help identify potential partners of the gCYTB5s. Since the localization of a protein would give insights into its function and interaction partners, immunofluorescence microscopy will be used to determine the subcellular locations of the three Giardia cytochrome isotypes. Finally, pull-down assays and co-immunoprecipitation experiments will be performed to identify the potential interacting partners of our proteins of interest.

ACKNOWLEDGEMENTS

First I would like to thank Dr. Janet Yee for the opportunity to complete my Ph.D. degree in her laboratory at Trent University and her constant help and support throughout the course of my project. I would also like to thank Dr. Steven Rafferty for co-supervising my project and answering my many biochemistry related questions. I am also grateful to the members of my supervisory committee, Dr. Barry Saville and Dr. Craig Brunetti, for their encouragement and advice, as well as our collaborator Dr. Jan Tachezy and Dr. Jan Pyrih for giving us some of the cell lines I used for my research.

I also would like to thank the members of the Dr. Yee and Dr. Rafferty labs with a special mention to Katie for her great help especially with IFM and for cell culturing, to Danielle, Liza and Eliza for their work on gCYTB5-III and to Meg, Adam, Jennifer, Elizabeth, Kelly and Julie for always be willing to give a hand when needed. Finally, I would like to thank my family and especially my parents and my brother for their moral and financial support during all these years as well as Rose and Adèle for their encouragements and Dalila and Sibylle for always making me smile.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

1.1 *Giardia intestinalis*

Giardia intestinalis is a binucleated intestinal protozoan parasite found in freshwaters worldwide that causes giardiasis (Adam 2001). The pathophysiology of giardiasis includes nausea, abdominal pain, and diarrhea (Sullivan et al. 1991; Adam 2001). Chronic diarrhea from giardiasis in children is linked to malnutrition or malabsorption and may play a role in their growth retardation (Sullivan et al. 1991; Gillin et al. 1996).

The life cycle of Giardia is divided into two main stages: the vegetative trophozoite, which causes the symptoms associated with the disease, and the infective cyst, which is the form found in the environment. The infectious cysts enter the host through ingestion of contaminated food or water (Adam 2001). When the cysts reach the stomach, the presence of gastric acid activates the excystation process, which generates two trophozoites per cyst. In the proximal small intestine, the trophozoites undergo proliferation by asexual binary fission and colonize the small intestine (Gillin et al. 1996; Adam 2001). While the underlining mechanism is not fully understood, it has been suggested that the symptoms of giardiasis are triggered by the attachment of the trophozoites through its ventral adhesive disc to the intestinal epithelia (Wolfe 1992). Trophozoites take advantage of adhesion to obtain nutrients from the intestinal lumen and to avoid being flushed out by peristaltic flow (Adam 2001). Eventually, trophozoites pass through the jejunum where exposure to biliary fluids and the lack of nutrients stimulates their encystation. The dormant cysts will then be excreted through the feces (Adam 2001).

1.2 Cytochrome b_5

Cytochromes b_5 are ubiquitous hemeproteins found in all kingdoms of life. The heme-binding domain is ~12 kDa (90 amino acid residues) and contains two invariant histidines as axial ligands to the heme iron (Fig.1). Well-studied examples of cytochromes b_5 are the mammalian proteins with a C-terminal hydrophobic extension that anchors the protein to the endoplasmic reticulum (CYTB5A) or the outer mitochondrial membrane (CYTB5B). A soluble form of CYTB5A is present in red blood cells, which arises as a result of alternative splicing that deletes the coding sequence for the C-terminal membrane anchor.

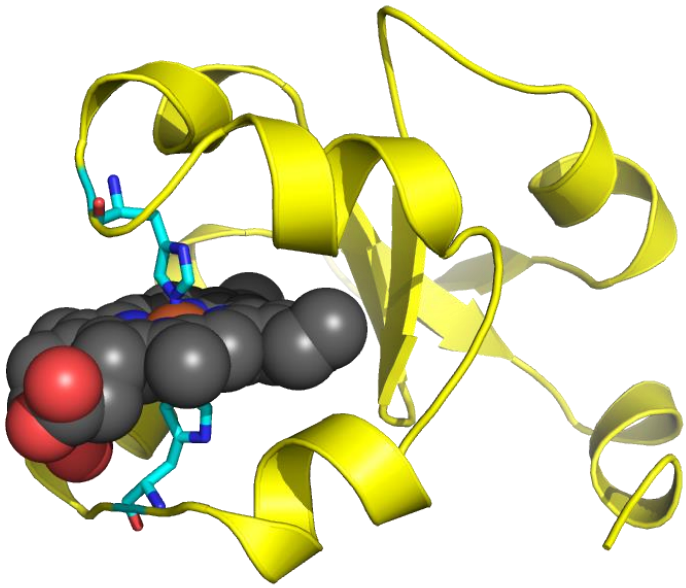


Figure 1: Structure of the bovine microsomal cytochrome b_5. Heme is presented in space filling mode, and the two invariant histidines are in cyan.

Sequence alignment of cytochrome b_5 sequences shows that these proteins can be classified into three types. Type I contains the well-characterized membrane-anchored microsomal (ER) and mitochondrial cytochromes b_5. Additional common features of Type I cytochromes b_5 are highly conserved consensus sequences surrounding the two

conserved histidine ligands, **H**PGG and G**H**S; in addition the former motif is typically flanked on each side by two pairs of acidic residues that help form the ring of negative charge about the exposed heme edge that promotes favourable electrostatic interactions with electron transfer partners (Salemme 1976). Type II proteins are soluble and possess the first **H**PGG motif but lack the acidic flanks, and the second histidine lies in another motif, **H**XWV(N/S) instead of GHS. Type III contains a protein found only in fungi such as *Neurospora crassa.*

The roles of Type I proteins are well-known, based on extensive studies on the mammalian cytochromes b_5. Microsomal/ER cytochrome b_5 (CYTB5A) is a versatile electron shuttle that acts in several metabolic pathways by transferring electrons from NADH-cytochrome b_5 reductases or NADPH-cytochrome P450 reductases to a wide range of acceptors, especially the cytochrome P450s monooxygenases (Schenkman and Jansson 2003). Outer mitochondrial cytochrome b_5 (CYTB5B) funnels electrons from sources within the mitochondria and donates them to mitochondrial cytochrome *c,* which in turn transfers them to cytochrome *c* oxidase, the final member of the respiratory chain (Doroshchuk and Dmitriev 2013). Soluble CYTB5A of red blood cells donates electrons to oxidized hemoglobin and restores hemoglobin to its active reduced state (Vergeres and Waskell 1995; Schenkman and Jansson 2003). In addition, cytochrome b_5 is also found as a domain within multidomain proteins. This includes fatty acid desaturases, nitrate reductase and sulfite oxidase (Lederer 1994; Schenkman and Jansson 2003) .

1.3 Giardia cytochromes b_5

As a highly derived protist, Giardia lacks certain organelles typically associated with eukaryotes (Adam 2001; Thompson and Monis 2012). Of particular relevance is

its lack of mitochondria, which is the site of oxidative phosphorylation, heme biosynthesis, and iron-sulfur cluster assembly. Giardia uses anaerobic glycolysis and substrate-level phosphorylation to fulfill its energy needs (Jarroll et al. 1989). Giardia does possess mitosomes, small organelles descended from mitochondria that retain the ability to synthesize iron-sulfur clusters but not the other two functions (Tovar et al. 2003). Furthermore, genes encoding mitochondrial respiratory chain proteins or heme biosynthetic enzymes appear to be absent in the Giardia genome. This, and the lack of biochemical evidence for the presence of common heme proteins such as catalase, led to the initial assumption that Giardia does not use heme (Alam et al. 2012).

Surprisingly, more recent studies of the *G. intestinalis* genome led to the identification of several heme-binding proteins including four members of the cytochrome b_5 family, all of which belong to Type II (Morrison et al. 2007; Alam et al. 2012; Pyrih et al. 2014). Three of the Giardia cytochrome b_5 isotypes (gCYTB5-I, II and III) are of similar size (14.5 to 15.5 kDa) and share all the characteristics of Type II members with the exception of an amino acid substitution in the **H**PGG motif to **H**PAG in isotypes I and III. Moreover, the regions that flank the conserved heme-binding domain of the gCYTB5s are not conserved with other members of this Type or amongst themselves, which suggests that each gCYTB5 may have specific interacting partners (Fig. 2). The gCYTB5-IV protein is distinct from the other isotypes as the motifs that include the axial histidine ligands are less well-conserved and the protein is much larger (29 kDa), owing to a longer N-terminal extension at the heme-binding domain. Despite these differences, gCYTB5-IV shares enough characteristics within its heme-binding domain to be classified as a Type II cytochrome b_5.

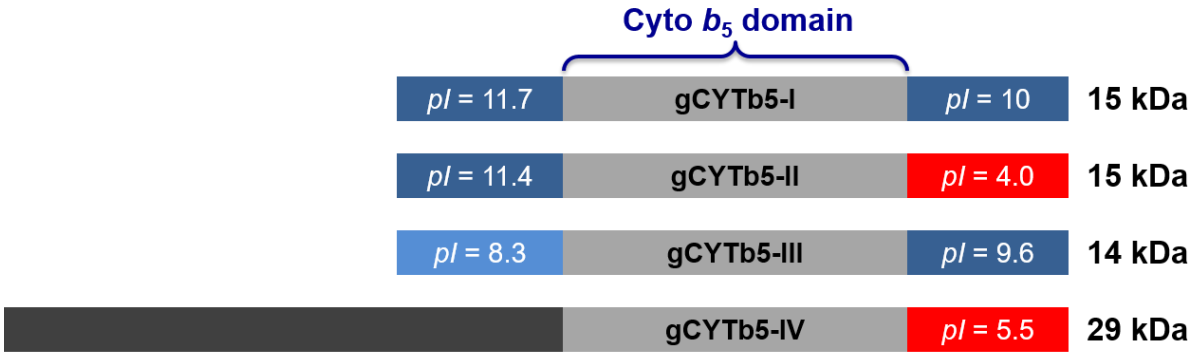


Figure 2: Schematic comparison of the gCYTB5s sequences.

The three smaller proteins (gCYTB5 I, II and III) were expressed as recombinant proteins in *E. coli*, which were isolated with heme bound to them and have spectroscopic properties expected for cytochromes b_5 (Alam et al. 2012; Pazdzior et al. 2015). They are also capable of stimulating NADH oxidation in Giardia cell lysates (Campanaro 2013), and are reduced by recombinant Giardia NADPH-dependent oxidoreductase GiOR-1 (Pyrih et al. 2014), which contains a cytochrome P450 reductase-like domain. Furthermore, *G. intestinalis* is able to take up heme from its environment and the heme group is incorporated in all gCYTB5 apoproteins (Pyrih et al. 2014). Collectively, these observations suggest that the genes for the gCYTB5s encode functional heme proteins that may have potential roles in electron transfer pathways.

However, with the possible exception of GiOR-1 the identity of these partners is presently unknown; the partners that are so well-characterized among mammalian Type I cytochromes b_5 are absent from the Giardia genome. These expected protein partners may be highly divergent in Giardia, or perhaps the Giardia cytochromes are used in different electron transfer pathways with novel partners. The identification of the interactome of the gCYTB5s – those proteins that are binding partners, which would

include electron donors and acceptors – is the aim of this thesis. To do so, both bioinformatics and physical experimental methods are used.

1.4 Bioinformatics and the Prediction of Interacting Partners

Although BLASTp searches of the Giardia genome did not identify proteins with sequence similarity to known electron transfer partners of cytochromes b_5, it still may be possible to identify prospective interactome members through other bioinformatics tools. STRING (Search Tool for the Retrieval of Interacting Genes/Proteins) is an online resource that contains the widest database of both direct and indirect protein interactions obtained from experimental and published data on the largest number of organisms (Franceschini et al. 2013). The latest release of the online resource (STRING 10) contains close to 10 million proteins distributed among more than 2000 organisms (http://string-db.org/). This tool also uses orthologous proteins to generate the network, which is very useful in cases where there is no direct information available for the species of interest. For example, if the database contains evidence that protein α and β interacts in yeast and the query protein from organism X has sequence similarity to the yeast protein α, then STRING predicts that proteins α and β from organism X would also interact.

For any query protein belonging to an organism within the database, the program generates a network of predicted physical interactions and functional interactions using information available in the different databases, called views:

1. The *Genomic Context view* includes three groups: the conserved neighbourhood group shows genes that occur on the same operon in prokaryotes; the co-occurrence group corresponds to the proteins that occur in the same KEGG (Kyoto

Encyclopedia of Genes and Genomes) metabolic pathway in several species; and the fusion group shows genes that are fused in some species.

2. The *Co-expression view* contains information about the genes that are co-expressed within the species of interest or in different species using homology analysis.
3. The *Experiment view* retrieves information from databases based on the results of large-scale, high-throughput experiments studies on whole cells from mostly yeast two hybrid screening experiments.
4. The *Databases view* corresponds to the database of an annotated pathway such as KEGG (Kyoto Encyclopedia of Genes and Genomes).
5. The *Text Mining view* searches the query protein alongside other proteins in different published papers available on PubMed (Franceschini et al. 2013)(www.string-db.org).

Each view has certain weaknesses. For example, the Experimental view is prone to a high level of false positives owing to its reliance of data from yeast two-hybrid system that share this drawback. Caution should be used in interpreting the Text mining view, as a retrieved protein may have been used as a control in an experiment involving a protein matching the query, or the two proteins are mentioned in unrelated contexts within the same publication.

Nonetheless, when used collectively the Views are informative. When a query protein is analysed using STRING, a score is calculated for each of the different views and a total score is provided. The final score obtained is between 0 to 1.0 where 0.9 and above is considered the highest confidence score, and 0.4 to 0.9 is considered as a medium confidence score.

As an example of a STRING analysis, consider the example where the beta subunit of *E. coli* tryptophan synthase (trpB) is the query or input and its predicted partners as the output (Figure 3).

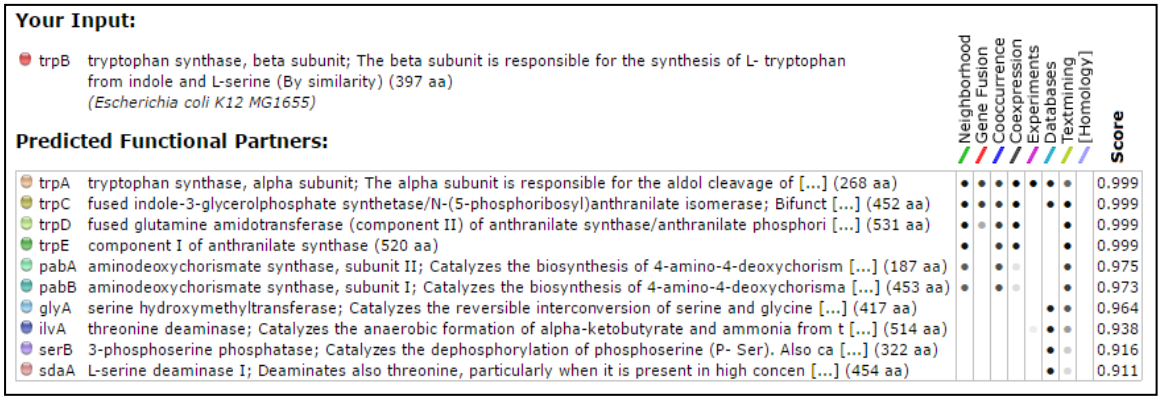


Figure 3: Snapshot of the STRING results for *E. coli* trpB protein. STRING yields several proteins with a confidence score higher than 0.900, with trpA protein common to all views. The degree of confidence for each view is representing by the intensity of the dot; the darkest dot corresponds to a high level of confidence.

The trpA protein is identified as a "Predicted Functional Partner" in all the Genetic Context views (Figure 4, 1st row). The Experiment view contains information retrieved from large-scale interactome studies that have been performed in different species including *E. coli*, *S. cerevisiae* and *H. pylori*. In this example for trpB, 77 sets of such data were used. The Databases view contains information from KEGG and other curated pathways, whereas Text Mining is based on 50 publications where trpA and trpB appears together.

In Figure 4 panel A, a schematic representation obtained after STRING analysis shows the genomic organization of trpB and trpA in multiple organisms. These two proteins are also found fused as two domains in a single protein in certain species such as the fungi *Thielavia terrestris* (Fig. 4, panel B). Finally, these proteins occur together in same metabolic pathway in several species (Fig. 4, panel C). Note that each scheme

represents only a small part of the graph generated. The proteins are also co-expressed in several species including plants (*Arabidopsis thaliana)* and yeast (*Saccharomyces cerevisiae)* (Fig. 4, panel C).

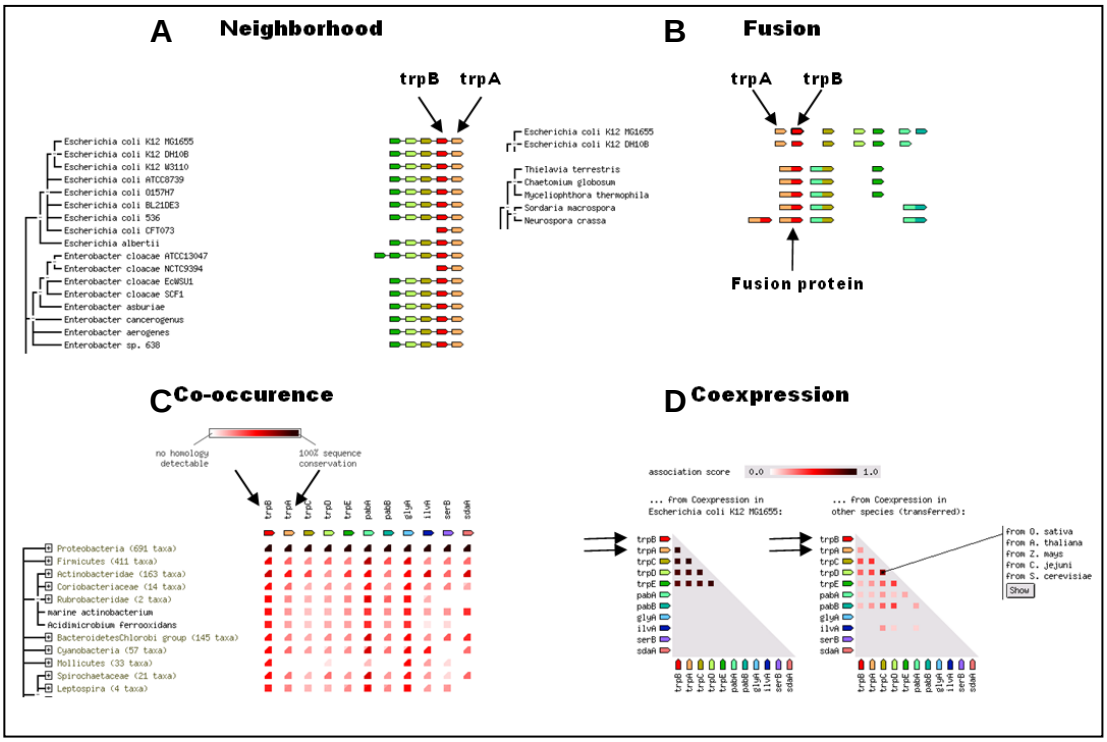


Figure 4: Genetic Context view and Coexpression view of the trpB protein. The figure shows different snapshots taken from the STRING results and gives a schematic representation of the Neighbourhood (Panel A), Fusion (Panel B) and Co-occurrence (Panel C) view of the trpB and trpA protein interaction. Panel D corresponds to the Coexpression view, where black arrows are pointing to trpB and trpA and the level of confidence for the Co-occurrence and Coexpression views are color coded (white = 0, dark red = 1).

Since the common partners of the canonical cytochrome b_5 are known and show consistency throughout multiple species (Vergeres and Waskell 1995; Schenkman and Jansson 2003), it is possible that this consistency also occurs with the partners of the Type II cytochromes b_5. Since the functions of these protein remain unknown (Pyrih et al. 2014), the first part of this thesis is to identify species other than Giardia that contain Type II cytochrome b_5 and to use the STRING tool to predict interacting partners in each species. Identification of common partners among the different species would allow us look for orthologues of these proteins in Giardia as potential interactors of gCYTB5s.

1.5 Experimental methods for interactome studies

Bioinformatics studies such a STRING can identify potential partners of a protein but it is also necessary to have experimental evidence of a protein-protein interaction (PPI). The study of these dynamic molecular interactions has become a fundamental topic of research in the understanding of cell activity, survival or control of the cell cycle as well as disease development (Cusick et al. 2005; Andreani and Guerois 2014). PPIs can be classified into two main categories: binary interactions (BI) and co-complex interactions (CCI) (Bonetta 2010). Methods for identifying binary interactions include the yeast-two-hybrid system, luminescence-based mammalian interactome mapping (Lumier) and protein microarray. Binary approaches have been extensively used in interactome studies where the complete network of protein interactions in a given species or metabolic pathway are analysed by testing every possible direct interaction between a pairwise combinations of bait and prey proteins (Cusick et al. 2005; Bonetta 2010). In contrast, co-complex interaction studies typically focus on a single bait protein used with a cell lysate that contains a complex mixture of

prey proteins, After affinity purification of the prey along with its interacting proteins, mass spectrometry analysis is used to identify all the recovered proteins in a single experiment (Cusick et al. 2005). Each approach has its own advantages and disadvantages. While the binary approach is more amenable to high throughput implementation, it is also more likely to generate false positives, that is, identifying prospective partners that actually do not interact. The co-complex approach is more suitable for small-scale analysis and has proven efficiency in discovering new interactions, but will generate more false negative results, in which a true interaction *in vivo* is missed during the analysis (Cusick et al. 2005; Bonetta 2010; De Las Rivas and Fontanillo 2010).

In this thesis, I am interested in studying the interactome of the Giardia cytochromes b_5. I used several co-complex approaches to identify their potential partners.

1.5.1 In vitro Pulldown assays

In a pulldown assay, the bait protein is anchored to an inert support, typically a chromatography resin. A mixture of proteins such as a total cell extract that contains potential partners (the prey) is applied to the resin. After washing the non-interacting proteins from the resin, the interacting proteins are released from the anchored bait by applying an elution buffer. Depending on the strength of the interaction between the bait and prey, the elution condition may be relatively mild (i.e. an increase in ionic strength) or harsh (i.e. a decrease in pH, or the application of protein denaturants such as SDS.)

The most well-known example of a pulldown assay involves tagging the bait protein with Glutathione S-Transferase (GST), which is also commonly used as an

affinity tag to express and purify recombinant proteins from *E. coli*. The method takes advantage of the strong affinity of GST for glutathione. In a GST pulldown, a plasmid construct is prepared that encodes a fusion protein between the bait and the GST tag. The recombinant fusion protein is bound to an affinity resin that has glutathione covalently attached to it, such as glutathione agarose. The interaction between the glutathione and the GST domain is very strong, and the effective result is that the bait is anchored to the resin. An extract of proteins that contains potential prey is then applied to the resin. After washing off the unbound proteins, the GST-bait and prey complex is eluted by the addition of glutathione, and the prey identified, usually by mass spectrometry (*see description below*) (Wang et al. 2016).

The use of GST-bait fusions in such experiments has two major disadvantages. First, as GST itself is a homodimer there is a tendency for GST fusion proteins to form oligomers, especially if the bait itself has quaternary structure. More importantly, as the GST tag is relatively large (MW 26 kDa), the likelihood increases that it could interfere with the ability of the bait protein to bind to its interacting partners (Medina et al. 2000). This has lead researchers to use smaller tags, which are less likely to have the same drawbacks. Chief among these is the hexahistidine tag (His_6-tag). Recombinant His_6-tagged proteins have a high affinity for immobilized metal affinity chromatography (IMAC) resins such as nickel (II)-nitrilotriacetic acid agarose (NiNTA agarose). The tagged bait-prey complex is selectively release from the resin by the addition of imidazole, which, having the same functional group as histidine, acts as a competing ligand for the bound metal. His_6-tagged pulldowns have been successfully used to identify the partners of the VimA protein in the bacterium, *Porphyromonas gingivalis*, which is implicated in periodontal disease (Aruni et al. 2012).

As an alternative to tagged bait proteins, one can also use an untagged bait protein that is linked directly to activated chromatography resins. Affi-Gel 10 resin is an agarose based resin that contains a reactive N-hydroxy-succinimide group on a 10-atom spacer arm. This group reacts with primary amines such as the lysine side chains of proteins to form covalent bonds, thereby linking the protein to the resin. In a pull-down assay using the Affi-gel resin, a purified bait protein is first bound to the resin, and the remaining reactive sites blocked using ethanolamine hydrochloride to prevent covalent attachment of other proteins to the resin. The baited resin is then added to a sample of potential prey proteins as described above (Markillie et al. 2005). Affi-Gel 10-baited resins have been used in pull-downs to identify the partners of the yeast type I arginine methyltransferase (Hsieh et al. 2007).

1.5.2 CO-IMMUNOPRECIPITATION

The most common method used to study protein interactions is co-immunoprecipitation (co-IP), which uses an antibody raised against the bait protein or against an epitope tag fused to the bait (Figure 5). If the endogenous protein or an epitope-tagged protein is expressed within the cell, the bait and prey interaction will occur *in vivo*. The interaction complex could be captured from the cell lysate by incubation with the antibody. The antibody itself is usually linked to a resin by their interaction with Protein A or Protein G-conjugated beads. A short centrifugation would pellet the beads. Washes are performed on the pellet to remove the unbound proteins and the bait/prey complexes are eluted under acidic conditions with 0.1 M glycine, pH 2.0. The eluted samples are subjected to SDS-PAGE to separate proteins based on their size. Gel slices are excised and analysed by mass spectrometry to identify prey proteins (Chang 2006; Miernyk and Thelen 2008; Free et al. 2009).

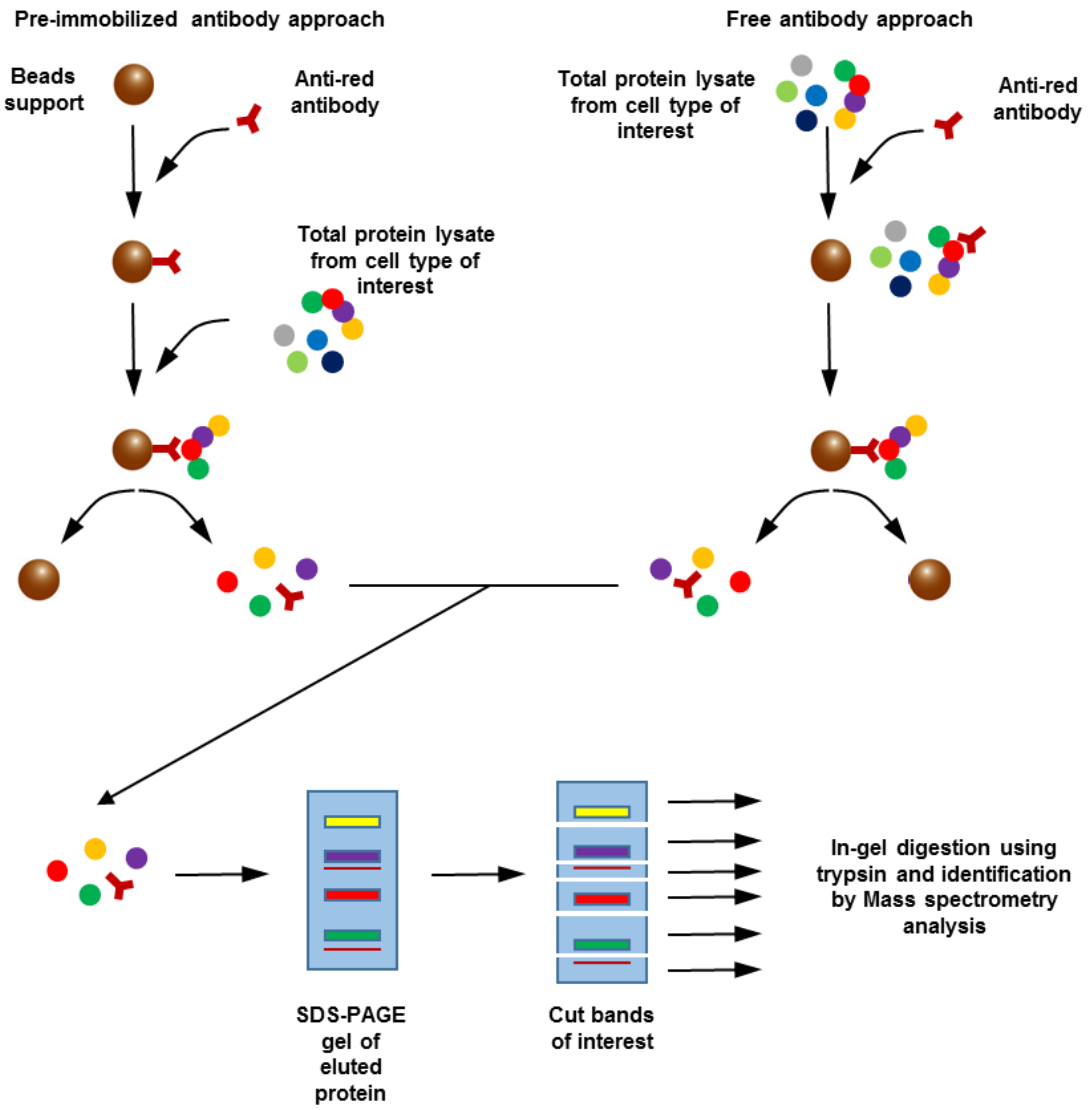


Figure 5: Schematic representation of Co-IP purification followed by MS analysis. The bait protein could be captured either using the pre-immobilized antibody approach or the free antibody approach. In the pre-immobilized approach the antibody is bound to the beads via the protein A or G (not show here) and the total protein lysate from the cell type of interest is added to allow the precipitation of the bait protein and its interacting partners. In the free antibody approach the antibody is mixed with the protein extract, and then the beads are then added to capture the antibody its bound target. Regardless of the approach used, the eluted proteins are then separated on a SDS-PAGE, the different protein-containing bands are then excised and send for mass spectrometry analysis. Note that during the elution the heavy and light chain of the antibody are usually separated and will yield two bands (the light chain usually found at around 25 kDa and the heavy chain at 50 kDa) easily observable on the polyacrylamide gel.

Co-IP offers several advantages over pulldown assays. In Co-IP the bait and prey are in their natural cellular environment where they retain their native conformation and post-translational modifications (Dwane and Kiely 2011). Furthermore, as these conditions may favour retention of any organelle specificity associated with the target protein complex, Co-IP can be coupled to subcellular fractionation experiments.

The main limitation of this approach is the necessity for an antibody specific to the bait of interest, which may be costly and time-consuming to prepare, and the possibility that the resultant antibody may not show strong specific binding to the target protein. A common solution to this challenge is to express within the cell type of interest, a recombinant version of the bait that contains an epitope tag that is recognized by commercially available antibody. Tandem repeats of the hemagglutinin (HA) tag of nine residues are commonly used for this purpose. Resin consisting of anti-HA antibody-conjugated beads for the co-IP experiments is also commercially available (Dastidar et al. 2012; Mellacheruvu et al. 2013; Muller et al. 2015).

1.5.3 Mass spectrometry analysis

The result of a pulldown or Co-IP experiment will be a mixture of proteins that include the bait, and prey, as well the antibody and non-interacting proteins (false positive), which tend to be highly-abundant proteins. To identify which proteins are recovered from the co-IP, mass spectrometry (MS) analysis is the method of choice, owing to its sensitivity, speed, and ability to simultaneously detect different proteins within a sample. The assignment of specific proteins in the sample is due to the usage of software that can correlate observed features on a mass spectrum with predicted features based on a knowledge of the cellular proteome if the complete genome for the

organism of interest is available (Meyer and Selbach 2015; Smits and Vermeulen 2016). MS analysis is typically done on samples that have been partially fractionated by SDS-PAGE, which separates proteins based on size. The gel is stained to detect proteins, and slices of the gel corresponding to a single protein or a specific size range can be excised. The gel slices are treated with specific proteases *in situ* to generate a mixture of peptides that are amenable to MS analysis. The most commonly used enzymes are trypsin (cuts after basic residues Lys and Arg), chymotrypsin (cuts after aromatic residues Trp, Tyr, Phe) and GluC (cuts after the acidic residue Glu).

Mass spectrometers differ in the details of their components but have in common three parts: an ion source, which generates gas phase ions of the peptides; a mass analyser, which separates these ions based on their mass to charge ratio, *m/z*; and an ion detector, which measures the abundance of each ion. The two most widely used ion sources suitable for biomolecules are Matrix-Assisted Laser Desorption Ionization (MALDI) and Electrospray Ionization (ESI). The mass analyser uses a combination of electrical and magnetic fields to separate the ions; examples include time-of-flight (TOF), quadrupole, ion trap, and orbitrap (Figeys et al. 2001; Smits and Vermeulen 2016).

The resulting mass spectrum is a plot of ion *m/z* on the abscissa and ion abundance on the ordinate, with each peak corresponding to a different ion. As each peptide can exist in multiple charge states the number of peaks observed is typically more than the number of peptides present. The central point is that the peaks for a particular peptide are *predicable.* Thus, the observed mass spectrum can be analysed against the expected digested peptides from a proteome database of the species of interest (based on the predicted proteins encoded by the genome). Several versions of software are available for this type of bioinformatics analysis, but the final dataset

generated contain similar information. For each protein identified, the analysis result includes its relative abundance in the sample based on the mass spectrum, the number of distinct peptides observed that correspond to this protein, the relative frequency of each peptide, as well as the molecular weight of the protein identified (Figeys et al. 2001).

1.5.4 PROTEIN STRUCTURE PREDICTION

More than half of the Giardia genome encodes proteins with no sequence similarity to known proteins. These fall under the category of putative uncharacterized proteins (PUPs) or hypothetical proteins, for which no function has yet been identified. Consequently it is likely that some of the interacting partners of the gCYTB5s will be PUPs, and it would be necessary to use other bioinformatic tools to obtain information about these proteins.

One of the first steps in characterizing a PUP is to use a protein Basic Local Alignment Search Tool (BLASTp) that compares the query sequence against all known protein sequences. This analysis returns a list of proteins with the highest level of sequence identity (exact sequence matches) and similarity (where residues do not match but have similar properties such as charge or hydrophobicity). It also allows the identification of conserved domain within the query protein. While this is fast and easy to do, it requires a certain level of similarity/identity between the query and the subject protein, which is not always reached, especially with PUPs (Altschul et al. 1990; Mount 2007). Given that a PUP is by its nature uncharacterized, the most likely outcome of a BLASTp search is to uncover other PUPs. Yet this itself is revealing as it implies that the significance of this protein goes beyond a single occurrence in one species.

Another approach is to make predictions on the fold or structure of a PUP. This is possible because a protein fold is more highly conserved than its amino acid sequence, (Krieger et al. 2003; Illergard et al. 2009). Several protein homology modelling tools that make predictions of structure based on the sequence have been developed to identify folds within a protein. These include Swiss-Model (Schwede et al. 2003) , Phyre2 (Kelley et al. 2015), RaptorX (Kallberg et al. 2014) and IntFOLD (McGuffin et al. 2015). These tools use different algorithms to generate a protein model from a query sequence (protein of interest) using a database of structurally characterized protein (template), but their strategies are similar.

In some algorithms such as Swiss Model, the query sequence is analysed by BLASTp and multiple sequence alignments against the sequences with known structures in the Protein Data Bank, (http://www.rcsb.org) to identify templates with sequence similarity. The query protein containing the conserved fold can then be modelled based on a template fold. The initial model structure can be energy-minimized to alleviate steric clash and electrostatic repulsion, and enhance electrostatic attractions. Note that different domains within the same query can be modelled separately. In this case a final step is to generate a stable structure after the of the assembled domains (Krieger et al. 2003).

The procedure above works well if sequence similarity is relatively high (≥ 30%) in which case it highly likely that the query and templates share the same fold. This situation does not apply for PUPs (see the comments on BLASTp, above). In principle it is possible to model a query PUP to a template PUP with a solved structure, but fewer than 0.6% (714) of the structures in the Protein Database (>125,000) are of uncharacterized proteins. In such cases, protein modelling programs based on protein threading will be more effective. These do not rely on sequence, but rely on the pattern

of the sequence properties (charge, size, tendency to form certain secondary structures) and statistical analysis of the properties of proteins of known structures to develop model structures. Raptor, Phyre and I-TASSER are examples of programs that predict structures based on protein threading. Certain bioinformatics sites permit one to submit a query sequence to multiple predictive programs, which has the advantage of allowing one to compare the various model structures obtained.

1.6 Research goals

Investigation of the Giardia genome has not identified any of the canonical cytochrome b_5 partners, suggesting that these proteins might have unique functions. My aim is to identify the interacting partners of the Giardia cytochromes b_5 isotypes I-III, which may reveal the roles of these proteins. Furthermore, although Type II cytochromes b_5 have been identified in most eukaryotes, they have no functions assigned so far (Pyrih et al. 2014). It can therefore be anticipated that a better understanding of the gCYTB5 partners and functions in Giardia will be relevant to understanding the entire class II cytochrome b5 in eukaryotes as a whole.

EXPERIMENTAL PROCEDURES

2.1 STRING Searches for interacting partners of Type II Cytochromes b_5

STRING was used to search for interacting partners for the Type II cytochromes b_5 in Giardia and in other species. The outcomes of these searches are described in the Results section. The protein UniProtKB number was entered as a query using the "Protein by name" search. If no results were obtained, the protein sequence and organism was inputted into the "Protein by sequence" search option. The results were analysed using the following Data setting options: all the active interaction sources were selected, a minimum required interaction score of 0.400 (medium confidence, maximum score is 0.999). If no results were obtained the analysis was performed using the lowest confidence score possible (below 0.150).

2.2 Giardia intestinalis cell culture

Giardia intestinalis trophozoites (ATCC#50803 WB clone C6) were cultured at 37°C in complete TYI-S-33 medium (Appendix A) in 16 mL glass culture tubes as previously described (Gillin et al. 1996). Immediately before use, the media was supplemented with cysteine (final concentration of 2 g/L in medium, Bioshop Canada) and ascorbic acid (final concentration 0.2 g/L in medium) and antibiotic-antimycotic (1X final concentration, Hyclone). Cell counts were measured with a ViCell XR® cell viability analyser (Beckman Coulter).

2.3 Immunofluorescence microscopy

The detailed procedure for preparation of samples for immunofluorescence microscopy is presented in Appendix B (Dr. Yee's laboratory protocol). In brief, glass microscope coverslips were pre-treated by addition of 100 µL of 0.1% polyethylenimine (PEI) diluted in PBS. The coverslips were rinsed with Millipore-grade water to remove excess PEI and were dried at room temperature. Giardia cells suspended in 75 µl of PBS were added to the coverslip, which was transferred onto a stack of wet paper towels into a Tupperware container that served as a humidity chamber. The lid was placed on the container and placed in an incubator at 37°C for 10 minutes. The coverslips were immersed in pre-chilled methanol at -20°C for 10 minutes. To permeabilize the cell membrane the coverslips were air-dried, then transferred, cell-side down, onto a droplet of 0.5% Triton X-100 in PBS for 10 minutes. The coverslips were transferred onto a drop of blocking buffer (200 µL, 50 mM Tris, 150 mM NaCl, 0.5% NP-40, 5 mg/mL bovine serum albumin) for an hour. The coverslips were subsequently incubated with the primary antibody and, after PBS washes, onto the secondary antibody. All the antibodies used in this experiment, their dilution and the buffer in which they were diluted are listed in Table I.

After the secondary antibody treatment, the coverslips were rinsed with PBS and placed on 200 µL of 3.7% paraformaldehyde in PBS for 10 minutes. After further rinses with PBS and Millipore-grade water, the coverslips were placed on slides bearing a drop of Vectashield mounting agent that contained the DNA stain DAPI for detection of the nuclei. The coverslips were sealed using nail polish, and were visualised the next day by fluorescence microscopy on a Leica DM 6000B microscope.

Table 1: Antibodies used in IFA and their conditions of use. Secondary antibodies are shaded grey. Antibodies were diluted in blocking buffer.

Antibody	Dilution	Incubation Time	Incubation Temperature	Provider
Rabbit gCYTB5-I	1:300	2 hours	Room temperature	GeneScript (custom order)
Rabbit gCYTB5-II	1:500	2 hours	Room temperature	GeneScript (custom order)
Rabbit gCYTB5-III	1:2000	2 hours	Room temperature	GeneScript (custom order)
Mouse Anti-HA	1:200	2 hours	Room temperature	Sigma (Cat # H9658)
Mouse Anti-HA	1:200	2 hours	Room temperature	Thermo Scientific (Cat # 26183)
Cy5-conjugated Goat anti-mouse antibody	1:200	1 hour	Room temperature	Jackson ImmunoResearch (Cat # 115-175-146)
Cy3-conjugated Goat anti-rabbit antibody	1:200	1 hour	Room temperature	Jackson ImmunoResearch (Cat # 115-165-144)

Due to the difficulty to obtained good results with the anti-HA antibodies, I followed the protocol from Lenka Cernikova in Dr. Adrian Hehl's lab (Appendix B) for the colocalization experiment of HA-PUP9861 and gCYTb5-II. Briefly, 10^7 Giardia trophozoites were resuspended and incubated in 1 mL of fixation solution (3% paraformaldehyde in PBS) for 1 hour at 37°C in a microcentrifuge tube. The cells were centrifuge at 900 *g* for 5 minutes at 4°C (all the following centrifugation were performed using this setting) and the supernatant was discarded. The pellet was resuspended in 0.1 M glycine in PBS for 15 minutes at room temperature and after centrifugation the buffer was removed. The permeabilization solution (0.2% BSA and 0.2% Triton X-100 in PBS) was added for 20 minutes at room temp. After centrifugation, the supernatant was discard and the cells were incubated at room temperature for an hour in the blocking buffer (0.2% BSA in PBS). Once the blocking buffer was removed, the primary

antibodies solution (Sigma mouse anti-HA 1:50 and rabbit anti-gCYTB5-II 1:500 diluted in the permeabilization buffer) was added. The cells were incubated with the antibodies overnight at 4°C on the nutator. The cells were centrifuge and the primary antibodies solution was discarded. Two 5 minutes washes were done using the wash buffer (1% BSA and 0.1% Triton X-100 in PBS). The secondary antibodies solution (Cy5-conjugated goat anti-mouse antibody and Cy3-conjugated goat anti-rabbit antibody 1:200 in permeabilization buffer) was added and the cells were incubated 2 hours at room temperature on the nutator. Two washes were performed as before and the cells were resuspended in 10 µL of Vectashield mounting media with DAPI. After a 10-minute incubation on ice, 5 µL of the suspension was loaded on two slides and the cells were allowed to dry for 30 minutes at 37°C. The coverslips were the added and fixed using nail polish.

2.4 Expression of recombinant Giardia cytochromes b5

Purified recombinant Giardia cytochrome b_5 isotypes I-III were obtained from other students (M. Mesbahuddin, D. Drake and L. Calhoun) in the laboratory of Dr. Steven Rafferty (Peterborough, ON). These were expressed as N-terminal His_6-tagged proteins in *E. coli* strain BL21 from a pJ401 vector background (DNA 2.0, Menlo Park CA). Cells were lysed by detergent solubilisation in buffer containing 1% octylthioglucoside. Proteins were purified by a combination of ammonium sulfate precipitation and immobilized metal affinity chromatography on cobalt-charged HisPur resin (Thermo-Fisher Canada, Mississauga ON). Purified proteins were stored in 50% glycerol at -20° C. To assess the protein purity, 100 µg of purified protein was loaded on a 10 % SDS-PAGE gel. After protein separation the gel was stained with PageBlue

Protein Staining Solution (Thermo Scientific, sensitivity of 5 ng, Cat# 24620). An example of the result obtained is shown in Appendix C.

2.5 Preparation of Giardia cell extracts

Giardia trophozoite cultures were centrifuged at 1,200 *g*, 4°C for 15 minutes. The media was removed and the cells were resuspended in PBS and transferred into a pre-weighed microfuge tube. The tubes were centrifuged as before, the PBS discarded, and the cell pellet weights recorded. Cell pellets were stored at -80°C.

Cells were thawed and lysed in CelLytic Y buffer (Sigma) supplemented with 10 mM DTT, 0.5 µg/µL Leupeptin, and 1X protease inhibitor cocktail (Bioshop Canada). The volume of lysis buffer used was 2.5 - 5.0 µL per milligram of cells (*i.e.* 250 - 500 µL of lysis buffer was added to 100 mg of cells). The samples were incubated on a nutator for 30 minutes at 4°C. Cell lysates were centrifuged at 14,000 *g*, 4°C for 10 minutes and the protein-containing supernatant was transferred to a clean microfuge tube. Protein concentrations were determined by Bradford assay and the protein extracts were either used immediately or stored at -80°C.

2.6 Western blot analysis

Giardia cytochromes b_5 are small (14.5 to 15.5 kDa) and such proteins require modification of Western blotting conditions to ensure their retention on the nitrocellulose membrane. The procedure used followed the protocol previously described (Nzengue et al. 2009). Protein samples were prepared by addition of an equal volume of 2X SDS loading buffer and boiled for 5 minutes. After cooling to room temperature SDS-PAGE was performed with 14% polyacrylamide gels at 80 V.

Following electrophoresis, the gel was placed in Millipore-grade water for 5 minutes, with gentle shaking.

For electroblotting, filter papers (Fisher Scientific), nitrocellulose membrane (AmershamTM ProtranTM 0.2 µm pore size, Cat# 1060044) and the gel were incubated for at least 20 minutes in Towbin buffer (25 mM Tris-HCl, 192 mM glycine, 20% methanol) supplemented with 2 mM $CaCl_2$. $CaCl_2$ improves retention of small proteins on nitrocellulose, as Ca^{2+} complexes with residual SDS that otherwise can cause small proteins to transfer through the membrane owing to its strong negative charge. Electroblotting was performed at 40 mA and 14 V for 45 minutes using a semi-dry transfer apparatus (HEP-1 series, OWL). The membrane was incubated in 2.5% glutaraldehyde for 1 hour with gentle shaking to crosslink the proteins to the membrane. The membrane was washed twice in PBS and once in 50 mM ethanolamine-HCl/PBS, which quenches residual glutaraldehyde. Proteins were detected by Ponceau staining (Bioshop) of the membrane to assess transfer efficiency and to verify equivalent protein loading in each lane. After recording the image of the Ponceau-stained blot, the stain was removed by repeated rinses with deionized water. The membrane was placed in blocking buffer consisting of 5% skim milk powder in Tris-buffer saline +Tween 20 (TBST: 20 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.01% Tween 20) with gentle shaking for one hour. The membrane was incubated overnight with gentle agitation in primary antibody diluted in blocking buffer (refer to Table 2 for antibody dilutions). To remove excess primary antibody, the membrane was rinsed three times in TBST. The membrane was incubated with the secondary antibody for one hour, and excess secondary antibody was removed by rinsing as before. Chemiluminescence reagents (Clarity Western ECL substrate, Bio-Rad Canada, Mississauga ON) were added to the membrane, which was

promptly placed in a Bio-Rad ChemiDoc^{TM} MP imaging system for capture of the image.

Table 2: Antibody conditions of use for Western blotting. Secondary antibodies are shaded grey.

Antibody	Dilution	Incubation Time	Incubation Temperature	Provider
Rabbit gCYTB5-I	1:5000	Overnight	4°C	GeneScript (custom order)
Rabbit gCYTB5-II	1:5000	Overnight	4°C	GeneScript (custom order)
Rabbit gCYTB5-III	1:5000	Overnight	4°C	GeneScript (custom order)
Mouse Anti-HA	1:1000	Overnight	4°C	Sigma (Cat # H9658)
Donkey anti-mouse HRP for HA antibody	1:100 000	1 hour	Room temperature	Jackson ImmunoResearch (Cat # 715-035-150)
Goat anti-rabbit HRP gCYTB5-I, II, and III	1:5000	1 hour	Room temperature	Jackson ImmunoResearch (Cat # 111-035-003)

2.7 Anti-gCYTB5s antibodies specificity

Antibodies raised in rabbits against peptides which are unique to each of the three gCYTB5 isotypes (I, II, and III) were custom ordered from GeneScript. The specificity of these antibodies were tested by other students in Dr. Yee's lab as follow:

First, Western blot analysis were performed with the pre-immune serum from rabbits used for the generation of the gCYTB5-I, II and III antibodies (G. William Batoff). The blots contained recombinant His_6-tagged gCYTB5-I, II, and III (10 ng/lane) as well as different amounts of proteins from a Giardia cell lysate. No bands were observed on the Western blots after hybridization with the pre-immune sera.

Second, Western blot hybridization with each of the gCYTb5 antibodies gave a signal only in the lane of the blot containing the corresponding recombinant protein (B. Sajer and G. William Batoff). For example, hybridization with the gCYTB5-I antibody resulted in a signal in the lane containing the recombinant protein of cytochrome isotype I and not the lanes containing recombinant versions of isotypes II and III.

Third, I used the gCYTB5-I and III pre-immune sera in immunofluorescent microscopy (IFM) assays. The cells on the slides were incubated overnight with the pre-immune sera diluted 1:200 followed by a 1-hour incubation with the secondary antibody diluted 1:200. Note that the gCYTB5s antibodies were used at dilutions between 1/200 to 1:2000, and the incubation time was for 2 hours. Only background fluorescence was observed (see results in Appendix D). Collectively, these results indicate that our antibodies do not give high background reactivity in Western blots and IFM, and are specific for each of the isotypes.

2.8 Interactome studies

2.8.1 Pull-down using Affigel-10 resin

Forty µL Affigel-10 resin (Bio-Rad Canada, Mississauga ON, Cat # 153-6099) was placed in a microcentrifuge tube and centrifuged at 1000 *g*, for 3 minutes. The isopropanol storage solution was removed leaving 20 µL of packed resin. One mg of purified recombinant His_6-gCYTB5-I diluted in 1 mL of 100 mM MOPS buffer pH 7.5 was added to the packed resin. The beads and gCYTB5-I were incubated for 4 hours at 4°C with gentle mixing to achieve maximum binding of the recombinant protein. Unbound protein was removed by centrifugation as before and was stored at -80°C. To block unreacted N-hydroxysuccinimide groups, the resin was resuspended in 1 mL of 1

M ethanolamine-HCl pH 8.0 and incubated with gentle mixing for one hour at 4°C. After centrifugation the solution was discarded and the resin was washed 3 times with Millipore-grade water. As a negative control, an equivalent amount of Affi-Gel 10 resin was processed in the same way but without the addition of gCYTB5-I bait.

The resin was resuspended with 500 µL of Giardia extract containing 1-2 mg of protein and was incubated overnight at 4°C on a nutator with gentle agitation. The tube was centrifuged and the unbound protein supernatant was removed and stored at -80°C. The resin was washed free of unbound proteins by resuspension in 0.5 mL 50 mM potassium phosphate buffer pH 7.5 followed by centrifugation; this step was repeated twice. To elute bound proteins the resin were sequentially eluted with 0.5 mL 1 M NaCl followed by 0.5 mL 8 M urea, and finally with SDS loading buffer.

Protein from the collected fractions (except the SDS loading buffer elution) were concentrated using StrataClean resin (Agilent Technologies Canada, Mississauga ON). StrataClean resin (5 µL) was added to each microfuge tube containing sample, and then were incubated with gentle shaking for 20 minutes at 4°C. The tubes were centrifuged at 4,500 *g* for 1 minute and the supernatant was removed. The resin containing the bound protein was resuspended in 20 µL SDS-PAGE loading buffer. SDS-PAGE was performed on a 10% polyacrylamide gel electrophoresed at 80 V. To detect proteins the gel was treated either with silver stain (Pierce Silver Stain for Mass spectrometry Cat# 24600, Thermo-Fisher Canada, Whitby ON) or with blue silver stain (Candiano et al. 2004) (see Appendix E).

2.8.2 PULL-DOWN ASSAY USING IMAC

Pull-down experiments were also performed using His_6-tagged gCYTB5s bound to Dynabeads (Life Technologies/Thermo-Fisher Canada, Cat# 10103D). This resin consists of chelated Co^{2+}, which has a high affinity for the His_6-tag, bound to a magnetic support. The magnetic properties of the beads permit rapid buffer exchanges when a tube containing suspended beads is placed next to a strong magnet; this replaces the centrifugation steps.

Dynabeads were resuspended by vortexing for 30 seconds and 25 µL (equivalent to 1 mg) of beads were transferred to a microfuge tube. The tube was placed next to the magnet (DynaMag^{TM}-2, Life Technologies, Cat # 12321D) to collect the beads on the side of the tube, and the supernatant was removed with a micropipette. The tube was removed from the magnet and the beads were resuspended in Binding/Wash buffer (50 mM sodium phosphate pH 8.0, 300 mM NaCl, 0.01% Tween 20) and incubated for 5 minutes on a nutator.

A saturating amount of at least 50 µg of recombinant His_6-tagged protein in 0.5 mL Binding/Wash buffer was added to 1 mg of Dynabeads. After a five minute incubation on the nutator the supernatant containing the unbound His_6-gCYTB5 was removed and stored for further analysis. To the beads was added 500 µg of Giardia protein extract in 1 mL of Pull-down buffer (3.25 mM sodium phosphate pH7.4, 70 mM NaCl, 0.01% Tween 20) supplemented with 10 µg/mL leupeptin and 1X Protease inhibitor cocktail (Bioshop Canada). After incubation for 30 minutes on the nutator, the supernatant was collected and the beads were washed 3 to 4 times with 500 µL of 2X Pull-down buffer (6.5 mM sodium phosphate pH7.4, 140 mM NaCl, 0.02% Tween 20). To elute bound His_6-gCYTB5 and its interacting partners the resin was treated sequentially with the following buffer for 5 minutes each:

1) 0.5 mL of 2X Binding/wash buffer (100mM sodium phosphate pH 8.0, 600 mM NaCl and 0.02% Tween 20), which increases the ionic strength.
2) 0.5 mL of His elution buffer (300 mM imidazole, 50 mM Sodium-phosphate pH 8.0, 300 mM NaCl and 0.01% Tween 20), which releases His_6-gCYTB5 from the resin
3) 20 µL of 4X SDS loading buffer, which unfolds proteins and releases all remaining proteins still bound to the beads.

For the negative control experiment, the same protocol was followed except that His_6-tagged gCYTB5 was not added to the beads. Washes and eluted samples were concentrated with StrataClean resin and analysed by SDS-PAGE as described above.

2.8.3 Pull-down assay using crosslinking reagents

Pull-down experiments using Dynabeads in the presence of crosslinking reagents (paraformaldehyde or glutaraldehyde) were performed with minor modifications to the protocol described above. Two sets of experiments were done with the crosslinker added at different stages.

For the first approach, recombinant His_6-gCYTB5 was bound to the beads and the Giardia protein extract was added as before. Paraformaldehyde or glutaraldehyde at a final concentration between 0.1-5% was then added followed by incubations for 5 to 20 minutes. The crosslinking reaction was terminated by the addition of 125 mM glycine/PBS (for paraformaldehyde) or 1M Tris-HCl (for glutaraldehyde). The washes were performed using a more stringent pull-down buffer of higher ionic strength (300 mM NaCl instead of 140 mM). The rest of the procedure was then performed as before.

For the second approach the His_6-tagged gCYTB5 was added to the Giardia protein extract and the sample was incubated for half an hour at room temperature. The crosslinking reagents were then added, the samples were incubated for 5 to 20 minutes and the reactions were quenched as before. The samples were then incubated with the beads for 10 to 30 minutes and the rest of the procedure completed as described above.

2.8.4 Co-immunoprecipitation experiments

We initiated a collaboration with Dr. Jan Tachezy of Charles University, Prague, Czech Republic, who provided us with Giardia cell lines stably transfected with plasmids for the expression of 2X-HA-tagged versions of each of the Giardia gCYTB5 isotypes (I, II and III). The vectors are maintained within Giardia under puromycin drug selection. As a negative-control, we produced a Giardia cell line transfected with the puromycin resistant plasmid without any gCYTB5-coding sequence. These four Giardia cell lines were used in co-immunoprecipitation experiments. Total protein extracts for all the cell lines were obtained as described previously, except that DTT, which can destabilize antibodies by the loss of disulfide bonds, was not included in the CelLytic lysis buffer. For the immunoprecipitation step, anti-hemagglutin immobilized on magnetic beads (Pierce Anti-HA Magnetic Beads, Thermo-Fisher Canada, Cat# 88836) were used. The procedure is similar to that used for the Dynabeads described above. The beads were gently vortexed and 50 µL (0.5 mg) were transferred to a microcentrifuge tube. The supernatant was removed, and 1 mg of soluble protein extract in 500 µL of CelLytic buffer (without DTT) was added. The sample was incubated for 30 minutes at 4°C with gentle agitation. The supernatant containing the unbound protein was collected and stored at -80°C. The beads with the immunoprecipitated HA-tagged bait and its interacting partners were washed three to four times with 500 µL TBST,

with the supernatants collected after each wash. The remaining bound proteins were eluted using 25 µL of SDS loading buffer. The washes as well as the unbound sample were concentrated using StrataClean resin as described previously.

To assess the efficiency of the 2xHA tagged bait immunoprecipitation, 20% of the final wash, the unbound sample and the eluted sample used for Western blot analysis as described in Section 2.5.

The balance of these samples, and 5 µL of the input protein extract and unbound fractions were subject to SDS-PAGE on a 14% polyacrylamide gel and 80 V. The gel was stained using Blue Silver staining, (Appendix E) which is compatible for subsequent analysis by mass spectrometry.

2.8.5 MASS SPECTROMETRIC ANALYSIS

Proteins recovered from co-immunoprecipitation experiment with each gCYTB5 isotype (I, II, and III) and the negative control (empty vector) were subjected to SDS-PAGE. Gel slices corresponding to whole lanes or portions thereof were excised and sent to the Alberta Proteomics and Mass Spectrometry Facility (University of Alberta, Edmonton AB) for analysis. The following procedure was used to prepare the samples (Jack Moore personal communication).

In-gel trypsin digestion was performed on the samples. Briefly, excised gel bands were destained twice in 100 mM ammonium bicarbonate/acetonitrile (50:50). The samples were then reduced using 10 mM beta-mercaptoethanol in 100 mM sodium bicarbonate and were alkylated using 55 mM iodoacetamide in 100 mM sodium bicarbonate. After dehydration, sufficient trypsin solution (6 ng/µL) was added to just cover the gel pieces and the digestion was allowed to proceed overnight (~16 hrs.) at room temperature. Tryptic peptides were first extracted from the gel using 97%

water/2% acetonitrile/1% formic acid followed by a second extraction using 50% of the first extraction buffer and 50% acetonitrile.

Fractions containing tryptic peptides were resolved and ionized by using nanoflow HPLC (Easy-nLC II, Thermo Scientific) coupled to an LTQ XL-Orbitrap hybrid mass spectrometer (Thermo Scientific). Nanoflow chromatography and electrospray ionization were accomplished by using a PicoFrit fused silica capillary column (ProteoPepII, C18) with 100 μm inner diameter (300Å, 5μm, New Objective). Peptide mixtures were injected onto the column at a flow rate of 3000 nL/min and resolved at 500 nL/min using linear gradients from 0 to 45% v/v aqueous ACN in 0.2% v/v formic acid. The mass spectrometer was operated in data-dependent acquisition mode, recording high-accuracy and high-resolution survey Orbitrap spectra using external mass calibration, with a resolution of 30,000 and m/z range of 400–2000. The fourteen most intense multiply charged ions were sequentially fragmented by using collision induced dissociation, and spectra of their fragments were recorded in the linear ion trap; after two fragmentations all precursors selected for dissociation were dynamically excluded for 60 s. Data was processed using Proteome Discoverer 1.4 (Thermo Scientific) and a non-reviewed Uniprot (uniprot.org) Giardia database was searched using SEQUEST (Thermo Scientific). Search parameters included a precursor mass tolerance of 10 ppm and a fragment mass tolerance of 0.8 Da. Peptides were searched with carbamidomethyl cysteine as a static modification and oxidized methionine and deamidated glutamine and asparagine as dynamic modifications.

2.9 PUP9861 vector construction

A prospective gCYTB5 partner identified in this work was the putative uncharacterized protein GL50803_9861 (PUP9861). To further explore this

relationship, we decided to generate a transgenic Giardia cell line expressing HA-tagged PUP9861. We used the Giardia expression vector pTG3039 provided by Dr. Jan Tachezy. The PUP9861 coding sequence, flanked by *Nde*I and *Apa*I restriction sites at the 5' and 3' ends respectively, was prepared by chemical synthesis (BioBasic, Markham ON) and was provided in the cloning vector pBluescript. The coding sequence was excised from pBluescript by digestion with *Nde*I and *Apa*I restriction enzymes (New England Biolabs Canada, Whitby ON) and was purified by agarose gel electrophoresis. The Giardia expression vector pTG3039 was likewise cut with these restriction enzymes and the digest was gel purified to obtain the vector backbone. The PUP9861 insert and vector pTG3039 were ligated with T7 Ligase (New England Biolabs) and the ligation reaction was used to transform chemically competent *E. coli* strain HB101, which were plated on an ampicillin-selective plate to screen for transformants. Two colonies from the plate were selected for plasmid mini-preparations (Qiagen Canada, Toronto ON). Portions of the purified plasmids were screened for the presence of the insert by digestions with *Nde*I and *Apa*I, which excised the insert, or with *Eco*RV, which cuts once within the vector backbone and once within the PUP9861 coding sequence. Restriction digests of both clones gave the expected pattern of DNA fragments after agarose gel electrophoresis, and one clone was verified by DNA sequencing (MOBIX Lab, McMaster University, Hamilton ON). The sequencing was performed using the forward primer PA/HA/F (5'-cgccaactaaacgctctaca3') located 91 bp upstream of the translation start site, and the M13(-20) primer located 145 bp downstream of the stop codon. See Appendix F for the result of DNA sequencing of the PUP9861 insert in the Giardia transfection plasmid.

2.10 Giardia transfection

Electroporation was used to generate two transgenic Giardia cell lines, one expressing HA-tagged PUP9861, and a negative control transfected with the empty plasmid. Both plasmids contain the puromycin resistance gene driven by the Giardia glutamate dehydrogenase (GDH) promoter. Trophozoites were prepared by growing Giardia in a 16 mL glass tube until reaching 75% coverage. The culture tube was centrifuged at 1200 *g* for 15 minutes at 4°C. The supernatant was discarded. The cell pellet was resuspended in 290 µL of Giardia media (final concentration 3.5 x 10^7 cells/mL) without cysteine-ascorbic acid and transferred to a pre-chilled 0.4 cm-gapped electroporation cuvette placed on ice. Plasmid DNA (30-40 µg in 10-15 µL) was added and the cuvette was placed in a BTX electroporator. The cells were electroporated with the following instrument conditions: low voltage setting; voltage =350 V, resistance =725 Ohms, capacitance =1025 µF. The electroporated cells were used to inoculate 15.5 mL of Giardia complete media (with cysteine-ascorbic acid and 1X antibiotic/antimycotic) and allowed to recover for 6 hours at 37°C after which puromycin (54 µg/mL final concentration) was added to select for trophozoites bearing the plasmid.

Since only a few initial cells were transfected with the plasmid, large clumps of dead cells were observed 1-2 days after the electroporation. Usually 1-5% of healthy cells could be seen attached to a single side of the tubes 48 hours after electroporation. After 48 hours, the attached cells are selected by removing the media containing the dead cells. Fresh media complemented with puromycin is added and the cells are allowed to grow for a further 3-4 days. Media exchange was then performed at least every 3 days to add fresh nutrients and drugs until the Giardia culture reach 75% coverage (1-2 weeks after electroporation). The cells are then maintained in culture

under drug selection. Finally, the cells reach their normal doubling time (6 hours) 3-4 weeks after electroporation and therefore could be used for further experiments. Frozen stocks of these cells could also be prepared at this time.

RESULTS

3.1 Bioinformatic searches for cytochrome b5 interacting partners in Giardia

I first used bioinformatics tools in an attempt to identify potential partners of the Giardia cytochromes b_5. The major partners of canonical cytochromes b_5 such as CYTB5A and CYTB5B are cytochrome P450 monooxygenases and fatty acid desaturases. However, the Giardia genome database (GiardiaDB) does not contain any genes with these annotations. Furthermore, BLAST searches of the GiardiaDB with full-length sequences of monooxygenases and desaturases from other species used as the queries failed to identify matches. Consequently the GiardiaDB was searched using minimal key sequence motifs characteristic of these partners rather than their full sequences. The motif used for cytochrome P450 was the highly conserved sequence FGXGX$_3$**C**XG, which contains the invariant cysteine residue (bold) required for heme binding (Nebert and Gonzalez 1987). For the fatty acid desaturases, the motifs used were the highly conserved eight histidines required for the desaturation reaction (HX$_{(3-4)}$H, HX$_{(2-3)}$HH, HX$_{(2-3)}$HH or QX$_{(2-3)}$HH) (Uttaro 2006). All the Giardia proteins retrieved using these motif searches were subject to the BLASTp and Swiss-Model searches to determine if some of them align, at least partially, to a cytochrome P450 or a fatty acid desaturase. No candidate Giardia cytochrome P450s or fatty acid desaturases were identified by these searches. It seems likely that these enzymes are absent from Giardia; if the gCYTB5s have partners these are not among those commonly associated with cytochromes b_5.

3.2 Bioinformatic searches for Type II cytochrome b5 proteins in other species

Since it was not possible to identify canonical partners of cytochrome b5 in Giardia by using BLAST searches, I decided to look for Type II cytochrome b5 in other species. The rationale for this is that it might be possible to find the interacting partners for these proteins in other species. Identification of common partners among the different species would allow us look for their orthologues in Giardia as potential interactors of gCYTB5s. Tachezy and colleagues noted that type II cytochromes, which include the gCYTB5s, occur in many other species and appear to be the only type of cytochrome b_5 found in anaerobic protozoa (Pyrih et al. 2014). As of the writing of this thesis there have been no reports in the literature assigning a function to any type II cytochrome b_5. To provide a larger context for studies on the gCYTB5 interactome I examined the distribution of type II cytochromes b_5 in other organisms and used the STRING online tool to predict their potential functions.

To seek the function or the partners of the cytochrome b_5 Type II protein, 86 organisms belonging to the animal, fungal and plant kingdoms were screened *in silico* by BLASTp searches for the presence of cytochrome b_5 in their proteome. The protists were selected based on the availability of a fully-sequenced genome for each organism − this includes all of protists available in the EupathDB. Finally, species that have been previously screened for cytochrome b_5 Type II proteins by Pyrih et al. (2014) were also added to the pool of organisms analysed. Table 3 shows all the organisms used in this search.

To identify which types of cytochrome b_5 protein were present in a given proteome, a search was made using a species-specific search of the UniProtKB database using the search term “species name + cytochrome b_5”. The results of this search gave all the proteins containing a cytochrome b_5 domain including multidomain proteins.

Characterized multidomain proteins such as fatty acid desaturases or cytochrome b_5 reductases that contained cytochrome b_5 domains were not kept for further analysis. The remainder, which included unnamed proteins (annotated as “uncharacterized protein”, “predicted protein” or “protein of unknown function” *etc.*), and those annotated as cytochrome b_5 were used in further analysis.

Table 3: Species used to screen for cytochromes b_5. For each species the red text indicates its kingdom and the green box its phylum. Anaerobic species are boxed in blue. Unranked are species with an unclear phylum description. For example, *Arabidopsis thaliana* belong to the Plantae kingdom but its phylum is unclear.

PROTOZOA			FUNGI	ANIMALIA
Apicomplexa	**Amoeboza**	**Heterokontophyta**	**Microsporidia**	**Chordata**
Babesia bovis	*Acanthamoeba castellanii*	*Albugo laibachii*	*Trachipleistophora hominis*	*Oikopleura dioica*
Cryptosporidium hominis	*Dictyostelium discoideum*	*Aureococcus anophagefferens*	*Encephalitozoon cuniculi*	**Nematoda**
Cryptosporidium parvum	*Entamoeba histolytica*	*Fragilariopsis cylindrus*	*Anncaliia algerae*	*Ascaris suum*
Cryptosporidium muris	*Polysphondylium pallidum*	*Hyaloperonospora arabidopsidis*	*Edhazardia aedis*	*Caenorhabditis elegans*
Eimeria tenella	*Pelomyxa*	*Nannochloropis gaditana*	*Enterocytozoon bieneusi*	**Mollusca**
Neospora caninum	*Euglenozoa/ Kinetoplastida*	*Phaeodactylum tricornutum*	*Nematocida*	*Mytilus edulis*
Plasmodium berghei	*Leishmania braziliensis*	*Phytophthora infestans*	*Nosema ceranae*	**Annelida**
Plasmodium chabaudi	*Leishmania infantum*	*Phytophthora ramorum*	*Spraguea lophii*	*Arenicola marina*
Plasmodium falciparum	*Leishmania major*	*Phytophthora sojae*	*Vavraia culicis floridensis*	**Sipuncula**
Plasmodium knowlesi	*Trypanosoma brucei*	*Pseudo-nitzschia multiseries*	*Vittaforma corneae*	*Sipunculus nudus*
Plasmodium vivax	*Trypanosoma cruzi*	*Pythium ultimum*	**Ascomycota**	**Unranked**
Plasmodium yoelii yoelii	*Euglena*	*Thalassiosira pseudonana*	*Neurospora crassa*	*Fasciola hepatica*
Theileria annulata	*Crithidia fasciculata*	**Cryptophyta**	*Saccharomyces cerevisiae*	**PLANTAE**
Theileria parva	**Metamonada**	*Chroomonas mesostigmatica*	**Unranked**	**Unranked**
Toxoplasma gondii	*Trichomonas vaginalis*	*Cryptomonas paramecium*	*Schizosaccharomyces pombe*	*Arabidopsis thaliana*
Gregarina niphandrodes	*Trichomonas tenax*	*Guillardia theta*	*Ustilago hordei*	*Chlamydomonas reinhardtii*
Theileria equi strain WA	*Tritrichomonas foetus*	*Hemiselmis andersenii*	*Piromyces sp. strain E2*	
Alveolate and Stramenophile	*Hexamita*	**Haptophyta**	*Fusarium oxysporum*	
Nyctotherus ovalis	*Trepomonas agilis*	*Emiliania huxleyi*		
Blastocystis	*Giardia intestinalis*	**Percolozoa**		
Unranked	*Spironucleus salmonicida*	*Naegleria gruberi*		
Monosiga brevicollis	**Unranked**	**Ciliate**		
	Dasytricha ruminantium	*Tetrahymena thermophila*		
		Dasytricha ruminantium		

The remaining proteins identified by the BLASTp searches were then placed into four different categories according to the criteria presented in Figure 6. This classification was applied to the 354 sequences selected and the detailed results are shown in Appendix G.

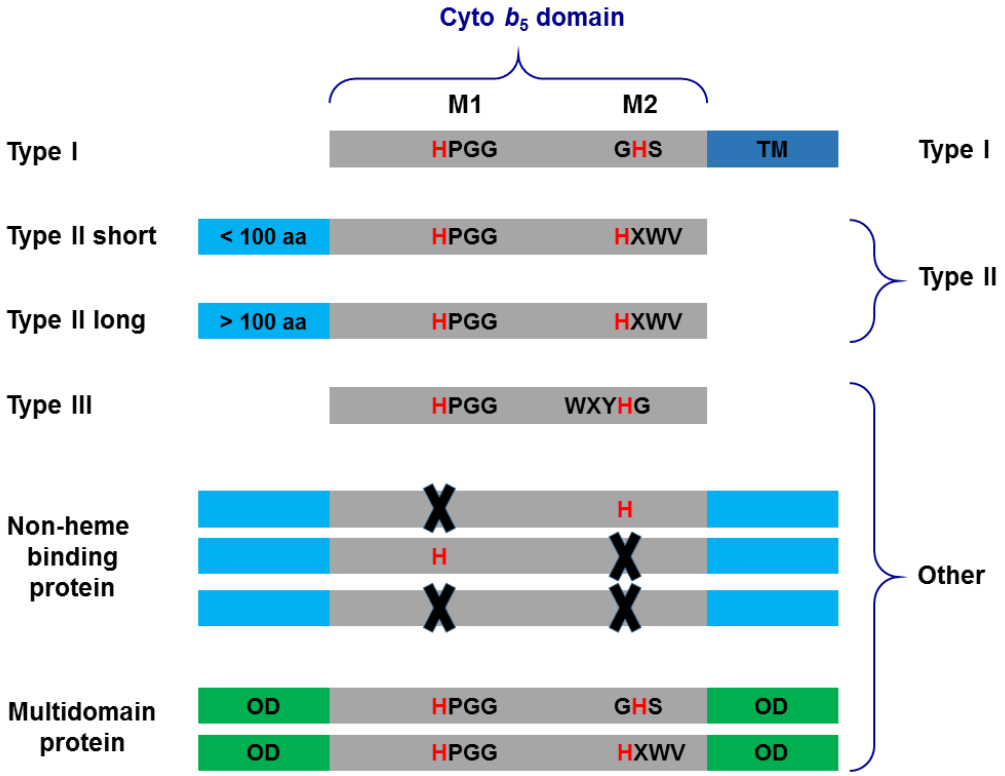


Figure 6: Schematic representation of the cytochrome b_5 proteins retrieved from the Uniprot database. M1 and M2 refer to Motif 1 or 2 around the two histidine ligands, TM corresponds to a transmembrane domain and OD is another domain such as a fatty acid desaturase domain. The first category contains all Type I cytochromes b_5; these have the first histidine ligand within an **H**PGG motif with the second histidine in the G**H**S motif, and a C-terminal TM. The second category contains the cytochrome b_5 Type II proteins with an N-terminal extension of less than 100 amino acid residues (similar to gCYTB5-I, II and III), the second invariant histidine motif is present in a more variable motif. The third category is similar to the second category but contains the type II cytochromes b_5 with an N-terminal extension between 100 to 500 amino acid residues. Finally, the fourth category contains all the other proteins including uncharacterized multidomain proteins that appear to contain a cytochrome b_5 domain, Type III cytochrome b_5 (only found in *Neurospora crassa*), and proteins lacking one (or both) of the axial histidine ligands (such as steroid binding proteins).

Of the 86 species screened, 34, including both aerobic and anaerobic organisms, lacked both Type I and II CYTB5, and were consequently excluded from further analysis (Appendix H). Seven species (all aerobic protozoa) had only type I CYTB5. Surprisingly *Plasmodium chabaudi*, the species that infects rodents, was found in this list whereas all the other *Plasmodium* analysed encoded at least one type II CYTB5 in their genome. The species that did not possess any Type II CYTB5 were also excluded from further analysis. Next, 14 species were found to have only Type II CYTB5s: ten species possess proteins with short N-termini, two species with long N-termini, and two species possess both (Table 4). Interestingly three of these species are aerobic organisms (*Babesia bovis*, *Gregarina niphandrodes* and *Thalassiosira pseudonana*). Finally, 31 species possess both Type I and Type II CYTB5s.

It was noted above that anaerobic protozoa only encode Type II CYTB5 in their genomes (Pyrih et al. 2014). I decided to test if that statement could be generalized to other anaerobic organisms and to aerobic species that could survive under anaerobic condition. In my analysis I found 26 species that are completely anaerobic (i.e. *Giardia intestinalis*) and I extended the investigations to 2 aerobe species that are also metabolically active under anaerobic conditions (i.e. *Chlamydomonas reinhardtii*) and 5 species that are anaerobic for only a part of their life cycle (*Ascaris suum*). Table 4 shows the distribution of the cytochrome b_5 proteins identified in these 33 organisms.

It is interesting to note that only two anaerobically active organisms encode Type I CYTB5. Moreover, these two organisms also possess anaerobic mitochondria that use a different terminal electron acceptor than oxygen under anaerobic conditions (Muller et al. 2012). One of these organisms is the parasitic worm *Ascaris suum,* which spends part of its life cycle under anaerobic conditions when it resides in the intestine of its host (human or pig). The other organism, *Chlamydomonas reinhardtii* (green

algae) produces hydrogen in anaerobic environments (Kita et al. 2002) (Terashima et al. 2010). However, based on this search it appears that not only do none of the anaerobic protozoa possess cytochrome b_5 Type I proteins, but that these proteins are absent from all amitochondriate species regardless of kingdom.

Table 4: Distribution of Type II cytochromes b_5 in anaerobic organisms. The red text shows the species Kingdom, the green boxes show the phylum and the blue boxes correspond to anaerobic organisms. For each of the species the mitochondria or mitochondria related organelle (MRO) status is labelled mitosome or hydrogenosome if the species possesses a mitochondria related organelle, anaerobic if they possess anaerobically functional mitochondria and present if they possess mitochondria. The cytochromes b_5 listed in the "other" column are those that lack one of the conserved histidine ligands found in heme binding cytochrome b_5 proteins. * Note that the Blastocystis mitochondria related organelle has retained proprieties of aerobic/anaerobic mitochondria as well as hydrogenosome (Stechmann et al. 2008).

PROTOZOA	Type 1	Short Type 2	Long Type 2	Other	Mitochondria or MRO status
Apicomplexa					
Cryptosporidium hominis	0	0	0	0	Mitosome
Cryptosporidium parvum	0	0	0	1	Mitosome
Cryptosporidium muris	0	0	1	0	Mitosome
Amoeboza					
Entamoeba histolytica	0	0	0	0	Mitosome
Pelomyxa	0	0	0	0	No Mitochondria/MRO
Metamonada					
Giardia intestinalis	0	3	1	1	Mitosome
Spironucleus salmonicida	0	5	0	2	Hydrogenosome
Trichomonas vaginalis	0	5	0	1	Hydrogenosome
Trichomonas tenax	0	0	0	0	Hydrogenosome
Tritrichomonas foetus	0	0	0	0	Hydrogenosome
Hexamita inflata	0	0	0	0	Mitosome
Trepomonas agilis	0	0	0	0	Mitosome
Alveolate and Stramenophile					
Nyctotherus ovalis	0	0	0	0	Hydrogenosome
Blastocystis	0	0	0	0	Hydrogenosome*
Unranked					
Isotricha	0	0	0	0	Hydrogenosome
FUNGI	**Type 1**	**Short Type 2**	**Long Type 2**	**Other**	**Mitochondria status**
Microsporidia					
Trachipleistophora hominis	0	1	0	0	Mitosome
Encephalitozoon cuniculi	0	1	0	0	Mitosome
Anncaliia algerae	0	1	0	0	Mitosome
Edhazardia aedis	0	1	0	0	Mitosome
Enterocytozoon bieneusi	0	0	0	0	Mitosome
Nematocida	0	1	0	0	Mitosome
Nosema ceranae	0	0	0	0	Mitosome
Spraguea lophii	0	0	0	1	Mitosome
Vavraia culicis floridensis	0	1	0	0	Mitosome
Vittaforma corneae	0	0	0	0	Mitosome
Unranked					
Piromyces sp. strain E2	0	0	0	0	Hydrogenosome
Fusarium oxysporum	0	1	1	9	Mitochondria present
ANIMALIA	**Type 1**	**Short Type 2**	**Long Type 2**	**Other**	**Mitochondria status**
Nematoda					
Ascaris suum	3	1	0	0	anaerobic mitochondria
Mollusca					
Mytilus edulis	0	0	0	0	anaerobic mitochondria
Annelida					
Arenicola marina	0	0	0	0	anaerobic mitochondria
Sipuncula					
Sipunculus nudus	0	0	0	0	anaerobic mitochondria
Unranked					
Fasciola hepatica	0	0	0	0	anaerobic mitochondria
PLANTAE	**Type 1**	**Short Type 2**	**Long Type 2**	**Other**	**Mitochondria status**
Unranked					
Chlamydomonas reinhardtii	3	2	0	0	Mitochondria present

Combining aerobic and anaerobic organisms, a total of 78 Type II CYTB5s were identified in 45 species (Table 5). Of these proteins, 53 possess a short amino-terminus and 25 have a long N-terminal sequence. The metamonads *Giardia intestinalis*, *Spironucleus salmonicida* and *Trichomonas vaginalis* are the species with the largest number of Type II CYTB5s (4, 5 and 5 respectively). Interestingly, none of the other metamonads (i.e. *Trichomonas tenax* or *Hexamita inflata*) examined possessed any Type II CYTB5s, although this might be partially explained by the incomplete sequencing coverage of their genomes. The aerobic protozoan *Dictyostelium discoideum* also possess five Type II cytochrome b_5.

A literature search on organisms that possess Type II CYTB5 provided no information on their possible functions. Nevertheless, some information has been found from proteomic studies on *Plasmodium falciparum*, which show that its Type II CYTB5 with a short N-terminal sequence is located in both the cytoplasm and nucleus. (Treeck et al. 2011; Oehring et al. 2012). It has also been reported that one of fatty acid desaturases (FAD) found in the protozoa *Thalassiosira pseudonana* lacks the cytochrome b_5 domain commonly found in these enzymes. It is possible that the Type II CYTB5 identified in this organism is the electron donor to this FAD (Tonon et al. 2005).

Table 5: Organisms encoding Type II CYTB5. The number of Type I CYTB5s in these organisms are also shown. Text annotation as in Table 4.

PROTOZOA	Type I	Short Type II	Long Type II
Apicomplexa			
Babesia bovis	0	1	0
Cryptosporidium muris	0	0	1
Neospora caninum	1	0	2
Plasmodium berghei	2	1	0
Plasmodium falciparum	1	1	1
Plasmodium knowlesi	1	1	1
Plasmodium vivax	1	1	1
Plasmodium yoelii yoelii	1	1	1
Toxoplasma gondii	2	0	2
Gregarina niphandrodes	0	0	1
Ciliate			
Tetrahymena thermophila	3	1	2
Amoeboza			
Dictyostelium discoideum	2	3	2
Polysphondylium pallidum	2	2	0
Heterokontophyta			
Albugo laibachii	1	0	1
Hyaloperonospora arabidopsidis	1	0	1
Phaeodactylum tricornutum	3	1	0
Pythium ultimum	2	0	1
Thalassiosira pseudonana	0	1	0
Haptophyta			
Emiliania huxleyi	3	0	1
Euglenozoa/Kinetoplastida			
Leishmania braziliensis	5	2	1
Leishmania infantum	2	2	0
Leishmania major	6	1	0
Trypanosoma brucei	3	1	0
Trypanosoma cruzi	6	2	2
Crithidia fasciculata	3	1	0
Percolozoa			
Naegleria gruberi	2	1	0
Metamonada			
Giardia intestinalis	0	3	1
Spironucleus salmonicida	0	5	0
Trichomonas vaginalis	0	5	0
FUNGI	**Type 1**	**Short Type 2**	**Long Type 2**
Microsporidia			
Trachipleistophora Hominis	0	1	0
Encephalitozoon Cuniculi	0	1	0
Anncaliia algerae	0	1	0
Edhazardia aedis	0	1	0
Nematocida	0	1	0
Vavraia culicis floridensis	0	1	0
Ascomycota			
Neurospora crassa	1	0	1
Saccharomyces cerevisiae	1	0	1
Unranked			
Schizosaccharomyces pombe	2	2	0
Ustillago hordei	1	1	0
Fusarium oxysporum	0	1	1
ANIMALIA	**Type 1**	**Short Type 2**	**Long Type 2**
Chordata			
Oikopleura dioica	1	1	0
Nematoda			
Ascaris suum	3	1	0
Caenorhabditis elegans	1	1	0
PLANTAE	**Type 1**	**Short Type 2**	**Long Type 2**
Unranked			
Arabidopsis thaliana	5	1	0
Chlamydomonas reinhardtii	3	2	0

3.3 STRING analysis for potential partners of CYTB5s

The aim of STRING analysis of the Type II CYTB5s was to determine if any partners or functions identified using this tool are shared among the different proteins used for the analysis. Among the 45 species (aerobic and anaerobic) possessing a Type II CYTB5, 20 species gave some results, 6 gave no results and 19 were not in the database. Unfortunately, no consensus was found among these protein partners/functions, which falls into two categories. The First category are prospective partners that are actually those of Type I CYTB5s. The second category of prospective partners are ribosomal proteins, which are frequently found as experimental contaminants due to their high abundance within the cell. As there were no common features across this wide search, the search was narrowed to focus on the anaerobic protozoa, which is more pertinent for Giardia.

STRING analysis gave results for three anaerobic protozoa species: *Giardia intestinalis*, *Trichomonas vaginalis* and *Cryptosporidium muris*. For Giardia, three of the isotypes (gCYTB5-I, II and III) yielded the same two predicted interacting partners based on Text Mining identification, a view that usually returns the least robust results. These two proteins are a glutamate synthase and an IMPACT-like protein (imprinted and ancient gene-like protein) with no known function in Giardia. While these scored moderately high (0.820 and 0.799, respectively) there were no publications that mentioned these proteins together with the gCYTB5s, so it is unclear why these proteins were identified as partners. Furthermore, no interactions of these proteins in the same pathway have been found in literature searches. STRING analysis of the gCYTB5-IV did not return any results.

One Type II CYTB5 with a long N-terminal sequence has been found in *Cryptosporidium muris*. STRING analysis of the protein yielded four proteins (Figure

7). The highest score (0.725) was obtained with a hypothetical protein identified by a BLASTp search as an IMPACT-like protein domain as found for Giardia. The carrier protein (score = 0.711) is a mitochondrial carrier protein/ADP/ATP transporter, based on BLAST domain identification. The TIM44-like domain-containing protein is a mitochondrial transporter (score = 0.423) and the ribosomal protein only gave medium confidence score (score = 0.508). However, the function that would involve these proteins with a cytochrome b_5 is unclear and no functional interactions have been found in the literature. Furthermore, this interaction might involve unidentified domains present in the long amino-terminus rather than the cytochrome b_5 domain itself.

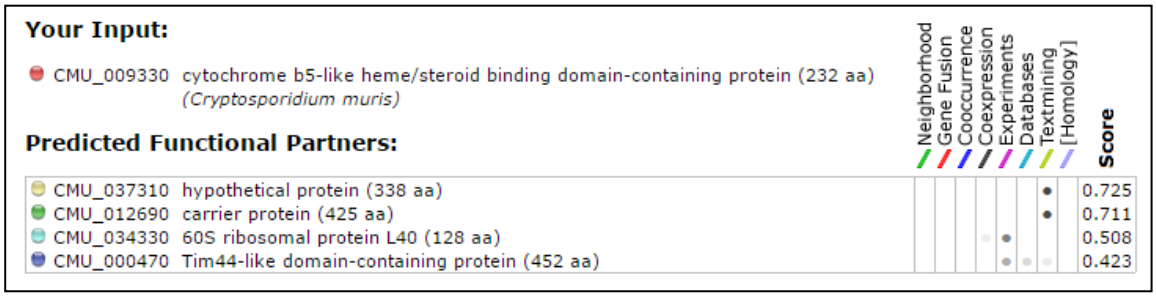


Figure 7: String results for the *C. muris* protein CMU_009330.

Finally, some results have been obtained for the three of the five Type II CYTB5s possessed by *Trichomonas vaginalis*. Two of proteins yielded the same set of results (Figure 8). As before there is no clear link between the Type II CYTB5s and these predicted partners. The hypothetical protein is also an IMPACT-like protein based on BLASTp analysis, but the confidence score for this protein is lower for *T. vaginalis* (0.561).

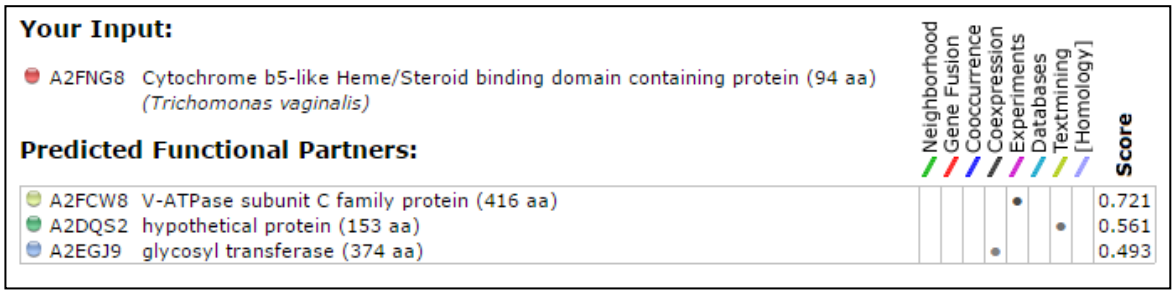


Figure 8: String results for the Type II CYTB5s of *T. vaginalis*, (UniProtKB A2FNG8 and A2DLT3).

For the third *T. vaginalis* protein several predicted partners were identified (Figure 9) which are predominantly ubiquitin family proteins and small GTP binding proteins, and two hypothetical proteins identified by BLASTp as Ras-GTPases. However, none of these proteins have been mentioned as cytochrome b_5 partners in the literature. Interestingly, two electron transfer flavoproteins are also present. The protein A2ENX0 corresponds to a flavodoxin, which are small proteins that contain the flavin mononucleotide redox-active cofactor. The protein A2EUJ7 also contains a flavodoxin domain as part of a larger protein and is likely a pyruvate-formate lyase enzyme. It is noteworthy that such enzymes also occur in Giardia.

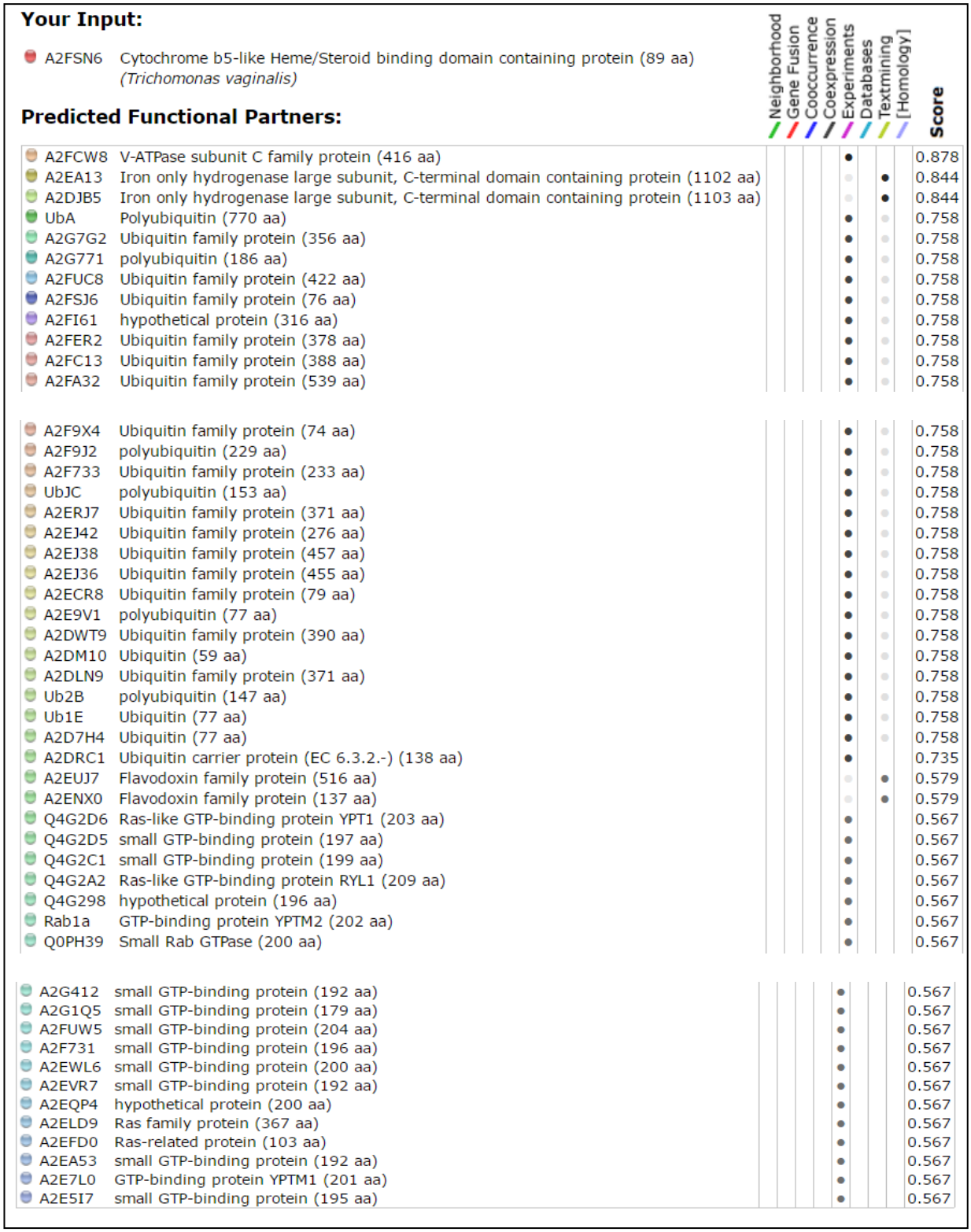


Figure 9: String results for the *T. vaginalis* type II CYTB5 (UniProtKB A2FNS6).

3.4 Immunolocalization of *G. intestinalis* gCYTB5-I, II and III

Since the STRING searches could not provide clues to the function and binding partners of the gCYTB5s, I tried experimental approaches. First, I used immunofluorescence microscopy to determine the cellular location of the gCYTB5s within Giardia trophozoites. Previous studies by Pyrih et al. (2014) used immunofluorescence microscopy on Giardia cells transfected with plasmids that express HA-tagged versions of each of the gCYTB5 isotypes. The use of a HA monoclonal antibody showed that all HA-tagged gCYTB5 isotypes were present mainly in the cytoplasm, with minor localization to the nuclei (Appendix H) (Pyrih et al. 2014).

In this study, I used antibodies generated against peptides that are unique to each gCYTB5 isotype in untransfected Giardia expressing only the endogenous proteins for immunofluorescence microscopy (Figure 10). I also obtained similar results with transfected Giardia expressing both the HA-tagged and endogenous gCYTB5s (Appendix J). Contrary to Pyrih et al.'s results, I found differences in the intracellular localization of the different gCYTB5 isotypes. I observed that gCYTB5-I has a predominantly perinuclear localization, while gCYTB5-III resides exclusively within the nucleus. Similar to the result from Pyrih et al. (2014), I also found gCYTB5-II in the nucleus and cytosol. However, the gCYTB5-II in the cytosol seems to be in punctate structures concentrated near the upper anterior portion of the cells. This pattern of staining is similar to that for proteins that are associated with the peripheral vesicles in Giardia (Zumthor et al. 2016). These observations are important for consideration of the prospective partners identified by mass spectrometry, as proteins with the same subcellular location as their gCYTB5 bait are more likely to be truly interacting proteins.

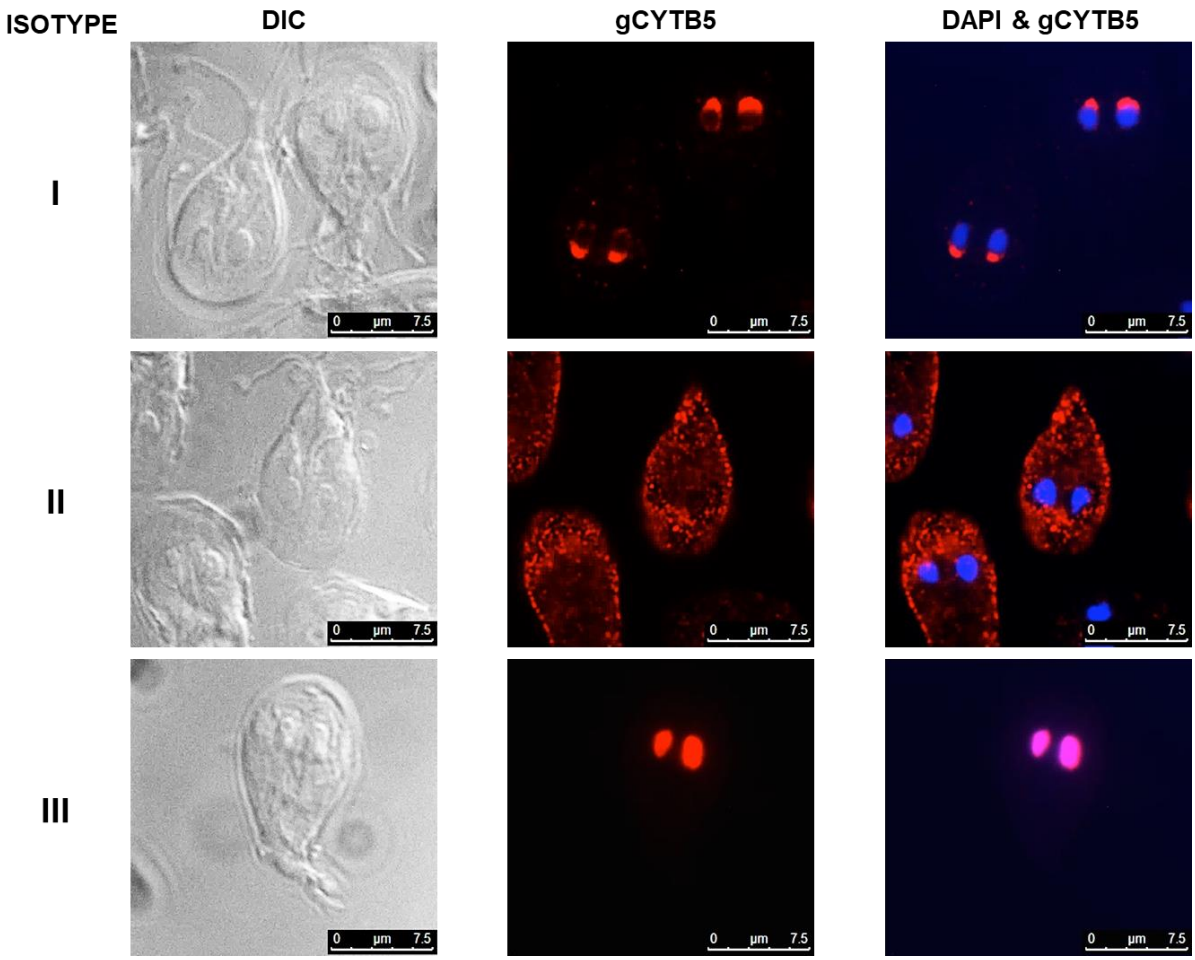


Figure 10: Immunofluorescence assay using untransfected Giardia trophozoites and protein specific antibodies. DIC (Differential Interference Contrast) was used to visualize the morphology of the cell, and DAPI (4', 6-diamidino-2-phenylindole) was used to stain the DNA and localize both nuclei. Anti-gCYTB5-I, II and III antibodies were used to localize the endogenous proteins.

3.5 Interactome study of *G. intestinalis* gCYTB5-I, II and III

3.5.1 Pull-Down Experiments with Dynabeads and Affi-Gel 10

Purified recombinant His_6-tagged gCYTB5-I (Appendix C) was used for pulldown assays with either cobalt coated magnetic beads (Dynabeads) or Affi-Gel 10 resin. Each support used has its own advantages and disadvantages; consequently, candidate partners common to both approaches are more likely to be robust. The Dynabeads bind His_6-tagged proteins in a single orientation through Co^{2+} coordination to the tag. Affi-Gel 10 is an agarose-based resin that forms covalent bonds to the lysine side chains on the bait protein, thus allowing the bait protein to bind via different sites and orientations to the resin. The binding capacity of the Affi-Gel 10 (30 µg/µL) is higher than the magnetic beads (1.6 µg/µL) and it has also been suggested that interactome studies that use Affi-Gel 10 are less prone to false positive results (Markillie et al. 2005). His_6-tagged gCYTB5-I was used in both procedures, as it was the more stable isotype among the three gCYTB5s when prepared as a highly purified recombinant protein in our laboratory. Conditions for each procedure were optimized to obtain the maximum amount of prey protein bound to the bait with the minimal amount of nonspecific binding of protein to the resin. This was determined by comparative SDS-PAGE of pull-downs with the bait bound to the resin (+bait) to pull-downs with resin alone (-bait). Note that the endogenous levels of the gCYTB5 partners in Giardia protein extracts may be too low to observe on a gel yet the sensitivity of mass spectrometry is such that proteins are nonetheless detectable. Negative control experiments (un-baited assays) was performed in parallel in which gCYTB5-I was omitted from the support before adding the Giardia protein lysate.

Figure 11, Panel A shows the results of a pull-down assay with gCYTB5-I on Affi-Gel 10 resin. Successful attachment of the bait to the resin is indicated by the bright-red colour of the resin due to the presence of the heme cofactor of the protein. After loading of the Giardia protein extract to the resin, three consecutive washes with the loading buffer are necessary to remove unbound protein. The presence of a prominent band near 35 kDa in the unbaited control (in the "-" lane of wash #3) is likely an artefact as this band was absent in other replicate experiments. Proteins eluted with 1 M NaCl that are unique to the baited resin are indicated by the red stars (24, 26, 42, 55, 100 kDa); these may correspond to gCYTB5-I interacting proteins. Some bait protein also appears in this lane at the bottom of the gel, which may be due to the observed tendency of this protein to form homodimers; if only one monomer is covalently bound to the resin the partner subunit would be eluted with an increase in ionic strength or with urea. With the exception of a small amount of gCYTB5-I no additional proteins were eluted from the resin by 8 M urea.

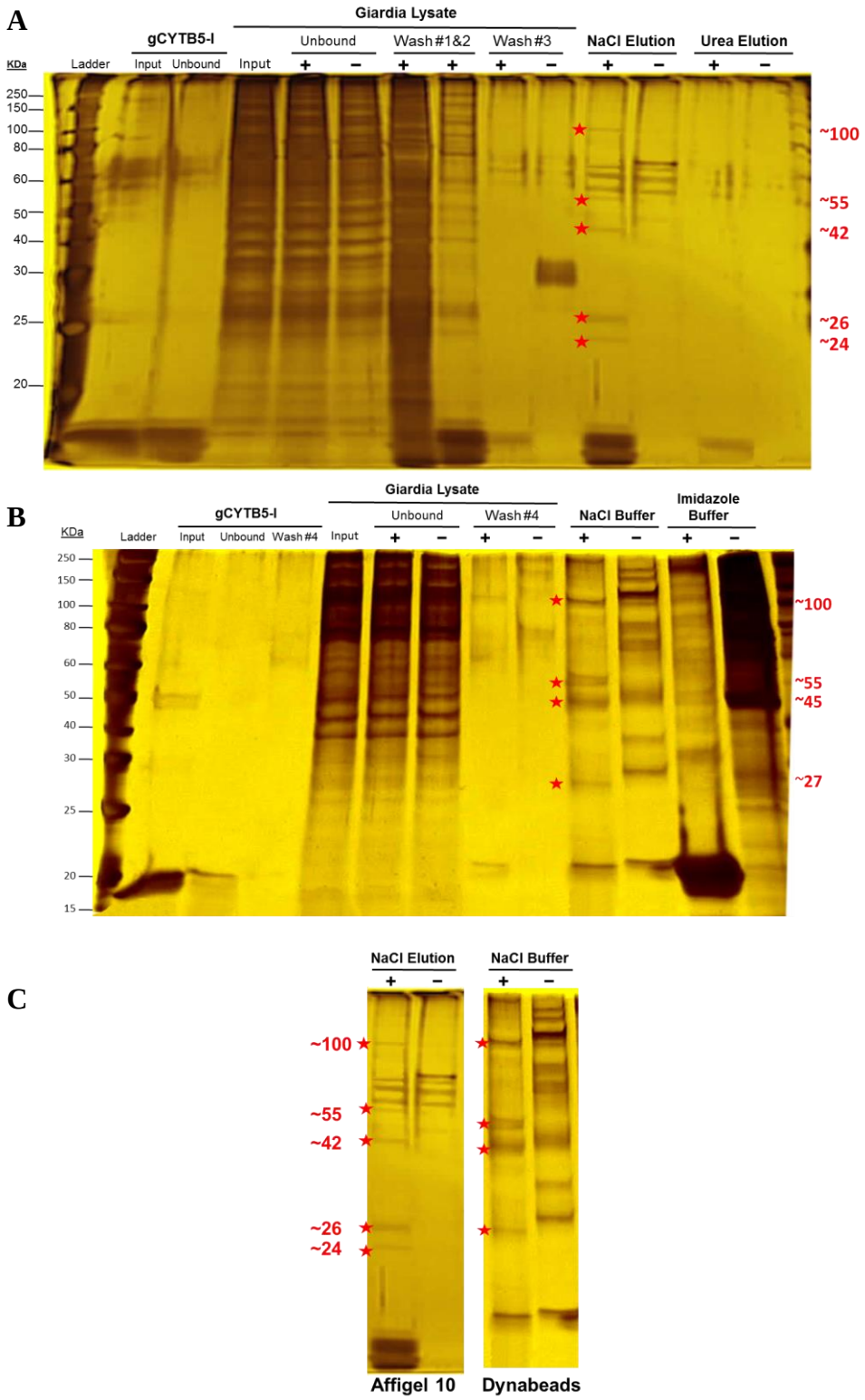


Figure 11: Pull-down experiments using gCYTB5-I as bait. (+) =pull-down assay with baited resin. **(-)** = pull-down assay with unbaited resin. Red stars indicate bands present in the baited resin and absent in the unbaited resin. Panel A: Affigel-10 pull-down. (**NaCl**) = elution with 1 M NaCl; (**Urea**) = elution with 8 M urea. Panel B: Dynabeads pull-down. (**NaCl**) = elution with 300 mM NaCl; (**Imidazole**) = 300 mM imidazole. Panel C: comparison of the results obtained with both approaches.

Similar experiments using His_6-gCYTB5-I bait and Dynabeads beads were also performed (Figure 11, Panel B). Bound proteins from the Giardia extract were eluted first by increasing the ionic strength of the elution buffer and then by the addition of imidazole, which will release the bait as well as any tightly-bound prey. Proteins eluted with 1 M NaCl that are unique to the baited resin and may be gCYTB5 partners are indicated by the red stars (27, 45, 55 kDa). Another possible partner may include the band around 100 kDa, which appears more intense than in the unbaited control. Interpretation of the elution with 300 mM imidazole was complicated by the release of a large amount of different proteins from the unbaited resin. This is likely a consequence of the binding of Giardia proteins with a weak affinity for the immobilized metal in the absence of saturating amounts of His-tagged bait proteins. In this regard, it is interesting to note the difference in the banding pattern between the baited and unbaited imidazole elution lanes.

Several experiments were performed with His_6-gCYTB5-I bait under different conditions using both supports. Interestingly, we detected several bands of similar sizes from both sets of resins when eluted with high concentrations of NaCl (Figure 11, panel C). This was expected, as cytochrome b_5 characterized from mammals tends to form weak electrostatic interactions with its redox partners (Meyer et al. 1995; Worrall et al. 2002). We decided to perform a Dynabeads pull-down experiment with a protein stain (Blue Silver) that was compatible with subsequent mass spectrometry analysis (Figure 12).

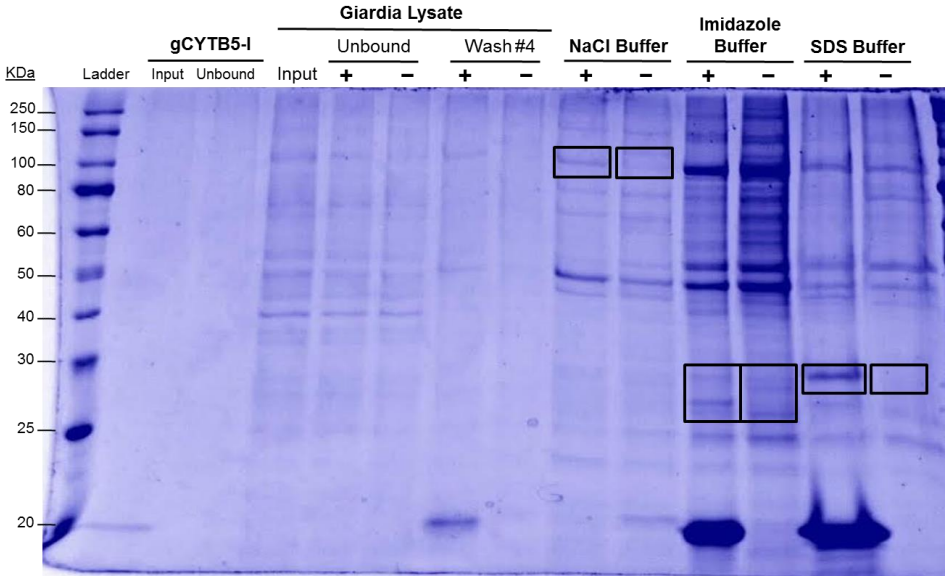


Figure 12: SDS-PAGE gel used for first MS analysis of gCYTB5-I. The pulldown assay was performed using Dynabeads and His_6-gCYTB5-I as bait. Blue Silver stain was used to reveal protein in the gel. Black boxes correspond to portions of the gel that were excised and analysed by mass spectrometry. (+) = baited resin; (-) = unbaited resin control.

Blue silver staining has slightly lower sensitivity than silver staining (0.5 ng protein compared to 0.2 ng for the latter). An additional elution with the detergent SDS was also included to strip the beads of any remaining bound protein. Bands were selected for MS analysis that were absent from the corresponding unbaited controls. These bands were excised along with the equivalent region from the un-baited pull-down assay.

Based on mass spectrometric analysis, proteins identified in the baited sample were classified as potential gCYTB5-I interacting proteins if they were absent from the negative control, or if their relative abundance was higher than in the negative control. The candidates identified by elution with 300 mM NaCl sample were promising, as several unique peptides were identified that are usually not considered to be contaminants (Table 6). The imidazole and SDS elution samples contained three candidate proteins but two of these (a ribosomal protein and a giardin) are usually

considered as contaminants in proteomic analysis of Giardia, owing to their high abundance in the cell. Mass spectrometry also detected a trace of bait protein that had been released from the resin. Based on this initial result from mass spectrometry we decided to focus on the 300 mM NaCl elution sample for further experiments (Table 6).

In preparation for our second mass spectrometry analysis, we tripled the amount of Giardia lysate applied to the resin and washed the resin more extensively (six washes) before increasing the ionic strength of the elution buffer. Using these parameters, a new pulldown experiment was performed and the entire lanes corresponding to the NaCl eluted samples from the baited (+) and un-baited (-) assays were excised and sent for MS analysis (Fig. 13). The potential gCYTB5-I interacting proteins are listed in Table 6.

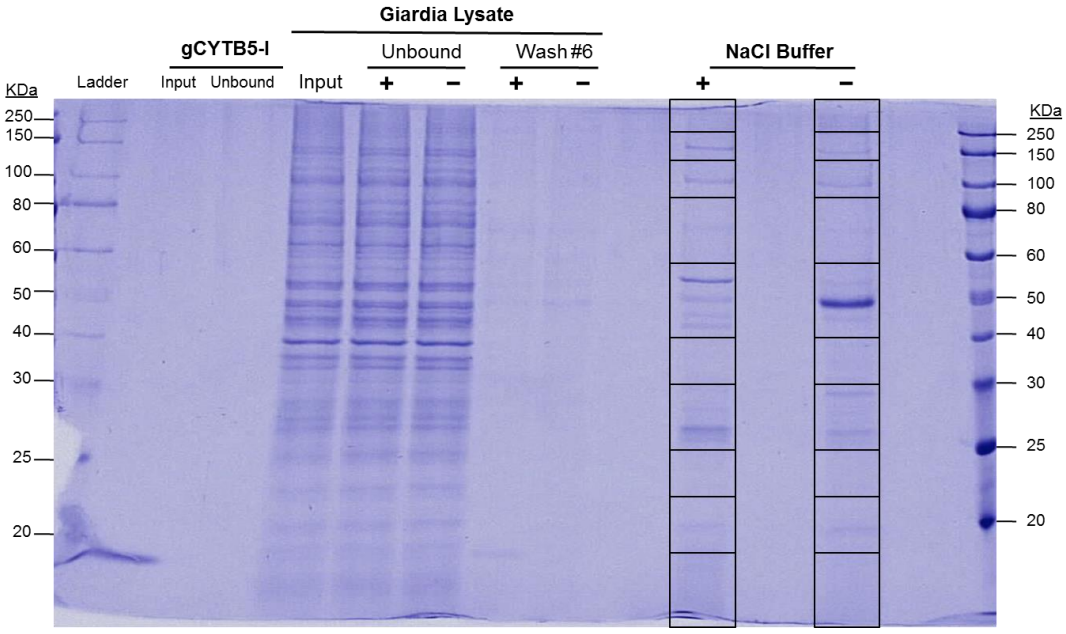


Figure 13: SDS PAGE gel used for the second MS analysis of gCYTB5-I. The pulldown assay was performed using Dynabeads and recombinant gCYTB5-I as bait. Blue Silver stain was used to reveal proteins in the gel. The full lane of the NaCl elution experiment and negative control were cut into slices indicated by the black boxes and analysed by spectrometry. (+) = baited resin; (-) = unbaited resin.

Table 6: List of the potential gCYTB5-I interacting proteins identified by MS from pull-down experiments.

	Area	Accession	Protein	Unique peptides +/-	Relative abundance +/-
MS 1 (Fig. 12)	NaCl	GL50803_21505	Protein 21.1	9/0	21/0
		GL50803_16568	Transcription factor p105	3/0	3/0
		DHA2_154376	Putative SUN domain protein	2/0	2/0
		GL50803_9909	Pyruvate, phosphate dikinase	22/15	35/17
	Imidazole	GL50803_16525	Ribosomal protein L3	2/0	3/0
		GL50803_11554	NEK kinase	2/0	2/0
	SDS	GL50803_9089	Cytochrome B5 I	7/0	87/0
		GLP15_1354	Alpha-7.3 giardin	3/0	3/0
MS 2 (results from Fig. 13)	Full lane analysis from NaCl elution	GL50803_94927	NEK kinase	10/0	18/0
		GL50803_8118	Ribosomal protein S2	6/0	14/0
		GL50803_5845	Ribosomal protein S8	4/0	18/0
		GL50803_12981	Ribosomal protein S5	4/0	6/0
		GL50803_9089	Cytochrome B5 I	3/0	15/0
		GL50803_5947	Ribosomal protein L35a	3/0	7/0
		GL50803_15930	RNA helicase, putative	3/0	5/0
		GL50803_15398	Chaperone protein dnaJ	3/0	4/0
		GL50803_92741	Kinase, CMGC CLK	3/0	3/0
		GL50803_10939	Putative uncharacterized protein	3/0	3/0
		GL50803_7870	Ribosomal protein L23A	3/0	3/0
		GL50803_113365	5'-3' exoribonuclease 2	2/0	5/0
		GL50803_221692	Putative uncharacterized protein	2/0	4/0
		GL50803_16887	ATP-dependent RNA helicase HAS1	2/0	4/0
		GL50803_14091	Ribosomal protein L14	2/0	4/0
		GL50803_40014	Protein 21.1	2/0	2/0
		GL50803_3993	Polyadenylate-binding protein	2/0	2/0
		GL50803_15042	Cleavage stimulation factor 50K chain	2/0	2/0
		GL50803_5533	DUB-1	2/0	2/0
		GL50803_1657	Translation initiation factor	2/0	2/0
		GL50803_5328	AP2 sigma adaptin	2/0	2/0
		GLP15_2373	Ribosomal protein S19e	2/0	2/0
		GL50803_9909	Pyruvate, phosphate dikinase	23/12	83/32
		GL50803_40067	Putative uncharacterized protein	10/3	36/4

While the same proteins were not identified in both experiments (which might be due to the changes in the protocol) proteins from the same family are present in each. First, a member of the Protein 21.1 family has been identified in both experiments. These proteins contain several ankyrin repeats, which are domains involved in mediating protein-protein interactions. While the roles of those interactions and the signalling pathways in which these proteins are involved in are not well understood (Manning et al. 2011), *Arabidopsis* Protein 2A, which contains ankyrin repeats has been shown to interact with membrane bound cytochrome b_5 (Shen et al. 2010).

Kinases are also present among the prospective partners of gCYTB5-I. This is notable as it appears that gCYTB5-I possesses a serine residue in the solvent-exposed C-terminal flanking region. This serine also has a high probability of phosphorylation, scoring a 0.992 (on a scale of 0 to 1) with the phosphorylation site predictor NetPhos2.0. Previous investigation of the Giardia kinome suggests that two of the prospective partners, CMGC CLK kinase (GL50803_92741) and a putative uncharacterized protein (GL50803_221692) are functional kinases (Manning et al. 2011). The two NEK kinases (GL50803_11554 and GL50803_94927) that were identified in this pulldown experiment are predicted to be inactive as they lack two essential residues in their catalytic domain. Such inactive kinases are common to most organisms, but their functions are unknown (Manning et al. 2011).

These pull-down experiments, while revealing the prospect of post-translational modification by phosphorylation of gCYTB5-I, did not identify any possible electron transfer partners. As complexes between redox proteins tend to be highly transient (Bashir et al. 2011), their absence from these experiments is perhaps not surprising, but nonetheless is disappointing.

One way to capture transient interactions is to employ covalent crosslinking agents within pull-down assays. We decided to use two different zero-length crosslinking reagents in our experiments. Paraformaldehyde is a crosslinker that is readily reversible when the sample is heated, while glutaraldehyde generates a crosslink that is effectively irreversible (Migneault et al. 2004). For these experiments the bait was preincubated with the Giardia protein extract in order for the bait and prey interactions to occur. Following the addition of the crosslinker, incubation and quenching, the solution was applied to the Dynabeads and processed as before. Figure 14 shows one of the results obtained using paraformaldehyde as a crosslinking reagent. Although two bands near 35 kDa (indicated by the black arrow) appeared in the imidazole elution of the baited lane, it is likely from their position and intensity that they represent homodimeric His_6-gCYTB5-I. To avoid the protein dimerization, a lower concentration of bait could be used in future experiments.

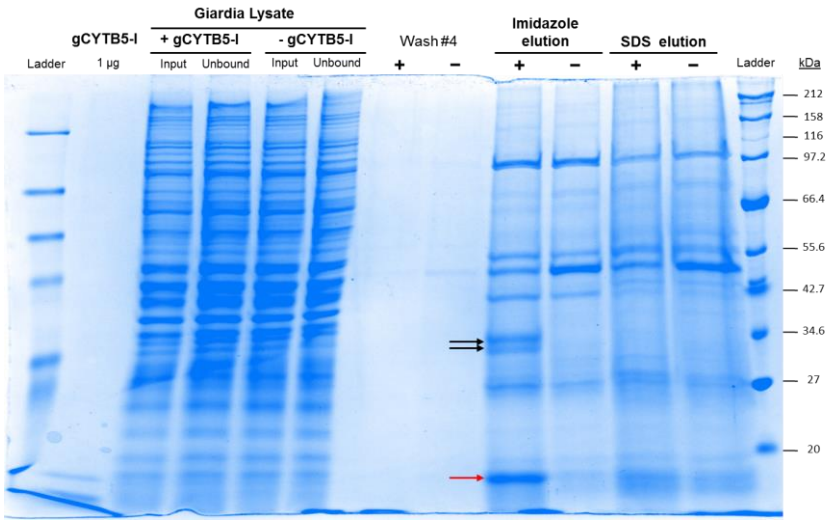


Figure 14: Pulldown assay using crosslinked protein. The pulldown assay was performed using Dynabeads, recombinant gCYTB5-I as bait and paraformaldehyde (1%, 5 minute incubation) as a crosslinker. The eluted samples were heat at 95°C for 20 minutes reverse the crosslinking. (+) = baited resin; (-) unbaited resin. The black arrow points out the two bands unique to the + imidazole lane and the red arrow points out the gCYTB5-I monomer protein.

3.5.2 Co-Immunoprecipitation

While conducting these crosslinking experiments we began a collaboration with the laboratory of Jan Tachezy at Charles University in Prague. Members of his laboratory have generated three different transgenic *G. intestinalis* cell lines, each expressing one of the gCYTB5 proteins with an N-terminal double hemagglutin tag (2xHA). These proteins are expressed from an episomal plasmid using an ornithine carbamoyltransferase promoter, and the plasmid is maintained under puromycin drug selection. Using a transgenic Giardia cell line expressing the tagged bait in co-immunoprecipitation experiments offers several advantages over the pull-down experiments. The bait proteins are endogenously expressed, allowing for any protein posttranslational modifications and interactions with the correct subcellular compartment to occur. Consequently, I decided to focus on using these cell lines for further experiments.

Total protein extracts were obtained from the cell lines expressing HA-tagged gCYTB5-I, II, and III, and from the negative control cell line that is transfected with a puromycin resistant vector that does not encode any HA-tagged proteins. A portion of each of the co-immunoprecipitation experiments was analysed by Western blotting with the anti-HA antibody to ensure that the tagged bait was successfully immunoprecipitated (Figure 15). The membrane was also stripped and reprobed with the different anti-gCYTB5s antibodies to confirm the presence of each isotype (Appendix K). We observe successful immunoprecipitation in all cases, with some slight degradation for the first two isotypes. The remaining 80% of each sample was also separated on a SDS-PAGE gel in preparation for the mass spectrometry analysis (Figure 16). Furthermore, the input and unbound samples were loaded on a second gel

to ensure equivalent loading between the tagged bait-containing experiments and the control (Appendix I).

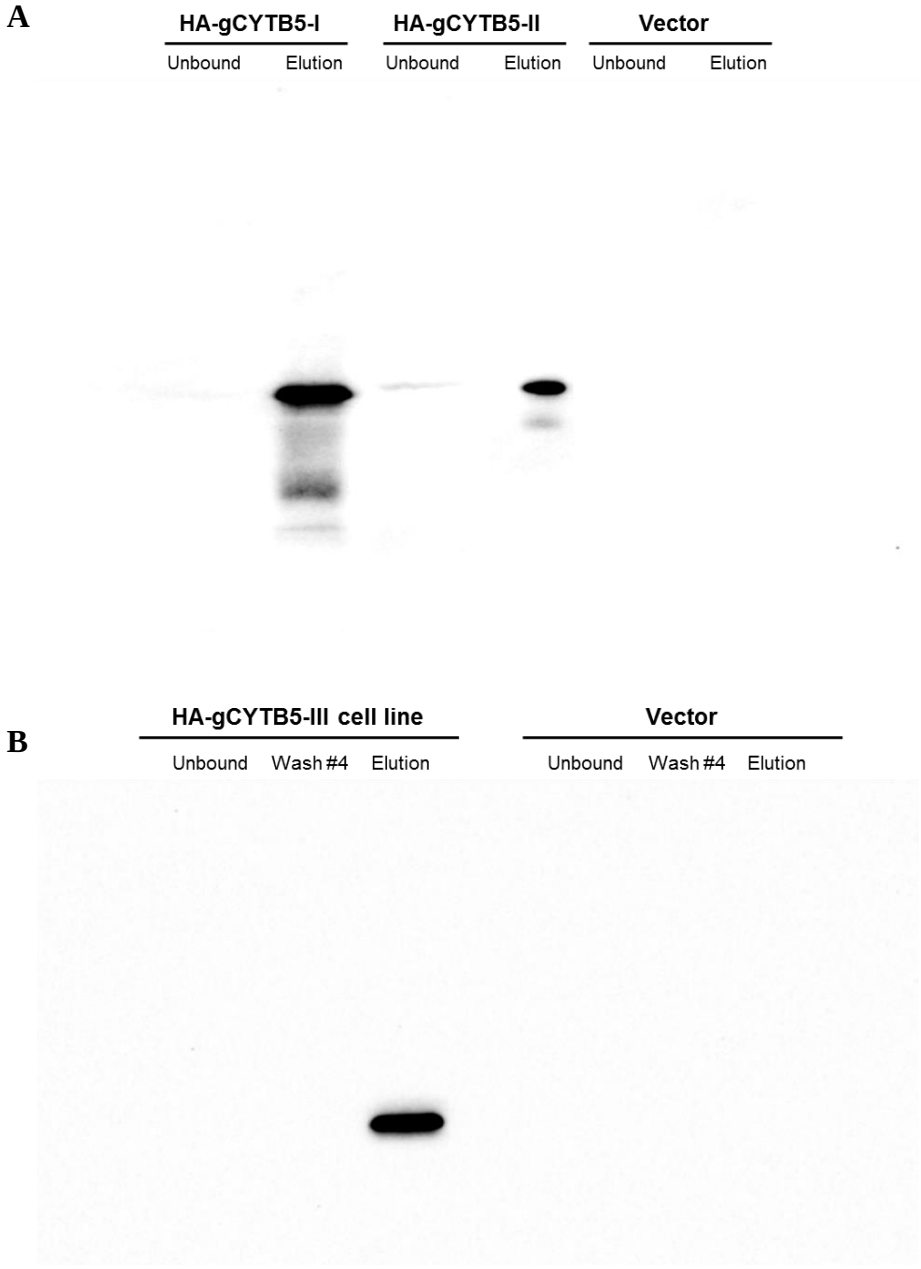


Figure 15: Western blot using anti-HA antibody. 20% of the unbound and eluted samples of each co-immunoprecipitation experiment were loaded on a SDS-PAGE gel for Western blot analysis. Panel A: HA-gCYTB5-I corresponds to the gCYTB5-I cell line, HA-gCYTB5-II corresponds to the gCYTB5-II cell line and the Vector is the control cell line. Panel B: HA-gCYTB5-III corresponds to the gCYTB5-III cell line.

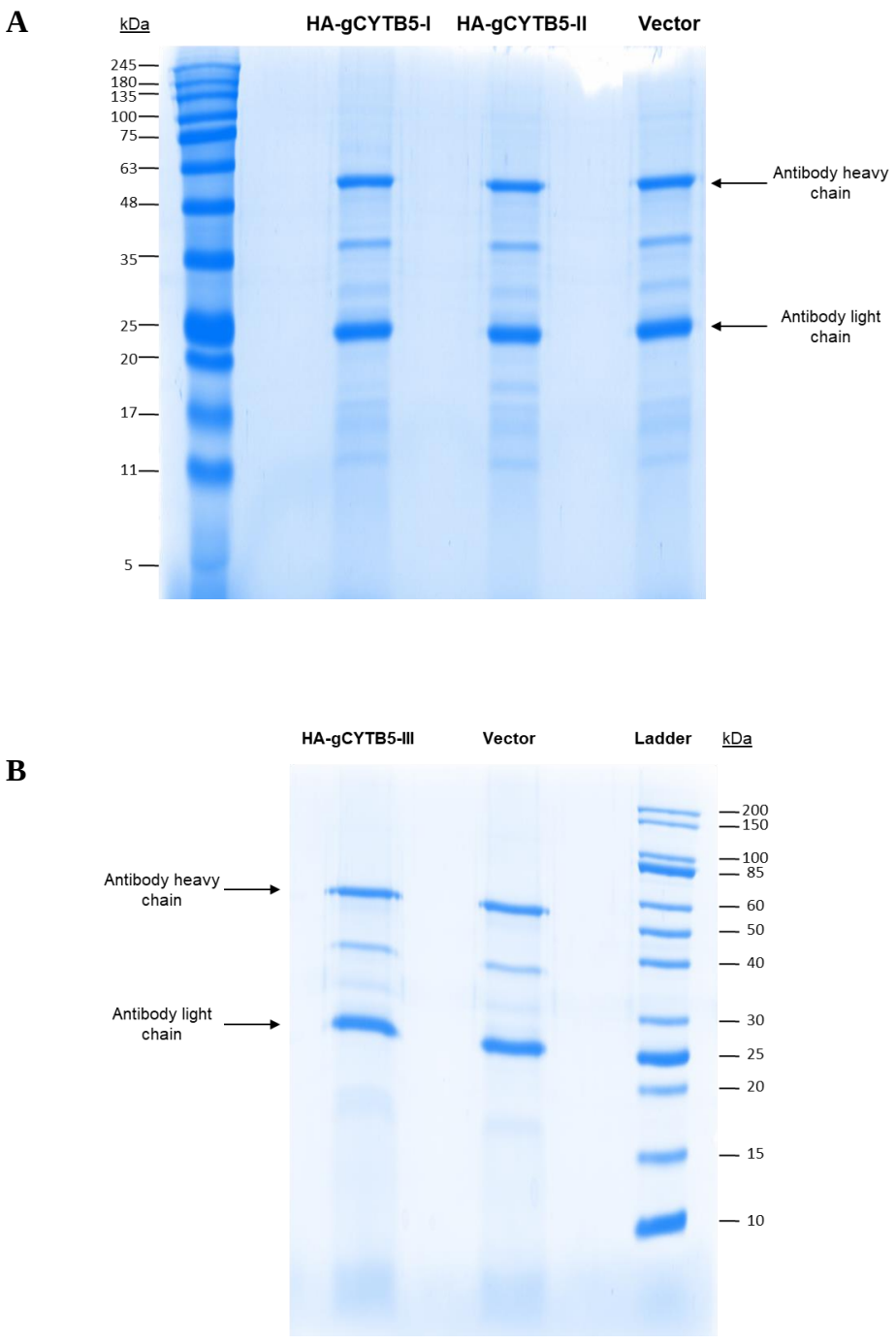


Figure 16: Co-IP experiments from Giardia trophozoites expressing HA-gCYTB5 I-III. Panel A: Co-immunoprecipitated samples for gCYTB5-I and II. Panel B: gCYTB5-III. Each lane was excised and analysed by mass spectrometry.

The full lane for each Co-IP experiment was excised and sent for MS analysis. Table 7 shows the proteins that were found only in the immunoprecipitate from cells expressing HA-tagged gCYTB5s. Each of the HA-gCYTB5 protein used as bait was identified in their respective immunoprecipitate and was not found in the control, indicating that the experiments were working properly.

Table 7: Potential interacting proteins identified by mass spectrometry from co-immunoprecipitations of Figure 16. Two prospective partners (blue) occur in co-immunoprecipitate of both isotypes I and II or in both isotypes II and III. Gene ID refers to the accession number in the GiardiaDB. Unique peptides are the number of peptide fragments of the protein detected by mass spectrometry.

gCYTB5-I		
Gene ID	**Protein name**	**Unique peptides**
GL50803_88765	Cytosolic HSP70	6
GL50803_9089	gCYTB5-I	4
GL50803_9861	Putative uncharacterized protein	3
GL50803_14373	Dynamin GTPase	2
gCYTB5-II		
Gene ID	**Protein name**	**Unique peptides**
GL50803_27747	gCYTB5-II	6
GL50803_9861	Putative uncharacterized protein	3
GL50803_14373	Dynamin GTPase	2
GL50803_7204	Putative uncharacterized protein	2
GL50803_15411	Kinase, NEK	2
GL50803_13864	Heat shock protein HSP 90-alpha	2
gCYTB5-III		
Gene ID	**Protein name**	**Unique peptides**
GL50803_33870	gCYTB5-III	5
GL50803_7204	Putative uncharacterized protein	2
GL50803_137716	Axoneme-associated protein GASP-180	2
GL50803_15409	Kinase, NEK	2

As expected, based on the Western blot results (Fig. 15), each of the immunoprecipitates resolved by mass spectrometry analysis contained the proper bait protein. In addition, some potential gCYTB5 interacting proteins were found in common among the different Co-IP results. A putative uncharacterized protein (GL50803_9861) and the dynamin GTPase were found in both gCYTB5-I and II precipitates, whereas another putative uncharacterized protein (GL50803_7204) was found in the samples corresponding to gCYTB5-II and III. Conversely, some potential interacting proteins were identified only in precipitates from gCYTB5-I (cytosolic HSP70), gCYTB5-II (Kinase GL50803_15411 and heat shock protein HSP90) or gCYTB5-III (axoneme-associated protein GASP-180 and kinase NEK GL50803_15409). These results are consistent with our hypothesis that each of the gCYTB5s would have common and unique partners. Interestingly, it appears that the gCYTB5 proteins are not co-immunoprecipitating each other, which suggest that these proteins do not interact directly with each other although they could be in the same metabolic pathway. Finally, only a few gCYTB5s interacting proteins have been identified by mass spectrometry. This may be due to the transient interactions of proteins involved in electron transfer reactions that are difficult to capture.

The mass spectrometry analysis identified a putative uncharacterized protein with the gene ID of GL50803_9861 as a potential interacting partner of both gCYTB5-I and gCYTB5-II. This protein, henceforth referred to as PUP9861, is a 385-residue protein of unknown structure. I considered PUP9861 as a priority candidate based on the following factors. First, the protein was identified in the MS data by the presence of three unique peptides in the immunoprecipitate for gCYTB5-I, which is considered a robust result. Second, immunofluorescent microscopy showed that PUP9861 is observed in punctate spots concentrated in the anterior of the cells (Dr. Staffan Svärd,

personal communication), which is similar to the localization of gCYTB5-II that I observed. Third, the protein level of PUP9861 increased during exposure of Giardia to nitrosative stress (Dr. Staffan Svärd, personal communication), which is similar to the increase in gCYTB5-II protein level during Giardia nitrosative stress that is observed by another student (B. Sajer) in our laboratory.

PUP9861 was used as a query in a BLASTp search but failed to identify any protein with a similar amino acid sequence in other species. Nevertheless, as proteins are more conserved at the folding level than at the sequence level (Marti-Renom et al. 2000), PUP9861 was used as the input for the structural homology-modelling Swiss-Model tool (Schwede et al. 2003). From this analysis, I found that a segment of the protein has a fold similar to the nitrite reductases and S-adenosylmethionine-dependent methyltransferases, which are redox-active enzymes (Figure 17). As this protein is the only one identified by mass spectrometry that seems to have a link with electron transfer reactions, its potential structure was investigated in more detail.

MPIIIKGVKLTEEFTRTLKEIIKGVDGFRSPYLANKARETDKG**VEIPCGYGKLLAAVRQKYPEADLLPL**
DQDEARKNCIDRNVSIFTKFKEYNADYLKTSKNVIITVTSAPGEFMDPEKVVAFLTKTLSKITPAQDHK
FKITAKYQVLDAIRGRINAGALKDASGSSVQGRCATMGIRYVEGRPPQRGTKVAVVVQSSEIEKLFAAL
SADNLFERVPAIRFINPEKRRKHIQELRERVGSAGDGTLGTKNKQRRHQEAKGGVKKPGAKKATGIRVQ
LADDQKAAKPKKKQGKKVQSQKERLPCLLTIAGIPEALAFDDIKENLDKDEHADILKALAESRLRRPKQ
PNPSEVSFYCTVENGKILRDAFGNMEINGAELRTTVSDVN

Figure 17: PUP9861 amino acid sequence. In bold is the portion of the protein with a nitrite reductase-like fold. This partially overlaps with the portion of the protein corresponding to an S-adenosylmethionine-dependent methyltransferases-like fold (underlined).

Application of the program Swiss Model generates a model for residues 44 to 209 of the PUP9861 that shares 16-17.5% sequence identity with a 171-residue portion of the nitrite reductases. Interestingly, cytochrome b_5 proteins are involved in protein complexes containing nitrite reductase activity (Sparacino-Watkins et al. 2014). The highest identity score is obtained with the tobacco leaf assimilatory nitrite reductase Nii3 (3vlz.pdb) (Nakano et al. 2012) with residues 160 to 330 of this enzyme used to generate the model of PUP9861 (Figure 18). Note that the I-TASSER protein structure prediction tool (Zhang 2008) yielded multiple results including a nitrite reductase (Probable ferredoxin-dependent nitrite reductase NirA from *Mycobacterium tuberculosis*, 1ZJ8.pdb with a Z-Score = 0.47 based on a maximum score of 1.00).

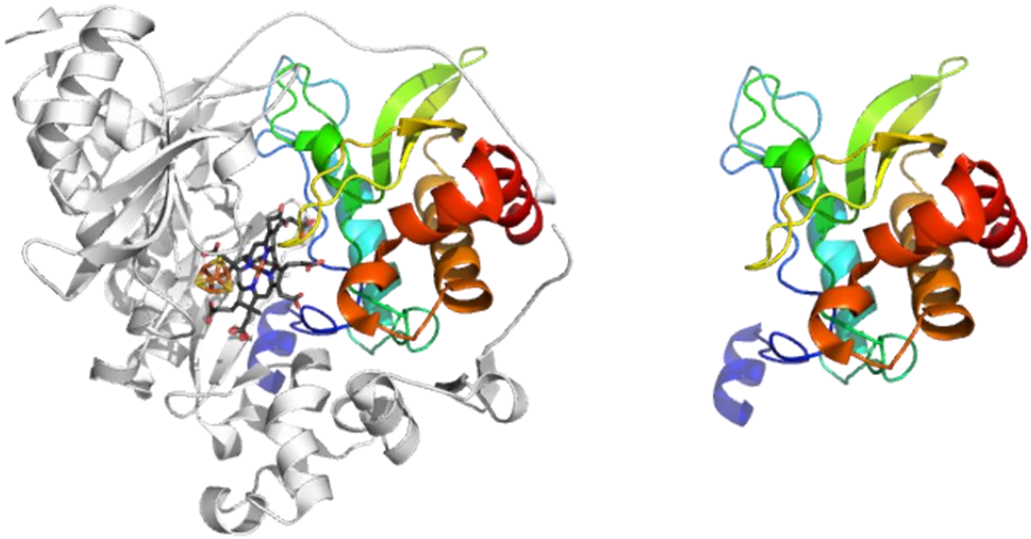


Figure 18: Structure of the PUP9861 generated using Swiss Model and Nii3 as a template. Left: Structure overlay of the Nii3 and PUP9861. The part of the PUP9861 protein modeled is shown in color and in white is the portion of the Nii3 absent in the protein of interest. The siroheme is show in the center of the protein. Right: Structure of the PUP9861. Note that the model is generated without its 43 residues at the N-terminus and 176 residues at the C-terminus.

Nitrite reductases use siroheme (a heme-like prosthetic group used by some enzymes to accomplish the six-electron reduction of sulfur and nitrogen) and an Fe_4S_4 clusters as cofactors (Murphy et al. 1974). The possibility that PUP9861 could bind such cofactors was examined with the RaptorX Binding Prediction program. This program first generates a three-dimensional structural model of the protein; for each predicted binding pocket, it analyses the frequency of which a given ligand is associated with that pocket. The score for each ligand found is given as the pocket multiplicity, which corresponds roughly to the number of proteins found in the database with a similar pocket as the query and with the same known ligand. If the score exceeds 40, there is a good chance that the ligand binds the protein (Kallberg et al. 2012). RaptorX Binding Prediction was used on segments of PUP9861 and Nii3 that were identified as similar in the Swiss-Model analysis. As expected, the most likely ligands for Nii3 are siroheme (score of 146) and iron-sulfur cluster (score of 43). However, neither ligand was a match to PUP9861; thus, it is unlikely to be a nitrite reductase. Nevertheless, the redox-active cofactor nicotinamide-adenine-dinucleotide (NAD) has a score of 77, which indicates a good probability that it is a ligand for PUP9861. This is of interest as the NADH and the related cofactor NADPH are the electron sources for cytochrome b_5 reductases and cytochrome P450 reductases (Schenkman and Jansson 2003) which indicates that the protein could be involved in electron transfer reactions.

The Swiss model tool also found resemblance between PUP9861 and some S-adenosyl methionine-dependent methyltransferases (SAM-methyltransferases) with an identity between 16 to 18%, but with a shorter coverage of the protein (Figure 17). The highest score is obtained with the SAM-dependent methyltransferase RosA (4d7k.pdb) from *Streptomyces davawensis.* The protein homology modelling tool, Phyre2, did not match PUP9861 to any nitrite reductases, but it did find matches with several SAM-

methyltransferases with low (8-20%) sequence identity. Similarly, I-TASSER identified the 5-methyltetrahydropteroyltriglutamate-homocysteine methyltransferase from *Candida albicans* (3PPC.pdb) as a potential match for PUP9861 (Z-Score = 0.58 out of a maximum score of 1.00). Using RosA as a template it was possible to provide a model for residues 41 to 142 of PUP9861 (Figure 19), which overlaps with the region modelled with Nii3 above. Using the RaptorX Binding Prediction function, a binding site for S-adenosyl-L-homocysteine was predicted for PUP9861 (score of 51).

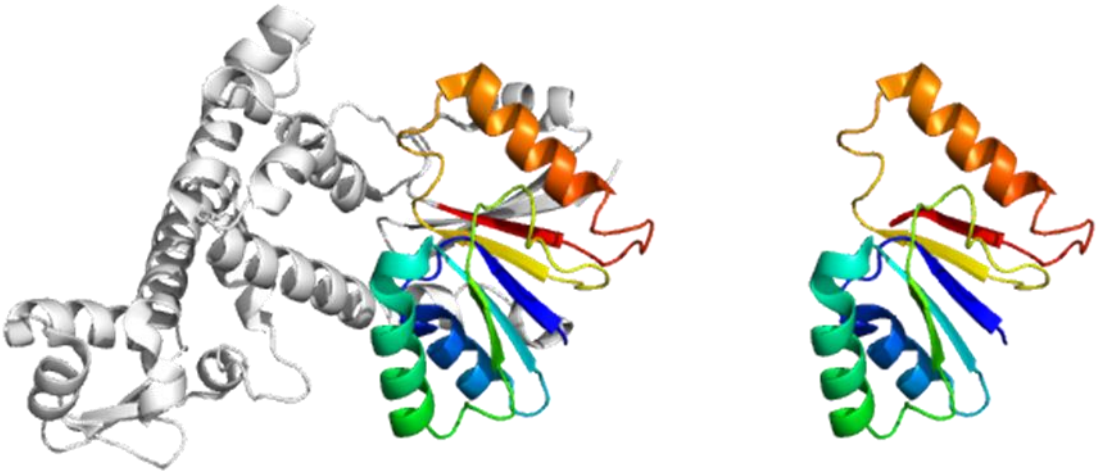


Figure 19: Structure of the PUP9861 generated using Swiss Model and RosA as a template. Left: Structure overlay of RosA and PUP9861. The part of the PUP9861 protein modeled is shown in green and in white is the portion of the RosA absent in the protein of interest. Note that only the chain of the RosA protein is shown here. Right: Structure of the PUP9861, note that the model is generated without 40 N-terminal and 243 C-terminal residues.

Although the N-terminal 40 residues and 176 C-terminal residues of PUP9861 have no matches to other known proteins at the amino acid sequence level (Fig. 17), several domains of very low identity were identified, especially on the C-terminal portion of the protein. This region has weak structural similarity to segments of other proteins, including, guanylate kinase, transferase, metallo-dependent hydrolases, RNA-

binding proteins, and an effector protein from Bartonella bacterium that causes cat scratch fever. However, none of these folds were large enough to be considered for further analysis.

In addition to PUP9861, two other proteins were identified in the gCYTB5-II immunoprecipitate: a dynamin protein and a specific NEK kinase with gene ID GL50803_15411. Both proteins have been shown to localize near the Giardia plasma membrane using GFP-tagged proteins (Hagen et al. 2011), and more recently it has shown that these protein localize specifically to the peripheral vacuole (PV)-plasma membrane (PM) interface in Giardia (Zumthor et al. 2016). It is intriguing that dynamin is identified as a potential interacting partner for gCYTB5-I, as dynamin has well-established functions in endocytosis, vesicle transport and membrane organization (De Camilli et al. 1995; Ferguson and De Camilli 2012). Since gCYTB5-I and II associate with specific membranes (perinuclear and peripheral vacuole respectively) it is possible that they do so through interactions involving this key membrane protein. Furthermore, dynamin is involved in control of cytochrome *c* release from mitochondria during apoptosis (Montessuit et al. 2010). While Giardia do not encode cytochrome *c* protein nor have mitochondria, both gCYTB5-I and II may be involved a cellular stress response as the level of gCYTB5-I mRNA increases following oxidative stress (Raj et al. 2014) while gCYTB5-II protein expression increases after exposure to nitrosative stress (Yee et. al, unpublished).

The putative uncharacterized protein with gene ID GL50803_7204, referred henceforth as PUP7204, was identified in the immunoprecipitate of both gCYTB5-II and gCYTB5-III. Analysis of this protein using BLASTp showed that it contains a central RNA Recognition Motif (RRM) superfamily domain (RRM_SF, cd00590), which is found in proteins involved in a broad range of post-transcriptional processes

such as mRNA processing and alternative splicing, as well as other process such as nuclear import of RNA binding proteins (Maris et al. 2005; Cassola et al. 2010). Proteins with this motif are also able to bind single-stranded RNA and single-stranded DNA, but they are usually accompanied by other domains to exert their functions (Maris et al. 2005). Modelling the possible structure of this protein with Swiss-Model and Phyre2 (Schwede et al. 2003; Kelley et al. 2015) suggested that PUP7204 has a ferredoxin-like fold (Figure 20). Ferredoxins are iron-sulfur cluster-containing proteins, which are part of several electron transfer processes (Mortenson et al. 1962; Valentine 1964). However, the ferredoxin fold is common to many proteins that do not bind metals, including many RNA-binding proteins (Wang and Li 2012). PUP7204 likely does not bind iron-sulfur clusters as it has only two cysteine residues whereas four are required to bind a minimal cluster (Fe_2S_2).

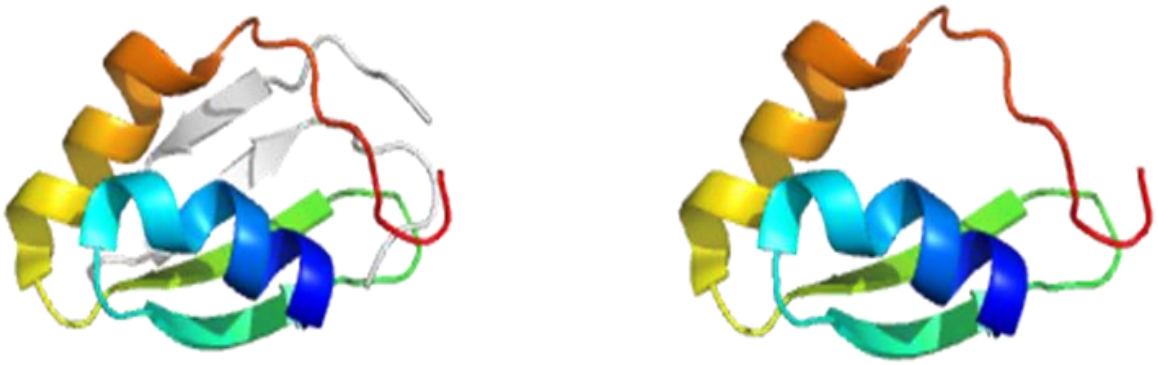


Figure 20: Structure of the P7204 generated using Swiss Model and the ferredoxin like fold domain of the Serine/arginine-rich splicing factor 1 (d3begb1.pdb). Left: Structure overlay of template ferredoxin-like fold domain and PUP7204. The part of the PUP7204 protein modeled is shown in color and in white is the portion of the template absent in the protein of interest. Right: Structure of the PUP7204 lacking 57 N-terminal and 48 C-terminal residues.

It is unclear why gCYTB5-III would interact with PUP7204 and which biological pathway would involve both these two proteins. As PUP7204 has also been

found in the gCYTB5-II co-immunoprecipitation, and both gCYTB5-II and III have at least partial localization to the nucleus, it is possible that PUP7204 has a role in the nuclear import of these cytochromes. Some heme proteins have functions in gene regulation (Kitatsuji et al. 2016), so that gCYTB5-II and III (more likely the latter since it is localized in the nucleus) may form heterodimers with PUP7204 to form a complex that binds nucleic acids. Such binding may use the cytochrome as a redox active sensor. However, such speculation would need to be supported by proven interactions between purified PUP7204, the gCYTB5s, and nucleic acids.

Several other proteins have been identified as potential interacting partners in the co-immunoprecipitation experiments for the different gCYTB5 isotypes but these are likely to be contaminants. This is probably the case for the heat shock proteins HSP70 and HSP90 identified in gCYTB5-I and II co-immunoprecipitations, respectively. Owing to their abundance in the cell, HSPs are usually considered contaminants although HSP60 has been shown to interact with cytochrome *c* in response to cell stress (Benjamin and McMillan 1998). The NEK kinases and the Axoneme-associated protein GASP-180 identified in the gCYTB5-III co-immunoprecipitation might also be contaminants, as these have been shown to localize to the plasma membrane and the cytoplasm (Hagen et al. 2011) while gCYTB5-III localizes to the nucleus.

3.6 Co-immunoprecipitation with PUP9861 as bait

Based on the information presented in the previous sections, PUP9861 appeared to be the most relevant of the potential gCYTB5-interacting candidates. I decided to confirm the interaction between PUP9861 and gCYTB5-I and II by performing a reverse co-immunoprecipitation. First, I cloned the coding region of PUP9861 into the 2xHA expression plasmid provided by the Tachezy lab. The resultant plasmid has the expression of the PUP9861 gene driven by the Giardia ornithine carbamoyltransferase gene promoter, which is the strongest known Giardia promoter (Jerlstrom-Hultqvist et al. 2012), and the gene for puromycin resistance driven by the GDH promoter. Giardia was transfected with this plasmid and this cell line was maintained by continuous exposure to puromycin. An HA-antibody was used to pulldown the HA-PUP9861 from the cell lysate, and then I looked for gCYTB5-I or II in the immunoprecipitate by Western blotting with the respective antibodies against each isotype.

In this experiment, HA-tagged PUP9861 was immunoprecipitated efficiently (Figure 21), although it appears to be degraded as shown by the presence of several lower bands below the band representing the expected size of the intact PUP9861 (42.7 kDa). To test for the presence of the gCYTB5-I and II in the eluted samples, the membrane was stripped and re-probed with the gCYTB5-I antibody, and then stripped again and re-probed with the gCYTB5-II antibody. No bands were observed for these last two Western blots.

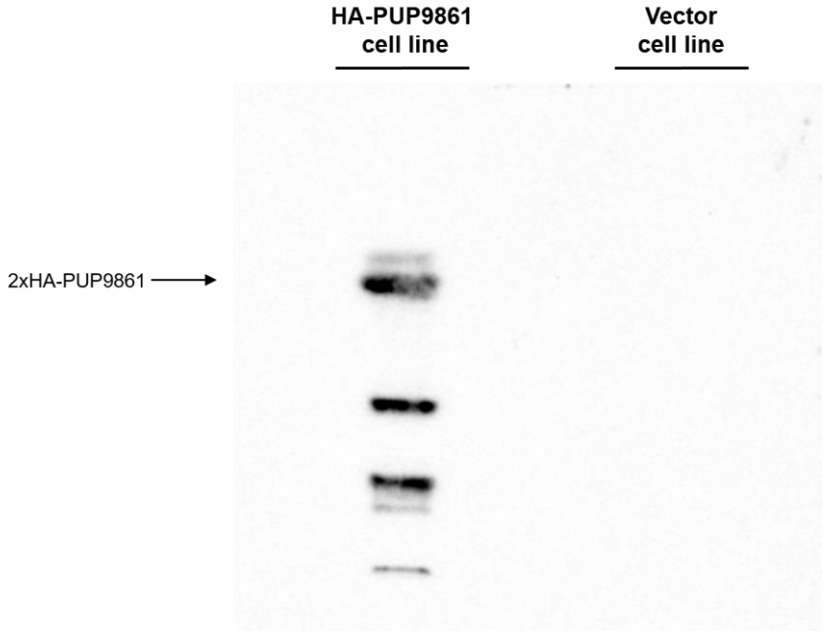


Figure 21: Western blot result of the 2xHA-PUP9861 immunoprecipitation. The plot was probed with ant-HA tag antibody. The vector cell line represents the Giardia cells stably transfected with the expression plasmid without any gCYTB5 sequences inserted. This vector cell line was also maintained by puromycin selection.

This could be explained by the low expression level of the endogenous gCYTB5 protein coupled to the transient nature of the electron transfer interaction. Moreover, some of the interaction might be lost during the washing of the resin so that insufficient protein remains in the immunoprecipitate to be detectable by Western blot analysis. Consequently, the same experiment should be performed using either an increased amount of total Giardia protein lysate or using a cross-linker reagent such as paraformaldehyde to covalently link the interacting proteins together.

It is unclear if the degradation of HA-tagged PUP9861 observed in Figure 21 occurred *in vivo* or during one of the incubation steps after the cells have been lysed. This degradation of the bait protein could impair the experiment, especially if the binding site of the gCYTB5s protein is on the cleaved C-terminal part of the PUP9861.

This could be resolved by adding an increased concentration of the protease inhibitor used in these experiments or by addition of other protease inhibitor. For example, ethylenediaminetetraacetic acid (EDTA), which is a metalloprotease inhibitor, was not added to this experiment as it could be potentially harmful for the iron-containing magnetics beads. EDTA could be used at a low concentration to determine if it helps maintain the protein's integrity.

3.7 Colocalization of PUP9861 and gCYTB5-II

To determine whether PUP9861 and gCYTB5-II are likely to be true interactors, I also used immunofluorescence microscopy to assess whether PUP9861 and gCYTB5-II are located in same cellular compartments. For this experiment, I used an anti-HA antibody to localize the HA-tagged PUP9861 in the Giardia cell line transfected with this plasmid (Fig. 22, in green fluorescent). I also used the gCYTB5-II antibody on the same cells to localize the cytochrome b_5 protein (Figure 22, in red fluorescent). The results showed that the cellular locations of PUP9861 and gCYTB5-II overlap as indicated by the yellow spots in the merged images of the two antibody hybridizations on Figure 22.

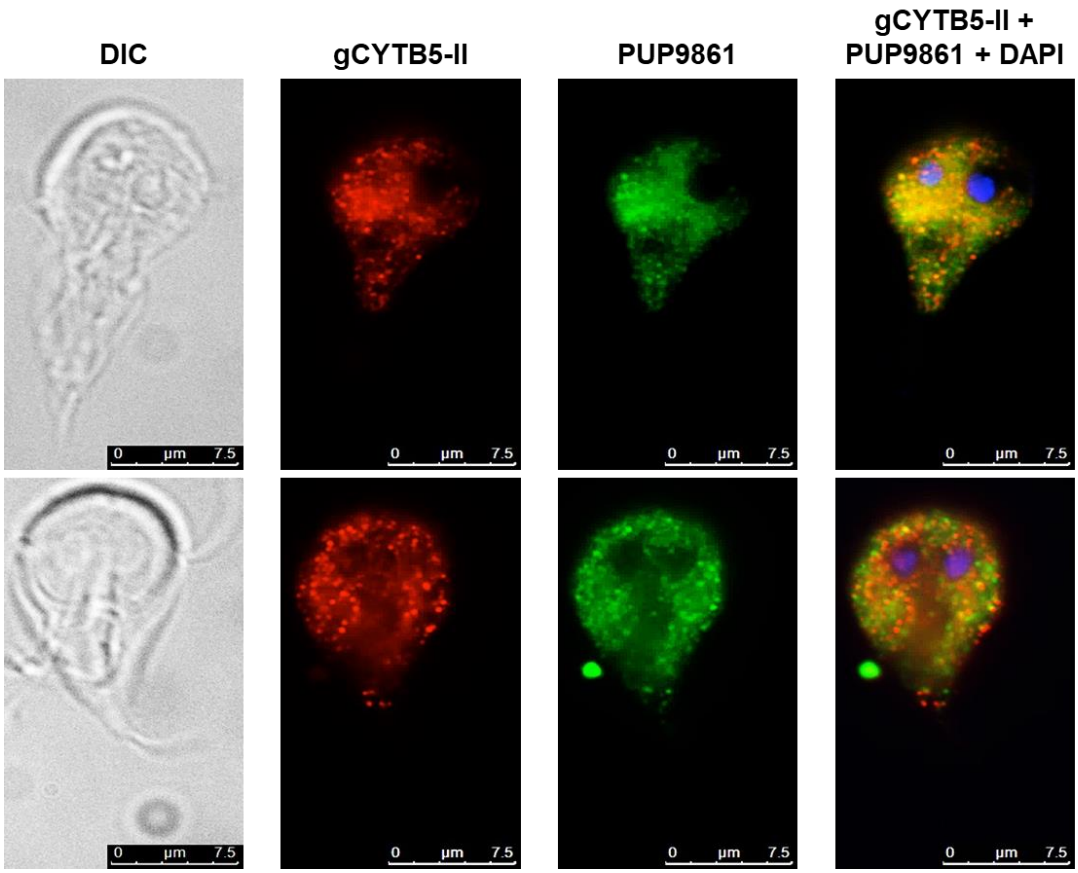


Figure 22: Colocalization of HA-tagged PUP9861 and gCYTB5-II using the Giardia cell line expressing HA-tagged PUP9861. Two representative cells are shown on this figure. The DIC is used to show the morphology of the cell, the localization of gCYTB5-II is shown in red while PUP9861 is shown in green. On the right, the overlay of the gCYTB5-II, PUP9861 and DAPI is shown. Each yellow spot on the overlay shows the interaction of the two protein.

DISCUSSION

4.1 Bioinformatics analysis of Type II CYTB5 and their potential partners

A previous study indicated that Type II CYTB5s are the only type of cytochrome b_5 found in anaerobic protozoa (Pyrih et al. 2014). I expanded the analysis to include anaerobic fungi, animals and plants. Of the 15 anaerobic protozoa species analysed, four species possess only Type II CYTB5s and the rest completely lack cytochrome b_5. My results showed that amitochondriate organisms encode exclusively Type II CYTB5s or encode no CYTB5s at all. However, more species would need to be screened to determine whether this observation is a general rule.

Although organisms expressing Type II CYTB5s are found in different eukaryotic kingdoms (animal, fungi, plant and protist), I was unable to discern the functions of any of these in literature searches, which included publications on the proteins that contain a cytochrome b_5 domain (fatty acid desaturase, cytochrome b_5 reductase) and the Type I cytochrome b_5. However, some information has been found on an ortholog of the Giardia gCYTB5s in the protozoan parasite responsible for malaria, *Plasmodium falciparum.* *P. falciparum* has one type-I and two type-II cytochrome b5s. One of the type-II CYTB5, PFI0885w ,is localized to the cytoplasm and the nucleus, whereas the type I CYTB5, PFL1555w, is phosphorylated (Treeck et al. 2011; Oehring et al. 2012). Furthermore, STRING analysis of this protein identified the protein kinase PfCK1 (*P. falciparum* casein kinase-1) as a potential partner. The interaction result of PFI0885w and PfCK1 is based on Coexpression view results in STRING and the confidence score was close to the medium cut-off (0.441). CK1 activity is required in a broad range of metabolic pathways as well as transcription, translation and regulation of the cell cycle (Schittek and Sinnberg 2014). The functions

of PfCK1 have not been fully characterized, but analysis of its interactome suggests that its partners are involved in a broad range of processes such as transcription, translation, splicing of pre-mRNA or quality control in mRNA biogenesis (Dorin-Semblat et al. 2015). Although PFI0885w CYTB5 was not identified in this interactome study of PfCK1 (Dorin-Semblat et al. 2015), it remains possible that the proteins are involved in the same signaling pathway without interacting directly but possibly through other kinases. These potential interactions are of interest as Giardia also encodes a CK1 (Manning et al. 2011). Moreover, immunofluorescence microscopy of Giardia locates the gCYTB5s to the cytoplasm and nucleus, and gCYTB5-I has a high-probability phosphorylation site in its carboxyl-terminal flanking region (Pyrih et al. 2014; Rafferty and Dayer 2015). The results of our pull-down experiments with gCYTB5-I as bait identified multiple kinases as potential interactors. Collectively these results suggest that one or more of the Giardia cytochromes are targets of post-translational modification by phosphorylation.

Fatty acid desaturases such as that of *Thalassiosira pseudonana*, an eukaryotic marine phytoplankton, lack an amino-terminal cytochrome b_5 domain and presumably have this domain on a separate protein (Tonon et al. 2005). Interestingly *Spironucleus salmonicida*, a Diplomonad closely related to Giardia that infects salmon, possesses a similar fatty acid desaturase that was likely obtained by horizontal gene transfer (Andersson et al. 2007). However, evidence for fatty acid desaturases in Giardia is lacking in spite of BLASTp searches for conserved motifs, literature searches, and searches of the protein family database (Pfam) of domains within the Giardia proteome (http://pfam.xfam.org/proteome/184922#tabview=tab2).

STRING analysis of the Type II CYTB5 of *Trichomonas vaginalis*, a protist responsible for a sexually transmitted disease in humans, identifies two possible redox

partners, both of which contain a flavodoxin domain that binds the redox-active cofactor flavin mononucleotide. A BLASTp search of these proteins against a Giardia sequence database identified Giardia oxidoreductase GiOR-2 as the closest match. GiOR-2 has sequence similarity to the NAD(P)-cytochrome P450 oxidoreductases, which contain two flavin-binding domains, and it localizes to vesicles within the Giardia cytoplasm. As with the gCYTB5s, the roles of GiOR-2 are unknown. A paralog of GiOR-2, GiOR-1, localizes to the mitosomes and is able to donate electrons to the gCYTB5s *in vitro* (Jedelsky et al. 2011; Pyrih et al. 2014).

STRING analysis of *G. intestinalis, C. muris and T. vaginalis* all identified an IMPACT-like protein as a potential partner, but the evidence is weak as this was solely through Text mining. The protein is a homolog of the mouse IMPACT protein (product of an Imprinted and Ancient gene) that also contains an RWD domain (RING finger and WD-domain-containing-proteins and DEAD-like helicases) at the amino terminus. The mouse protein is a regulator of translation and maintains translation at a high level in stressed cells, and it also acts as a negative regulator of some kinases (Pereira et al. 2005).

Bioinformatics searches and STRING analysis have provided no clear partners for Type II CYTB5s in any of the species examined, and a few leads that are speculative. This speaks to the limitations of *in silico* techniques in addressing questions about an unknown protein's functions and partners – although in the case of the Type II CYTB5s, their structural similarity to the well-characterized type I proteins would naturally lead one to expect that this approach would have been more productive. Consequently, more direct experimental effort is needed to understand the function of these enigmatic cytochromes.

4.2 Immunolocalization of the *G. intestinalis* gCYTB5-I, II and III

The goal of this research was to gain insight into the function(s) of the Giardia cytochrome b_5 proteins. Due to the absence of canonical cytochrome b_5 interacting proteins in Giardia together with the lack of information about the putative functions and partners of the Type II cytochromes b_5, little could be anticipated about the functions of these proteins in the anaerobic protozoa. In an effort to reveal the roles of these proteins, their intracellular localization was investigated. My results showed that gCYTB5-I was mainly found in the perinuclear space, gCYTB5-II may be associated with the peripheral vesicles, and gCYTB5-III was found exclusively in the nucleus.

In Giardia, the perinuclear space is contiguous with the endoplasmic reticulum and a very rudimentary Golgi complex. The association of gCYTB5-I with the ER is interesting because in higher eukaryotes, the membrane-bound cytochrome b_5 is found in the ER where the protein transfers an electron to several cytochrome P450s (Vergeres and Waskell 1995). However, due to the absence of cytochrome P450s in Giardia, it is likely that the gCYTB5-I protein is involved in a different metabolic pathway. To confirm the perinuclear localization of gCYTB5-I co-localization immunofluorescent microscopy experiments can be done with the 7-nitro-2-(1,3-benzoxadiazol-4-yl) amino-ceramide, which is a labelled lipid that is specific for these organelles (Lanfredi-Rangel et al. 1999; Lanfredi-Rangel et al. 2003).

The possible gCYTB5-II localization to the peripheral vesicle (PV) is also of interest as this organelle represents another Giardia peculiarity. In other eukaryotes, endocytosis is a well characterized mechanisms for the uptake of extracellular molecules and lipids (Samaj et al. 2004; Miaczynska and Stenmark 2008; Zumthor et al. 2016). Although two of the key proteins involved in this process, clathrin and dynamin, have been identified in Giardia, endocytosis in the parasite shows some

peculiarities (Zamponi et al. 2016; Zumthor et al. 2016). For example, the parasite use two different mechanisms to uptake nutrient for the environment. Lipids such as ceramide and low density lipids are internalized via receptor mediated endocytosis. However, in contrast to higher eukaryotes, the newly formed vesicles do not mature to form late endosome nor lysosome but rapidly transfer their content to the perinuclear space where protein degradation occurs (Hernandez et al. 2007; Abodeely et al. 2009). The parasite also possesses polarized peripheral vesicles that function as an endosomal-lysosomal system that encompasses the processes of endocytosis, recycling and degradation (Lanfredi-Rangel et al. 1998). These PV are static organelles mainly localized underneath the plasma membrane facing the intestinal lumen when the parasite is attached to the epithelial cells through its ventral disc (Zumthor et al. 2016). At some places, invagination of the plasma membrane reaches the peripheral vesicle creating an interface between the extracellular environment and the PV content. In addition, it has been suggested that the membrane separating the intestinal media to the PV is non-selective allowing the uptake or release of molecules, drugs or protein (Zamponi et al. 2016; Zumthor et al. 2016). Certain PV structures are also directly connected to the endoplasmic reticulum (Zumthor et al. 2016). The localization of gCYTB5-II is similar to the distribution of PVs throughout the cells, which suggests that the protein may have a role in the selective entry of compounds into the cell or a role in the response to drugs or harmful compounds. Furthermore, since the gCYTB5s proteins acquire their heme cofactor from the extracellular environment it is possible that their proximity to the PVs allows the gCYTB5 proteins access to the heme that may enter the cells through these structures. The peripheral vacuole localization of gCYTB5-II can be verified in co-localization experiments using fluorescently-labelled dextran (Zumthor et al. 2016).

The localization of the gCYTB5-III is clearly in the nucleus. Furthermore, Western blot analysis performed in our laboratory by another student (B. Sajer) with Giardia cellular fractions showed gCYTb5-III in the nuclear fraction but not in the cytosolic fraction. This result was surprising as the canonical Type I cytochromes b_5 are not nuclear proteins. However, as noted previously, mass spectrometry analysis of *Plasmodium falciparum* nuclear proteins identified the Type II cytochrome b_5 PFI0885w. These observations suggest that nuclear Type II cytochromes b_5 may not be restricted to Giardia, and may have role other than participating in metabolic pathways. Their function may resemble that of another heme-containing protein found in the nucleus such as heme oxygenase-1, which is involved in gene regulation under stress conditions (Lin et al. 2007).

Localization of gCYTB5s has been previously reported by the Tachezy laboratory where they used immunofluorescence microscopy with an anti-HA antibody on transgenic Giardia cell lines that expressed HA-tagged proteins episomally (Pyrih et al. 2014). Their results showed that all four isotypes localize mainly to the cytoplasm, with minor localization to the nucleus. These results were at odds with what we expected, as we hypothesized that the gCYTB5 isotypes would have different subcellular locations based on significant differences in the charge properties of the sequences that flank their more well-conserved heme-binding cores (Fig. 2). Tachezy's results are also different than my results where I used antibodies generated against peptides that are unique to each gCYTB5 isotype to detect the localization of the endogenous proteins in untransfected (Figure 10) or transfected Giardia trophozoites (Appendix J). My results showed that gCYTB5-I has a predominantly perinuclear localization, gCYTB5-II was in vesicles within the cytosol, and gCYTB5-III resides exclusively within the nucleus.

The differences in my results and the results from the Tachezy laboratory may be due to several factors. First, they used an antibody against the HA-tag for all three gCYTb5 isotypes rather than the custom peptide antibody against each isotype that I used. Their result showing the same predominately cytoplasmic localization of all three gCYTb5 isotypes may be an artifact from background staining of the HA-antibody. Second, the preparation of Giardia cells for IFM were slightly different in the Tachezy lab than the procedure that I used. For example, I fixed the Giardia cells in methanol whereas they used paraformaldehyde. In some case, it has been observed that methanol fixation could alter the cell structure that could affect the localization of proteins as viewed by IFM (Hoetelmans et al. 2001). However, my IFM results did not change when I replaced methanol with paraformaldehyde as the primary fixation step in the preparation of the cells (data not shown).

4.3 Interactome study of *G. intestinalis* gCYTB5s

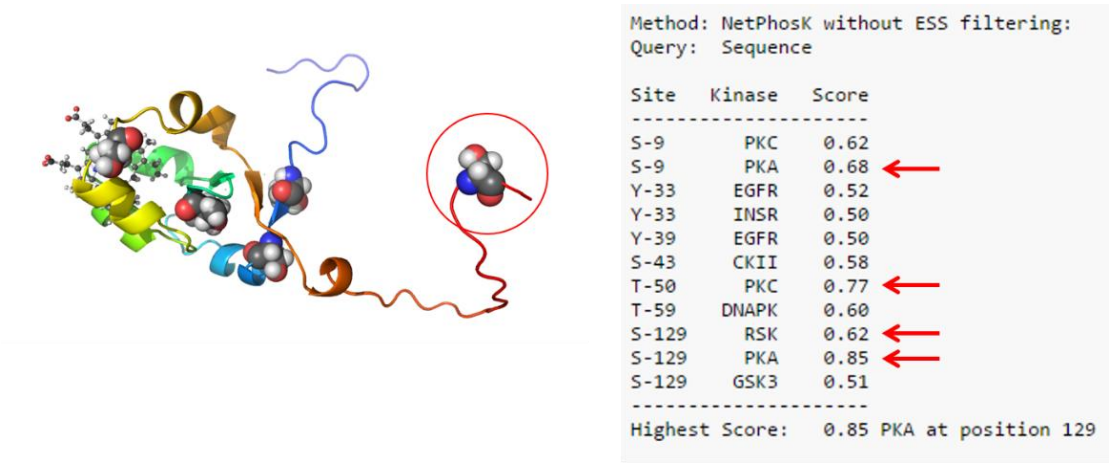
Different supports (Affigel-10 and Dynabeads) were used in pull-down experiments to identify possible interacting partners of gCYTB5-I. The results obtained using both methods were comparable (Fig 11). The potential interacting proteins identified in these experiments were mainly kinases and kinase-related proteins such as Protein 21.1, which have a similar domain organization as the NEK kinase but lacks the kinase domain. Although Protein 21.1s are abundant in Giardia, they are not usually considered as contaminants in mass spectrometric analysis. Kinases and Protein 21.1 members are found in all Giardia organelles (Manning et al. 2011) and are proposed to be involved in a broad range of metabolic pathways (Jedelsky et al. 2011; Manning et al. 2011; Faso et al. 2013; Raj et al. 2014). Furthermore, several hemeproteins are phosphorylated; for example, rat liver cytochrome P450 LM2 activity is regulated by

cAMP-dependent protein kinase phosphorylation, which could be inhibited by cytochrome b_5 (Epstein et al. 1989). This raises the possibility that some of these proteins identified in the MS data are true gCYTB5-I interacting proteins, especially as it bears a high-probability phosphorylation site in its carboxyl-terminal flanking region.

One of the putative partners of gCYTB5-I identified from the MS data is the protein with the ID of GL50803_221692 (Table 6). This is the only protein belonging the phototrophin and flippase kinase (PTF kinase) family found in Giardia (Manning et al. 2011). Phototrophins are receptors for certain flavin-containing electron-transfer proteins. Members of the PTF kinase family occur in yeast, fungi, and plants, as well as the free-living protozoan Naegleria. Such enzymes are functionally diverse. In yeast the PTF kinases FPK1 and FPK2 regulate the maintenance of membrane phospholipid asymmetry by activating lipid flippases (Nakano et al. 2008). In green plants, PTF kinases are mainly blue-light dependent kinases with a phototrophin domain, also known as a light-oxygen and voltage (LOV) domain (Aihara et al. 2012). Interestingly, LOV domains belong to the family of Per-ARNT-Sim (PAS) domains, which are involved in several reactions including sensing of oxygen or redox reactions (Taylor and Zhulin 1999; Ogura et al. 2008). Naegleria also encodes for a PTF kinase but no information has been found in the literature about this protein.

Despite the fact that no direct links between cytochrome b_5 and PTF kinase appear in the literature, some of the functions of the PTF kinase might involve interaction with electron transfer proteins. Bioinformatics tools were used to assess the likelihood of gCYTB5-I phosphorylation by this protein. As noted above, gCYTB5-I possesses an exposed serine at position 129, which has a high probability of phosphorylation. The NetphosK phosphorylation site predictor tool was used to determine which kinases are most likely to phosphorylate this residue. The score

obtained predicts that gCYTB5-I would most likely to be phosphorylated by a PKA kinase (Figure 23); interestingly GL50803_221692 possesses several functional domains including a PKA catalytic domain among others (PTF, PKC and RSK) suggesting gCYTB5-I could be a substrate and hence a partner of this kinase.



```
Method: NetPhosK without ESS filtering:
Query:  Sequence

Site   Kinase   Score
--------------------
S-9       PKC    0.62
S-9       PKA    0.68 <--
Y-33     EGFR    0.52
Y-33     INSR    0.50
Y-39     EGFR    0.50
S-43     CKII    0.58
T-50      PKC    0.77 <--
T-59    DNAPK    0.60
S-129     RSK    0.62 <--
S-129     PKA    0.85 <--
S-129    GSK3    0.51
--------------------
Highest Score:   0.85 PKA at position 129
```

Figure 23: gCYTB5-I homology model and NetphosK analysis. The modelled protein structure of the gCYTB5-I protein is shown on the left with the exposed Serine (129) circled in red. The list of the phosphorylation sites and their kinase preferences is shown on the right. The red arrows point at the kinase corresponding to the potential catalytic domains also found in GL50803_221692.

The results obtained using His_6-gCYTB5-I in pull-down assays were useful in revealing potential post-translational modification interactions but did not identify any electron transfer proteins. Therefore, I decided to use co-immunoprecipitation of HA-tagged gCYTB5s protein expressed in Giardia cells from a plasmid, and to use protein crosslinkers, with the hope of capturing the transient protein interactions that characterize biological electron transfer complexes.

Interestingly none of the proteins identified from the pull-down experiments with the His_6-gCYTB5-I were identified in these co-immunoprecipitation experiments. The lack of overlap in interactomes identified by different techniques is a common issue (Cusick et al. 2005) and may be explained by several factors. Perhaps most important

is the difference in the buffer conditions employed, especially in the tacit assumption of the pulldown experiments that the predominant interactions involving the gCYTB5s were electrostatic in nature, being preserved at low ionic strength and being release at higher ionic strength. While this may be true for the most well-studied type I cytochromes b_5, it may not be the case for the type II proteins of Giardia. Our bias in the pull-down experiments for selecting based on electrostatic interactions may therefore have missed protein-protein interactions that were driven by other considerations such as hydrophobic interactions. Consequently, the different conditions of capture for the protein-protein interactions of gCYTB5-I may have favoured the identification of different sets of partners.

To analyse the results of the co-immunoprecipitation experiments with the HA-tagged gCYTB5s, I first considered if the different potential partners identified are found in the same cellular compartment as their respective baits. PUP9861 was identified in gCYTB5-I and II immunoprecipitates, and its localization is very similar to that of gCYTB5-II (Dr. Staffan Svärd, personal communication) and this information was confirmed by the colocalization experiment I performed. PUP9861 is found throughout the cytoplasm with punctuated localization at the plasma membrane that is consistent with the peripheral vesicles (PVs). Recently, co-immunoprecipitations were performed with the Giardia clathrin, a protein found in the PV, to identify its interacting proteins. PUP9861was one of the proteins identified in the immunoprecipitate, thus supporting the idea that PUP9861 is associated with PV (Zumthor et al. 2016). Moreover, the plasma membrane, peripheral vesicles, endoplasmic reticulum and perinuclear space form a contiguous interacting network in Giardia trophozoites (Zumthor et al. 2016). This explains how PUP9861 could interact with both gCYTB5-

I (endoplasmic reticulum, perinuclear space) and gCYTB5-II (plasma membrane, peripheral vacuole, cytosol).

Dynamin is another protein found in the interactome of both gCYTB5-I and II. Localization studies with a GFP fusion with the Giardia dynamin showed that it is mainly found at the plasma membrane (Hagen et al. 2011) and is also associated with peripheral vesicles (Zumthor et al. 2016). This localization is very similar to that of NEK kinase (Hagen et al. 2011; Zumthor et al. 2016) that was identified in the gCYTB5-II co-immunoprecipitation, which suggests that these three proteins could interact. The main difference between the dynamin and NEK kinase localization is that dynamin also localizes to an area close to the basal bodies that could be linked to the ER. This localization might allow the dynamin to interact with gCYTB5-I

The final protein identified in the gCYTB5-I immunoprecipitate is cytosolic HSP70. Although its name suggests that it is in the cytosol, there is no experimental data on its localization in Giardia. Similarity, the putative uncharacterized protein GL50803_7204 and heat shock protein HSP90 identified in the gCYTB5-II immunoprecipitate also do not have any experimental data on their localization in Giardia.

The gCYTB5-III co-immunoprecipitation contained three interacting candidates: PUP 7204 (GL50803_7204), a NEK kinase (GL50803_154094) and the axoneme-associated protein GASP-180 (GL50803_137716). BLASTp analysis showed that PUP 7204 contains a potential RNA-binding domain, so it is possible that this protein is located in the nucleus, and thus, is a true interactor of gCYTB5-III. In contrast, previous work on the localization of the other two interacting candidates of gCYTB5-III showed that a GFP-tagged version of this specific Giardia NEK kinase is in the plasma membrane, and a GFP-tagged version of this GASP-180 is in the

cytoplasm (Hagen et al. 2011). The lack of these proteins in the nucleus suggests that they are unlikely to be real gCYTB5-III partners. Alternatively, these results could indicate that rather than interacting with other proteins, gCYTB5-III interacts directly with DNA or RNA. Such an interaction has not been observed for the cytochromes b_5 before, but the highly basic flanking regions of the protein and its nuclear localization lead us toward this hypothesis. Furthermore, interaction between proteins with heme binding capacity and DNA are known to occur (Singleton et al. 2010). Indeed, the *Rhizobium leguminosarum* iron regulatory protein which functions as a repressor of multiple genes involved in heme biosynthesis has its DNA binding capacity altered upon binding of the heme cofactor (Singleton et al. 2010). As the HA-tagged gCYTB5-III can be immunoprecipitated very efficiently it may be possible to use it in chromatin immunoprecipitation assays to test this hypothesis and if needed, identify the DNA binding site of the protein.

Among the different proteins identified using co-immunoprecipitation, the most promising is PUP9861 due to: 1) its presence in the immunoprecipitates of both gCYTB5-I and II; 2) the colocalization of PUP9861 and gCYTB5-II; 3) its predicted protein fold places it among those proteins that bind nucleotide cofactors such as NAD(P)H. Furthermore, expression of the gCYTB5-I mRNA increases in response to oxidative stress (Raj et al. 2014) while protein expression level of gCYTB5-II and PUP9861 increase after nitrosative stress exposure. However co-immunoprecipitation with HA-tagged PUP9861 failed to support this hypothesis, as neither gCYTB5-I nor gCYTB5-II was identified. It may be possible that with improvements of the methodology, such as the inclusion of a crosslinking step, that these interactions could be detected.

Due to the low sequence identity between PUP9861 and the templates used to model its partial structure, the function of the protein could not be determined with certainty. Nevertheless, redox roles of PUP9861 in oxidative stress are possible, and comparison to known systems provide promising avenues of speculation. The *S. cerevisiae* protein Dre 2 contains an N-terminal nucleotide binding domain (S-adenosylmethionine-dependent methyltransferase-like domain) and a C-terminal domain, which binds two iron-sulphur clusters. Together with Tah18 protein, Dre 2 forms a complex that is proposed to function as a sensor of oxidative stress (Soler et al. 2012). Based on the PUP9861 sequence, it is unlikely that it binds iron-sulfur clusters in its C-terminal half as it lacks a sufficient number of cysteine residues. However, its homology models are consistent with its potential binding to redox active nucleotide cofactors. The complex of PUP9861 and gCYTB5-I or II may have a similar function to Dre 2, which does not appear to have a homolog in the Giardia genome based on sequence similarity. Furthermore the Giardia flavoprotein GiOR-1 resembles Tah18 in its domain structure and organization and is able to donate electrons to gCYTB5-I and III *in vitro* (gCYTB5-II was not evaluated in this study) (Pyrih et al. 2014). Since the gCYTB5-I mRNA expression increased in response to oxidative stress (Raj et al. 2014) and gCYTB5-II expression increased due to nitrosative stress, it may be that gCYTB5 I/II, PUP9861 and GiOR-1 form a ternary stress complex. To investigate this possibility, our laboratory is currently performing localization studies of the gCYTB5s and PUP9861 under stress conditions.

The I-TASSER prediction tool identified methyltransferase as templates for PUP9861. Methionine synthase is a multidomain protein with methyltransferase activity that methylates homocysteine to produce methionine. Soluble Type I cytochrome b_5 is able to reduce methionine synthase using membrane-bound

cytochrome P450 reductase as an electron source (Figure 24) (Chen and Banerjee 1998). Giardia does not encode any methionine synthase and this metabolic pathway is likely to be absent in the parasite (Tekwani and Mehlotra 1999; Morrison et al. 2007). However, it could represent a "model" of the interaction occurring at the peripheral vacuolar membrane. Indeed, it has been established that the Giardia A-type flavoprotein (GL50803_10358) localizes to the peripheral vacuole. This protein contains a NAD(P)H-dependent FMN reductase domain and could be a potential electron donor to a gCYTB5. In a speculative model shown in Figure 24, an electron transport chain would be formed by the flavoprotein to gCYTB5-II and PUP9861.

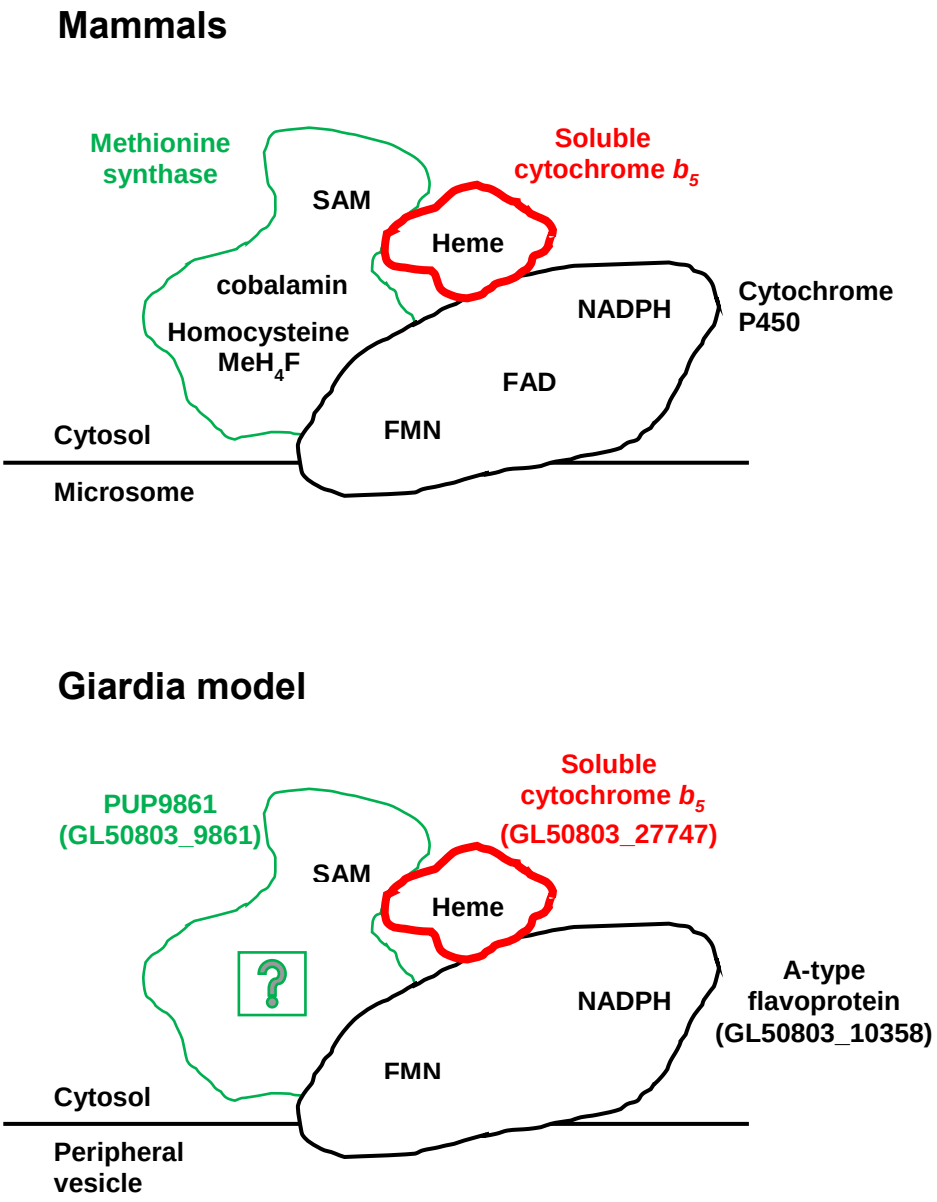


Figure 24: Hypothetical model of the gCYTB5-II interaction at the PV. Top: Reductive activation system for the mammalian methionine synthase based on Chen et al. (Chen and Banerjee 1998). Note that in the mammalian cytochrome b_5 is a Type I cytochrome b_5. MeH_4F= methyl tetrahydrofolate. Bottom: proposed model for the gCYTB5-II interaction at the PV membrane. Each protein is shown with their respective cofactor. For the PUP9861 the SAM cofactor remains to be confirmed and the function of the protein could depend on the function of the C-terminal part of the protein.

CONCLUSIONS

I used several approaches to determine the function of the three isotypes of the Giardia cytochrome b5s (gCYTB5-I, II, III). Literature searches and bioinformatics analysis of Type II cytochromes b_5 (the class that all three gCYTB5s belong to) was not useful for predicting the likely partners of these proteins. However, my bioinformatic analysis showed that Pyrih et al.'s (2014) observation that the Type II is the sole type of cytochrome b5 found in anaerobic protozoa could be extended to include amitochondriate organisms from other kingdoms. Immunofluorescence microscopy shows that each of the gCYTB5s localize to different cellular structures, which is consistent with our hypothesis that these proteins would have both common and unique partners as well as different functions. The gCYTB5-I and II proteins seem to associate with specific membrane structures (perinuclear space and peripheral vesicles), which is interesting as both isotypes co-immunoprecipitated the membrane structural protein dynamin. The results obtained also show that gCYTB5-III localizes to the nucleus. This localization is unusual for a cytochrome b_5 protein but it is interesting that another member of the type II cytochrome b_5 group, the *P. falciparum* protein PFI0885w, whose function is also unknown, is also found in the nucleus.

Although the function of PUP9861 protein is unknown, its localization in the cell and colocalization with gCYTB5-II, its potential capacity to bind nucleotide cofactor involve in redox reaction and structure similarity to proteins involved in electron transfer suggests that it is a promising interaction partner for gCYTB5-I and II. Characterizing the metabolic pathways involving gCYTB5s I and II will require determining the partners of the PUP9861 protein and the availability of the Giardia cell line that expresses HA-tagged PUP9861 can be used in future co-immunoprecipitations experiments.

In addition to the PUP9861, isotype I and II also co-immunoprecipitate the dynamin protein. Interestingly, the subcellular localization of dynamin (PV and an area in between the two nuclei which could be linked to the ER) suggests that it would be able to interact with both isotypes. The function of the dynamin in vesicle formation or reorganization of proteins at the membrane level is well characterized, although the protein is not known to interact with the canonical cytochrome b_5 protein. More work would be needed to determine if the dynamin has a function in maintaining the gCYTB5-partner interaction in the membrane neighbourhood or if the gCYTB5-I and II have roles in vesicle formation. Multiple proteins were also identified in the gCYTB5-III immunoprecipitate but only the PUP7204 seems to represent a valid candidate. Based on its amino acid sequence, the PUP7204 protein is likely to bind DNA or RNA, which would indicate that gCYTB5-III might also interact indirectly with DNA/RNA. This would represent a new function for the cytochrome b_5 protein.

The results obtained in this thesis have revealed avenues for further work to identify the functions of the Giardia cytochrome b_5 proteins. The main improvement that I suggest is to perform co-immunoprecipitation of the HA-tagged proteins in the presence of a crosslinking agent, preferably one that can be used within an intact cell. This will allow us to maintain the transient interactions between our bait and prey proteins. Finally, since gCYTB5-I and II have their expression increased after exposure to oxidative and nitrosative stressors, the experiment described could also been performed under these stress conditions.

REFERENCES

Abodeely, M., K. N. DuBois, et al. (2009). "A contiguous compartment functions as endoplasmic reticulum and endosome/lysosome in Giardia lamblia." Eukaryotic Cell **8**(11): 1665-1676.

Adam, R. D. (2001). "Biology of Giardia lamblia." Clin Microbiol Rev **14**(3): 447-475.

Aihara, Y., T. Yamamoto, et al. (2012). "Mutations in N-terminal flanking region of blue light-sensing light-oxygen and voltage 2 (LOV2) domain disrupt its repressive activity on kinase domain in the Chlamydomonas phototropin." J Biol Chem **287**(13): 9901-9909.

Alam, S., J. Yee, et al. (2012). "Cytochrome b5 from Giardia lamblia." Metallomics **4**(12): 1255-1261.

Altschul, S. F., W. Gish, et al. (1990). "Basic local alignment search tool." J Mol Biol **215**(3): 403-410.

Andersson, J. O., A. M. Sjogren, et al. (2007). "A genomic survey of the fish parasite Spironucleus salmonicida indicates genomic plasticity among diplomonads and significant lateral gene transfer in eukaryote genome evolution." BMC Genomics **8**: 51.

Andreani, J. and R. Guerois (2014). "Evolution of protein interactions: from interactomes to interfaces." Arch Biochem Biophys **554**: 65-75.

Aruni, A. W., J. Lee, et al. (2012). "VimA-dependent modulation of acetyl coenzyme A levels and lipid A biosynthesis can alter virulence in Porphyromonas gingivalis." Infect Immun **80**(2): 550-564.

Bashir, Q., S. Scanu, et al. (2011). "Dynamics in electron transfer protein complexes." FEBS J **278**(9): 1391-1400.

Benjamin, I. J. and D. R. McMillan (1998). "Stress (heat shock) proteins: molecular chaperones in cardiovascular biology and disease." Circ Res **83**(2): 117-132.

Bonetta, L. (2010). "Protein-protein interactions: Interactome under construction." Nature **468**(7325): 851-854.

Campanaro, K. (2013). Seeking the Redox Partners of *Giardia lamblia* cytochrome b5. Peterborough, Ontario, Trent University.

Candiano, G., M. Bruschi, et al. (2004). "Blue silver: a very sensitive colloidal Coomassie G-250 staining for proteome analysis." Electrophoresis **25**(9): 1327-1333.

Cassola, A., G. Noe, et al. (2010). "RNA recognition motifs involved in nuclear import of RNA-binding proteins." RNA Biol **7**(3): 339-344.

Chang, I. F. (2006). "Mass spectrometry-based proteomic analysis of the epitope-tag affinity purified protein complexes in eukaryotes." Proteomics **6**(23): 6158-6166.

Chen, Z. and R. Banerjee (1998). "Purification of soluble cytochrome b5 as a component of the reductive activation of porcine methionine synthase." J Biol Chem **273**(40): 26248-26255.

Cusick, M. E., N. Klitgord, et al. (2005). "Interactome: gateway into systems biology." Hum Mol Genet **14 Spec No. 2**: R171-181.

Dastidar, E. G., G. Dayer, et al. (2012). "Involvement of Plasmodium falciparum protein kinase CK2 in the chromatin assembly pathway." BMC Biol **10**: 5.

De Camilli, P., K. Takei, et al. (1995). "The function of dynamin in endocytosis." Curr Opin Neurobiol **5**(5): 559-565.

De Las Rivas, J. and C. Fontanillo (2010). "Protein-protein interactions essentials: key concepts to building and analyzing interactome networks." PLoS Comput Biol **6**(6): e1000807.

Dorin-Semblat, D., C. Demarta-Gatsi, et al. (2015). "Malaria Parasite-Infected Erythrocytes Secrete PfCK1, the Plasmodium Homologue of the Pleiotropic Protein Kinase Casein Kinase 1." PLoS One **10**(12): e0139591.

Doroshchuk, A. D. and L. F. Dmitriev (2013). "Changes in the redox state of cytochrome b5 in the outer mitochondrial membrane as a result of interaction with lipid intermediates: Role of cytochrome c." Dokl Biochem Biophys **453**: 292-296.

Dwane, S. and P. A. Kiely (2011). "Tools used to study how protein complexes are assembled in signaling cascades." Bioeng Bugs **2**(5): 247-259.

Epstein, P. M., M. Curti, et al. (1989). "Phosphorylation of cytochrome P450: regulation by cytochrome b5." Arch Biochem Biophys **271**(2): 424-432.

Faso, C., S. Bischof, et al. (2013). "The proteome landscape of Giardia lamblia encystation." PLoS One **8**(12): e83207.

Ferguson, S. M. and P. De Camilli (2012). "Dynamin, a membrane-remodelling GTPase." Nat Rev Mol Cell Biol **13**(2): 75-88.

Figeys, D., L. D. McBroom, et al. (2001). "Mass spectrometry for the study of protein-protein interactions." Methods **24**(3): 230-239.

Franceschini, A., D. Szklarczyk, et al. (2013). "STRING v9.1: protein-protein interaction networks, with increased coverage and integration." Nucleic Acids Res **41**(Database issue): D808-815.

Free, R. B., L. A. Hazelwood, et al. (2009). "Identifying novel protein-protein interactions using co-immunoprecipitation and mass spectroscopy." Curr Protoc Neurosci **Chapter 5**: Unit 5 28.

Gillin, F. D., D. S. Reiner, et al. (1996). "Cell biology of the primitive eukaryote Giardia lamblia." Annu Rev Microbiol **50**: 679-705.

Hagen, K. D., M. P. Hirakawa, et al. (2011). "Novel structural components of the ventral disc and lateral crest in Giardia intestinalis." PLoS Negl Trop Dis **5**(12): e1442.

Hernandez, Y., C. Castillo, 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**(1): 21-32.

Hoetelmans, R. W., F. A. Prins, et al. (2001). "Effects of acetone, methanol, or paraformaldehyde on cellular structure, visualized by reflection contrast

microscopy and transmission and scanning electron microscopy." Appl Immunohistochem Mol Morphol **9**(4): 346-351.

Hsieh, C. H., S. Y. Huang, et al. (2007). "Expression of proteins with dimethylarginines in Escherichia coli for protein-protein interaction studies." Protein Sci **16**(5): 919-928.

Illergard, K., D. H. Ardell, et al. (2009). "Structure is three to ten times more conserved than sequence-A study of structural response in protein cores." Proteins-Structure Function and Bioinformatics **77**(3): 499-508.

Jarroll, E. L., P. Manning, et al. (1989). "Biochemistry and metabolism of Giardia." J Protozool **36**(2): 190-197.

Jedelsky, P. L., P. Dolezal, et al. (2011). "The minimal proteome in the reduced mitochondrion of the parasitic protist Giardia intestinalis." PLoS One **6**(2): e17285.

Kallberg, M., G. Margaryan, et al. (2014). "RaptorX server: a resource for template-based protein structure modeling." Methods Mol Biol **1137**: 17-27.

Kallberg, M., H. Wang, et al. (2012). "Template-based protein structure modeling using the RaptorX web server." Nat Protoc **7**(8): 1511-1522.

Kelley, L. A., S. Mezulis, et al. (2015). "The Phyre2 web portal for protein modeling, prediction and analysis." Nat Protoc **10**(6): 845-858.

Kita, K., H. Hirawake, et al. (2002). "Role of complex II in anaerobic respiration of the parasite mitochondria from Ascaris suum and Plasmodium falciparum." Biochim Biophys Acta **1553**(1-2): 123-139.

Kitatsuji, C., K. Izumi, et al. (2016). "Protein oxidation mediated by heme-induced active site conversion specific for heme-regulated transcription factor, iron response regulator." Sci Rep **6**: 18703.

Krieger, E., S. B. Nabuurs, et al. (2003). "Homology modeling." Methods Biochem Anal **44**: 509-523.

Lanfredi-Rangel, A., M. Attias, et al. (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**(3): 225-235.

Lanfredi-Rangel, A., M. Attias, et al. (2003). "Fine structure of the biogenesis of Giardia lamblia encystation secretory vesicles." J Struct Biol **143**(2): 153-163.

Lanfredi-Rangel, A., W. M. Kattenbach, et al. (1999). "Trophozoites of Giardia lamblia may have a Golgi-like structure." Fems Microbiology Letters **181**(2): 245-251.

Lederer, F. (1994). "The cytochrome b5-fold: an adaptable module." Biochimie **76**(7): 674-692.

Lin, Q., S. Weis, et al. (2007). "Heme oxygenase-1 protein localizes to the nucleus and activates transcription factors important in oxidative stress." J Biol Chem **282**(28): 20621-20633.

Manning, G., D. S. Reiner, et al. (2011). "The minimal kinome of Giardia lamblia illuminates early kinase evolution and unique parasite biology." Genome Biol **12**(7): R66.

Maris, C., C. Dominguez, et al. (2005). "The RNA recognition motif, a plastic RNA-binding platform to regulate post-transcriptional gene expression." FEBS J **272**(9): 2118-2131.

Markillie, L. M., C. T. Lin, et al. (2005). "Simple protein complex purification and identification method for high-throughput mapping of protein interaction networks." J Proteome Res **4**(2): 268-274.

Marti-Renom, M. A., A. C. Stuart, et al. (2000). "Comparative protein structure modeling of genes and genomes." Annu Rev Biophys Biomol Struct **29**: 291-325.

McGuffin, L. J., J. D. Atkins, et al. (2015). "IntFOLD: an integrated server for modelling protein structures and functions from amino acid sequences." Nucleic Acids Res **43**(W1): W169-173.

Medina, D., N. Moskowitz, et al. (2000). "Rapid purification of protein complexes from mammalian cells." Nucleic Acids Res **28**(12): E61.

Mellacheruvu, D., Z. Wright, et al. (2013). "The CRAPome: a contaminant repository for affinity purification-mass spectrometry data." Nat Methods **10**(8): 730-736.

Meyer, K. and M. Selbach (2015). "Quantitative affinity purification mass spectrometry: a versatile technology to study protein-protein interactions." Front Genet **6**: 237.

Meyer, T. E., K. Shirabe, et al. (1995). "Transient kinetics of intracomplex electron transfer in the human cytochrome b5 reductase-cytochrome b5 system: NAD+ modulates protein-protein binding and electron transfer." Arch Biochem Biophys **318**(2): 457-464.

Miaczynska, M. and H. Stenmark (2008). "Mechanisms and functions of endocytosis." J Cell Biol **180**(1): 7-11.

Miernyk, J. A. and J. J. Thelen (2008). "Biochemical approaches for discovering protein-protein interactions." Plant J **53**(4): 597-609.

Migneault, I., C. Dartiguenave, et al. (2004). "Glutaraldehyde: behavior in aqueous solution, reaction with proteins, and application to enzyme crosslinking." Biotechniques **37**(5): 790-796, 798-802.

Montessuit, S., S. P. Somasekharan, et al. (2010). "Membrane remodeling induced by the dynamin-related protein Drp1 stimulates Bax oligomerization." Cell **142**(6): 889-901.

Morrison, H. G., A. G. McArthur, et al. (2007). "Genomic minimalism in the early diverging intestinal parasite Giardia lamblia." Science **317**(5846): 1921-1926.

Mortenson, L. E., R. C. Valentine, et al. (1962). "An electron transport factor from Clostridium pasteurianum." Biochem Biophys Res Commun **7**: 448-452.

Mount, D. W. (2007). "Using the Basic Local Alignment Search Tool (BLAST)." CSH Protoc **2007**: pdb top17.

Muller, J., S. Rout, et al. (2015). "Comparative characterisation of two nitroreductases from Giardia lamblia as potential activators of nitro compounds." International Journal for Parasitology-Drugs and Drug Resistance **5**(2): 37-43.

Muller, M., M. Mentel, et al. (2012). "Biochemistry and evolution of anaerobic energy metabolism in eukaryotes." Microbiol Mol Biol Rev **76**(2): 444-495.

Murphy, M. J., L. M. Siegel, et al. (1974). "Siroheme: a new prosthetic group participating in six-electron reduction reactions catalyzed by both sulfite and nitrite reductases." Proc Natl Acad Sci U S A **71**(3): 612-616.

Nakano, K., T. Yamamoto, et al. (2008). "Protein kinases Fpk1p and Fpk2p are novel regulators of phospholipid asymmetry." Mol Biol Cell **19**(4): 1783-1797.

Nakano, S., M. Takahashi, et al. (2012). "Structure-function relationship of assimilatory nitrite reductases from the leaf and root of tobacco based on high-resolution structures." Protein Sci **21**(3): 383-395.

Nebert, D. W. and F. J. Gonzalez (1987). "P450 genes: structure, evolution, and regulation." Annu Rev Biochem **56**: 945-993.

Nzengue, Y., E. Lefebvre, et al. (2009). "Metallothionein expression in HaCaT and C6 cell lines exposed to cadmium." Journal of Trace Elements in Medicine and Biology **23**(4): 314-323.

Oehring, S. C., B. J. Woodcroft, et al. (2012). "Organellar proteomics reveals hundreds of novel nuclear proteins in the malaria parasite Plasmodium falciparum." Genome Biol **13**(11): R108.

Ogura, Y., S. Tokutomi, et al. (2008). "PAS/LOV proteins: A proposed new class of plant blue light receptor." Plant Signal Behav **3**(11): 966-968.

Pazdzior, R., Z. A. Yang, et al. (2015). "Low reduction potential cytochrome b5 isotypes of Giardia intestinalis." Exp Parasitol **157**: 197-201.

Pereira, C. M., E. Sattlegger, et al. (2005). "IMPACT, a protein preferentially expressed in the mouse brain, binds GCN1 and inhibits GCN2 activation." J Biol Chem **280**(31): 28316-28323.

Pyrih, J., K. Harant, et al. (2014). "Giardia intestinalis incorporates heme into cytosolic cytochrome b(5)." Eukaryotic Cell **13**(2): 231-239.

Rafferty, S. P. and G. Dayer (2015). "Heme proteins of Giardia intestinalis." Exp Parasitol **159**: 13-23.

Raj, D., E. Ghosh, et al. (2014). "Differential gene expression in Giardia lamblia under oxidative stress: significance in eukaryotic evolution." Gene **535**(2): 131-139.

Salemme, F. R. (1976). "An hypothetical structure for an intermolecular electron transfer complex of cytochromes c and b5." J Mol Biol **102**(3): 563-568.

Samaj, J., F. Baluska, et al. (2004). "Endocytosis, actin cytoskeleton, and signaling." Plant Physiol **135**(3): 1150-1161.

Schenkman, J. B. and I. Jansson (2003). "The many roles of cytochrome b5." Pharmacol Ther **97**(2): 139-152.

Schittek, B. and T. Sinnberg (2014). "Biological functions of casein kinase 1 isoforms and putative roles in tumorigenesis." Mol Cancer **13**: 231.

Schwede, T., J. Kopp, et al. (2003). "SWISS-MODEL: An automated protein homology-modeling server." Nucleic Acids Res **31**(13): 3381-3385.

Shen, G., S. Kuppu, et al. (2010). "ANKYRIN REPEAT-CONTAINING PROTEIN 2A is an essential molecular chaperone for peroxisomal membrane-bound ASCORBATE PEROXIDASE3 in Arabidopsis." Plant Cell **22**(3): 811-831.

Singleton, C., G. F. White, et al. (2010). "Heme-responsive DNA binding by the global iron regulator Irr from Rhizobium leguminosarum." J Biol Chem **285**(21): 16023-16031.

Smits, A. H. and M. Vermeulen (2016). "Characterizing Protein-Protein Interactions Using Mass Spectrometry: Challenges and Opportunities." Trends in Biotechnology **34**(10): 825-834.

Soler, N., C. T. Craescu, et al. (2012). "A S-adenosylmethionine methyltransferase-like domain within the essential, Fe-S-containing yeast protein Dre2." FEBS J **279**(12): 2108-2119.

Sparacino-Watkins, C. E., J. Tejero, et al. (2014). "Nitrite reductase and nitric-oxide synthase activity of the mitochondrial molybdopterin enzymes mARC1 and mARC2." J Biol Chem **289**(15): 10345-10358.

Stechmann, A., K. Hamblin, et al. (2008). "Organelles in Blastocystis that blur the distinction between mitochondria and hydrogenosomes." Curr Biol **18**(8): 580-585.

Sullivan, P. B., M. N. Marsh, et al. (1991). "Prevalence and treatment of giardiasis in chronic diarrhoea and malnutrition." Arch Dis Child **66**(3): 304-306.

Taylor, B. L. and I. B. Zhulin (1999). "PAS domains: internal sensors of oxygen, redox potential, and light." Microbiol Mol Biol Rev **63**(2): 479-506.

Tekwani, B. L. and R. K. Mehlotra (1999). "Molecular basis of defence against oxidative stress in Entamoeba histolytica and Giardia lamblia." Microbes Infect **1**(5): 385-394.

Terashima, M., M. Specht, et al. (2010). "Characterizing the anaerobic response of Chlamydomonas reinhardtii by quantitative proteomics." Mol Cell Proteomics **9**(7): 1514-1532.

Thompson, R. C. and P. Monis (2012). "Giardia--from genome to proteome." Adv Parasitol **78**: 57-95.

Tonon, T., O. Sayanova, et al. (2005). "Fatty acid desaturases from the microalga Thalassiosira pseudonana." FEBS J **272**(13): 3401-3412.

Tovar, J., G. Leon-Avila, et al. (2003). "Mitochondrial remnant organelles of Giardia function in iron-sulphur protein maturation." Nature **426**(6963): 172-176.

Treeck, M., J. L. Sanders, et al. (2011). "The phosphoproteomes of Plasmodium falciparum and Toxoplasma gondii reveal unusual adaptations within and beyond the parasites' boundaries." Cell Host Microbe **10**(4): 410-419.

Uttaro, A. D. (2006). "Biosynthesis of polyunsaturated fatty acids in lower eukaryotes." IUBMB Life **58**(10): 563-571.

Valentine, R. C. (1964). "Bacterial Ferredoxin." Bacteriol Rev **28**: 497-517.

Vergeres, G. and L. Waskell (1995). "Cytochrome b5, its functions, structure and membrane topology." Biochimie **77**(7-8): 604-620.

Wang, J., C. Dupuis, et al. (2016). "Sterile alpha Motif Domain Containing 9 Is a Novel Cellular Interacting Partner to Low-Risk Type Human Papillomavirus E6 Proteins." PLoS One **11**(2): e0149859.

Wang, R. and H. Li (2012). "The mysterious RAMP proteins and their roles in small RNA-based immunity." Protein Sci **21**(4): 463-470.

Wolfe, M. S. (1992). "Giardiasis." Clin Microbiol Rev **5**(1): 93-100.

Worrall, J. A., Y. Liu, et al. (2002). "Myoglobin and cytochrome b5: a nuclear magnetic resonance study of a highly dynamic protein complex." Biochemistry **41**(39): 11721-11730.

Zamponi, N., C. Feliziani, et al. (2016). "Endocytosis in Giardia: Evidence of Absence." Trends Parasitol **32**(11): 838-840.

Zhang, Y. (2008). "I-TASSER server for protein 3D structure prediction." BMC Bioinformatics **9**: 40.

Zumthor, J. P., L. Cernikova, et al. (2016). "Static Clathrin Assemblies at the Peripheral Vacuole-Plasma Membrane Interface of the Parasitic Protozoan Giardia lamblia." PLoS Pathog **12**(7): e1005756.

APPENDIX A. GIARDIA CULTURE MEDIA

Materials for Giardia 5x basic media

Component	**Amount (/L)**
Casein Peptone Digest (N-Z Case Plus)	100 g
Glucose	50 g
Yeast Extract	50 g
NaCl	10 g
0.5 M Phosphate buffer solution pH 7.2*	100 mL
Millipore Water	To 1 L

* 0.5 M phosphate buffer pH 6.8-7.2: dissolve 15 g KH_2PO_4 (monobasic) and 25 g of K_2HPO_4 (dibasic) into 500 mL of distilled water. Ensure that the pH is in the correct range and autoclave the buffer.

Material for Giardia unsupplemented media

Component	Volume (/L)
Giardia 5x Basic media	200 mL
Bovine serum	100 mL
$FeNH_4$ citrate**	10 mL
6.5% Bovine bile***	8 mL
Millipore water	To 1 L

** $FeNH_4$ citrate: dissolve 114 mg of powder in 50 mL distilled water. Filter sterilized with 0.22 mm syringe filter and stored protected from light at room temperature.

*** 6.5% Bovine bile: dissolved 3.25 g of powder (Sigma B8381) in 50 mL of distilled water. Filter sterilized as before and store at 4^{o}C.

The unsupplemented media was pH adjusted to 7.2 and sterilised using a 0.2 µm PES filter unit inside a Class 2, A2 biological safety cabinet and store at 4^{o}C.

Giardia complete media

For cell culture the unsupplemented media was completed with 320 µL of 50x Cysteine and Ascorbic acid (2 g L-cysteine and 0.1 g ascorbic acid diluted in 20 mL Millipore water and adjusted to 7.2, store in aliquots at -80^{o}C) and 40 µL (10’000units/mL penicillin, 10’000 µg/mL streptomycin and 25 g/mL amphotericin B or fungizone, store at – 20^{o}C).

APPENDIX B. IMMUNOFLUORESCENCE MICROSCOPY

Dr. Yee's laboratory protocol:

Solutions:

a) Absolute methanol- stored at RT and used in -20°C freezer

b) 0.1% polyethylenimine (PEI) in PBS

c) 0.5% Triton X-100 in PBS

d) Immunofluorescence blocking buffer (50 mM Tris-HCl pH 6.8, 150 mM NaCl, 0.5% NP-40, 5 mg/mL BSA, in dH_2O)

e) Cyst wall protein 1 Antibody (stored in 4°C fridge- covered in tinfoil)

f) 3.7% paraformaldehyde in PBS

g) Vectashield mounting media with DAPI

1. Pre-treat coverslips by applying 100 µL of 0.1% polyethylenimine in PBS for 5-10 min then rinse well with dH_2O, allow to air dry.

2. Pre-warm coverslips in a Tupperware container in the 37 °C incubator (10 min) on a pre-wet paper towel with the PEI treated side up.

3. Place 100 µL of cells onto the coverslip ensuring that the entire surface is covered.

4. Close the lid on the container and place in 37°C incubator for 10 minutes.

5. Remove from incubator and place coverslip in coverslip holding rack, cell side facing the notch on the rack. Place rack into a pre-chilled beaker of methanol and place into the -20°C freezer for 10 minutes.

6. Place methanol treated coverslips on a piece of paper towel to air dry for 5 min. (Can stop here and continue staining the next day if doing time point experiments).

Perform remaining steps in a tray lined with parafilm.

7. Permeabilize cells by floating coverslips cell side down on 100 µL of 0.5% Triton X-100 (diluted in PBS) for 10 min.

8. Block for 1 hour at room temperature by floating coverslips on 200 µL of blocking buffer.

9. Incubate the cells on 100 µL of CWP1 antibody (1:20 dilution in blocking buffer) for 1-2 hours in the 37°C incubator.

10. Wash 4 times by placing slides on 200 µL of 1X PBS (5 min each wash).

11. Postfix in 100 µL 3.7% paraformaldehyde (diluted in 1X PBS) for 10 min.

12. Rinse 2 times in 1 X PBS (200 µL) and a final time in dH_2O only.

13. Place a small drop of mounting medium (Vectashield with DAPI) on microscope slide and put coverslip cell side down.

14. Allow to air dry.

15. Seal coverslip to slide with clear nail polish and store overnight to 24 h in a dark place at room temperature. Store for short term (1-3 days) in 4°C fridge.

Dr. Hehl laboratory protocol:

Solutions:

a) 3% paraformaldehyde in PBS
b) 0.1 M glycine in PBS
c) 2% bovine serum albumin (BSA) and 0.2% Triton X-100 in PBS
d) 2% BSA in PBS
e) 1% BSA and 0.1% Triton X-100 in PBS

1. Incubate 10^6-10^7 cells in 1 mL fixation solution (3% paraformaldehyde in PBS) for at least 30 minutes at room temperature or at 4°C overnight in a microcentrifuge tube.

2. Centrifuge at 900 g for 2-3 minutes at 4°C, remove supernatant and wash once with PBS.

3. Centrifuge and remove supernatant, add 1 mL of 0.1 M glycine in PBS for 5 minutes.

4. Centrifuge and remove supernatant. Incubate in 1 mL of 2% bovine serum albumin and 0.2% Triton X-100 in PBS solution freshly prepared for 20 minutes.

5. Centrifuge and remove supernatant. Incubate for at least 2 hours at 4°C in 1 mL of 2% BSA in PBS.

6. Prepared primary antibody solution. Dilute primary antibody in 2% BSA and 0.2% Triton X-100 in PBS. Spin down at maximal speed for 15 minutes at 4°C, use the supernatant.

7. Centrifuge and discard supernatant. Resuspend the cells in 100 µL of primary antibody solution and incubate 30 minutes to 1 hour at room temperature or 4°C overnight.

8. Spin down the cells and discard supernatant. Wash twice (5 minutes) in 1% BSA and 0.1% Triton X-100 in PBS.

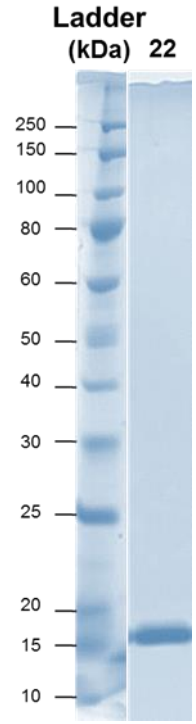
9. Incubate the cells in 100 µL of secondary antibody solution prepared as described for the primary antibody solution.

10. Centrifuge and discard supernatant. Wash twice in 1% BSA and 0.1% Triton X-100 in PBS as before.

11. Centrifuge and carefully remove all the buffer.

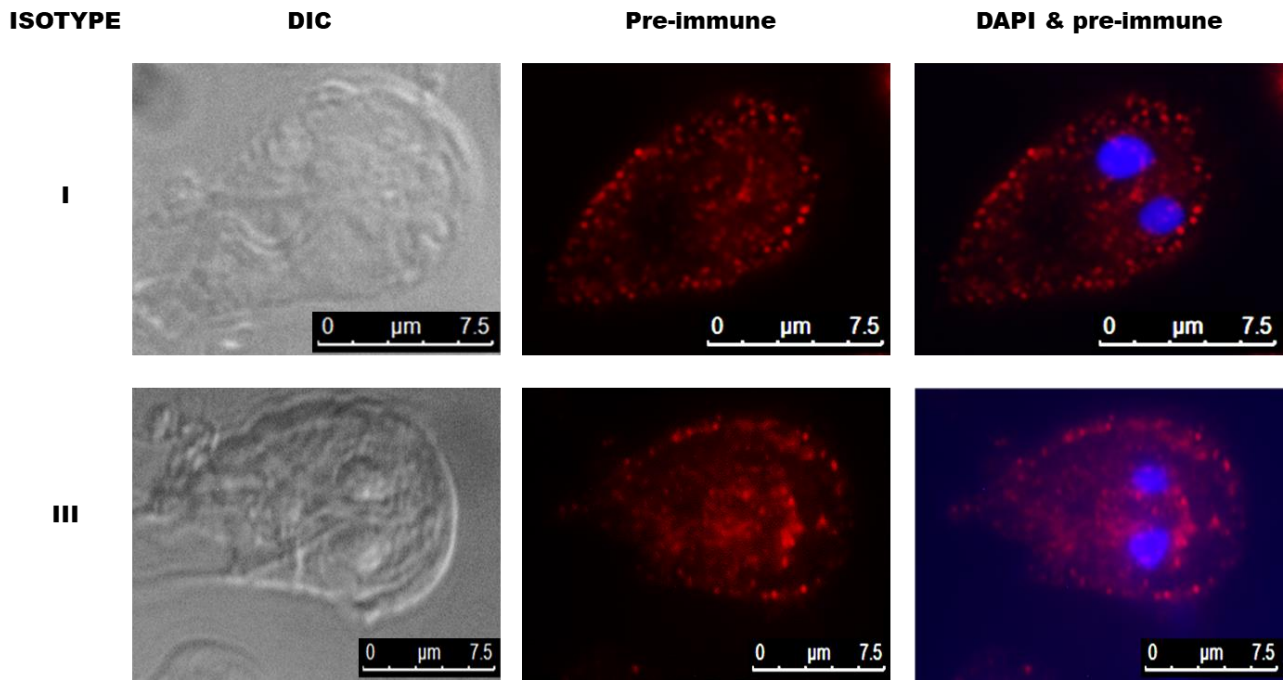
12. Gently resuspend cells in 10-30 µL of Vectashield with DAPI depending on the size of the cell pellet. Pipet up and down to resuspend the cells and let sit on ice for 10 minutes. Use 3-4 µL of this suspension on a slide for a 22 x 22 mm coverslip. Remove excess Vectashield by gently pressing the coverslip with a paper tissue. Fix the coverslip with clear nail polish.

APPENDIX C. RECOMBINANT GCYTB5-I PURITY



The protein purity was assessed by loading the different his_6-gGYTb5-I eluted fraction on a SDS-PAGE. After electrophoreses the gel was stained using PageBlue Protein Staining Solution (Thermo Scientific, sensitivity: 5 ng). Ladder is shown on the left alongside the fraction 22 showing the tagged protein (expected molecular weight 17.5 kDa).

APPENDIX D. IFM USING ISOTYPES I AND II PRE-IMMUNE SERUME



Immunofluorescence assay using untransfected cell line and pre-immune serum. The cells were incubated overnight with the pre-immune serum for isotype I (top) and III (bottom) (dilution 1:200) in order to detect a signal corresponding to background fluorescence. In comparison, the anti-gCYTB5-I antibody was used with a 1:300 dilution for 2 hours and the anti-gCYTB5-III antibody dilution was 1:2000 for 2 hours.

APPENDIX E. BLUE SILVER STAINING

Buffer composition: 0.12% Coomassie Blue G-250, 10% ammonium sulfate, 10% phosphoric acid and 20% methanol

For 1 L of Staining Solution:

· To 100 ml water add phosphoric acid (enough to obtain 10% in the final 1L).

· Add 100 g ammonium sulphate.

· Add 1.2g Coomassie Blue G-250.

· Add water to 800 mL.

· Add 200 mL of 100% methanol.

1. Fix gel in 50% ethanol and 2% phosphoric acid , 2 times 20 minutes.

2. Wash the gel for 2x 20 minutes in ddH2O.

3. Add staining solution and stain overnight or longer.

4. Rinse the gel with ddH2O and store in ddH2O at 4°C until further use.

APPENDIX F. SEQUENCING RESULT OF THE PUP9861 INSERT

```
CLUSTAL O(1.2.3) multiple sequence alignment

query        TAGTATGtATCCTTATGACGTGCCTGACtaTGCCTATCCTTATGACGTGCCTGACTATGC   60
PUP9861      ------------------------------------------------------------   0
query        CCATATGCCGATCATTATTAAGGGTGTGAAGCTTACAGAGGAGTTCACAAGGACCCTGAA   120
PUP9861      ----ATGCCGATCATTATTAAGGGTGTGAAGCTTACAGAGGAGTTCACAAGGACCCTGAA   56
                 ************************************************************
query        GGAGATCATCAAAGGTGTCGACGGCTTTCGCTCGCCGTACCTAGCCAACAAGGCCCGTGA   180
PUP9861      GGAGATCATCAAAGGTGTCGACGGCTTTCGCTCGCCGTACCTAGCCAACAAGGCCCGTGA   116
             ************************************************************
query        AACTGACAAGGGCGTCGAAATACCTTGTGGCTATGGGAAGCTGCTTGCCGCGGTTAGGCA   240
PUP9861      AACTGACAAGGGCGTCGAAATACCTTGTGGCTATGGGAAGCTGCTTGCCGCGGTTAGGCA   176
             ************************************************************
query        GAAGTACCCAGAAGCTGATCTGCTTCCCCTCGACCAGGATGAAGCACGCAAGAACTGCAT   300
PUP9861      GAAGTACCCAGAAGCTGATCTGCTTCCCCTCGACCAGGATGAAGCACGCAAGAACTGCAT   236
             ************************************************************
query        CGACCGCAATGTTTCCATTTTTACTAAGTTCAAGGAGTACAATGCAGACTACCTCAAGAC   360
PUP9861      CGACCGCAATGTTTCCATTTTTACTAAGTTCAAGGAGTACAATGCAGACTACCTCAAGAC   296
             ************************************************************
query        AAGCAAAAATGTCATCATCACGGTTACCTCCGCACCAGGAGAGTTCATGGACCCGGAGAA   420
PUP9861      AAGCAAAAATGTCATCATCACGGTTACCTCCGCACCAGGAGAGTTCATGGACCCGGAGAA   356
             ************************************************************
query        GGTCGTCGCCTTTTTGACTAAGACCCTCTCGAAGATCACGCCAGCTCAGGATCACAAGTT   480
PUP9861      GGTCGTCGCCTTTTTGACTAAGACCCTCTCGAAGATCACGCCAGCTCAGGATCACAAGTT   416
             ************************************************************
query        CAAGATCACCGCGAAGTATCAAGTGCTGGATGCTATCCGTGGAAGGATCAATGCGGGTGC   540
PUP9861      CAAGATCACCGCGAAGTATCAAGTGCTGGATGCTATCCGTGGAAGGATCAATGCGGGTGC   476
             ************************************************************
query        TCTCAAGGATGCCAGCGGCTCTTCTGTGCAAGGCAGGTGCGCGACAATGGGGATTAGATA   600
PUP9861      TCTCAAGGATGCCAGCGGCTCTTCTGTGCAAGGCAGGTGCGCGACAATGGGGATTAGATA   536
             ************************************************************
query        TGTTGAAGGTCGTCCGCCGCAGAGGGGGACGAAGGTCGCCGTTGTTGTTCAGTCTTCTGA   660
PUP9861      TGTTGAAGGTCGTCCGCCGCAGAGGGGGACGAAGGTCGCCGTTGTTGTTCAGTCTTCTGA   596
             ************************************************************
query        GATAGAGAAGCTTTTTGCTGCTCTCTCTGCCGATAATCTTTTTGAACGCGTTCCCGCCAT   720
PUP9861      GATAGAGAAGCTTTTTGCTGCTCTCTCTGCCGATAATCTTTTTGAACGCGTTCCCGCCAT   656
```

```
            ************************************************************
query       AAGGTTCATTAACCCGGAGAAGCGCCGTAAGCATATCCAGGAGCTTCGCGAGAGGGTTGG   780
PUP9861     AAGGTTCATTAACCCGGAGAAGCGCCGTAAGCATATCCAGGAGCTTCGCGAGAGGGTTGG   716
            ************************************************************
query       GTCTGCAGGCGATGGTACCCTGGGCACAAAAAATAAGCAGAGGAGACACCAAGAGGCGAA   840
PUP9861     GTCTGCAGGCGATGGTACCCTGGGCACAAAAAATAAGCAGAGGAGACACCAAGAGGCGAA   776
            ************************************************************
query       GGGTGGTGTGAAGAAGCCTGGAGCCAAGAAGGCCACCGGAATTAGAGTACAGCTTGCTGA   900
PUP9861     GGGTGGTGTGAAGAAGCCTGGAGCCAAGAAGGCCACCGGAATTAGAGTACAGCTTGCTGA   836
            ************************************************************
query       TGACCAGAAGGCCGCCAAGCCTAAGAAGAAGCAGGGCAAGAAGGTGCAATCTCAGAAGGA   960
PUP9861     TGACCAGAAGGCCGCCAAGCCTAAGAAGAAGCAGGGCAAGAAGGTGCAATCTCAGAAGGA   896
            ************************************************************
query       AAGGCTCCCATGTCTCCTTACGATCGCAGGGATTCCTGAGGCTCTAGCATTCGACGATAT   1020
PUP9861     AAGGCTCCCATGTCTCCTTACGATCGCAGGGATTCCTGAGGCTCTAGCATTCGACGATAT   956
            ************************************************************
query       CAAGGAAAACCTCGATAAAGATGAGCACGCAGACATTCTAAAAGCTCTTGCAGAGTCTCG   1080
PUP9861     CAAGGAAAACCTCGATAAAGATGAGCACGCAGACATTCTAAAAGCTCTTGCAGAGTCTCG   1016
            ************************************************************
query       CTTGAGGAGACCCAAGCAGCCCAATCCTTCCGAAGTGTCGTTCTACTGCACTGTGGAGAA   1140
PUP9861     CTTGAGGAGACCCAAGCAGCCCAATCCTTCCGAAGTGTCGTTCTACTGCACTGTGGAGAA   1076
            ************************************************************
query       TGGCAAGATCCTTAGGGATGCCTTCGGTAACATGGAGATTAACGGGGCCGAGCTCCGCAC   1200
PUP9861     TGGCAAGATCCTTAGGGATGCCTTCGGTAACATGGAGATTAACGGGGCCGAGCTCCGCAC   1136
            ************************************************************
query       TACGGTGTCGGATGTGAACTAAGGGCCCGATCCACTAGTTCTAGAGCGGCCGCCAC       1256
PUP9861     TACGGTGTCGGATGTGAACTAA---------------------------------       1158
            **********************
```

The result of the plasmid insert sequencing (query) was compared to the PUP9861 gene sequence retrieved from the database www.giardiadb.org. The stars indicate identity.

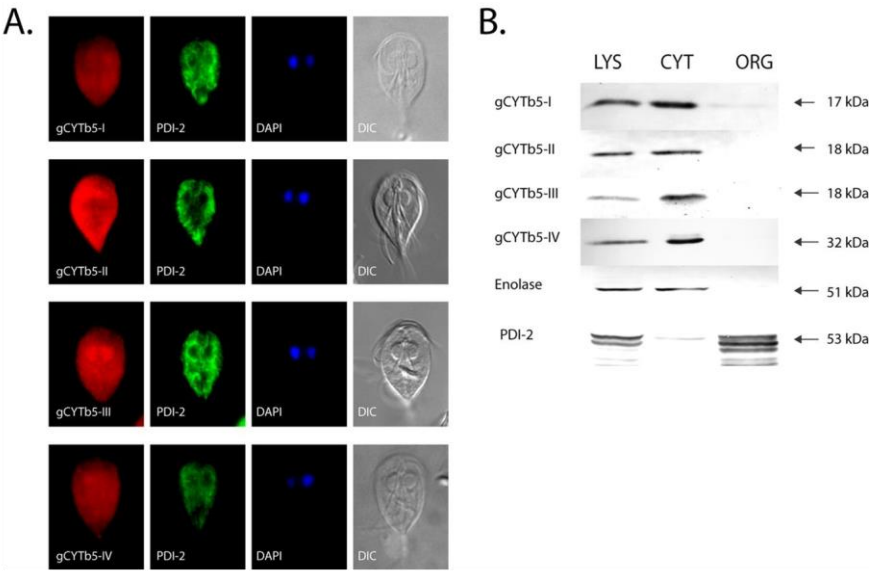
APPENDIX G. DISTRIBUTION OF CYTOCHROMES B_5 FOR ALL THE SPECIES INVESTIGATED.

Green boxes show the phylum and the blue boxes correspond to anaerobic organisms.

Organisms	Type I	Short Type II	Long Type II	Other
Apicomplexa				
Babesia bovis	0	1	0	0
Cryptosporidium hominis	0	0	0	0
Cryptosporidium parvum	0	0	0	1
Cryptosporidium muris	0	0	1	0
Eimeria tenella	0	0	0	4
Neospora caninum	1	0	2	4
Plasmodium berghei	2	1	0	3
Plasmodium chabaudi	1	0	0	2
Plasmodium falciparum	1	1	1	0
Plasmodium knowlesi	1	1	1	0
Plasmodium vivax	1	1	1	0
Plasmodium yoelii yoelii	1	1	1	0
Theileria annulata	0	0	0	0
Theileria parva	0	0	0	0
Toxoplasma gondii	2	0	2	7
Gregarina niphandrodes	0	0	1	0
Theileria equi strain WA	0	0	0	2
Ciliate				
Tetrahymena thermophila	3	1	2	11
Amoeboza				
Acanthamoeba castellanii	0	0	0	3
Dictyostelium discoideum	2	3	2	6
Entamoeba histolytica	0	0	0	0
Polysphondylium pallidum	2	2	0	2
Pelomyxa	0	0	0	0
Heterokontophyta				
Albugo laibachii	1	0	1	1
Aureococcus anophagefferens	1	0	0	17
Fragilariopsis cylindrus	0	0	0	0
Hyaloperonospora arabidopsidis	1	0	1	5
Nannochloropis gaditana	0	0	0	3
Phaeodactylum tricornutum	3	1	0	6
Phytophthora infestans	1	0	0	3
Phytophthora ramorum	1	0	0	7
Phytophthora sojae	1	0	0	9
Pseudo-nitzschia multiseries	0	0	0	0
Pythium ultimum	2	0	1	7
Thalassiosira pseudonana	0	1	0	7
Unranked				
Bigelowiella natans	0	0	0	0
Cryptophyta				
Chroomonas mesostigmatica	0	0	0	0
Cryptomonas paramecium	0	0	0	0
Guillardia theta	1	0	0	13
Hemiselmis andersenii	0	0	0	0
Haptophyta				
Emiliania huxleyi	3	0	1	25
Euglenozoa/Kinetoplastida				
Leishmania braziliensis	5	2	1	2
Leishmania infantum	2	2	0	4

Organisms	Type I	Short Type II	Long Type II	Other
Leishmania major	6	1	0	3
Trypanosoma brucei	3	1	0	1
Trypanosoma cruzi	6	2	2	2
Euglena	0	0	0	0
Crithidia fasciculata	3	1	0	0
Percolozoa				
Naegleria gruberi	2	1	0	4
Metamonada				
Giardia intestinalis	0	3	1	1
Spironucleus salmonicida	0	5	0	2
Trichomonas vaginalis	0	5	0	1
Trichomonas tenax	0	0	0	0
Tritrichomonas foetus	0	0	0	0
Hexamita	0	0	0	0
Trepomonas agilis	0	0	0	0
Unranked				
Monosiga brevicollis	1	0	0	3
Microsporidia				
Trachipleistophora Hominis	0	1	0	0
Encephalitozoon Cuniculi	0	1	0	0
Anncaliia algerae	0	1	0	0
Edhazardia aedis	0	1	0	0
Enterocytozoon bieneusi	0	0	0	0
Nematocida	0	1	0	0
Nosema ceranae	0	0	0	0
Spraguea lophii	0	0	0	1
Vavraia culicis floridensis	0	1	0	0
Vittaforma corneae	0	0	0	0
Unranked				
Dasytricha ruminantium	0	0	0	0
Isotricha	0	0	0	0
Fasciola hepatica	0	0	0	0
Nematoda				
Ascaris suum	3	1	0	0
Caenorhabditis elegans	1	1	0	3
Mollusca				
Mytilus edulis	0	0	0	0
Annelida				
Arenicola marina	0	0	0	0
Sipuncula				
Sipunculus nudus	0	0	0	0
Fungus				
Piromyces sp. strain E2	0	0	0	0
Fusarium oxysporum	0	1	1	9
Alveolate and Stramenophile				
Nyctotherus ovalis	0	0	0	0
Blastocystis	0	0	0	0
Archaeplastida				
Chlamydomonas reinhardtii	3	2	0	0
Unranked				
Arabidopsis thaliana	5	1	0	0
Schizosaccharomyces pombe	2	2	0	0
Ustilago hordei	1	1	0	0
Chordata				
Oikopleura dioica	1	1	0	6
Ascomycota				
Neurospora crassa	1	0	1	8
Saccharomyces cerevisiae 283	1	0	1	0

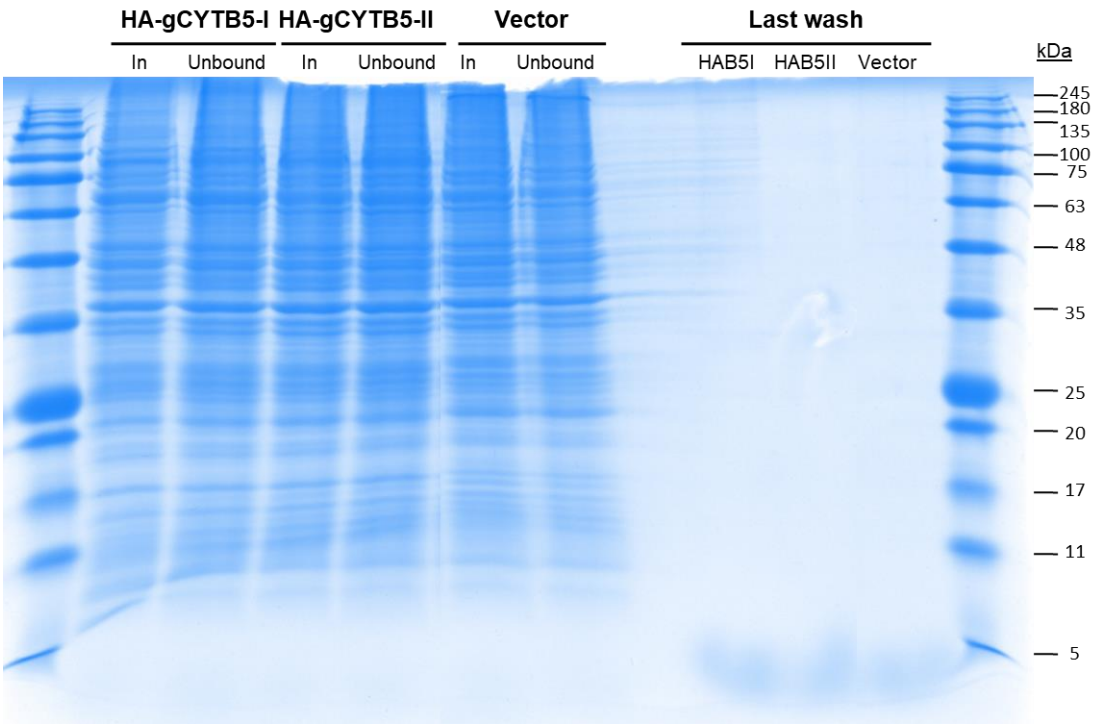
APPENDIX H. LOCALIZATION OF GCYTB5 IN GIARDIA BY PYRIH ET AL., 2014.



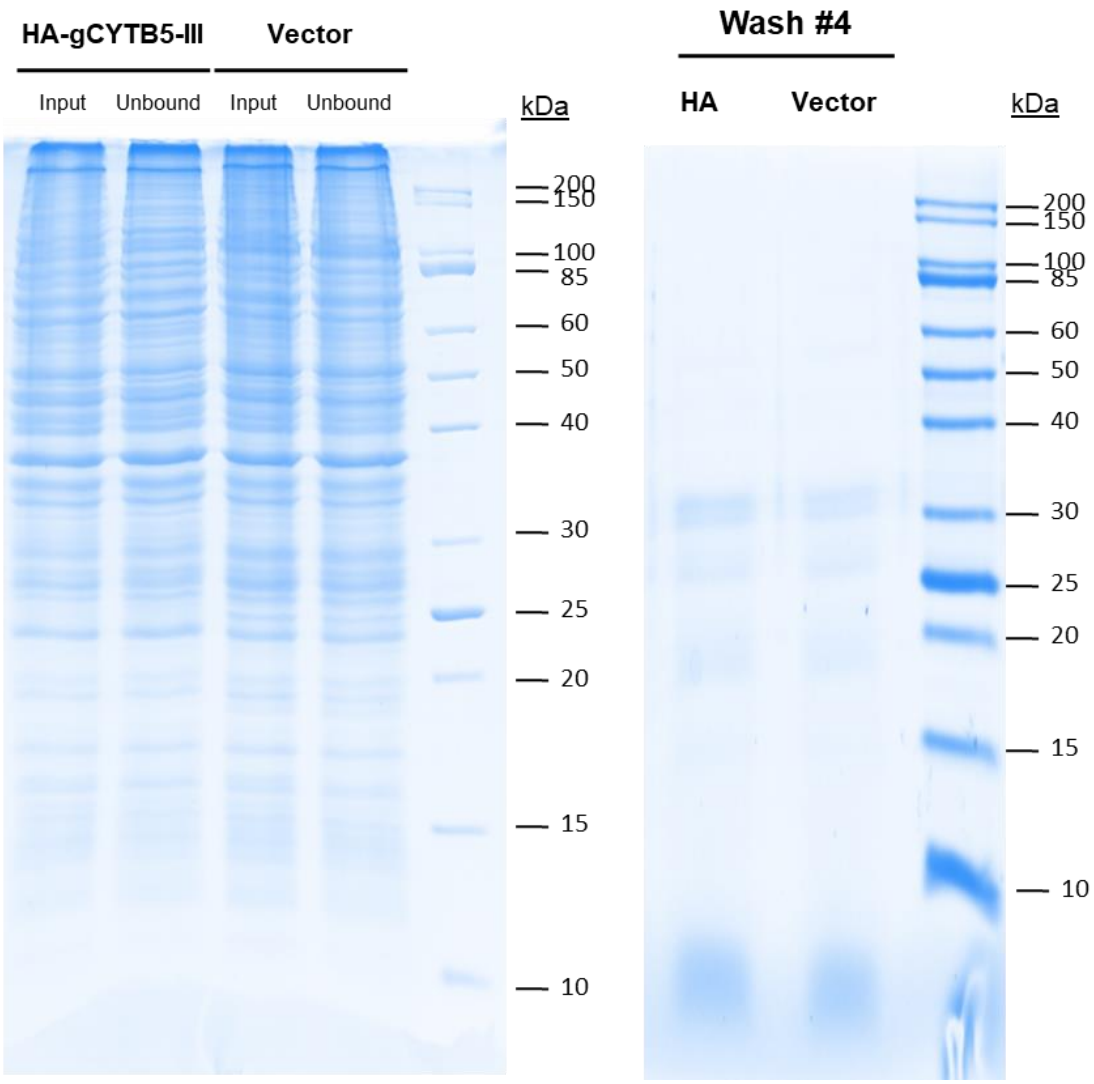
Jan Pyrih et al. Eukaryotic Cell 2014;13:231-239

Localization of gCYTb5 proteins in Giardia. (A) The cytosolic localization of gCYTb5-I to -IV visualized using immunofluorescence microscopy. The gCYTb5 proteins were visualized using rat anti-HA tag and anti-rat Alexa Fluor 594 (red) antibodies. PDI-2 was detected using mouse anti-PDI-2 and anti-mouse Alexa Fluor 488 (green) antibodies. The nuclei were stained with DAPI (4′,6-diamidino-2-phenylindole) (blue). DIC, differential interference contrast. (B) Localization of gCYTb5 in subcellular fractions of Giardia using immunoblot analysis. LYS, cell lysate; CYT, cytoplasm; ORG, organellar fraction; Enolase, a cytosolic marker protein; PDI-2, an endoplasmic reticulum marker protein (Pyrih et al. 2014).

APPENDIX I. CO-IMMUNOPRECIPITATION LOADING CONTROLS

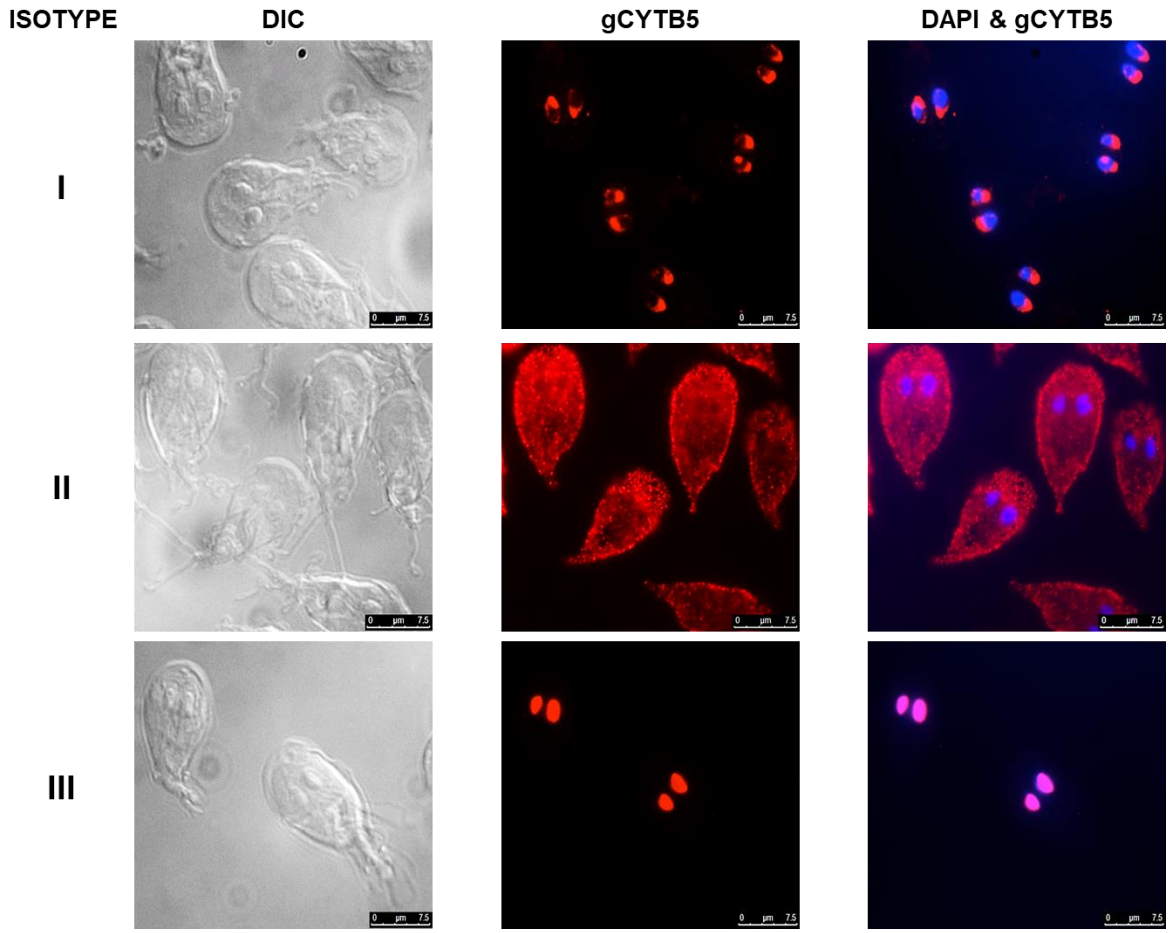


SDS-PAGE gel from the pull-down assay from the HA-gCYTb5-I, Ha-gCYTb5-II and vector cell line. Blue Silver stain was used to visualize the gels. The input, unbound and wash sample were loaded to control that the same amount of protein were used in each condition and to control that most of the nonspecific binding proteins were removed.



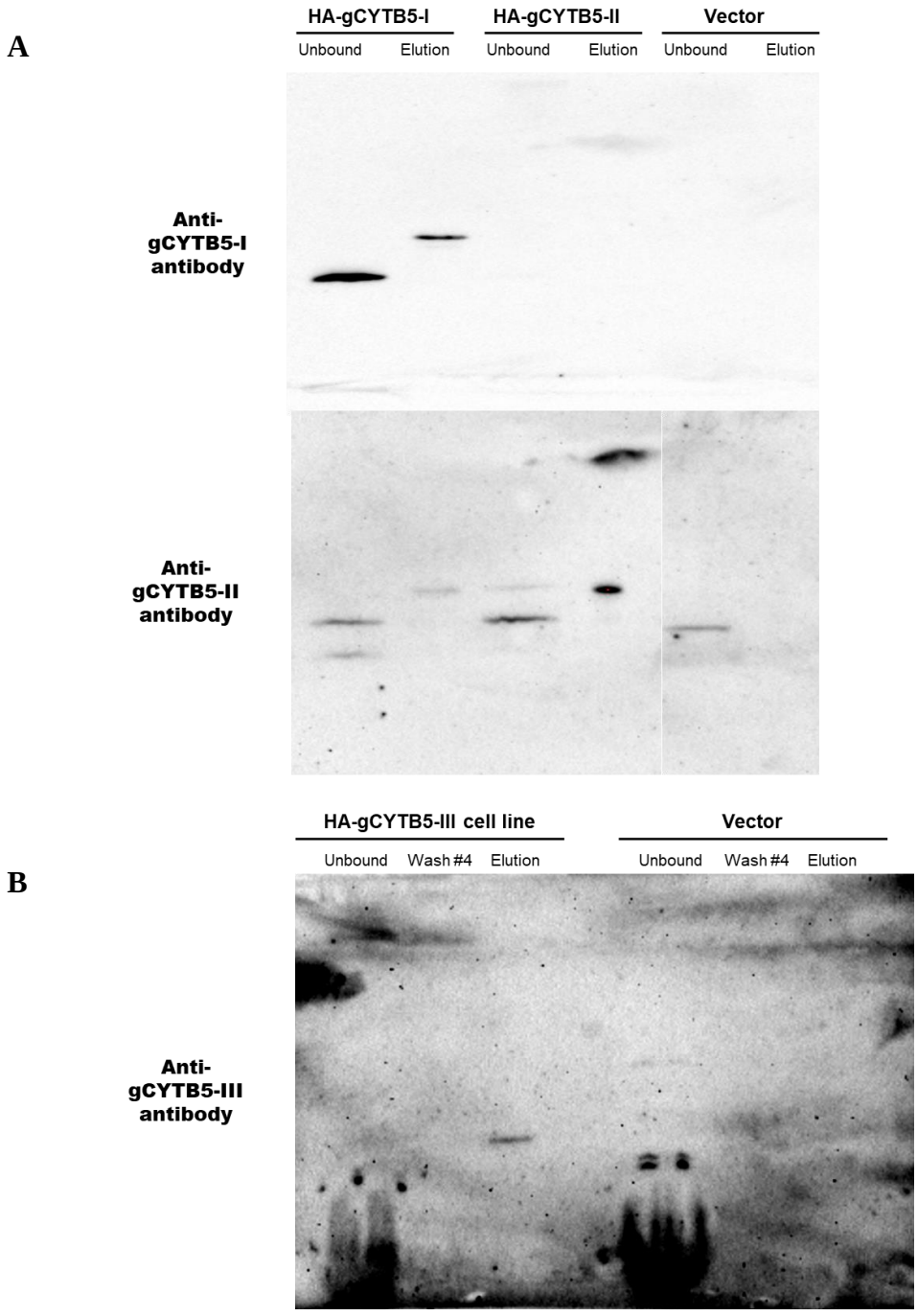
SDS-PAGE gel from the pull-down assay from the HA-gCYTb5-III and vector cell line. Blue Silver stain was used to visualize the gels. The input, unbound and wash sample were loaded to control that the same amount of protein were used in each condition and to control that most of the nonspecific binding proteins were removed.

APPENDIX J. IMMUNOLOCALIZATION OF THE GIARDIA CYTOCHROMES USING TRANSFECTED GIARDIA



Immunofluorescence assay using the HA-gCYTB5-I, II and III cell lines and protein specific antibodies. DIC (Differential Interference Contrast) was used to control the morphology of the cell, DAPI (4', 6-diamidino-2-phenylindole) was used to stain the DNA and localize both nuclei and anti-gCYTB5-I, II and III antibodies were used to localize the endogenous and tagged proteins.

APPENDIX K. CONFIRMATION OF THE GCYTB5S IMUNOPRECIPITATION



Panel A: the membraned was first reprobed using anti- gCYTB5-I antibody (top) and the membrane was stripped and reprobed with anti- gCYTB5-II antibody (bottom). Panel B: the blot was re-hybridized with anti- gCYTB5-III antibody.