5-2019

Protein Interaction Testing of TgCycY Putative Cyclin Protein of Toxoplasma gondii for a Novel Drug Target Identification

Wadumesthrige Dinusha Niranjali Peiris
Saint Cloud State University, wnpeiris@stcloudstate.edu

Protein Interaction Testing of TgCycY Putative Cyclin Protein of *Toxoplasma gondii* for a Novel Drug Target Identification

by

Wadumesthrige Dinusha N. Peiris

A Thesis

Submitted to the Graduate Faculty of

St. Cloud State University

in Partial Fulfillment of the Requirements

for the Degree

Master of Science in

Cell and Molecular Biology

May 2019

Thesis Committee:
Christopher Kvaal, Chairperson
Timothy Schuh
Nathan Bruender

Abstract

The apicomplexan protozoan parasites cause critical health problems and limitation of anti-parasitic drugs is a major problem. Among the apicomplexan parasites, *Toxoplasma gondii* (*T. gondii*) is a highly prevalent obligate intracellular protozoan parasite, which seems to rely mostly on proteins that are defined as cyclin and cell cycle kinases that needs to regulate the cell cycle of the tachyzoites. These proteins regulate DNA replication during the cell cycle and lead to moderate cell division which results toxoplasmosis disease. Before targeting the proteins in *T. gondii*, it is essential to identify these proteins and define their function in the cell cycle via protein-protein interactions. In vitro, protein-protein interaction testing of *Toxoplasma gondii* ME49 putative TgCycY cyclin via yeast two-hybrid screen was conducted to identify a protein interaction partner. Even though previously tested in vivo analyses showed no direct interaction between TgCycY and TgCrk2, TgCrk2 hypothesized as a potential interacting partner for the TgCycY with the association of a bridging protein. Hypotheses were made based on the ortholog TgCycY-TgCrk2 complex interactions in *Homo sapiens*, *Drosophila melanogaster*, and *Saccharomyces cerevisiae*. Among the identified 8 protein interaction partners from a cDNA library of asynchronous tachyzoite transcriptome that used in the Y2H screen, only TgDJ-1 protein was introduced as the potential interacting partner for the TgCycY. Other protein partners were excluded due to limitations of the yeast-two hybrid screen and lack of information. Based on the previous studies, TgDJ-1 protein has shown function involved in the micronemes secretion and it has introduced as the regulator of the *Toxoplasma gondii* secretion, motility, and invasion by interacting with TgCDPK1.TgDJ-1 inhibition did not affect the cell cycle of the tachyzoite. A CDK interacting partner for the TgCycY was not detected in this screen. Previously known fact of Cyclin Y function as a substrate for the CDK and the function of the mediating the ortholog CycY-CDK protein-protein interaction via a third protein was applied to the TgCycY-TgDJ-1-TgCDPK1 predicted complex. Due to limitations of the current screen and obtained results, indicated hypotheses 1 & 2 cannot be proven and the results direct the conclusion towards the TgCycY having TgDJ-1 as a new protein interacting partner that has existing literature and role of outside the cell cycle regulation. Overall, our project aims to map at the molecular level interactions of putative TgCycY cyclin protein in vitro and identify its function in the *T. gondii* tachyzoites to introduce it as a novel drug target. Further study regarding the function of TgCycY-TgDJ-1-TgCDPK1 complex is needed to understand the TgCycY role in micronemes secretion and involvement in the tachyzoites.

Acknowledgements

I would first like to thank my husband for continuous support, motivation, and encouragement throughout the program. Special thanks to my parents for guiding, supporting during the whole my educational career and letting me follow my dreams. Without their support, I would not be able to achieve the education that I dreamed.

Second, I would also like to give my regards to my advisor Dr. Christopher Kvaal for the guidance, vision, and knowledge that he provided throughout the research project. He is a great scientist and a mentor that I met during my educational career and because of his support I could complete the program even I had to face the setbacks in my life.

Special thanks to Dr. Timothy Schuh and Dr. Nathan Bruender not only for their comments, guidance and their knowledge for my thesis project but also for their support to write the thesis. Thanks to St Cloud State University, Graduate school, Department of Biology and Center for international studies for allowing me to complete the degree and experience a whole new level of research work that will build up steps for my future research career.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

History of *Toxoplasma gondii*

Toxoplasma gondii is an obligate intracellular protozoan and apicomplexan parasite that can infect most warm-blooded animals (Araujo & Silfer, 2003; Howe & Sibley, 1995). This is one of the most successful protozoa that can infect warm-blooded animals and manipulate their immune system to create a chronic infection. *T. gondii* is one of the species within the apicomplexan phylum, and it is diversified by having three types of strains; Type I (RH & GT), Type II (ME49), and Type III (VEG). The ME49 strain has been reported as the type most frequently associated with human disease (Araujo & Silfer, 2003; Howe & Sibley, 1995). This parasite was discovered in 1908 and named in 1909 (Dubey, 2008). First identification occurred in human tissues of a congenitally infected infant (Dubey, 2008). In 1948 Sabin-Feldman dye test was used to recognize the *T. gondii* as a common parasite in warm-blooded animals (Dubey, 2008).

Uncontrollable *Toxoplasma gondii*

Toxoplasma gondii is ubiquitous throughout the world and estimated to infect approximately 10 – 90% of the world's population (Hu, et al., 2002). Of those who are infected with *T. gondii*, a tiny percentage shows the symptoms because a healthy person's immune system keeps the parasite hidden via cell differentiation. The life cycle of the *T. gondii* has three stages. They are sexual cycle, environmental stage, and the asexual cycle (Figure 1). In cats, which is the dominant host, the parasite undergoes a sexual cycle. In other warm-blooded animals, which are intermediate hosts, the parasite undergoes an asexual cycle. The rapid asexual growth of *T. gondii* in warm-blooded animals, including the humans, results in significant issues such as cell lysis and inflammation. The hindrance of growth with medication is crucial in disease pathology,

and in this case, drug resistance is a major concern (Butler, el at., 2014). During the sexual cycle, the micro and macro gametes fused to create a zygote which is defined as the oocyst. The oocysts are released to the environment and turn to matured oocysts. Ingestion of matured oocysts into intermediate host starts the asexual cycle. Matured oocysts differentiate into tachyzoites. The tachyzoite stage of *T. gondii* cell cycle has a rapid cell division and during this phase, tachyzoites which are motile and quickly multiplying, are responsible for expanding the population of the parasite in the host via proliferation. This process is called endodyogeny and the cell cycle proteins such as cyclins and cyclin-dependent kinases are involved in the generation of second messenger molecules important for upregulating cell cycle progression, transcription, cell differentiation and proliferation (Figure 1; Goncz & Rothman, 1996).

T. gondii tachyzoite is a model organism to study the cell cycle progression and proliferation since it utilizes endodyogeny to produce two progenies from a mother cell (Hu, et al., 2002). However, bradyzoite is a biologically inactive form that rarely replicates in the host cell. It has a unique ability to persist within the host's cells by creating a tissue cyst. When a barrier is formed by bradyzoites due to the immune stress occurred by the tachyzoite rapid cell division, antigens are produced for long-term viability of the bradyzoites in the host's tissues. The bradyzoite number within the cyst increases with the increasing size of the cyst. However, the differentiation of tachyzoites to bradyzoites and the exact molecular triggers in this process are not well studied (Kim, 2015). How the parasite detects the environmental signals to differentiation and creates changes in its intracellular environment are also not known. Stress-induced signaling pathways are involved in altering the cAMP and cGMP levels to participate in the role of stimulating the *T. gondii* differentiation (Skariah, Mcintyre, & Mordue, 2010). Also, interferon-gamma (IFN-γ) dependent cell-mediated immune response in the host is identified as

immune stress that kills off most the tachyzoites and lead the tachyzoites to convert into bradyzoites that reside in the cysts which covered by the cysts wall (Figure 2; Suzuki, 1989).

Immunosuppression of the host can directly differentiate bradyzoites back to tachyzoites and start invading the host cells which replicate to continue the asexual cycle (Figure 1 & 2). This makes the *T. gondii* an opportunistic parasite. Releasing the tissue cysts to the environment can infect a definitive host for the continuation of the sexual cycle (Blader, Coleman, Chen, & Gubbels, 2015). Furthermore, the cell differentiation from tachyzoite to bradyzoite helps the organism to avoid treatments that disrupts its cell cycle by clumping bradyzoites together and creating a cyst wall around the bradyzoites. This cyst wall acts as a barrier for the host immune activity. Also, bradyzoites does not replicate at all. The environmentally resistant stage, tissue cysts is the reason for the worldwide prevalence of *T. gondii* to be an uncontrollable parasite (Dubey, 2009).

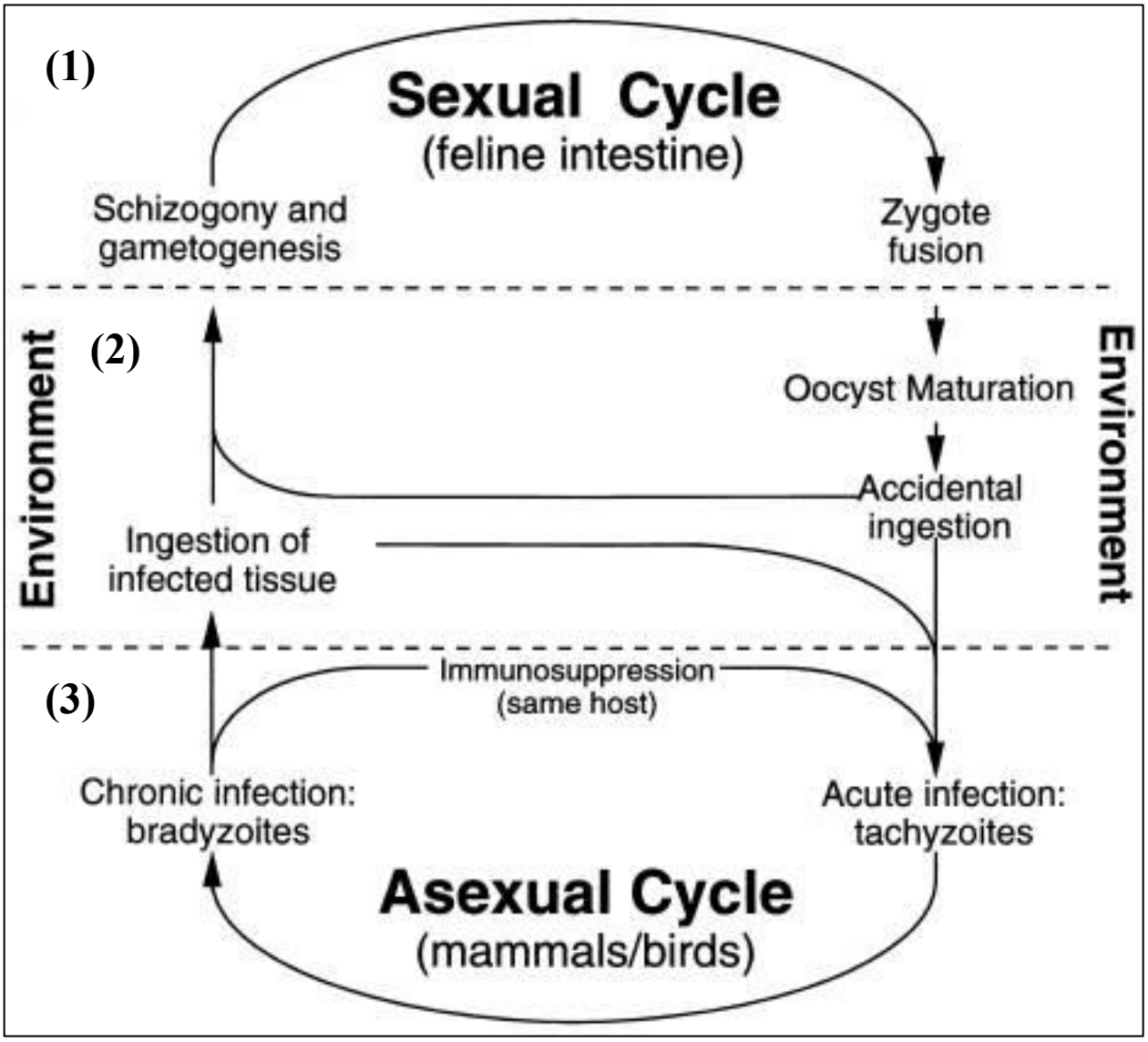


Figure 1: Diagram of *T. gondii* life cycle. (1) Sexual cycle in the definitive host, cat. Micro & macro gametes fusion to create the zygote and differentiate into oocyst. (2) Environmental stage for oocyst maturation. Ingestion of the tissues with cysts by the definitive host continue the sexual cycle (3) Asexual cycle in intermediate host, mammals/warm blooded animals. Ingestion of matured oocysts by intermediate host release tachyzoites rapid growing cells. Immuno-stimulation differentiates tachyzoite into bradyzoite which create cysts by creating a cyst wall around the clumped bradyzoites. Immuno-suppression differentiates bradyzoites back to tachyzoites. Ingestion of the cysts in the environment by intermediate host continue the asexual cycle of *T. gondii*. (Black & Boothroyd, 2000)

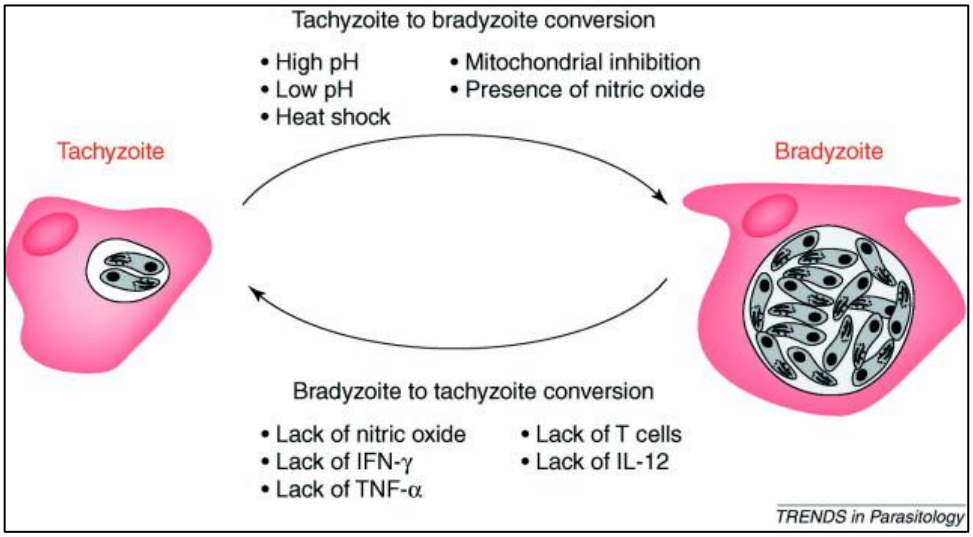


Figure 2: Tachyzoite to bradyzoite interconversion. The process of differentiation is reversible in *T. gondii*. Chronic infection is associated with the bradyzoites and acute infection associated with the tachyzoite. Immuno-stimulation in the host activates the tachyzoite differentiation and immune-suppression of host trigger the differentiation of tachyzoite to bradyzoite (Lyons, Mcleod, & Roberts, 2002).

Infection Cycle

Definitive host, the cat, is continuing the sexual life cycle of the *T. gondii* by feeding on infected rats, and birds. The intermediate host are all the warm-blooded animals including mammals, birds, etc., are needed for the continuation of the asexual cycle (Dubey, 1996). Oocysts released with the cat feces into the soil and water contaminate the vegetables, fruits, and meat. Ingestion of these contaminated water, vegetable, and undercooked meat can initiate the oocysts differentiation into the tachyzoites and bradyzoites. The development of *T. gondii* in humans can transmit into placenta if the infected human is pregnant (Hill & Dubey, 2002). Transmission of the *T. gondii* can occur through organ or bone marrow transplant or blood transfusion (Esch & Petersen, 2013; Figure 3).

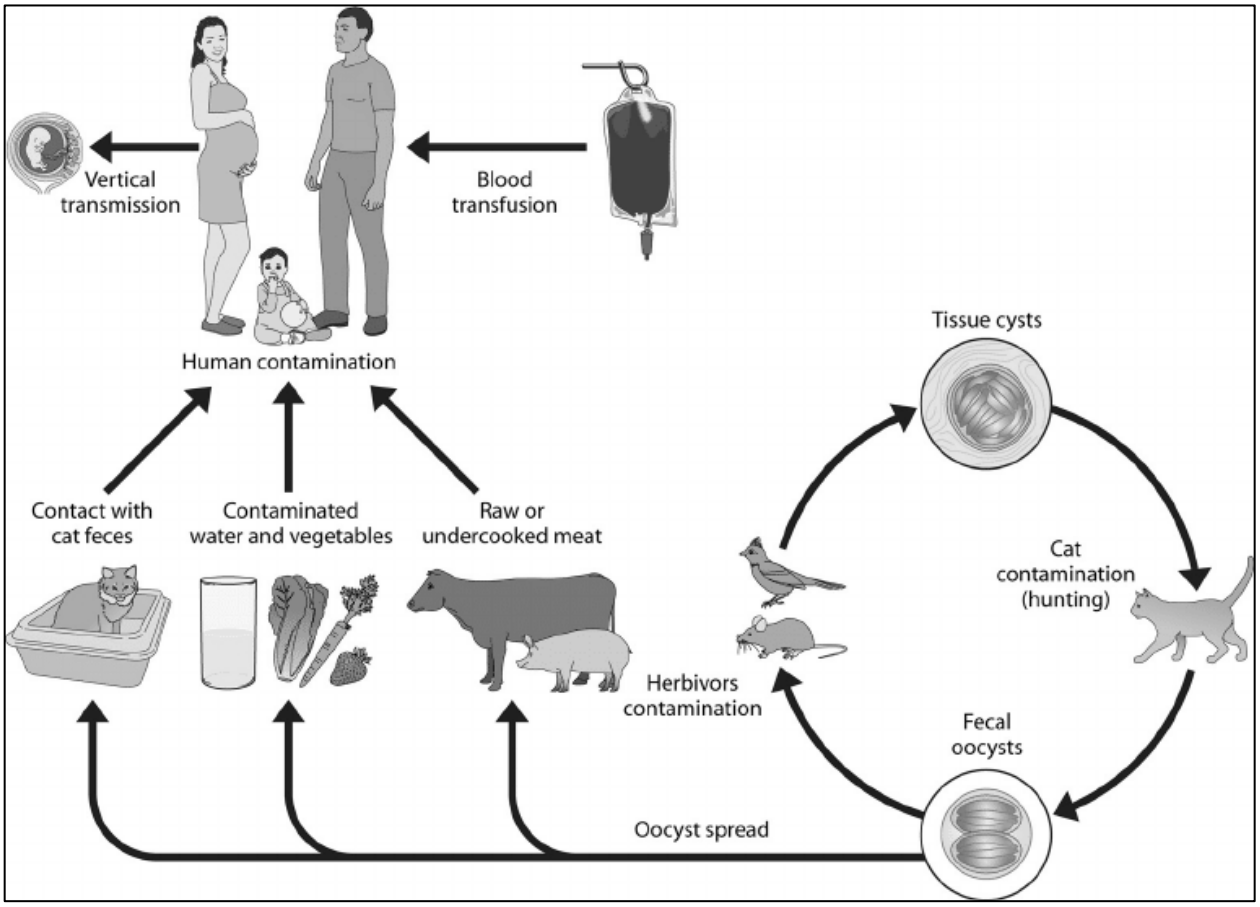


Figure 3: Infection cycle of the *T. gondii*. Infection cycle for the continuation of sexual cycle in the definitive host and continuation of asexual cycle in the intermediate host. (Esch & Petersen, 2013).

Medical Importance

T. gondii is the causative agent of the toxoplasmosis disease where the general infection is asymptotic. People who developed symptoms may experience a fever, swollen lymph nodes, muscles aches, sore throat, and a headache. Toxoplasmosis can get severe in the people who have a weak immune system such as HIV AIDS patients. The chronic toxoplasmosis can be in three major infections. They are three subclinical infections called ocular toxoplasmosis, cerebral toxoplasmosis, and congenital toxoplasmosis. Ocular toxoplasmosis (OT) causes the white retinal lesions with a vigorous inflammatory reaction in the focal area. This results in ‘headlight

in the fog' appearance. These lesions are occurring due to parasite invasion and immune activity against the parasite invasion. The risk of occurrence of OT is higher in the age group 40 or older. This inflammatory activity in the retina causes scarring which can develop into blindness (Furtado, Smith, Belfort, Gattey, & Winthrop, 2011).

The most severe type of the infection is cerebral toxoplasmosis. This infection is the most common cause of secondary CNS infections in patients with AIDS (Patil, Rajmane, Raje, & Patil, 2011). This infection is defined as toxoplasmic encephalitis (TE) that cause the death, coma and brain abscess in AIDS patients which activate the latent infection. The nonspecific symptoms are dementia, ataxia, seizures, and lethargy. These symptoms make it harder to diagnose. Cerebral toxoplasmosis manifest as multiple lesions. Severe infection can lead to death or comma (Furtado, Smith, Belfort, Gattey, & Winthrop, 2011).

Congenital toxoplasmosis (CT) occurs in newborn babies if the infection occurs during the placenta stage in the pregnant women. The CT can be asymptotic in babies and it can be developed into the retinochoroiditis or Central Nervous System (CNS) damage. The prevalence is 1 to 10 per 10,000 live births in the United States, one per 770 live births in southeast Brazil and one per 3,000 live births in France. The vertical transmission of this infection is higher at later stages of pregnancy, but it can be severe if the transmission occurs at early gestation period. As for the clinical symptoms, the delayed mental development, epilepsy, and hydrocephalus can occur. Severe development of CT can be causing the miscarriages during the pregnancy, stillbirths or babies with abnormalities (Furtado, Smith, Belfort, Gattey, & Winthrop, 2011). Due to these severe types of toxoplasmosis, it is important to study about the *T. gondii* to control their progression.

Treatments

Current treatments for toxoplasmosis are Daraprim which is a folic acid antagonist and sulphadiazine which is an antibiotic. Daraprim contains pyrimethamine and sulphadiazine contains triple sulfonamide but these treatments can only target the tachyzoite stage of the *T. gondii* and they do not eradicate the tissue cysts. Daraprim is a competitive inhibitor of dihydrofolate reductase (DHFR) which is a key enzyme in the redox cycle for production of tetrahydrofolate. This DHFR is essential co-enzyme for the nucleic acid synthesis. Daraprim has higher affinity for *T. gondii* DHFR than human DHFR hence daraprim can use for the toxoplasmosis without effecting human DHFR. Daraprim can work by killing *T. gondii* tachyzoites or preventing their cell growth (Brown, 1984). Sulphadiazine is a synthetic pyrimidinyl sulfonamide derivative and it inhibits bacterial folic acid synthesis by competing with para amino benzoic acid. This drug used together with daraprim to treat toxoplasmosis (PubChem).

These drugs have little impact on subclinical toxoplasmosis infections but the growth of the *T. gondii* in mice is restrained by the Sulphonamides. There are other drugs such as diaminodiphenylsulphone, atovaquone, spiramycin and clindamycin that are used for the toxoplasmosis in difficult situations (Hill & Dubey, 2002). There is no vaccine available to treat *T. gondii* infected humans. One reason for the need of new treatments for the toxoplasmosis is current drugs cannot completely eradicate the tissue cysts of the *T. gondii* and make it less impactful. The side effects of the drugs are another issue that patients are facing during the treatment sessions. Daraprim intake can prevent the human body from absorbing the folic acid. Also, it can cause the bone marrow suppression and liver toxicity. Current mostly used drug for the toxoplasmosis is daraprim. Daraprim has a controversy regarding its price hike from $13.00

to $750.00 per pill. The high price makes it less available, which makes toxoplasmosis a life-threatening disease, especially for HIV patients. Because of the limitation of drugs, a greater understanding of molecular mechanisms of proteins that control the rapid *T. gondii*, is vital to identify new therapeutic drug targets within the parasite.

Drug Targets

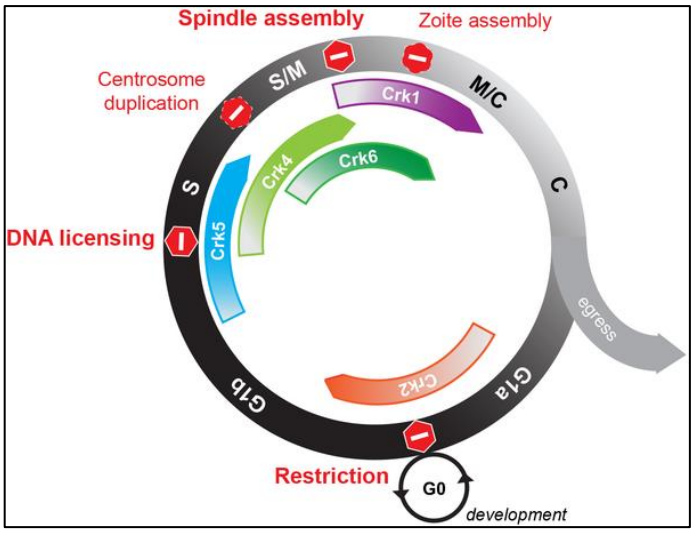


Figure 4: Cell cycle of *T. gondii*. Cyclin dependent kinases (CDK) involved in the *T. gondii* cell cycle and their corresponding cell cycle phase is shown along with the activity at each cell cycle phase (Alvarez & Suvorova, 2017).

Cyclin-dependent kinases (CDKs) are family of Ser/Thr protein kinases that share a highly conserved PSTAIRE motif which is essential for the binding to cyclin, and their catalytic activity is conducted by the interaction with the regulatory subunit cyclins and the CDK inhibitors (CKI). These CDKs can be regulated by both cyclin abundance and cyclin localization (Moore, 2013). The interactions between these three molecules are essential to ensure the *T. gondii* cell cycle progression and cell proliferation. Various cyclin/CDK complexes are responsible for promoting the transition between the major cell cycle phases: Growth (G1), Synthesis (S), Mitotic (M) and Cytokinesis (C) and overlap phases: S/M, M/C (Lim & Kaldis,

2013). Furthermore, there are three conserved restriction points in the *T. gondii* cell cycle where one checkpoint lies in the G1 phase to differentiate cells by the TgCrk2 in complex with TgPHO80 cyclin (TGME49_267580), the second lies in the S phase to control the licensing of DNA by TgCrk5, and the third lies in the M phase to operate spindle assembly by TgCrk6 (Figure 4). Even though they are identified, these checkpoint mechanisms are poorly understood (Alvarez & Suvorova, 2017). Cyclins are other major protein that control the cell cycle progression by interacting with the CDK to activate them to phosphorylate the target protein. Cyclins are involved in not only kinase-dependent transcriptional functions but also in kinase-independent transcriptional functions (Naumov, et al., 2017). After the phosphorylation of the target protein, cyclin is degraded by ubiquitin. So, identification of interactions between Cyclin, CDK, and CDK substrate are important to understand their function in the cell cycle and it will open a window to introduce them as novel drug targets for the *T. gondii*.

Cyclin Y and CDK5

Cyclin Y (CycY) is a highly-conserved protein of the cyclin superfamily in metazoans. They are famous for the role of regulating cell cycle and transcription but there is little known about the Cyclin Y. Also, Cyclin Y involved in the developmental stages of higher eukaryotes. The only Cyclin Y protein that expressed in the *T. gondii* cell cycle is TGME49_266900 but the function of this Cyclin Y and its interacting partner are not identified yet. So, it is important to understand and the identify the Cyclin Y function in *T. gondii*. Yeast two-hybrid screening results of previous research (Liu, Guest, & Finley, 2010) shown the Cyclin-Dependent Kinase 14 (CDK14) which is CDK5 family protein in both vertebrates and *Drosophila melanogaster* is the interacting partner for the CycY. The CycY ortholog in the yeast is PCL1 cyclin. The partner for this PCL1 cyclin is Pho85 which is a CDK5 and it is similar to the CDK14 in human and

Drosophila (Figure 5, Table 7 & 8; Liu, Guest, & Finley, 2010). In vertebrates, this CycY-CDK5 protein-protein interaction is important for the phosphorylation of the Wg/Wnt co-receptor which is needed for the Wnt signaling pathway that regulates cell fate determination, cell migration, cell polarity, and neural patterning during embryonic development (Komiya & Habas, 2008). In *Drosophila*, this protein-protein interaction is needed to recruit the CDK14 to the plasma membrane and tether it to the membrane. The Cyclin Y in *Homo sapiens* can be a substrate for the CDK14/Cyclin Y complex. The knockdown of the *Drosophila* CycY does not create defects in cell cycle which showed the CycY is not essential for its cell cycle (Liu, Guest, & Finley, 2010). In yeast, PHO85-PCL1 CDK/Cyclin interaction can phosphorylate CDK5 substrate (Huang, et al., 2007). Comparison of Cyclin Y sequence high conservation raised the possibilities for CDK5/Cyclin Y interaction to be ancient and these complexes might have novel properties which are separate from the other CDK-Cyclin complexes (Liu, Guest, & Finley, 2010).

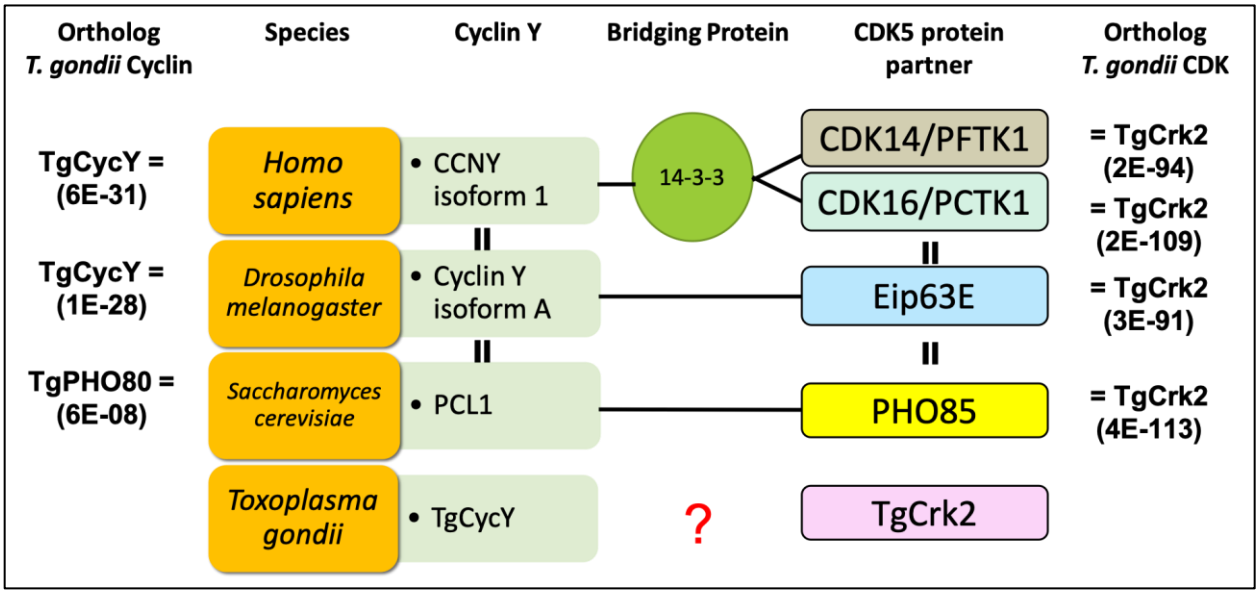


Figure 5: Ortholog Cyclin Y & CDK5 interaction. Cyclin Y and CDK5 protein – protein interactions in *Homo sapiens*, *Drosophila melanogaster*, and *Saccharomyces cerevisiae* comparison with *T. gondii* Cyclin Y and CDK5. BlastP values of ortholog Cyclin Y and CDK5 in *T. gondii* indicated the similarity to the Cyclin Y and CDK5 proteins in these three higher eukaryotes. Two CDK5 proteins partners are identified for Cyclin Y in *Homo sapiens*. CDK14 contains PFTAIRE motif hence named as PFTK1 and CDK16 contains PCTAIRE motif hence it is named as PCTK1. CDK5 partner in *Drosophila melanogaster* is Ecdysone-induced protein encoded by a gene at chromosomal position 63E. PHO85 is CDK5 in *Saccharomyces cerevisiae*.

Most of the CDK and cyclin interactions are identified as critical and conserved cell cycle regulators, but they also represent a regulatory motif that is employed in other cell processes. Sequence alignment of the cyclin creates two major functional categories where the first category is for the cell cycle regulators and the second category is for the RNA polymerase II regulators. Surprisingly, CycY does not fit into either category. So, the sequence of the CycY does not provide any clues of their cellular functions. CycY is the only member of the cyclin family that contains myristoylation signal. This signal is required for a localized activity of CDK5-myristoylated p35 and p39 complex to the plasma membrane in humans. This localization is needed for the direct phosphorylation of the membrane-associated substrates. (Liu, Guest, & Finley, 2010). In *T. gondii*, S-palmitoylation is important for the *T. gondii* lytic cycle. S-

palmitoylation is occurred by the attachment of palmitate through a thioester linkage to a cystine residue. Palmitoylation is important in *T. gondii* for the regulation of protein membrane localization, alter protein-protein interactions, protein stability and protein trafficking. This signal is dynamic and reversible. There is no literature that shows the presence of Myristylation signal in *T. gondii* Cyclin Y (Foe, et al., 2015).

TGME49_266900

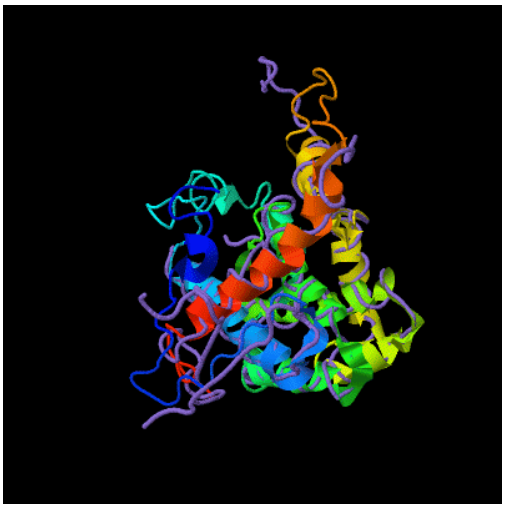


Figure 6: Predicted TGME49_266900 structural model. (i-Tasser structural modeling software)

TGME49_266900 is a 257 amino acids length putative cyclin protein with an N-terminal domain, and it is the protein of interest in this study. In previous *T. gondii* cyclin Y studies, the cyclin Y protein has named as TgCycY and this name will be used in this study to keep the cyclin Y naming consistent (Alvarez & Suvorova, 2017). TgCycY is a member of the Cyclin Y family. The structural homology modeling of the TgCycY was identified by inserting the protein sequence into the i-Tasser online platform. This online platform showed a high-quality model prediction of the 3D protein structure which is similar to the Pho85-Pho80 CDK/Cyclin complex of a phosphate-responsive signal transduction pathway (Figure 6). This protein has a mostly

similar structure to a cyclin and has a predicted function of binding protein dependent kinase to initiate phosphorylation of the residues in the carboxyl-terminal domain repeats of CDK (ToxoDB.org).

Also, the Phyre2 online platform predicted transmembrane helix in TgCycY using its protein sequence but it does not indicate the location of the transmembrane helix in the TgCycY protein sequence. Even though TgCycY does not have identified membrane localization function yet, in humans CDK14 localized to the plasma membrane by Cyclin Y and tethered to the plasma membrane (Liu, Guest, & Finley, 2010). Prior immunofluorescence analysis showed that the TgCycY was moderately expressed and localized to the nucleus and cytoplasm in tachyzoites (Alvarez & Suvorova, 2017). Importantly, among the other cell cyclins, TgCycY was the only oscillating protein that had a peak expression in the G1 phase of the cell cycle (Alvarez & Suvorova, 2017). Even though in vivo testing TgCycY was expressed in the G1 phase, it is not essential for the tachyzoite cell division which was tested by the conditional knockdown using the tet-OFF mutants of TgCycY and cyclin related kinases (Crks), nor the TgCycY interact directly with the TgCrk1, TgCrk2, TgCrk3, TgCrk4, TgCrk6, TgCrk7, & TgCrk8 which are essential Crks for the tachyzoite cell division (Alvarez & Suvorova, 2017).

These Crks are localized in the nucleus except TgCrk4 which is localized in the cytoplasm. Even though most of the Crks are localized in the nucleus only TgCrk2 has expressed during the G1 phase of the cell cycle where the TgCycY also has a peak expression (Figure 4). This indicates TgCycY might have a possibility to be an interacting partner for the TgCrk2. Also, TgCycY should have some important role in *T. gondii* since it is expressed in tachyzoites with a peak value in G1 phase and with an oscillation. This previous study (Alvarez & Suvorova, 2017) did not test the interaction between TgCrk5 though we could suggest that the interaction between

TgCycY-TgCrk5 will not be occur since TgCrk5 expressed in the S phase of the cell cycle (Alvarez & Suvorova, 2017). Though our research group has conducted a direct test between TgCycY and TgCrk5-L1 (TGME49_285160), which is an apicomplexan adaptation of TgCrk5 that is restricted to the environmental stage of merozoite & sporozoite in sexual cycle along with TgCrk2-L1, adaptation of TgCrk2. This previous test showed that TgCycY didn't create an interaction with TgCrk5-L1 (Schorr, 2018). Even though, the previous tested results did not identify direct interaction between TgCycY and TgCrk2, no one has tested for the indirect interactions between these two proteins. So, there is a possibility for the indirect interaction between TgCycY and TgCrk2 since they are expressed in the same G1 phase of the cell cycle.

TgCrk2 is responsible for the cell differentiation and dormancy in the G1 phase while TgCrk5 is involved in the licensing of DNA replication during the S phase of the cell cycle. The arrest of the G1 phase can occur by the inhibition of cytoplasmic TgCrk2 interaction with a P-type TgPHO80 cyclin. The TgCrk2 grouped with the cell cycle CDK5 family kinases and clustered with the eukaryotic cell cycle regulators family such as neuronal HmCDK5 based on the phylogenetic analysis of *T. gondii* Crks. Also, in contrast to other TgCrks, TgCrk2 expressed throughout the cell while most of other TgCrks expressed in the nucleus. TgCrk2 predominantly interacts with the TgPHO80 cyclins and it has a weak interaction with TgCycH (TGME49_260250) while no complexes formed with the TgCycY. TgCrk2 is one of the three Crks that has orthologs in other eukaryotes. This evidence might indicate atypical mechanism controls the half of the *T. gondii* cell cycle (Alvarez & Suvorova, 2017).

Computational predictions of protein-protein interactions can be very useful backup evidence for the experimental evidence. STRING database is one of the databases that gives the known and predicted protein-protein interaction based on the protein sequence. The interactions

include direct and indirect associations. The STRING protein-protein interaction database predicted the interactions between TgCycY and TgCrk1, TgCrk2, TgCrk4, TgCrk5, TgCrk6, TgCycH, thioredoxin protein and one unnamed CDK based on the protein sequence of the TgCycY (Figure 7 and Table 1). These interactions predictions are derived by the STRING database using the genomic context predictions, High-throughput experiments, co-expression, automated text-mining, and previous knowledge in databases.

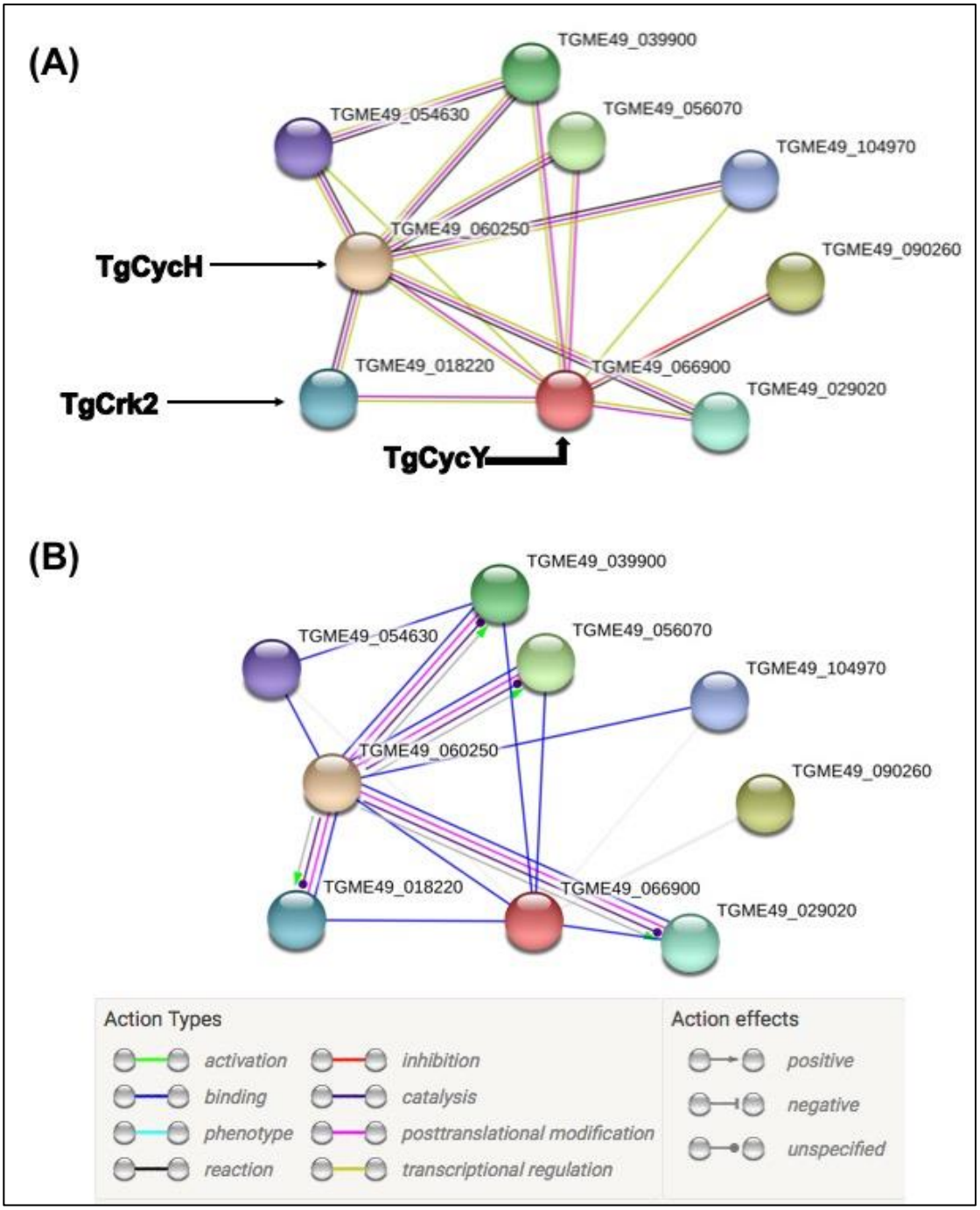


Figure 7: Protein – protein interaction predictions of TgCycY. (A) based on the experimental data, text-mining, co-expression and gene fusion evidence. (B) Predicted mode of action for the predicted protein interaction with TgCycY. String software was used to identify the possible interactions and mode of the actions of the TgCycY. TgME49_066900 is the protein of interest and the TGME49_018220 represents the TgCrk2. TGME49_060250 represents the TgCycH cyclin which has experimentally proven weak interaction with TgCrk2. Six other potential interactions of TgCycY were also included along with the TgCrk2 & TgCycH in the Table 1.

Table 1: STRING predicted protein interaction partners of TgCycY. In this Table 1, old Toxo ID in STRING database, new Toxo ID, corresponding naming in Alvarez & Suvorova, 2017, protein product, number of amino acids and their interaction score were included where 0.7 shows a higher confident interaction, 0.4 shows moderate confident interaction and 0.1 or lower score indicate lower confident interaction.

STRING ID (TGME49_#)	TOXODB ID (TGME49_#)	Naming in Alvarez & Suvorova 2017	Protein product	# of Amino Acid	Interaction Score
066900	266900	TgCycY	Cyclin Y	257	-
060250	260250	TgCycH	Cyclin H	600	0.635
090260	290260	-	thioredoxin family Trp26 protein	208	0.613
056070	256070	TgCrk4	CMGC Kinase	916	0.586
039900	239910	-	Cyclin Dependent Kinase	2325	0.586
029020	229020	TgCrk5	Putative Cell cycle associated protein Kinase CDK	1266	0.586
018220	218220	TgCrk2	Putative cell-cycle-associated protein Kinase CDK	300	0.586
104970	304970	TgCrk1	Putative Cell cycle associated protein Kinase CDK	1372	0.400
054630	254630	TgCrk6	CMGC Kinase	807	0.400

In Figure 7A, colored lines represent the existence of the types of evidence in the predicted protein-protein interactions. The red lines indicate the presence of fusion evidence, the purple lines indicate experimental evidence, yellow lines indicate text mining evidence, and the black lines indicate co-expression evidence. The TgCycY - TgCrk2, TgCycY - TgCycH predicted interactions indicated binding activity while TgCycH – TgCrk2 predicted complex indicated positive and unspecified activation, binding, and post-translational modification activities (Figure 7B). The probability of predicted interaction exists between two proteins measured by the interaction score. The confidence of these predicted associations has medium confidence limit of 0.4 and confidence between medium and higher confidence of 0.7. Lower the interaction score gives false positive predicted interactions.

Even though interactions between TgCycY-TgCrk1, TgCycY-TgCrk2, TgCycY-TgCrk4, TgCycY-TgCrk5, and TgCycY-TgCrk6 predicted to occur (Figure 7 & Table1), the experimentation on direct interaction between TgCycY and these TgCrks except TgCrk5 were not able to prove these interactions (Alvarez & Suvorova, 2017). TgCrk2 is included among these predictions. Even though there is no direct interaction between TgCycY and TgCrk2, there is no experimentation that has been conducted to identify the indirect interaction of TgCycY and TgCrk2. This may suggest that TgCycY might have an indirect interaction or it may have a non-canonical function (Hydbring, Malumbres, & Sicinski, 2016). Another point of view might be that TgCycY might interact with a regulatory protein to localize the TgCrk2 to the plasma membrane based on the human Cyclin Y/CDK5 localization to plasma membrane (Liu, Guest, & Finley, 2010).

Upregulation of TgCycY with the bradyzoite induction indicates that there is a possibility that TgCycY is involved in the bradyzoites cell growth. Interestingly, TgCycY has a similar

expressing pattern as the AP2IX-4 nuclear protein that is exclusively expressed in the tachyzoites and bradyzoites of the *T. gondii* except for the expression cell cycle phases. Surprisingly, AP2IX-4 is not essential for the tachyzoite replication, though it is required for bradyzoite development and it is necessary for the tachyzoite – bradyzoite cell differentiation. These results can be used to predict that TgCycY is a cyclin that involves in the bradyzoite cell development and it is an essential protein for the tachyzoite – bradyzoite cell differentiation because AP2IX-4 acts as a transcriptional repressor in the tachyzoites (Huang, et al., 2017).

CDK share highly conserved PSTAIRE motif and this motif is involved in cyclin binding. TgCrk2 (TGME49_218220) is included in the cdc2-related kinase subfamily and it is grouped under CDK5 protein family (Alvarez & Suvorova, 2017). According to the motif finder, TgCrk2 has a PSTAIRE motif, ATP binding site, and polypeptide substrate binding site which indicates that TgCrk2 can bind to a Cyclin. Even though the in vivo interaction testing between TgCrk2 and TgCycY did not show any interaction, the data mining software, STRING predicted TgCrk2 as the potential binding partner for the TgCycY (Figure 7 and Table 4). This gives the idea of indirect interaction between TgCycY and TgCrk2. Also, TgCrk2 is a homolog to the Pho85 in *Saccharomyces cerevisiae*, Eip63E in *Drosophila melanogaster*, PFTK1/CDK14 and PCTK1/CDK16 in *Homo sapiens*. The TgCycY is ortholog to the interacting Cyclin Y which is protein partner for Eip63E, PFTK1, PCTK1 but it is not ortholog to the PCL1 which is the Cyclin Y protein and interacting partner of Pho85 in *S. cerevisiae* (Figure 5; Liu, Guest, & Finley, 2010). This suggest that TgCycY is the most ortholog and has the more function similarity to the Cyclin Y in *Homo sapiens* and *Drosophila melanogaster*.

Yeast Two-Hybrid System

Testing for protein-protein interactions with a *T. gondii* cDNA library via a yeast two-hybrid system can be used to identify an interacting protein partner for TgCycY which confirms the putative interaction and defines a novel interacting domain. Here the fusion of GAL4 domains, which are GAL4-BD (binding domain) and GAL4-AD (activating domain), in separate expression vectors occurs by interacting the prey and bait proteins in vitro to create a complete GAL4 transcription factor (TF). This TF can recognize the upstream activating sequence (UAS) and binds to a promoter to activate the transcription of the reporter genes ADE2, HIS3, and MEL1. The final identified interaction proteins can be confirmed via co-immunoprecipitation or sequencing and sequencing results can be defined via in silico analysis (Lai & Lau, 2017). pGBKT7 and pGADT7 expression vectors are the most popular and widely used yeast two-hybrid vectors. The pGBKT7 is used as a bait expression vector which is designed to express a fusion protein of the GAL4 DNA- BD and a bait protein while pGADT7 is using as a prey expression vector which is designed to express a fusion protein of the GAL4 DNA-AD and a protein of interest. In this study, TgCycY is fused to the multiple cloning site (MCS) of pGBKT7 in frame with the GAL4- BD and cDNA library of tachyzoite transcriptome fused to the MCS of pGADT7 in frame with the GAL4 – AD.

Even though yeast two-hybrid system is the most widely used method to identify the protein-protein interactions, it has limitations. Y2H system can only detect protein-protein interactions that occur in the nucleus and due to this limitation, some protein-protein interactions can be missed such as interactions involving in the membrane proteins, self-activating proteins or proteins that require post-translational modifications (Brückner, Polge, Lentze, Auerbach, & Schlattner, 2009). This can be avoided in this screening since the TgCycY is localized in the

nucleus but if TgCycY has an interaction with a protein that localized out of the nucleus, it cannot be detected. TgCycY has evidence to show that it has an expression in both tachyzoites and bradyzoites life stages where the peak expression is at G1 phase in tachyzoites and upregulated expression with the induction of bradyzoites (Alvarez & Suvorova, 2017; Huang, et al., 2017). This study is only doing screening on tachyzoites TgCycY and cDNA library interaction hence it cannot detect the TgCycY interactions in the bradyzoites. If the hypothesized protein interaction between TgCycY and TgCrk2 needs more than one protein to construct the interaction, it will not detect in this screen hence the yeast two hybrid system can only detect direct interaction between bait and prey two proteins. In humans, binding of Cyclin Y to14-3-3 protein which is a molecular chaperone protein that can interact with cell signaling proteins and has ability to alter the localization, stability, phosphorylation state or molecular interactions of a target protein, is needed to enhance the association between Cyclin Y and CDK14 or CDK16 kinases (Figure 5). Binding of 14-3-3 to Cyclin Y exposes the cyclin box which allows access to PFTAIRE motif of CDK14. Also, two 14-3-3 binding sites on Cyclin Y is highly conserved in all metazoans and some apicomplexan (Li, Jiang, Wang, & Chen, 2014, & Mikolcevic, et al., 2011). The 14-3-3 binding sites on TgCycY and literature review regarding the 14-3-3 in *T. gondii* are not identified.

Y2H system is prone to give a lot of false negative and false positive protein interactions. Membrane proteins interactions that cannot be detected due to localization or fused yeast reporter proteins that causes steric hindrance create false negative interactions. Lacking post-translational protein modifications in the yeast system can create false negative interactions. Y2H screens also have a problem in reproducibility and discrepancy which can arise due to the difference in selection stringency. Overexpression can reduce false negative results, but it can

also induce false positive results. Another reason for the false positive is the nuclear localization of the bait and prey that may not be their natural cellular environment. Also, the interaction between prey and reporter proteins or membrane anchors can create an interaction with bait and lead for false positives. Some proteins can overcome the nutritional selections when it is overexpressed. Sticky proteins are major false positive creators in this screen by unspecific interactions due to incorrect folding (Brückner, Polge, Lentze, Auerbach, & Schlattner, 2009).

The previous experimental evidence, computational data, assumptions, and suggestions are useful to conduct the interaction testing of putative TgCycY with the goal of identifying an interacting partner for the TgCycY via yeast two-hybrid system and to introduce it as a novel drug target for the *T. gondii*. Studying TgCycY protein-protein interactions can also provide better explanations for the tachyzoite-bradyzoite differentiation mechanism or other mechanisms that might be involved in the tachyzoites.

Chapter 2: Hypotheses

As stated before, TgCycY has a peak expression at the G1 phase of tachyzoites, but it is not essential for the cell division of tachyzoites. By contrast, TgCycY has upregulated expression during the bradyzoite induction. This expression is a shred of evidence for the involvement of expressed tachyzoite TgCycY cyclin in bradyzoite cell development. The predicted molecular function of TgCycY is to bind protein kinase and the biological process is to regulate CDK activity (ToxoDB.org). Cell localization and expression cell cycle phase evidence of TgCycY and TgCrk2 lead to predict TgCycY interacts with the TgCrk2. Also, the STRING computational protein-protein interaction analysis of the TgCycY has shown the potential to interact with TgCrk2.

Ortholog Cyclin Y/CDK5 interaction comparison and the BLASTp results in *Saccharomyces cerevisiae*, *Homo sapiens*, and *Drosophila melanogaster,* led to the prediction that TgCycY and TgCrk2 interaction is mediated by a third protein since the pre-test evidence showed that there is no direct interaction between TgCycY and TgCrk2. Furthermore, the yeast two-hybrid system can only measure the reporter gene activity, it is possible that a third protein is bridging the TgCycY and TgCrk2 interaction partners since reporter gene expression clarify whether it is interacting with the TgCrk2 or the interaction bridging protein (Criekinge & Beyaert, 1999). Ability of acting as a substrate by itself for the Cyclin Y-CDK5 complex in human lead to predict that TgCycY can act as a target protein by creating an interaction with the TgCrk2 along mediation of a third protein.

If the TgCycY does not interact with the TgCycY with a help of mediating protein, this lead to predict that TgCycY is not involved with the TgCrk2 to control the cell cycle of tachyzoite and TgCycY might have a novel function that represent a non-canonical pathway in

tachyzoite. Also, if the TgCycY involved in a non-canonical pathway, it might interact with a new unidentified protein or protein that is known in previous literature and in Toxodb database. If this prediction is proved, the previous predictions will be ruled out. Based on the above evidence and predictions, the purpose of this project was to test protein-protein interactions of TgCycY putative cyclin protein via yeast two-hybrid system and to define the role of TgCycY in tachyzoite for a novel drug target identification for *T. gondii*. The project was designed based on the following hypothesis:

H_1: TgCycY putative cyclin creates a protein-protein interaction with TgCrk2 to regulate the *T. gondii* cell cycle.

H_2: TgCycY putative cyclin interacts with a third protein that creates a bridge between TgCycY and TgCrk2 protein interaction.

H_3: TgCycY has a novel, undiscovered protein interaction or interaction with a new protein with an existing literature.

H_0: TgCycY does not interact with TgCrk2 and it will interact with a new protein that does not regulate the cell cycle of tachyzoites.

Chapter 3: Materials & Methods

The protein of interest, TgCycY (TGME49-266900) provided by Dr. White's research group at the University of South Florida, was cloned into the pGADT7 prey expression vector plasmid. The cloned restriction enzyme sites were unknown.

pGBKT7.TgCycY Bait Expression Clone Construction

Primer design for TgCycY isolation. Isolation of the TgCycY gene of interest from the pGADT7 vector plasmid was conducted by polymerase chain reaction (PCR) to clone it into the pGBKT7 vector plasmid. To do this, 28 to 31 nucleotides long forward and reverse primers were designed with an oligo tail including the NdeI and EcoRI Restriction Enzyme (RE) sites in the 5' end.

Figure 8: Forward and Reverse Primer design of TgCycY. Nde I and EcoRI restriction enzyme sites are highlighted in yellow color and the oligo tails are shown in blue color.

Sub – cloning of TgCycY into pGBKT7. Amplification of TgCycY along with a positive control via PCR was programmed by initial denaturing at 94 ^{0}C for 30 seconds, the secondary denaturing at 94 ^{0}C for 30 seconds, annealing at 55 ^{0}C for 30 seconds, and extending at 68 ^{0}C for 1 minute per kb of PCR product. The isolated PCR product was run on a 1.2 % agarose gel at 150 V for 1 hour in parallel with the pGADT7 expression vector and a 1 kb quanti-marker. TgCycY PCR product was extracted from the gel using a Gel/PCR DNA fragments extraction kit from IBI/MidSci. Extracted DNA was quantified via Nanodrop spectrophotometer.

Subsequently, 100 ng of the pGBKT7 expression vector and 60 ng of the amplified TgCycY insert were digested using NEB-Nde I and NEB-EcoRI RE enzymes overnight at 37^0C. Linearized PGBKT7 and the TgCycY insert were run through a 0.8% (w/v) agarose gel at 200 V for 80 minutes along with the 1Kb Quanti-marker and extracted from the gel using a razor blade. The concentration of this linearized pGBKT7 was quantified using a Nanodrop spectrophotometer. Shrimp alkaline phosphatase (rSAP) was used to dephosphorylate the 5' end of linear pGBKT7. Before ligation, dephosphorylated pGBKT7 and linearized TgCycY were heat- inactivated at 65 ^{0}C for 5 minutes. 3 moles linearized pGBKT7 vector to 1 mole of the amplified, linearized insert ratio was used to create the bait plasmid. 100 ng of pGBKT7 vector and 60 ng of the linearized TgCycY insert were ligated using T4 DNA ligase at 25 ^{0}C for 15 minutes and again heat-inactivated at 65 ^{0}C for 5 minutes. Clarification of sub-cloning of the pGBKT7 plasmid with TgCycY insert was done by running a 1% (w/v) agarose gel at 100V for 80 minutes along with the pGBKT7 empty vector, EcoRI/NdeI digested sub-cloned pGBKT7-TgCycY plasmid and 1kb Quanti marker.

Bacterial transformation and RE digest of clone construct. Immediately, the 500pg of ligation was transformed into 100 uL of High-efficiency DH5α *E. coli* competent cells in parallel with the pGBKT7. XPMC2 positive control. For the bacterial transformation, 5 uL of 100 pg/uL diluted ligation and 5 uL of 100 pg/uL diluted positive control were added to two separate chilled falcon tubes along with the 100 uL of DH5α *E. coli* competent cells. Then incubated both tubes in an ice water bath for 30 minutes and heat shocked at 42 ^{0}C for 30 seconds. Re-incubated the tubes in an ice water bath for another 5 minutes. After that 950 uL of super optimal broth (SOC) media was added and incubated in 37 ^{0}C shaker at 250 RPM for 1 hour and warm the LB (Luria-Bertani) + 50 mg/mL Kanamycin media plates at 37 ^{0}C. 1:10, 1:100, and 1:1000 serial dilutions

of the transformations in SOC were spread on warmed up LB + 50 mg/mL Kanamycin medium and incubated overnight at 37 ^{0}C. After confirming the transformation efficiency, an *E. coli* - (pGBKT7.TgCycY) colony was inoculated in the LB + Kanamycin liquid medium. Then pGBKT7. TgCycY expression clone was isolated via a High-speed Plasmid mini kit (IBI High-Speed Plasmid mini kit protocol). Purification and quantification were conducted for the purified cloned construct using a nanodrop 2000 spectrophotometer at a wavelength of 260 nm and 280 nm. Nde I, EcoRI single and double RE digests of quantified clone construct were used to clarify the success of cloning by a running a 1.0% agarose gel at 100 V for 80 minutes in parallel with pGBKT7 empty vector and 1 kb Quanti marker.

pGBKT7.TgCycY bait transformation into Y2HGold strain. Y2HGold yeast cells were grown on YPDA agar for three days and inoculated 2-3mm sized three colonies in 50 mL of YPD liquid media. Inoculation was incubated at 30 ^{0}C with shaking at 250 RPM overnight. After confirming the concentration of the overnight culture, 30 mL of the starter culture was transferred into fresh 300 mL of YPD media and confirmed the cell density via hemocytometer before the incubation. The new culture was incubated at 30 ^{0}C with shaking at 230 RPM. After 3 hours of incubation, we confirmed the cell density of 1* 10^6 CFU/mL by using a hemocytometer to clarify the healthiness of the cells before making them competent. Yeast cells were harvested by centrifugation at 1000 x g for 5 minutes at room temperature and re-suspended the pellet in 1X TE. Discarded the supernatant by centrifugation at 1000 x g for 5 minutes and room temperature. The pellet was re-suspended with 1.5 mL of 1X TE/LiAc.

In a 1.5 mL microcentrifuge tube, mixed the 100 ng of pGBKT7.TgCycY, 100 ng of Clontech career DNA, 0.1 mL of Competent Y2HGold yeast cells, and 0.6 mL of 1X PEG/LiAc solution. The mixture was vortexed for 10 seconds at high speed and incubated at 30 ^{0}C for 30

minutes with shaking at 200 RPM. 100% DMSO was added into the mixture after the incubation and mixed by gentle inversion. Then heat shocked at 42 ^{0}C water bath for 15 minutes. After the chilling the cells on ice for 2 minutes, centrifugation was done for 5 seconds at 14000 rpm and room temperature. Separated cell pellet was re-suspended in 0.5 mL of 1X TE buffer. The transformed bait plasmid was selected in SDO (Single Drop Out) -Trp, SDO-Leu, DDO (Double Drop Out)-Trp/-Leu, and QDO (Quadruple Drop Out) -Leu/-Trp/-His/-Ade minimal medium using 1:1 to 1:10000 serial dilutions. Additionally, the Y2HGold – (pGBKT7.TgCycY) was streaked on DO (Drop Out) – Trp/-His, DO – Trp/-Ade, and DO-Trp/-Ade/-His along with the positive control (Y2HGold – [pGBKT7. P53 + pGADT7. T]) and the negative control (Y2HGold – [pGBKT7. lam + pGADT7. T]) to test the auto-activation ability of the bait.

Matchmaker Gold Yeast Two-Hybrid Screening

Mating of pGBKT7. TgCycY with cDNA library. We prepared a concentrated overnight culture of bait strain (Y2HGold [pGBKT7. TgCycY]) by inoculating one 2 mm sized colony in 50 mL of SDO-Trp/Kanamycin liquid medium in a baffled flask and incubating at 30 ^{0}C with shaking at 250 rpm until the OD600 reached 0.8. The bait cells were pelleted using the JA-14 rotor at 1,000 x g for 5 minutes, and the separated pellet was re-suspended in 5mL of the SDO-Trp liquid medium. Before the mating, 10 µL of bait culture and the prey cDNA library were used to count the cell density via hemocytometer to ensure the cell density was > 1 x 10^8 CFU/mL for bait culture and > 2 x 10^7 CFU/mL for the prey library. Also, another 10 µL of the prey library which is Y187 yeast strain with pGADT7 plasmids of cDNA library provided by the University of South Florida was used in 1:10, 1:100, and 1:1000 serial dilutions to plate in SDO-Leu agar medium to check the prey library viability. They were incubated overnight at 30 ^{0}C. The bait re-suspension was mixed with 45 mL of 2X YPDA (Yeast extract Peptone Dextrose

Adenosine hemisulfate), 1 mL of prey library with washing twice and 50 μL of 50 μg/mL kanamycin in a 2 L flask. The mating culture was incubated at 30 ^{0}C with shaking at 50 rpm for 24 hours.

After confirming the presence of zygotes under a phase contrast microscope, cells were pelleted at 1,000 x g for 10 minutes. The 2L flask was rinsed with 50 mL of 2X YPDA and the rinse was used to re-suspend the pellet. The cells were pelleted for the second time, and the pellet was re-suspended in the 10 mL of 0.5X YPDA/5μg/mL kanamycin liquid medium. The final volume was measured after the re-suspension. The 10 uL of mated culture was serially diluted into 1:10, 1:100, 1:1000, and 1: 10000. 100 μL of each dilution was plated in SDO-Trp, SDO-Leu, and DDO-Trp-Leu. Rest of the culture was plated on 56 plates that are 150 x 15 mm sized QDO-Leu/-Trp/-Ade/-His plates using 200 μL per plate. Incubation of the plates at 30 ^{0}C for 3 days was used to calculate viability, mating efficiency and the number of screened clones.

testing of protein-protein interaction strength: The colonies were transferred onto secondary patch QDO plates, and more colonies were transferred onto these plates after an additional one day of incubation. Tertiary QDO streak plates were created using the secondary QDO patch plates to confirm the positive interactions. The quaternary streak plates (QDO/X/α/Gal) were created from tertiary streak plates along with the positive and negative controls. These plates were incubated at 30 ^{0}C for three days to estimate the protein-protein interaction strength based on the blue color strength of the colonies. The color of the yeast growth was scaled from 1-3 where #3 scale was given for bright blue color which was confirmed by the bright blueness of the positive control, and the #1 scale was provided for non-growing or weak white color colonies which was confirmed by the no growth or no blueness of the negative control.

Isolation of Protein Interaction Partner

Colonies ranked greater than #1, were selected for plasmid isolation via inoculating in 5 mL of YPD (Yeast extract Peptone Dextrose) liquid medium and incubated at 30 ^{0}C, 250 rpm for 18 hours. The yeast cells were pelleted in a 1.5 mL microcentrifuge tube at 14,000 x g for 1 minute. Yeast pellet was lysed by vortexing for 2 minutes with a mixture of 0.2 mL of lysis solution (2% v/v Triton X-100, 1g (w/v) SDS, 100 mM NaCl, 10 mM Tris-HCl, pH 8.0, 1 mM EDTA), 0.2 mL of phenol-chloroform-isoamyl alcohol and 0.3 g of acid-washed glass beads. Afterward, each supernatant containing plasmid DNA, separated by centrifugation at 14,000 x g for 10 minutes, was transferred into new 1.5 mL microcentrifuge tubes. DNA in the supernatant were pelleted by adding 1/10 volume of 3M Sodium acetate (NaOAc) pH 5.2 and 5/2 volume of ethanol and centrifuging at 14,000 x g for 10 minutes. The obtained DNA pellets were washed with the 1 mL of 70% ethanol and vacuum dried at 14,000 x g for 10 minutes. Extracted DNA was dissolved in 50 μL of 1X TE (Tris EDTA [Ethylene diamine tetra acetic acid]) buffer and transformed into high-efficiency DH5α *E. coli* competent cells. Selected prey plasmids on LB + Ampicillin serial dilution plates were inoculated in LB + 1g/mL Ampicillin liquid medium to amplify the plasmid concentration. Then the prey plasmids were purified from *E. coli* according to the high-speed Plasmid mini preparation protocol and using an IBI mini prep kit. Isolated prey plasmids were quantified for concentration & purity via Nanodrop2000 spectrophotometer.

Mapping Insert size via Hind III Restriction Enzyme Digest

500 ng of isolated prey plasmids were digested via Hind III restriction enzyme for 1 hour at 37 ^{0}C to measure the approximate insert size in the prey plasmids. Fragments of the digestion were visualized using E-Z vision one dye in a 2% agarose gel that ran at 150 V for 90 minutes with 0.5X TBE buffer. Also, 200 ng of uncut plasmids and 1 kb Quanti marker were included in

parallel with the digests. Based on the banding pattern, clones were categorized into 10 groups, and the approximate insert size was calculated.

Testing for Auto-Activating & False Positive Interacting Protein Partners

Only 30 clones that represent categorized 10 groups were co-transformed into Y2H Gold – (pGBKT7.TgCycY) and selected for diploids on DDO medium. Selected clones were transferred onto the QDO/X/α/Gal plates to re-scale the interaction strength based on the blue color. Co-transformation results were compared with the primary interaction strength results, as well as positive, and negative controls to select the false positive interactions. Also, auto-activation ability of the co-transformed 30 clones was tested by streaking each Y2H Gold strain containing these 30 clones on Drop Out (-Leu/-His), (-Leu/-Ade), and (-Leu/-His/-Ade) media along with the positive control (Y2HGold- [pGBKT7.p53 + pGADT7. T]) and the negative control (Y2HGold- [pGBKT7. lam + pGADT7.T]).

Sequencing of Isolated Protein Interaction Partners

Based on the RE digest banding pattern, interaction strength, and purity, these 30 clones from 10 different insert groups were selected for the sequencing. Sequencing was processed by the GenScript company using the 5’ T7 and 3’ AD sequencing primers. In silico, the received sequencing data was aligned to Contig by using the SeqMan Pro software to identify the similarity of the sequences. The consensus sequence of each Contig was used as the query sequence to conduct a BLASTn search against a transcript database in the Toxodb.org. This search was done to identify the highest matching cDNA for the consensus nucleotide sequences with the highest bits score and the lowest E-value.

Computational Analysis of TgCycY Interactions & Predictions

Expression profiles of the isolated and identified protein partners along with the TgCycY, TgCrk2, and TgCrk5 was done to compare their expression in G1 cell cycle phase. Also, the consensus sequences of isolated interacting cDNA inserts were aligned using Snapgene software with their corresponding full-length protein sequences to identify the aligned cDNA region in the full-length sequences. Computational analysis of interactions of TgCycY (TGME49_266900) protein of interest using String software was done to identify the possible interaction predicted partners for the TgCycY. Further computational analysis using Toxodb.org and BLASTp was used to compare the similar CycY cyclin protein in *Homo sapiens*, *Drosophila melanogaster*, and *Saccharomyces cerevisiae* and their interactions with interacting CDK5 protein partner to relate these results to TgCycY interactions in *T. gondii* since CycY is a highly-conserved protein with an essential function in these organisms. Also, the phylogenetic relationship of these ortholog proteins was analyzed using the UniProt database.

Chapter 4: Results

Auto-Activation of pGBKT7.TgCycY Bait Vector

In contrast to the positive control, Y2HGold - (pGBKT7.TgCycY) bait strain did not grow on the DO-Trp/-His, DO-Trp/-Ade, DO-Trp/-His/-Ade media, which proved that the bait strain does not have the auto-activation ability and that TgCycY acts as a binding domain in the yeast two-hybrid system.

Matchmaker Gold Yeast Two-Hybrid Screening

Table 2: Pre-mating data of bait and prey library. Describes the cell density of the bait and prey library prior to the mating. Cell density measured in CFU/mL. The bait cell density was greater than the $1 * 10^8$ CFU/mL and the prey library cell density was greater than the $2 * 10^7$ CFU/mL.

Prior to mating	CFU/mL
Bait Cell Density	$9.9 * 10^8$
Prey Library Cell Density	$2.8 * 10^8$

Table 3: Post mating results of the yeast two-hybrid screening. Describes the viability of the haploids (bait and prey) and the diploids obtained via yeast two-hybrid screening. Viability measured in CFU/mL. The limiting partner is the prey library.

After Mating	CFU/mL
Prey Viability	$2.25 * 10^6$
Bait Viability	$3.12 * 10^7$
Diploid Viability	$3.21 * 10^5$

The measured final volume of resuspension of the cells was 12.2 mL. The average number of clones screened (diploids) on serial dilutions of DDO media was 3.9 million. In this

screening, the limiting partner was the prey and the mating efficiency of the Y2H Gold yeast two-hybrid screening was 14.2%. The number of screened colonies in QDO was 95 which indicated that 95 colonies containing diploid cells had the interacting proteins with the TgCycY protein.

The secondary QDO patch plates (Figure 9) were showed strong and weak interaction strength based on the activation of ADE2 gene transcription gene and used to clarify the interaction occurred in the initial colonies in QDO screen plates and tertiary QDO streak plates (Figure 10B). All the colonies on the tertiary QDO streak plates grew on the media and they were ranked based on the blueness of the colonies (Figure 6B & Table 4). According to the blueness scaling, only 15 colonies were excluded for the inoculation due to their weak interaction. 80 colonies were selected from 95 colonies for inoculation in YPD medium. From inoculations, 68 plasmids were isolated successfully.

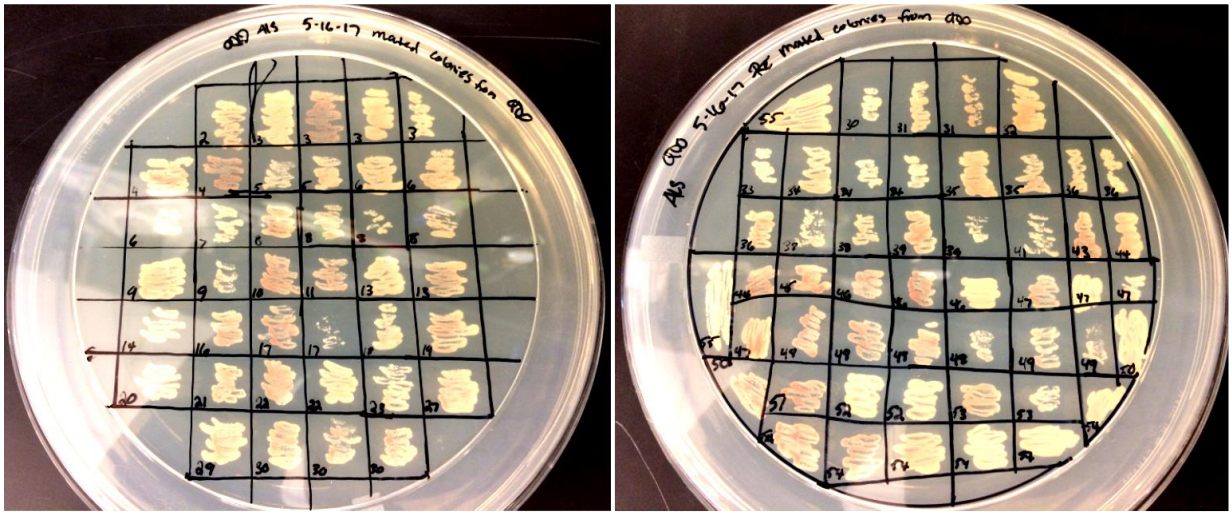


Figure 9: Secondary QDO patch plates of the yeast two-hybrid screening. The color of the patches was scaled 1 to 5. Scale 1 referred to strong white color and scale 5 referred to strong pink color. Strong white indicated a strong interaction and strong pink color indicated a weak interaction.

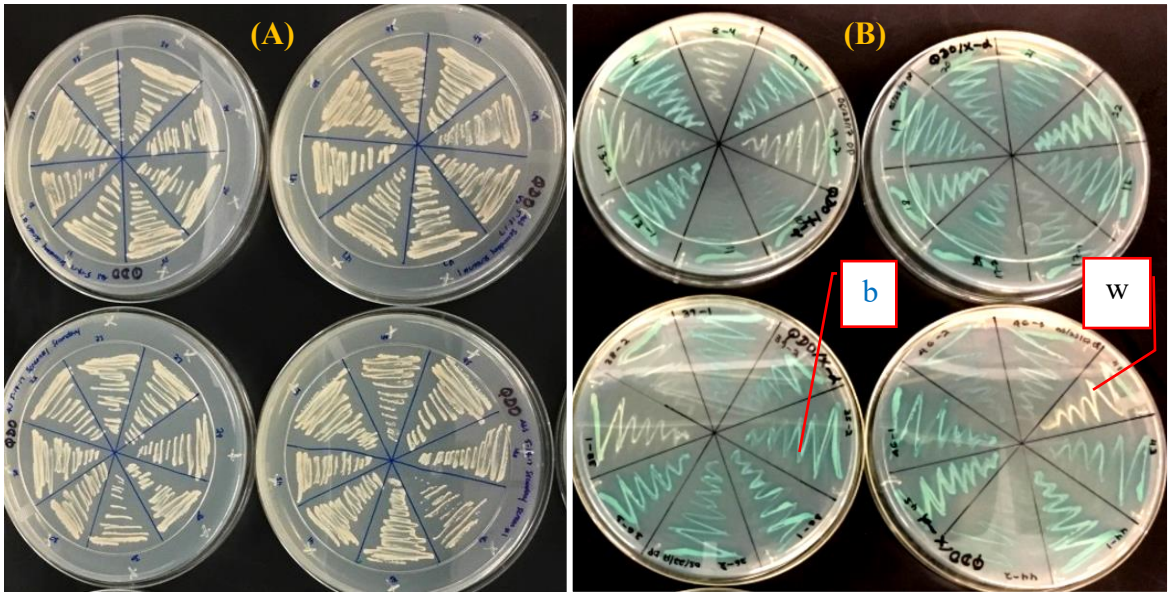


Figure 10: Testing of interaction strength. (A) Tertiary QDO streak plates and (B) QDO/X/α/Gal tertiary plates to test interaction strength. (A) Tertiary QDO plates recreates the positive interaction. (B) QDO/X/α/Gal results measures the interaction strength qualitatively by scaling 1 to 3. 1 = white color (w)/weak interaction, 2 = light blue color/medium interaction & 3 = strong blue color (b)/strong interaction.

Table 4: Scaling of blueness of QDO/X/α/Gal. Shown the intensity of the blueness scale of the 95 colonies isolated from the Y2HGold screening. #1 = weak interaction, #2 = medium interaction & #3 = strong interaction.

Scale	# of colonies
1	15
2	23
3	57

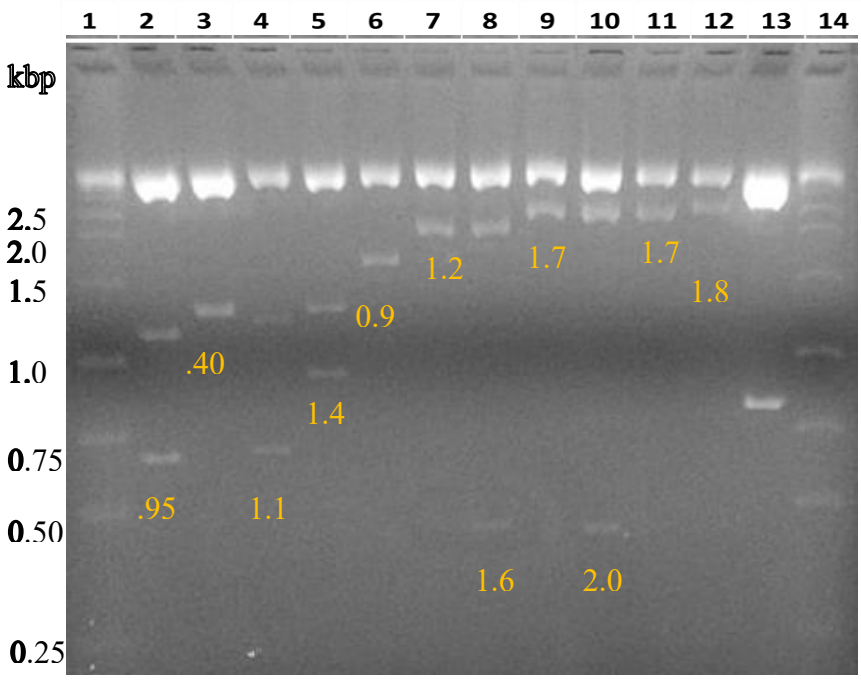
Mapping Insert Size via Hind III Restriction Enzyme Digest

From the inoculated 80 plasmids, 12 plasmids were excluded for further analysis due to their lower purity & concentration compared to the high purity & concentrated plasmids. These successfully isolated 68 plasmids were visualized in a 1% (w/v) agarose gel. These 68 plasmids

were grouped to 10 groups based on their banding pattern received by the Hind III RE digest which was visualized by the 2% (w/v) agarose gel (Figure 11 & Table 5).

Table 5: Grouping of isolated proteins interacting partners. The 10 groups were conducted by the fragment sizes of the Hind III RE digest on 2% (w/v) Agarose gel. In this table, the fragment sizes, calculated insert size, the number of the samples that included in each group, and interaction strength based on 1-3 scale were included for the comparison. The insert size was calculated by reducing the 8000 bp (base pair) of empty pGADT7 prey plasmid from the total base pairs of the digest fragments. (Rank$^{\#}$ = # represent the number of samples in corresponding interaction strength)

Group	Fragment sizes (kbp)	Calculated insert size (kbp)	# of samples	Interaction strength (Scale 1-3)
1	1.2	0.4	5	2^1, 3^4
2	1.7	0.9	6	2
3	1.1 + 0.65	0.95	2	3
4	1.2 + 0.7	1.1	1	2
5	2.0	1.2	7	2^4, 3^3
6	1.3 + 0.9	1.4	1	3
7	2.0 + 0.4	1.6	1	2
8	2.5	1.7	37	1^1, 2^2, 3^{34}
9	2.6	1.8	4	3
10	2.4 + 0.4	2.0	4	2^3, 3^1



Lane	Sample ID	Group (kbp)
1	1 kbp Marker	
2	TG21	1.1+0.65
3	TG12-2	1.2
4	TG46-2	1.2+0.7
5	TG17-2	1.3+0.9
6	TG5-1	1.7
7	TG48-4	2.0
8	TG6-3	2.0+0.4
9	TG7	2.5
10	TG56	2.4+0.4
11	TG50-1	2.5
12	TG5-2	2.6
13	pGADT7 vector	7.2+0.8
14	1 kbp Marker	

Figure 11: RE Mapping of prey plasmids. 2% (w/v) gel electrophoresis of Hind III Restriction Enzyme digest of the selected prey plasmid from 10 insert groups. The gel lane number included on the top and the band sizes of the 1 kbp Quanti Marker shown on the left side in kbps. Approximate calculated insert size of each group presented in the bottom of each gel lane Gel map with each group band size included in the right of the Figure 11.

Initial Hind III digest of the TG7 clone showed higher banding pattern than the group 9 clones which is the largest group in the digest results but re-digest of the TG7 showed that it has same banding pattern as the group 9 clones. This clone has re-grouped with the group 9.

Testing for Auto-Activation and False Positive Interacting Protein Partners

Isolated clones did not have a binding domain in them to auto-activate, which was proven by no growth, except for positive control on the DO-Leu/-His, DO-Leu/-Ade, and DO-Leu/-His/-Ade media. If these clones had the ability of auto-activation in similar to the positive control, they should grow in these minimal media (Figure 12).

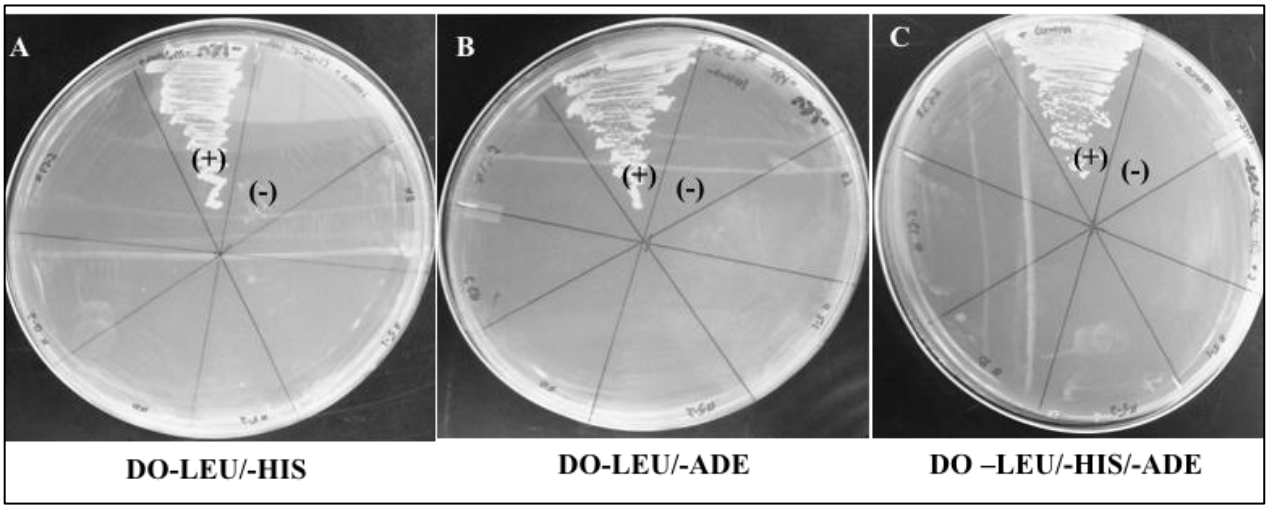


Figure 12: Auto-activation testing of prey cDNA. (A) DO-Leu/-His, (B) DO-Leu/-Ade, and (C) DO-Leu/-His/-Ade media. Positive control (+) and negative control (-).

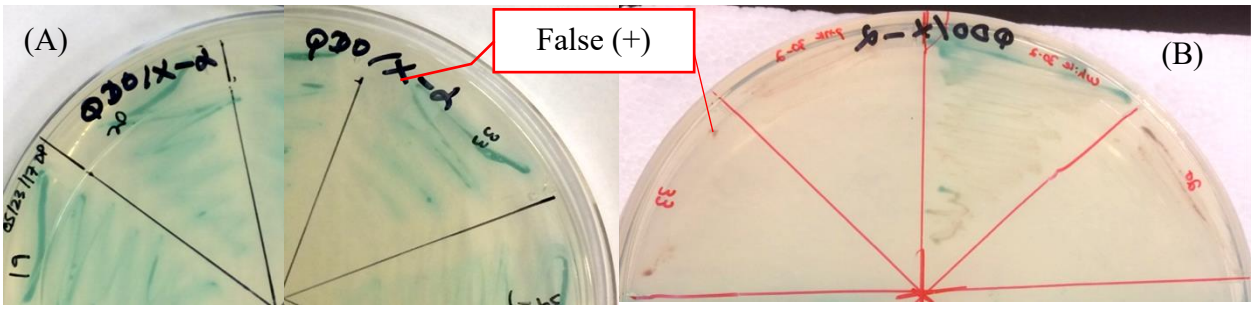


Figure 13: False positive interaction testing in QDO/X/α/Gal. (A) Initial QDO/X/α/Gal streak plates & (B) Co-transformation QDO/X/α/Gal streak plates. The blue color shown in #33 clone on the QDO/X/α/Gal streak plate disappeared after Co-transformation. Among the 30-tested clones, 4 of the plasmids gave false positive results. This was shown by the non-growth on the QDO/X/α/Gal medium.

After the co-transformation of the isolated insert in order to confirm the interaction, it appeared that 4 of the co-transformed plasmids from insert groups were false positives. This reduced the possible interaction partners to 64 (Figure 13). These four inserts were grouped under the group 8 which also have more likely half of the insert in the Hind III digest (Table 5).

Sequencing Results and Analyze of the Interaction Partners

Sequencing results of 30 samples that represent the 10 digest groups and obtained using the 5' T7 and 3' AD sequencing primers were aligned in the Seqman pro software. Reverse complementary sequencing results were created for the sequences resulted from the 3' AD using EditSeq software and used to align with 5' T7 sequencing results. Aligned sequences created 8 Contigs. Among these Contigs, 15 of the inserts were included in the Contig 6 representing the fragment of one full-length DNA sequence (Table 6). The sequences in the Contig 6 were also grouped under 8th Hind III digest group which has the largest inserts (Figure 11 & Table 5).

Table 6: Contig and identification of consensus sequences. Sequence alignment of 30 inserts created 8 Contigs. Consensus sequence of each Contig was used to conduct for the nucleotide BLAST (BLASTn) to find exact protein match that gives zero E-value.

Contig #	Insert Group #	# of Sequences	Insert size (kbp)	Interaction strength	# of Amino acids	Toxo ID (TGME49_)	Predicted protein product
1	1	3	0.4	3	120	286450	GRA5
2	5	3	1.2	2, 3	139	226570	Hypothetical
3	5, 7, 10	3	1.2, 1.6, 2	2, 3	629	215220	GRA22
4	3, 6	2	0.95, 1.4	3	395	203358	Hypothetical
5	9	1	1.8	3	1676	266690	Hypothetical
6	8	15	1.7	3	2933	314890	ThiF Family protein
7	4	1	1.1	2	310	320050	Ribosomal RPL5
8	2	2	0.9	2	256	214290	DJ-1 Family protein

From the sticky protein testing, sequences which were aligned in the Contig 6 resulted as sticky proteins (Figure 14). Hence, the Contig 6 where mostly half of the sequencing samples

was considered as a sticky protein group. All the insert in this contig were grouped under the group 8 of Hind III digest along with the four false positive interacting partners. Interaction strength of Contig 2 & 3 were varied. So, the possible interaction partners for the TgCycY were reduced to 7. Further investigation of these 7 partners in the discussion section is carried to identify true possible interacting partner for TgCycY.

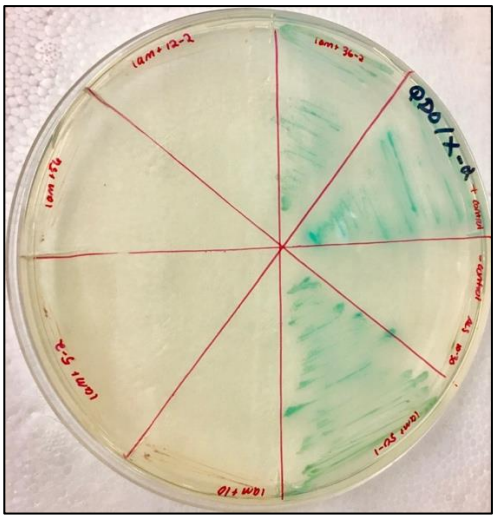


Figure 14: Sticky protein testing of sequenced inserts. Interaction of pGADT7. Insert with pGBKT7.lam tested on QDO/X/α/Gal along with positive control and negative control. Contig 6 insert protein interactions with lam gave bright blue color in the QDO/X/α/Gal and considered as sticky proteins.

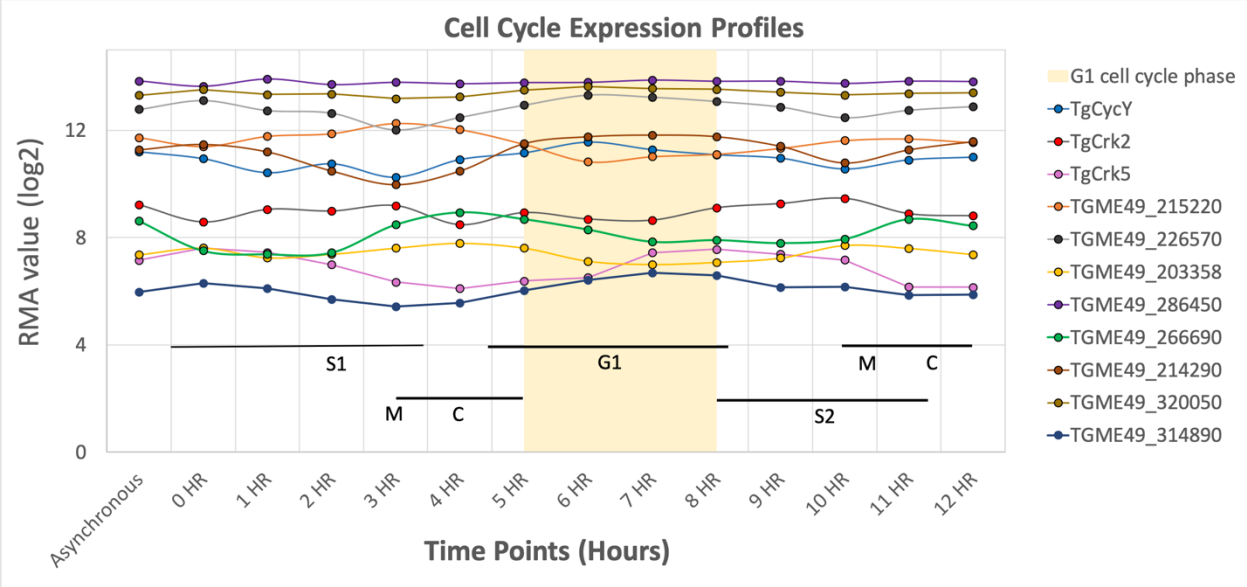


Figure 15: Cell cycle expression profiles comparison with TgCycY. Compared with predicted interacting partner TgCrk2, TgCrk5 and identified protein partners from yeast two hybrid screen. The expression profiles of these proteins were determined hourly and expression data were taken from ToxoDB.org. TgCycY peak expression occurs at G1 phase. Similarly, peak expression of TgME49_214290 (TgDJ-1 family protein), TGME49_226570 (Hypothetical protein), and TGME49_320050 (Ribosomal RPL5 protein) were occurred at the G1 cell cycle phase.

RMA is an algorithm used to create an expression matrix from Affymetrix data. It is a form of quantile normalization that can apply to gene expression or microarray experiments. So, RMA value is an algorithmic value to measure the intensity of the gene expression. In Figure 15, the expression intensity of isolated and identified cDNA compared along with the expression intensity of the TgCycY, TgCrk2, and TgCrk5 in different time points of the cell cycle of tachyzoites. TgCycY peak expression occurred in the G1 phase with a cyclic pattern. TgDJ-1 also indicated a similar expression pattern as TgCycY and the expression levels were close to each other. TgDJ-1 had a peak expression in G1 phase of the tachyzoite cell cycle. Also, TGME49_226570 hypothetical protein and TGME49_320050 ribosomal RPL5 protein had a peak expression at G1 phase of the tachyzoite cell cycle. TGME49_226570 hypothetical protein has similar expression pattern as TgCycY and TgDJ-1 (Figure 15).

TgCrk5 cell expression profile was compared and it had opposite and low level of expression pattern from TgCycY (Figure 15). Even though, TGME49_314890 (ThiF family protein or thioredoxin binding protein) identified as a sticky protein (Figure 14 & Table 6), this false positive protein has similar expression pattern as TgCrk5 and slightly lower expression level than TgCrk5. Also, this ThiF family protein has a peak expression at S phase of tachyzoite cell cycle (Figure 15). Since the interacting partner for the TgCrk5 was not identified in the previous literature, can TGME49_314890 (ThiF family protein) be a potential interacting partner for the TgCrk5? Surprisingly, ThiF family proteins are a thioredoxin family proteins. In the STRING protein interaction predictions, thioredoxin family Trp26 protein predicted as a possible interacting partner for the TgCycY (Figure 7 & Table 1). But the thioredoxin family protein was not identified as a sticky protein in the previous literature. This creates a path to future research on possible sticky protein activity of thioredoxin proteins in *T. gondii*.

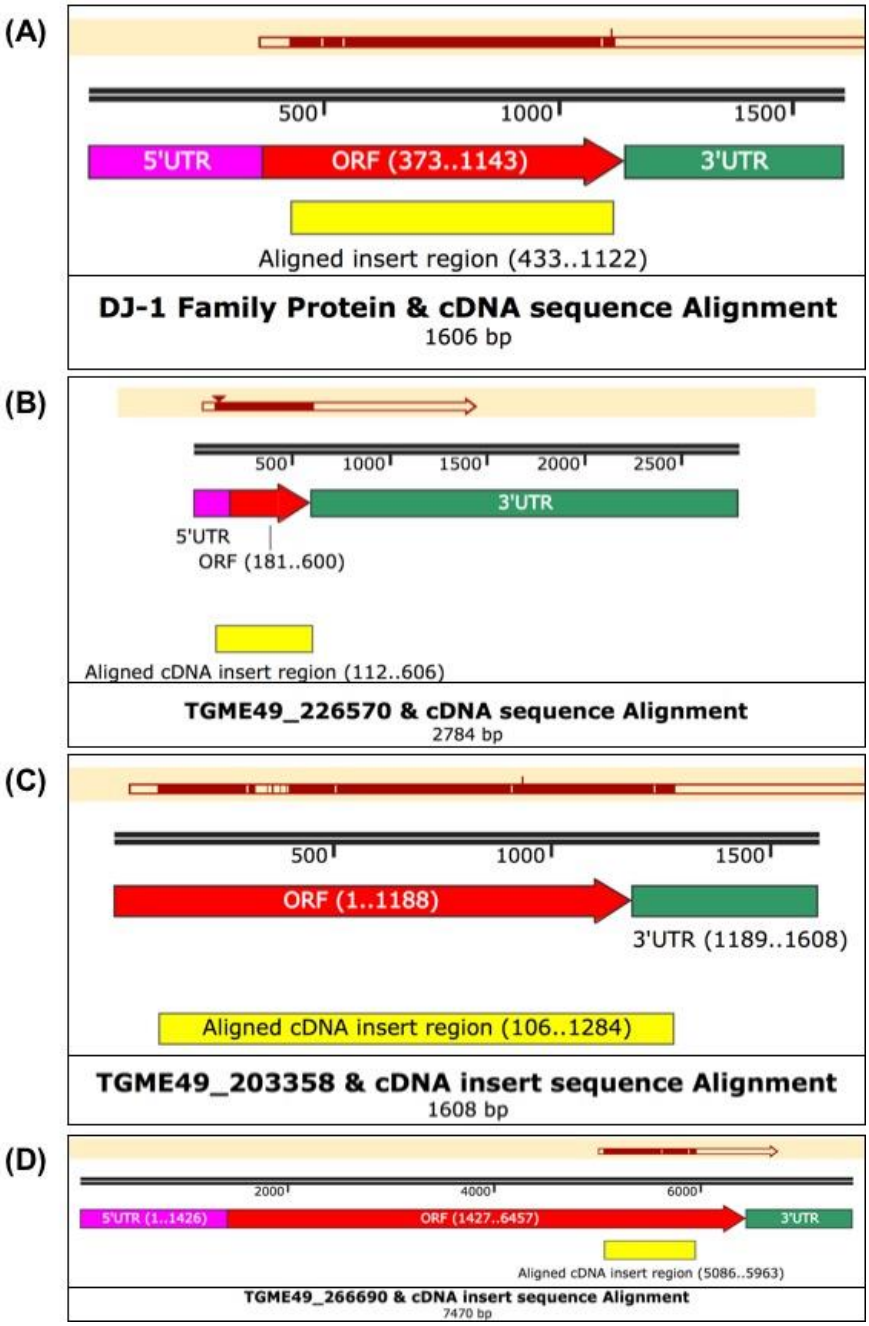


Figure 16: Contig consensus cDNA sequences alignment. Alignment was done with full length nucleotide sequences. (A) Alignment of Contig 8 consensus sequence with DJ-1 family full length sequence. (B) Alignment of Contig 2 consensus sequence with TGME49_226570, (C) Alignment of Contig 4 consensus sequence with TGME49_203358, (D) Alignment of Contig 5 consensus sequence with TGME49_266690, Hypothetical proteins full length sequences. Red area represents the ORF region, purple color area represents 5'UTR, green color area represents the 3'UTR and yellow color area represents the aligned consensus sequence region. Brown arrow is the aligned consensus sequence or cDNA insert.

The cDNA that represents Contig 8 consensus sequence has 1499 bp and only 690 bp of it aligned with the 433 – 1122 bp region of DJ-1 family protein ORF. It did not cover the whole ORF of DJ-1 family protein because it was missing first 60 bp (first 20 amino acids) and last 22 bp of the DJ-1 family protein ORF. Hence, this cDNA insert in Contig 8 was mostly aligned in the C-terminus of the DJ-1 family protein ORF (Figure 16A).

Figure 16B, 16C, & 16D represented the consensus sequence alignment of the three hypothetical proteins identified from the sequencing. Contig 2 consensus sequence has 1837 bp and only 495 bp were aligned with the 112 - 606 bp region of the TGME49_226570 hypothetical protein full-length sequence. This alignment covered the all the bp of TGME49_226570 ORF. Also, it covered the last 69 bp of 5'UTR and first 6 bp of the 3'UTR (Figure 16B). When comparing the cDNA alignment with the full sequence, it appeared that this cDNA was mostly aligned in the N-terminus of the full-length sequence since the ORF region was small compared to the long 3' UTR.

Contig 4 consensus sequence with 2917 bp was aligned with the TGME49_203358 hypothetical protein full length sequence and only 1179 bp of Contig 4 consensus sequence aligned in the 106-1284 bp region of the full sequence. It covered the first 96 bp of the 3'UTR but it did not cover the first 105 bp (first 35 amino acids of the protein sequence) of the ORF. The insert was mostly aligned to the c-terminus of the full sequence (Figure 16C).

Contig 5 consensus sequence with 1761 bp was aligned with the TGME49_266690 hypothetical protein full length sequence and only 878 bp of the Contig 5 sequence aligned in the 5086-5963 bp region of the full-length sequence. It did not cover the first 3659 bp and last 494 bp of the ORF. This cDNA sequence mostly aligned in the 3'terminus of the full-length

sequence. Also, the ORF of this protein is too large when compared to the 5' and 3' UTR. This can be helpful identify the TgCycY interacting region of TGME49_266690. (Figure 16D).

Computational Analysis of TgCycY, TgCrk2 and Ortholog Proteins

Table 7: BLASTp results of ortholog *T. gondii* ME49 Cyclin Y. BLASTp results of *Homo sapiens*, *Drosophila melanogaster*, and *Saccharomyces cerevisiae* using the NCBI Blast database. *T. gondii* ortholog with the lowest E-value included in this Table 7. Red color highlighted ortholog Cyclin Y in *T. gondii* represents TgCycY.

Species	Cyclin Y protein	Ortholog *T. gondii* ME49 Toxo ID	E-value	Score
Homo sapiens	CCNY-isoform 1	TGME49_266900 (TgCycY)	6E-31	116
Drosophila melanogaster	Cyclin Y Isoform A	TGME49_266900 (TgCycY)	1E-28	111
Saccharomyces cerevisiae	PCL1	TGME49_267580 (TgPHO80)	6E-08	52.8

Table 8: BLASTp results of ortholog *T. gondii* ME49 CDK5. In this table, BLASTp results of *Homo sapiens*, *Drosophila melanogaster*, and *Saccharomyces cerevisiae* using the NCBI BLASTp database. These CDK5 protein in these species are interacting partners of their Cyclin Y mentioned in the Table 7. *T. gondii* ortholog with the lowest E-value to the CDK in these species included in this Table 8. Red color highlighted ortholog CDK5 in *T. gondii* represents TgCrk2.

Species	CDK5 protein	Ortholog *T. gondii* ME49 Toxo ID	E-value	Score
Homo sapiens	PFTK1/CDK14	TGME49_218220 (TgCrk2)	2E-94	286
Homo sapiens	PCTK1/CDK16	TGME49_218220 (TgCrk2)	2E-109	326
Drosophila melanogaster	Eip63E Isoform H	TGME49_218220 (TgCrk2)	3E-91	280
Saccharomyces cerevisiae	Pho85	TGME49_218220 (TgCrk2)	4E-113	328

Red color highlighted ortholog *T. gondii* cyclin Y represent the TgCycY cyclin which is the protein of interest in this study (Table 7) while red color highlighted ortholog CDK5 in *T.*

gondii represent the predicted Hypothesized TgCycY interacting partner TgCrk2 (Table 8). Cyclin Y in *Homo sapiens, Drosophila melanogaster, and Saccharomyces cerevisiae* are ortholog to each other while CDK5 in these organisms are ortholog to each other. According to the E-values, Pho85 is the most similar to the TgCrk2 and CCNY Isoform 1 in *Homo sapiens* is the most similar to TgCycY since they have lowest E-value (Table 7 & 8).

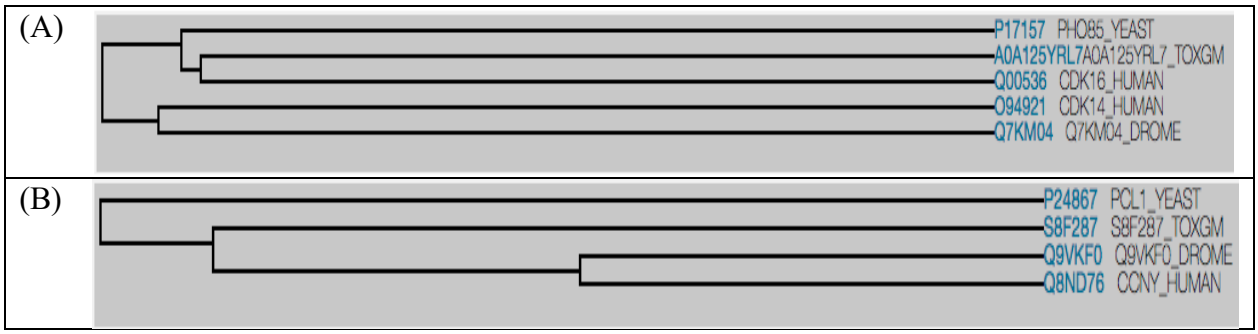


Figure 17: Phylogenetic tree analysis of Cyclin Y and CDK5. Comparison done with *T. gondii, Homo sapiens, Drosophila melanogaster, and Saccharomyces cerevisiae*. (A) Phylogenetic tree analysis of CDK5 in above mentioned species. These CDK5 proteins are interacting protein partners of their corresponding Cyclin Y (Fig.16B) except Toxo CDK5. *T. gondii* CDK5 is predicted as the interacting partner of TgCycY based on the String analysis (Figure 7 & Table 1). (B) Phylogenetic tree analysis of Cyclin Y in *T. gondii, Homo sapiens, Drosophila melanogaster, and Saccharomyces cerevisiae*.

According to the Figure 17(A), TgCrk2 is most similar to the Pho85 in *Saccharomyces cerevisiae* and CDK16 in *Homo sapiens* Cyclin Dependent Kinases while TgCycY is most similar to the *Homo sapiens* and *Drosophila melanogaster* Cyclin Y (Figure 17B). When comparing TgCycY and TgCrk2 with Cyclin Y and CDK5 in unicellular *Saccharomyces cerevisiae* and multicellular *Drosophila melanogaster* and *Homo sapiens*, TgCycY is not ortholog to the Cyclin Y in *Saccharomyces cerevisiae* while TgCrk2 ortholog to the Pho85 in *Saccharomyces cerevisiae*. TgPHO80 which is *T. gondii* ortholog of the PCL1 protein is the only cyclin that localized to the tachyzoite cytoplasm and it is predominantly interact with the TgCrk2 for the continuation of G1 phase (Alvarez & Suvorova, 2017).

Results Summary

Yeast two hybrid screen of TgCycY identified only DJ-1 family protein as true potential interacting protein partner since from rest of the identified partners, ThiF family protein excluded as a sticky protein, GRA5, GRA22, and RPL5 proteins excluded due to their cellular localization and other three hypothetical proteins excluded due to their lack of identification or information (Table 6). The DJ-1 family protein cell cycle expression profile indicated a similar pattern and similar level expression as the TgCycY throughout the cell cycle which might indication of possible interaction between TgCycY and DJ-1 family protein (Figure 15).

Computational data of predicted protein interactions of TgCycY point to TgCycH, TgCrk1, TgCrk2, TgCrk4, TgCrk5, TgCrk6, thioredoxin family Trp26 protein, and a putative protein cyclin dependent kinase as 8 possible interactions with different interaction scores (Figure 7 & Table 1). Based on the previous experimental results of direct interaction with TgCrk1, TgCrk4, and TgCrk6 eliminate the TgCrk1, TgCrk4, TgCrk5 and TgCrk6 as potential interacting partners for TgCycY due to expression cell cycle phase and non-nucleus localization (Figure 4). Interaction with TgCrk5 was not tested it is expressed in the S phase (Alvarez & Suvorova, 2017). Interaction with TgCycH, thioredoxin family Trp26 protein, and other putative CDK were disregarded from the interaction partners list due to lack of information (Figure 7 & Table 1). Based on the analysis of homology of TgCycY and TgCrk2 in higher eukaryotes and computational predictions, TgCrk2 can suggest as a potential indirect interacting partner for the TgCycY since TgCrk2 expressed in G1 phase and not having a direct interaction with the TgCycY (Figure 5, Figure 17, Table 7, & Table 8).

Chapter 5: Discussion

Expression Vector Construction

Constructing a pGBKT7.TgCycY expression vector was essential to conduct the yeast two-hybrid screening to find an interacting partner for the TgCycY. PCR cloning with designed primers were used due to the incompatibility of identified approximate Cla I and XhoI RE clone sites in the pGADT7-Multiple Cloning Site (MCS) with the pGBKT7-MCS. Prevention of concatemers by rSAP de-phosphorylation of pGBKT7, usage of the compatible buffers in the sub-cloning procedure, and compatibility of reading frames and incompatibility of the sticky ends of RE digests co-operated for a successful pGBKT7.TgCycY clone construction. The success of the clone construction was proved by not creating auto-activation and Nde I, EcoRI RE digest. In general, fusion protein loses the auto-activation properties when it is fused in frame with a vector plasmid. If the cloned DNA nucleotide sequence is not in frame with the ORF of the plasmid, it will change the ORF and as a result the conformation of the protein translated from the newly constructed plasmid will be different and its functions will be changed. Cloning TgCycY in both pGBKT7 and pGADT7 enables an establishment panel to score new targets in the mating experiment. Single transformation of pGBKT7.TgCycY fusion protein in clarification of constructed pGBKT7.TgCycY bait plasmid method proved its inability to initiate transcription of HIS3 and ADE2 reporter genes which proved that constructed pGBKT7.TgCycY protein does not have the auto-activating function. (Criekinge & Beyaert, 1999).

Yeast Two-Hybrid Method

In Y2H Gold screen, the fusion of diploid yeast strains that contained bait & prey plasmids encoding for GAL4-binding and activating domains are eukaryotic split transcription factors. The GAL4 - binding domain binds to the transcriptional activator binding site. If the prey

protein in the GAL4-activating domain interact with the bait protein, GAL4-activating domain can initiate transcription of HIS3 and ADE2 reporter genes (Figure 9). They allow the diploids to grow in QDO reporter medium by activating the transcription of the HIS3, and ADE2 reporter genes (Stephens & Banting, 2000). This system is testing interactions of TgCycY against a large library of proteins (cDNA) at once, but there are limitations to this method. Improper protein folding and post-translational modifications that cause false positive and negative interactions are the concerns (Figure 13 & 14). In order to minimize the false negative results, yeast two hybrid screen was replicated two more times, but these two additional repetitions of the screen were not successful. Due to time deficiency, we continued the experiments with the results that we obtained from the first screen. Also, prey plasmids were constructed using a tachyzoite cDNA library and it will minimize the false negative results since there are no full length ORFs in the prey library (Brückner, Polge, Lentze, Auerbach, & Schlattner, 2009). PGBKT7. P53 + pGADT7. T positive control assayed the fusion proteins with correct conformations since artificial fusion proteins embody a potential risk by changing the conformation of bait and prey with altering functionalities. These artificial proteins may not be in proximity within the cell, nor are they expressed at the same time point of the cell cycle or localization in the same subcellular compartments. Screening over 1 million haploids for the prey cDNA library was evidence that one out of six cDNAs in the library is in the correct reading frame that can create a true interaction (Criekinge & Beyaert, 1999). The success of true interaction screening depended on the limiting partner/cDNA library viability, resulting in 3.9 million clones in the screen (Criekinge & Beyaert, 1999; Table 2 & 3). Improper folding and steric hindrance of short proteins may create false negative & false positive interactions, but replication of yeast two-

hybrid screen can reduce the probability of false negatives and false positive (Coates & Hall, 2003).

Therefore, it is expected that the number of screened interactions from initial colony screening to cDNA insert isolation can be decreased step by step by identifying false positives (Figure 13), sequence similarity (Table 6), and sticky protein testing in Figure 14. Hind III RE digest was one of the steps that narrowed down the potential interacting cDNA clones into 10 groups based on the size-similarity of interacting clones (Figure 11 & Table 5). Only two Hind III RE sites located in the pGADT7 prey vector plasmid and multiple cloning site (MCS) of pGADT7 is located in between these two Hind III RE sites. Reducing the size of the pGADT7 from the total size of the fragments identified in the Hind III RE enzyme gel electrophoresis can identify the approximate size of the cDNA insert cloned in the pGADT7 prey plasmid. The size grouping of the isolated inserts in the Y2H screen is done based on calculated approximate insert size (Figure 11 & Table 5). These results cannot a that the insert in a same group have a same protein sequence. The RE enzyme digest procedure was conducted to narrow down the isolated insert from the Y2H screen based on the approximate cDNA insert size. The confirmation of sequence similarity was conducted by the sequencing the inserts that represent the digest groups (Table 6).

The strength of the protein-protein interaction predicted by the yeast two-hybrid screen correlates with the discrimination of high, intermediate, and low-affinity interactions (Figure 10 & Table 4). The issue is an indication of strong sticky protein interactions with TgCycY among the actual interactions, which was shown by the strong blue color which was occurred due to hydrolyzation of X-alpha-Gal by alpha-galactosidase that encoded by the transcription of MEL1 reporter gene (Figure 14). Weak interactions can be detected in the yeast two-hybrid screen since

the genetic reporter gene strategy results in significant amplification. Also, the size of the colonies can be evidence for the strength of interaction though it cannot be assumed that weak interactions are biologically less significant than strong interactions. Since transcription initiation due to auto-activation is present in 5% of proteins, including the cDNA libraries, it is important to test the auto-activation activity of cDNA inserts to eliminate potential false positives. However, obtained cDNA inserts that interacted with TgCycY did not contain any auto-activation activity in them (Criekinge & Beyaert, 1999; Figure 12).

A gradual decrease in cDNA clones in the yeast two-hybrid screen process ensures an accurate screening of true interactions since the yeast two-hybrid system is prone to create a high percentage of false positives. Even though RE digest mapping of the cDNA is a method that can be used to identify the similarity of the isolated cDNA clones' sizes, it cannot provide the actual sequence similarity between the cDNA clones to identify the sequences that represent the same protein translated from one mRNA (Figure 11 & Table 5). Therefore, DNA sequencing is a method that can be used to derive much more information about the cDNA sequences by classifying the inserts into Contig (Table 6).

Protein Interacting Partners

According to the BLASTn results of the sequencing data that obtained from Toxodb database, which is the genome database of genus *Toxoplasma*, seven possible protein interaction partners were identified. Among these interacting partners, two proteins are dense granule protein (GRA5 & GRA22), three are hypothetical proteins, one is a ribosomal protein (RPL5), and the last one is a DJ-1 family protein (Table 6).

TGME49_286450 & TGME49_215220: In Toxodb, TGME49_286450 (GRA5) protein is mentioned as a dense granule protein. Even though TGME49_215220 protein is mentioned as

a hypothetical protein in Toxodb.org, it is considered as a GRA22 protein in a peer-reviewed publication (Guiton, Sagawa, Fritz, & Boothroyd, 2017). GRA5 protein is associated with the parasitophorous vacuole membrane. It can delineate the PV membrane and lead the inhibition of host cell apoptosis for the long-term residence of the intracellular parasite. Also, based on the motif finder, it has bacterial Ig-like domain 1 and an arginine-rich region. It is predicted to interact with the *T. gondii* p30 surface antigen which is very abundant and important for the surface localization (Tomavo, 1996). GRA22 protein contains a TOBE-2 domain which is involved in recognition of small ligands such as molybdenum and sulfate. This domain is important for the transmembrane movement of substance and transport activity. Also, it is predicted to interact with the TGME49-305860 (Calcium-dependent protein kinase/CDPK3). When considering the cell cycle expression profiles of these two proteins, TGME49_215220 has the lowest expression in the G1 phase while TGME49_286450 expression mostly remained consistent throughout the cell cycle (Figure 15). But these proteins were removed from the list of true interacting partners due to their localization in the parasitophorous vacuole membrane which is far away from the nucleus where TgCycY localization has occurred. These two proteins may be false positive interactions since false positive interactions in the yeast two-hybrid screen are mostly due to inconsistent localization of the proteins.

TGME49_226570, TGME49_203358, & TGME49_266690: According to the toxodb.org, TGME49_226570, TGME49_203358, and TGME49_266690 are hypothetical proteins. TGME49_226570 does not have any predicted protein interactions and nor it does have any predicted motifs in the structure. So, it is hard to identify TGME49_226570 function in the *T. gondii*. In silico analysis shows TGME49_203358 protein has TFIIE beta subunit core domain in it and it is one of the subunits of the transcriptional factor TFIIE which is needed for the

initiation of eukaryotic mRNA transcription from RNA polymerase II promoter. Also, STRING database prediction showed GTF2E1 and TAF7L transcription initiation factors as predicted interaction partners for TGME49_203358 hypothetical protein. This indicates that TGME49_203358 might have a role involved in transcription initiation. According to the Pfam database, a Golgi subfamily A member 5 domain is identified in the TGME49_266690 hypothetical protein. This domain is involved in the Golgi organization and intra-Golgi retrograde transport. Predicted protein interactions are not identified for this protein. There is no experimental or computational evidence for the interactions of these proteins with the CDK and according to the lack of current evidence, it is hard to show the function of these proteins with the interaction of TgCycY.

TGME49_320050: TGME49_320050 is the Ribosomal Protein L5 and it has regions of ribosomal large subunit proteins 60s L5 and 50S L18. This protein is involved in binding 5S RNA to the large ribosomal subunit. Apart from protein synthesis, many ribosomal proteins involved in other cellular functions and these proteins are shown to affect the mechanism of the development and apoptosis. Ribosomal proteins can alter cell cycle fate by interacting with the Cyclin/CDK complexes. As examples, human L34 inhibits the cell cycle proteins CDK4, CDK5 while L26 binds to 5' UTR of p53 mRNA to increase translation of p53. Also, eukaryotic ribosomal protein L5 functions in p53 mediated neuronal apoptosis (Bhavsar, Makley, & Tsonis, 2010). P53 is a tumor suppressor and it responds to various stress signals that induce apoptosis. CDK5 interacts with the p53 and increases stability through post-translational regulation. CDK5 disrupts the interaction between p53 and Mdm2 (E3 ubiquitin-protein ligase). This interruption prevents Mdm2 induced p53 ubiquitylation and downregulation (Lee, Kim, Lee, & Kim, 2007). This RPL5 protein localized in the cytosol and based on this localization we can exclude RL5

protein form the protein interaction partners list since yeast two-hybrid screen can detect the interactions that occur in the nucleus.

TGME49_214290: TGME49_214290 is a *T. gondii* DJ-1 family protein (TgDJ-1). It has a localization in the cytoplasm, mitochondria, and the nucleus (Junn, Jang, Zhao, Jeong, & Mouradian, 2009). This protein has B12 binding domain, DJ-1 family protein domain, TATA box-binding protein (TBP) domain, CobB/CobQ-like glutamine aminotransferase domain which involved in regulation of cell division and a papain-like protease domain. The potential proteolytic function of this papain–like domain in TgDJ-1 is not well understood (Hall, et al., 2011).

In higher eukaryotes, DJ-1 protein has a proposed role in the regulation of oxidative stress since it interacts with the master regulator of the oxidative stress response, Nrf2 transcription factor. Protein-specific cellular function is still unclear in higher eukaryotes, but it is associated with the Parkinson disease and cancer in humans. In plants, Calcium-dependent kinase and abscisic acid (ABA) play a major role not only in host cell invasion pathways but also in stress response. The relationship between TgDJ-1 and ABA is not explained yet. Previous study identified that there is a plant-like pathway in *T. gondii* for controlling cyclic ADP ribose (cADPR) production through ABA (Nagamune, et al., 2008). This ADPR is important for Ca^{2+} signaling to initiate the micronemes secretion. ABA accumulated during intracellular replication triggers the increase of cADPR to elevate the Ca^{2+} that needed for the signaling pathway of micronemes secretion and egress. Blocking of this pathway turn on the development switch to slow growing tissue cysts (Nagamune, et al., 2008). The TgDJ-1 controlled microneme secretion in *T. gondii* is needed for the egress of the tachyzoites. Egress is a crucial phase for the *T. gondii* lytic cycle. Egress from the host cell is hypothesized as the largest stress in the *T. gondii* and

other intracellular parasites life cycles (Hall, et al., 2011). *T. gondii* egress and differentiation of tachyzoites to bradyzoites can induced by the IFN-γ in human foreskin fibroblasts. However, the egress event is still poorly understood even though the different molecules shown to involved in this process (Caldas & Souza, 2018).

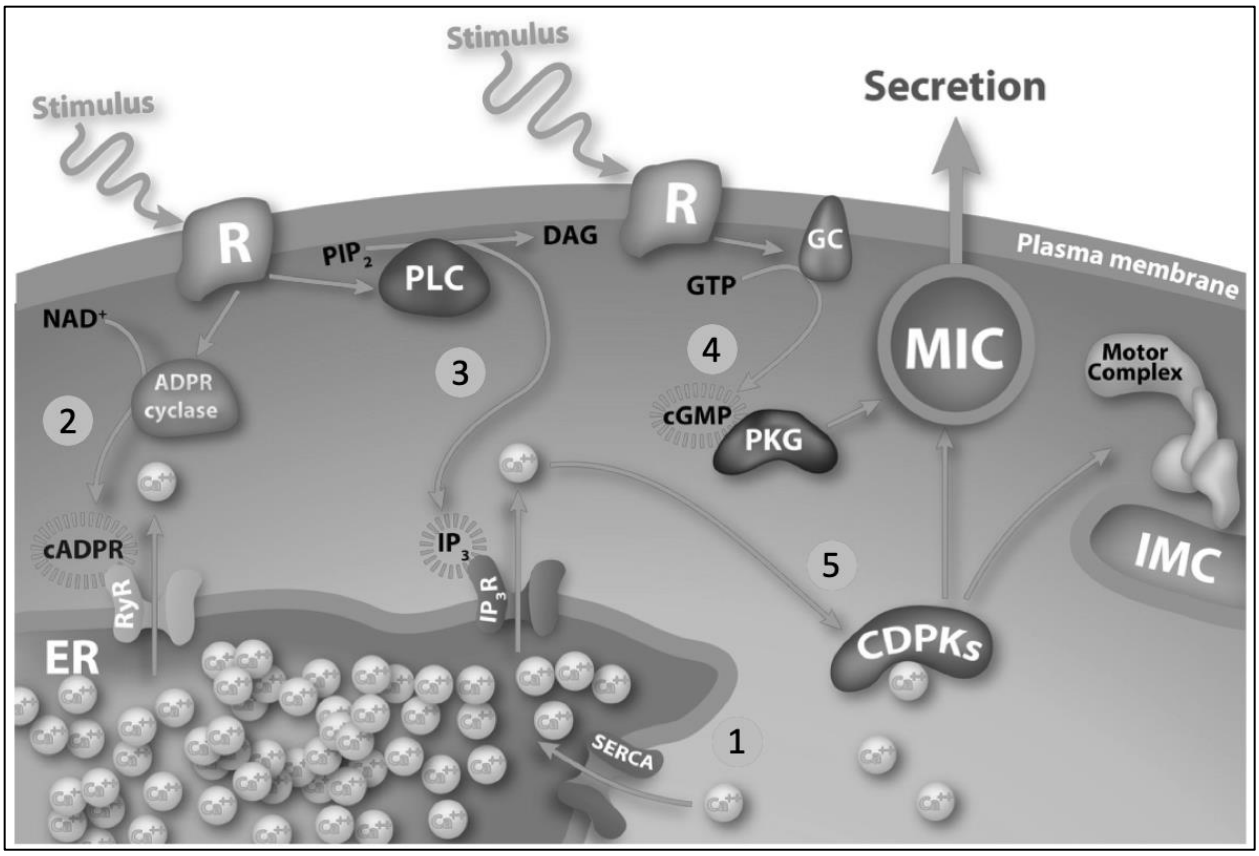


Figure 18: Model *T. gondii* Ca^{2+} signaling pathway for microneme secretion. (1) Storing cytoplasmic Ca^{2+} into the mobile Ca^{2+} storage, Endoplasmic reticulum (ER) through SERCA ATPAse. (2) Generation of cyclic ADP ribose (cADPR) to activate Ca^{2+} from ER by activation of ADP ribose cyclase (ADPR cyclase) due to identification of stimulation from the host. (3) Generation of Inositol triphosphate (IP_3) and diacyl glycerol (DAG) from phosphatidyl inositol bisphosphate (PIP_2) to release Ca^{2+} from ER. (4) Activation of guanylyl cyclase (GC) to generate cyclic GMP (cGMP). cGMP activates protein Kinase G (PKG) to activate microneme secretion without binding to the Ca^{2+}. (5) TgDJ-1-TgCDPK1 activates with the binding of Ca^{2+} to activate microneme secretion from microneme (Billker, Lourido, & Sibley, 2009).

Based on the previous studies, TgDJ-1 protein is involved in regulating exocytosis of micronemes by joining with a virulence-associated organelle in *T. gondii*. *T. gondii* micronemes secretion powered by the glideosome protein complex which modulated by Ca^{2+} dependent signaling and phosphorylation. Glideosome complex is a macromolecular complex that consists adhesive proteins which released apically and translocated to the posterior pole of the parasite (Keeley & Soldati, 2004). Not only micronemes secretion is tightly controlled by a Ca^{2+} dependent pathway which involved TgCDPK1, but also control of micronemes secretion has a Ca^{2+} independent pathway which controlled by *T. gondii* cyclic-GMP dependent kinase (TgPGK) (Figure 18). These two pathways are independent from each other. Mostly, this Ca^{2+} dependent pathway needed Ca^{2+} signaling cascade for the micronemes secretion and the releasing of Ca^{2+} to the cytoplasm mostly occurred through ryanodine receptors (RyR) and inositol triphosphate receptors (IP_3R) in the endoplasmic reticulum (ER) which is the primary mobilize store for Ca^{2+} in *T. gondii*. This mobilized store filled by the cytoplasmic Ca^{2+} through the action of Ca^{2+} ATPase called SERCA. Also, *T. gondii* has other Ca^{2+} mobilizing stores. The releasing of Ca^{2+} from ER initiate by the cADPR and inositol triphosphate (IP_3). These molecules are generated from the stress stimulations identified by the unidentified receptors in the plasma membrane. In the Ca^{2+} independent pathway, PKG activated by cGMP upon receiving the stimulation to the unidentified receptors in the plasma membrane (Figure 18). This means that even the Ca^{2+} dependent pathway is inhibited, micronemes secretion can happen using the Ca^{2+} independent pathway. This Ca^{2+} mobilizing and its fluctuations in the *T. gondii* were not very well studied (Billker, Lourido, & Sibley, 2009).

TgDJ-1 is interacting with the *T. gondii* Calcium-Dependent Protein Kinase 1 (TgCDPK1) to regulate the TgCDPK1 mediated signaling that is necessary for the micronemes

secretion regulation (Child, et al., 2017). The micronemes secretion initiation from TgCDPK1 needed intracellular Ca^{2+} binding to it. The model molecular pathway for this process mentioned that the TgDJ-1 associates with the phosphorylated TgCDPK1 under the reducing conditions. This inactive complex activated by binding of cytoplasmic Ca^{2+} with the TgCDPK1/TgDJ-1 complex (Figure 18). This complex is needed to maintain normal micronemes secretion of *T. gondii*. Dissociation of this complex occurs via the reactive oxygen species (ROS) mediated signal which is possibly be H_2O_2 and leave the TgCDPK1 to re-associate with a pool of TgDJ-1 protein under reducing conditions. High Ca^{2+} levels can knockdown the association of TgCDPK1 and TgDJ-1 but also the Ca^{2+} needs to function synergistically to potentiate the TgCDPK1 activity (Child, et al., 2017; Billker, Lourido, & Sibley, 2009). Inhibition of micronemes secretion can affect the motility and egress but there is no evidence that micronemes inhibition effect *T. gondii* cell growth by arresting the cell growth. TgDJ-1 has a Cys127 residue which can result in inhibition of micronemes secretion, motility and invasion via the modification of cys127 residue by WRR-086 which is covalent inhibitor compound, even in the presence of the direct stimulators of Ca^{2+} release. This has shown that TgDJ-1 as a regulator of *T. gondii* secretion, attachment and invasion and suggested that TgDJ-1 is downstream of Ca^{2+} signaling cascade that leads to invasion of host cells (Hall, et al., 2011).

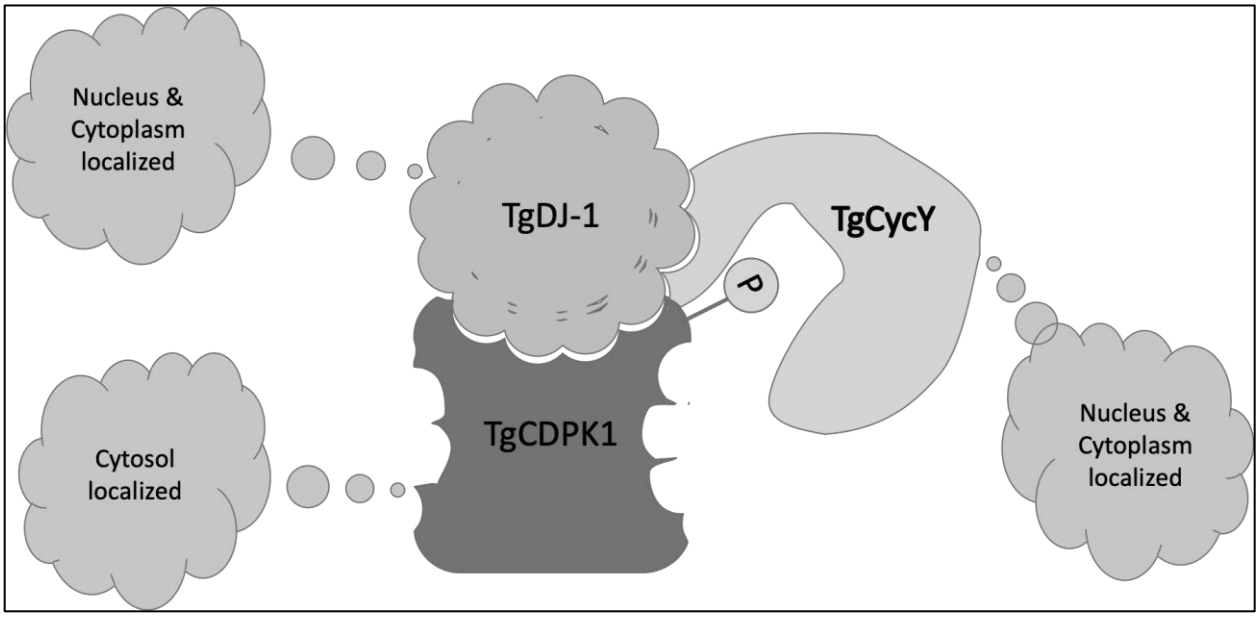


Figure 19: Suggested TgCycY/TgDj-1/TgCDPK1 interaction model. TgCycY interacts with the TgDJ-1 protein to localize the TgDJ-1 to the TgCDPK1. TgCycY acts as a substrate for the TgCDPK1 and get phosphorylated.

Also, there is a suggestion that TgDJ-1 protein has a "non-canonical kinase-regulatory scaffold" that can integrate multiple intracellular signals to carry on the exocytosis of micronemes and motility. Even though the TgDJ-1 protein has a role in regulating parasite secretion, attachment and invasion in *T. gondii*, higher eukaryotes do not provide any clues to support the possible mechanism of the DJ-1 protein (Hall, et al., 2011). The results of this study have shown that a protein interacting partner of TgCycY is TgDJ-1 (Figure 15 & Table 6). This can be an indication that TgCycY might involve in the micronemes secretion pathways. This makes the suggestion that TgCycY protein interact with the TgDJ-1 to localize TgDJ-1 to the *T. gondii* Calcium dependent protein kinase 1 (TgCDPK1). TgCycY protein acts as a target protein in this protein complex and phosphorylated by the TgCDPK1 in TgDJ-1/TgCDPK1 complex (Figure 19). So, TgCycY interacts with the TgCDPK1 while TgDJ-1 is acting as mediating

protein of the interaction. If the TgCycY is phosphorylated according to this study, what will happen to the phosphorylated TgCycY? Whether this phosphorylated TgCycY will have another role *in T. gondii*? These are some questions that need to answer in a future study by proving this suggested model. When considering the TgCycY ortholog value to the *Homo sapiens* Cyclin Y (Figure 5), TgCycY is most similar to *Homo sapiens* Cyclin Y. If this suggested protein interaction model is proved and designed a drug based on the model, it should make sure that designed drug will not inhibit the *Homo sapiens* Cyclin Y by having low affinity to human Cyclin Y and DJ-1 proteins. Also, inhibition of ABA/cADPR pathway turn on the switch of tachyzoite differentiation to bradyzoite. During the bradyzoite induction, TgCycY expression is upregulated (Huang, et al., 2017). When tachyzoite sense the immune stress, it needs to start differentiating into the bradyzoite to remain alive in the infected cells and also it needs to invade other cells to keep spreading the infection to healthy cells. So, this can suggest that TgCycY/TgDJ-1 complex and ABA signaling pathways might have a relationship to control the microneme secretion for the egress and differentiation into the bradyzoites. This insight can suggest that TgDJ-1/TgCycY upregulation can affect the ABA pathway in cADPR upregulation to switch on the bradyzoite induction.

Previous studies showed that knockdown of TgCycY cyclin does not create the cell growth arrest which leads to decide that TgCycY is not essential for the tachyzoites cell cycle (Alvarez & Suvorova, 2017). There is no evidence that knockdown of TgCycY arrests the secretion, motility, and invasion. Also, TgCycY does not have prior evidence involving in the micronemes secretion process. Furthermore, there is no experimental or computational evidence that TgCDPK1 interacts with the TgCycY. If the TgCycY/TgDJ-1 complex involved in a Ca^{2+} signaling cascade, the knockdown of TgCycY might not affect the micronemes secretion since

there are Ca^{2+} independent pathway to initiate the micronemes secretion. But this Ca^{2+} independent signaling pathway is also inhibited when the TgDJ-1 is inhibited because TgDJ-1 inhibition can inhibit microneme secretion even in presence of direct stimulators for Ca^{2+} release (Child, et al., 2017; Billker, Lourido, & Sibley, 2009). Hence, these facts raise the questions whether knockdown of TgCycY affects the micronemes secretion, motility, and invasion of *T. gondii* and if it does how the Ca^{2+} independent signaling pathway inhibited in this situation. If there is a relationship between TgCycY and micronemes secretion, motility, and invasion, the TgCycY might have a role of outside the regulating cell cycle. When considering the higher eukaryotic homology Cyclin-CDK5 to the *T. gondii* (Figure 5, Figure 17, Table 7, & Table 8), an interaction between TgCycY and TgCrk2 can be predicted. Also, PCL1 and PHO85 Cyclin Y/CDK5 complex is able to phosphorylate the substrate of CDK5 protein. According to the human Cyclin Y ability to get phosphorylate by the human Cyclin Y/PFTK1 complex, human Cyclin Y can function as a substrate for its own complex. The interaction between TgCrk2 and TgPHO80 is already proven (Alvarez & Suvorova, 2017). They are the ortholog Cyclin Y and CDK5 for the yeast. By combining these facts, we can suggest that TgCycY is interacting with the TgPHO80 to localize it to the TgCrk2 in the nucleus. So, TgCycY can interact with the TgCrk2 with the mediation by the TgPHO80 protein and TgCycY can get phosphorylated. But in order to confirm this suggestion the direct interaction between TgCycY and TgPHO80 needs to be tested.

TgCycY and TgCrk2 expressed in the same G1 phase of the tachyzoite cell cycle and these two proteins does not have a direct interaction (Alvarez & Suvorova, 2017). Based on the ortholog indirect cyclin Y and CDK5 protein interaction mediated by a third protein in *Homo sapiens* (Liu, Guest, & Finley, 2010), TgCycY might interact with TgCrk2 indirectly with a help

mediating protein. Even though there is no direct interaction between these two proteins, the indirect interactions between TgCycY and TgCrk2 need to be tested to prove that TgCycY does not have function in tachyzoite cell cycle regulation. Also, if there is an indirect interaction between TgCycY and another protein that interact with TgDJ-1, depth research is needed to identify a relationship of this new complex in microneme secretion. The mechanisms of micronemes secretion and egress are poorly studied in *T. gondii* and if we could relate the TgCycY- (TgDJ-1) interaction to the micronemes secretion, it will be a window to the understanding of *T. gondii* micronemes secretion and egress mechanisms. Also, we can identify the role of TgCycY/TgDJ-1 in tachyzoites and introduce this complex as a drug target to prevent the microneme secretion, motility, and invasion of tachyzoite.

Computational Insights into the TgCycY Interactions

TgCycY computational interaction predictions (String Database) indicated the TgCrk1, TgCrk2, TgCrk4, TgCrk5, TgCrk6, a putative CDK (TGME49_239910) and thioredoxin family Trp26 protein as interacting partners (Figure 7 & Table 1). Even though there are no direct interactions between TgCycY and these Crks except TgCrk5, the indirect interactions were not identified. TgCrk5 interaction with TgCycY was not tested but this TgCrk5 form a stable complex that is essential for chromosome replication 1 (ECR1) during tachyzoites replication and this complex localized into the centrocone. TgCrk5 has a maximum expression level at S phase and during the G1 phase expression level is undetectable. It is also suggested that periodic expression of TgCrk5 may cause a lack of cyclin partner (Naumov, et al., 2017). The interaction between TgCrk1, TgCrk4, and TgCrk6 might not occur with the TgCycY since they are expressed during different cell cycle phases such as S and M/C and their expression levels are low at G1 phase. Even though the interaction between TgCrk5 was not tested, this logic can be

applied to the TgCrk5 since it is expressed during the S phase. Only the TgCrk2 is expressed throughout the cell and there are no huge differences in the expression throughout the cell cycle. Also, the loss of TgCrk2 blocks the G1 phase (Alvarez & Suvorova, 2017).

Computational predictions indicated that TgCycY interacts with the TgCycH which has a weak interaction with the TgCrk2 (Alvarez & Suvorova, 2017). This might be an indication of indirect interaction between TgCycY and TgCrk2. Thioredoxin family Trp26 protein is an important protein in the reduction of ribonucleotides to deoxy-ribonucleotides. This protein is localized in the cytosol. Hence, the interaction between this protein and TgCycY cannot be detected in this yeast two-hybrid screen. The gene of this protein is not mapped to any metabolic pathways in KEGG. In the TDR target database, the Thioredoxin family Trp26 protein was indicated as a non-essential protein. There is no experimental evidence for the interaction between TgCycY and Thioredoxin family Trp26 protein but TGME49_314890 ThiF family protein was identified as a sticky protein interacting partner for TgCycY in this screen since ThiF family proteins are thioredoxin proteins (Table 6).

When comparing cell cycle expression profiles of TgCycY and TgDJ-1 proteins, they show similar expression pattern throughout the cell cycle and the peak expression occurs at the G1 phase for both proteins. TgDJ-1 has slightly higher expression levels than TgCycY. This can be proof for the possibility of interaction between TgCycY and TgDJ-1 proteins. Among the three hypothetical interacting partners, only TGME49_226570 has a peak expression at the G1 phase and the expression pattern similar to the TgCycY and TgDJ-1. Also, the expression level is higher than both TgCycY and TGDJ-1. Other two hypothetical protein interacting partners have peak expression at the M/C phase. When considering the cell cycle expression phases, the possible true interacting partners reduced to TgDJ-1 and TGME49_226570 hypothetical protein.

TgCrk2 expression level is mostly kept consistent throughout the cell cycle (Figure 15). The expression profiles of these proteins are not enough to conclude the relationship between TgCycY, TgDJ-1, and TgCDPK1.

Results from Contig consensus sequences alignment with their corresponding full-length sequence using Snapgene was done to understand the coverage of cDNA in the full-length sequence. These results revealed that the cDNA library used in the experiment has 3' end biased inserts because of the oligo-dT primer used to create the cDNA library. TgDJ-1 cDNA insert alignment missed the first 60 bps (20 amino acids of the ORF which revealed that interacting region of DJ-1 cDNA with TgCycY lies in the C-terminus. But TGME49_226570 hypothetical cDNA insert has covered the whole ORF and last 69 bp of 5'UTR and first 6 bp of 3'UTR. This is due to the small size of the ORF of TGME49_226570. This cDNA insert is mostly aligned in the N terminus. cDNA of TGME49_203358 hypothetical protein also aligned in the c-terminus of the sequence and it did not cover the first 105 bp of ORF. cDNA of TGME49_266690 also C-terminus aligned because it did not cover the first 3659 bp and last 494 bp of the ORF. TGME49_266690 has the largest ORF among these four proteins. So, the interacting region of this hypothetical protein lies close to the C-terminus (Figure 16). This yeast two-hybrid screen of TgCycY did not give the hypothesized TgCrk2 protein as an interacting partner. This may be since the interacting region of TgCrk2 lies in the N-terminus of the sequence or the cDNA insert library might not have the TgCrk2 cDNA insert with this N-terminus region.

Homology analysis of TgCycY and TgCrk2 in Figure 5, Figure 17, Table 7, & Table 8 revealed that other eukaryotes have ortholog Cyclin Y that interact with the CDK5. *Drosophila melanogaster* Cyclin Y Isoform A which is ortholog cyclin Y to the TgCycY interacts with the Ecdysone-induced protein 63E (Eip63E Isoform H) which is ortholog of the TgCrk2. Even

though in *Saccharomyces cerevisiae*, PCL1 Cyclin Y type protein ortholog to the Cyclin Y Isoform A in *Drosophila melanogaster* and CCNY in *Homo sapiens*, it does not ortholog to the TgCycY in *T. gondii*. PCL1 ortholog in *T. gondii* indicated to be TgPHO80 P-type cyclin (Figure 5 & Table 7). This TgPHO80 has a dominant interaction with TgCrk2 and needed for the control of checkpoint at the G1 phase of the cell cycle (Alvarez & Suvorova, 2017). In this analysis, TgCycY is most close with the *Homo sapiens* CCNY isoform 1 cyclin Y. This *Homo sapiens* CCNY interact with the PFTAIRE protein kinase 1 (PFTK1/CDK14) and PCTAIRE protein kinase 1 (PCTK1/CDK16) which included in the family of CDK5 kinases. Also, these CDK14 and CDK16 are ortholog to the TgCrk2. The Cyclin Y isoform 1 not only interact with the PFTK1 but also it can act as a substrate for the PFTK1 protein (Liu, Guest, & Finley, 2010). In *Saccharomyces cerevisiae*, the PHO85/PCL1 complex can phosphorylate the CDK5 substrate (Huang, et al., 2007). From these CDK5 kinases, Pho85 in *Saccharomyces cerevisiae* is most ortholog to the TgCrk2 (Figure 5 & Table 8). This shows that TgCycY might act similar to CCNY while TgCrk2 acts similar to Pho85. Phylogenetic tree analysis of these proteins backed up the BLASTp results (Figure 17). When considering these facts, we could suggest that in *T. gondii* TgCrk2/TgPHO80 complex interact with the TgCycY to phosphorylate it. Since the strong interaction between TgCrk2 and TgPHO80 is already proven. So, the TgCycY might be acting as a substrate for the TgCrk2/TgPHO80 CDK/Cyclin complex.

This CDK16 or PCTK1 kinase activated by the membrane-bound CCNY and it is important for the membrane-associated processes, vesicle formation, fusion, and transport. For the binding of CCNY to CDK16, CDK16 requires having a kinase domain as well as part of the N-terminal extension containing S153, S119. Also, binding of 14-3-3 protein to a binding site of CDK16 that created by these two serine residues needed for the CCNY binding. In conclusion,

CDK16 with N-terminus extension and 14-3-3 proteins binding are required for the interaction between CCNY and CDK16 (Figure 5; Mikolcevic, et al., 2011). Also, CDK14 interaction with CCNY was mediated via 14-3-3 protein binding to the CDK14 (Li, Jiang, Wang, & Chen, 2014). TgCrk2 sequence alignment with the CDK14 and CDK16 indicated that TgCrk2 is lacking N-terminal extension. These evidences can suggest that TgCrk2 needed an N-terminus extension or a bridging protein that binds to TgCrk2 to associate the interaction between TgCycY and TgCrk2. This might be a reason that this screen did not isolate TgCrk2 cDNA since cDNA library is 3' biased and unable to TgCycY to interact with the N-terminus region of the TgCrk2. Otherwise, this screen might not detect the interaction due to a bridging protein association between TgCycY and TgCrk2.

Chapter 6: Conclusion

In summary, Yeast two-hybrid screening of TgCycY with a cDNA library identified TgDJ-1 as the protein interaction partner for the TgCycY. Even though there are three hypothetical protein partners (TGME49_226570, TGME49_203358, & TGME49_266690) isolated, they cannot be identified as a potential interaction partner due to the lack of understanding and information. TgCrk2 kinase can be suggested as a potential indirect interacting partner for the TgCycY based on the computational predictions, homology analysis and cell cycle expression, and TgCrk2/TgCycY ortholog interactions in higher eukaryotes. The suggestion from this gathered knowledge of TgCycY and TgCrk2 interactions is TgCycY locate the TgPHO80 to the TgCrk2 and TgCycY act as a substrate for this complex by getting phosphorylated. Involvement of the TgPHO80 between TgCycY and TgCrk2 cannot be detected in this yeast two-hybrid screen due to limitations and to prove this suggestion direct interaction between TgCycY and TgPHO80 is needed to prove the second hypothesis. Due to the limitations of the current yeast two-hybrid screen, TgCycY- TgCrk2 interaction was not detected and there is no experimental evidence that supports the hypothesis of TgCycY interaction with TgCrk2.

The function of the TgDJ-1 protein involved in controlling micronemes secretion and Ca^{2+} signaling pathways. TgDJ-1 is the interacting partner for the TgCycY and TgCDPK1. According to the obtained results in this study, we have found a new interacting protein (TgDJ-1) with its Toxodb information and existing literature. Hence, this study has proved the third hypothesis in this study. By adapting the concept discussed in the second hypothesis, TgDJ-1 protein can suggest as the bridging protein for the interaction between TgCycY and TgCDPK1 to phosphorylate the TgCycY by the TgCDPK1. TgCycY non-essentiality for the tachyzoite cell growth and the interaction with TgDJ-1 protein create a prediction that TgCycY involved in a

non-canonical biological pathway rather involving in the cell growth or cell cycle of the tachyzoites. Further interaction testing between these three proteins in this complex can confirm the prediction and proving this suggested interaction model can prove the null hypothesis of this study. It will open a window to TgCycY role in the regulation of microneme secretion. Also, bradyzoites have an upregulation of TgCycY expression with bradyzoite induction which can be related to the bradyzoite growth (Huang, et al., 2017). TgCycY involvement in micronemes secretion and regulation of motility in the tachyzoites needs to be tested by testing the effect of TgCycY gene knockdown in tachyzoite motility since TgDJ-1 has shown that the mutation in TgDJ-1 does affect the secretion, motility, and invasion but not in the cell growth.

The purpose of this study is to identify a protein interaction partner for the putative TgCycY cyclin in *T. gondii*. TgDJ-1 protein identified as the potential interacting partner for the TgCycY and the cellular mechanism that lies behind this interaction cannot be well explained due to insufficient understanding of micronemes secretion process and the molecules that involved in this process in *T. gondii*. Clarification of TgCrk2 cyclin-dependent kinase and bridging protein associated interaction with TgCycY via further modified screen can support the hypothesis 1 and 2 in this study and they cannot prove based on the current evidence. Until then, TgCycY cannot be classified as a cyclin that does not interact with a cyclin-dependent Kinase. Further study in TgDJ-1/TgCycY complex can be a window to prove the TgCycY involvement in the micronemes secretion which can relate to the non-canonical function of TgCycY while disproving the hypothesis 1 and 2. If the TgCycY plays a primary role in micronemes secretion by involving with TgDJ-1 protein, TgCycY/TgDJ-1 complex can introduce as a potential drug target that can control the *Toxoplasma gondii* tachyzoite secretion, motility, and invasion which can eventually relate to the *T. gondii* differentiation.

References

Alvarez, C. A., & Suvorova, E. S. (2017). Checkpoints of apicomplexan cell division identified in Toxoplasma gondii. *PLOS Pathogens, 13*(7). doi:10.1371/journal.ppat.1006483

Araujo, F. G., & Slifer, T. (2003). Different strains of Toxoplasma gondii induce different cytokine responses in CBA/Ca mice. *Infection and Immunity, 71*(7), 4171-4174. doi:10.1128/iai.71.7.4171-4174.2003

Bhavsar, R. B., Makley, L. N., & Tsonis, P. A. (2010). The other lives of ribosomal proteins. *Human Genomics, 4*(5), 327. doi:10.1186/1479-7364-4-5-327

Billker, O., Lourido, S., & Sibley, L. D. (2009). Calcium-dependent signaling and kinases in apicomplexan parasites. *Cell Host & Microbe, 5*(6), 612-622. doi:10.1016/j.chom.2009.05.017

Black, M. W., & Boothroyd, J. C. (2000). Lytic cycle of Toxoplasma gondii. *Microbiology and Molecular Biology Reviews, 64*(3), 607-623. doi:10.1128/mmbr.64.3.607-623.2000

Blader, I. J., Coleman, B. I., Chen, C., & Gubbels, M. (2015). Lytic cycle of Toxoplasma gondii: 15 years later. *Annual Review of Microbiology, 69*(1), 463-485. doi:10.1146/annurev-micro-091014-104100

Brown, D. (1984). Pyrimidines and their benzo derivatives. *Comprehensive Heterocyclic Chemistry,*57-155. doi:10.1016/b978-008096519-2.00035-7

Brückner, A., Polge, C., Lentze, N., Auerbach, D., & Schlattner, U. (2009). Yeast two-hybrid, a powerful tool for systems biology. *International Journal of Molecular Sciences, 10*(6), 2763-2788. doi:10.3390/ijms10062763

Butler, C. L., Lucas, O., Wuchty, S., Xue, B., Uversky, V. N., & White, M. (2014). Identifying novel cell cycle proteins in apicomplexa parasites through co-expression decision analysis. *PLoS ONE, 9*(5). doi:10.1371/journal.pone.0097625

Caldas, L., & Souza, W. D. (2018). A window to Toxoplasma gondii egress. *Pathogens, 7*(3), 69. doi:10.3390/pathogens7030069

Child, M. A., Garland, M., Foe, I., Madzelan, P., Treeck, M., Linden, W. A., . . . Bogyo, M. (2017). Toxoplasma DJ-1 regulates organelle secretion by a direct interaction with calcium-dependent protein kinase 1. *MBio, 8*(1). doi:10.1128/mbio.02189-16

Coates, P., & Hall, P. (2003). The yeast two-hybrid system for identifying protein-protein interactions. *The Journal of Pathology, 199*(1), 4-7. doi:10.1002/path.1267

Criekinge, W., & Beyaert, R. (1999). Yeast two-hybrid: state of the art. *Biological Procedures Online, 2*(1), 1-38. doi:10.1251/bpo16

Dubey, J. P. (2009). History of the discovery of the life cycle of Toxoplasma gondii. *International Journal for Parasitology, 39*(8), 877-882. doi:10.1016/j.ijpara.2009.01.005

Dubey, J. P. (2008). The history of toxoplasma gondii - the first 100 years. *Journal of Eukaryotic Microbiology, 55*(6), 467-475. doi:10.1111/j.1550-7408.2008.00345.x

Dubey, J. P. (1996). Toxoplasma gondii. In Baron S (Ed.), *Medical Microbiology* (4th ed., chapter 84). S.l.: University of Texas Medical Branch at Galveston.

Esch, K. J., & Petersen, C. A. (2013). Transmission and epidemiology of zoonotic protozoal diseases of companion animals. *Clinical Microbiology Reviews, 26*(1), 58-85. doi:10.1128/cmr.00067-12

Foe, I., Child, M., Majmudar, J., Krishnamurthy, S., Van Der Linden, W., Ward, G., . . . Bogyo, M. (2015). Global analysis of palmitoylated proteins in Toxoplasma gondii. *Cell Host & Microbe,18*(4), 501-511. doi:10.1016/j.chom.2015.09.006

Furtado, J., Smith, J., Belfort, R., Gattey, D., & Winthrop, K. (2011). Toxoplasmosis: a global threat. *Journal of Global Infectious Diseases, 3*(3), 281. doi:10.4103/0974-777x.83536

Goncz, K. K., & Rothman, S. S. (1996). Membrane protein transport in eukaryotic secretion cells. *Membrane Protein Transport,* 279-293. doi:10.1016/s1874-592x(96)80011-4

Guiton, P. S., Sagawa, J. M., Fritz, H. M., & Boothroyd, J. C. (2017). An in vitro model of intestinal infection reveals a developmentally regulated transcriptome of Toxoplasma sporozoites and a NF-κB-like signature in infected host cells. *Plos One, 12*(3). doi:10.1371/journal.pone.0173018

Hall, C. I., Reese, M. L., Weerapana, E., Child, M. A., Bowyer, P. W., Albrow, V. E., . . . Bogyo, M. (2011). Chemical genetic screen identifies Toxoplasma DJ-1 as a regulator of parasite secretion, attachment, and invasion. *Proceedings of the National Academy of Sciences*, *108*(26), 10568-10573. doi:10.1073/pnas.1105622108

Hill, D., & Dubey, J. (2002). Toxoplasma gondii: transmission, diagnosis and prevention. *Clinical Microbiology and Infection, 8*(10), 634-640. doi:10.1046/j.1469-0691.2002.00485.x

Howe, D. K., & Sibley, L. D. (1995). Toxoplasma gondii comprises three clonal lineages: correlation of parasite genotype with human disease. *Journal of Infectious Diseases, 172*(6), 1561-1566. doi:10.1093/infdis/172.6.1561

Hu, K., Mann, T., Striepen, B., Beckers, C. J., Roos, D. S., & Murray, J. M. (2002). Daughter cell assembly in the protozoan parasite Toxoplasma gondii. *Molecular Biology of the Cell, 13*(2), 593-606. doi:10.1091/mbc.01-06-0309

Huang, K., Ferrin-Oconnell, I., Zhang, W., Leonard, G., Oshea, E., & Quiocho, F. (2007). Structure of the Pho85-Pho80 cdk-cyclin complex of the phosphate-responsive signal transduction pathway with bound ATP-gamma-s. doi:10.2210/pdb2pmi/pdb

Huang, S., Holmes, M. J., Radke, J. B., Hong, D., Liu, T., White, M. W., & Sullivan, W. J. (2017). Toxoplasma gondii AP2IX-4 regulates gene expression during bradyzoite development. *MSphere, 2*(2). doi:10.1128/msphere.00054-17

Hydbring, P., Malumbres, M., & Sicinski, P. (2016). Non-canonical functions of cell cycle cyclins and cyclin-dependent kinases. *Nature Reviews Molecular Cell Biology, 17*(5), 280-292. doi:10.1038/nrm.2016.27

Junn, E., Jang, W. H., Zhao, X., Jeong, B. S., & Mouradian, M. M. (2009). Mitochondrial localization of DJ-1 leads to enhanced neuroprotection. *Journal of Neuroscience Research, 87*(1), 123-129. doi:10.1002/jnr.21831

Keeley, A., & Soldati, D. (2004). The glideosome: A molecular machine powering motility and host-cell invasion by apicomplexa. *Trends in Cell Biology, 14*(10), 528-532. doi:10.1016/j.tcb.2004.08.002

Kim, K. (2015). A bradyzoite is a bradyzoite is a bradyzoite? *Trends in Parasitology, 31*(12), 610-612. doi:10.1016/j.pt.2015.10.005

Komiya, Y., & Habas, R. (2008). Wnt signal transduction pathways. *Organogenesis, 4*(2), 68-75. doi:10.4161/org.4.2.5851

Lai, M., & Lau, Y. (2017). Screening and identification of host proteins interacting with Toxoplasma gondii SAG2 by yeast two-hybrid assay. *Parasites & Vectors, 10*(1). doi:10.1186/s13071-017-2387-y

Lee, J., Kim, H., Lee, S., & Kim, K. (2007). Stabilization and activation of p53 induced by Cdk5 contributes to neuronal cell death. *Journal of Cell Science, 120*(13), 2259-2271. doi:10.1242/jcs.03468

Li, S., Jiang, M., Wang, W., & Chen, J. (2014). 14-3-3 binding to cyclin Y contributes to cyclin Y/CDK14 association. *Acta Biochimica Et Biophysica Sinica, 46*(4), 299-304. doi:10.1093/abbs/gmu005

Lim, S., & Kaldis, P. (2013). Cdks, cyclins and CKIs: roles beyond cell cycle regulation. *Development, 140*(15), 3079-3093. doi:10.1242/dev.091744

Liu, D., Guest, S., & Finley, R. L. (2010). Why cyclin Y? a highly conserved cyclin with essential functions. *Fly, 4*(4), 278-282. doi:10.4161/fly.4.4.12881

Lyons, R. E., Mcleod, R., & Roberts, C. W. (2002). Toxoplasma gondii tachyzoite–bradyzoite interconversion. *Trends in Parasitology, 18*(5), 198-201. doi:10.1016/s1471-4922(02)02248-1

Mikolcevic, P., Sigl, R., Rauch, V., Hess, M. W., Pfaller, K., Barisic, M., . . . Geley, S. (2011). Cyclin-dependent kinase 16/PCTAIRE kinase 1 is activated by cyclin Y and is essential for spermatogenesis. *Molecular and Cellular Biology, 32*(4), 868-879. doi:10.1128/mcb.06261-11

Moore, J. D. (2013). In the wrong place at the wrong time: does cyclin mis-localization drive oncogenic transformation? *Nature Reviews Cancer, 13*(3), 201-208. doi:10.1038/nrc3468

Nagamune, K., Hicks, L. M., Fux, B., Brossier, F., Chini, E. N., & Sibley, L. D. (2008). Abscisic acid controls calcium-dependent egress and development in Toxoplasma gondii. *Nature, 451*(7175), 207-210. doi:10.1038/nature06478

Naumov, A., Kratzer, S., Ting, L., Kim, K., Suvorova, E. S., & White, M. W. (2017). The Toxoplasma centrocone houses cell cycle regulatory factors. doi:10.1101/122465

Patil, V., Rajmane, V., Raje, V., & Patil, H. (2011). Successful treatment of cerebral toxoplasmosis with cotrimoxazole. *Indian Journal of Sexually Transmitted Diseases and AIDS, 32*(1), 44. doi:10.4103/0253-7184.81255

Schorr, A. L. (2018). Yeast two hybrid screen of a putative Toxoplasma gondii cyclin*, TGME49_266900*. *Culminating Projects in Biology*. 31. https://repository.stcloudstate.edu/biol_etds/31

Skariah, S., Mcintyre, M. K., & Mordue, D. G. (2010). Toxoplasma gondii: determinants of tachyzoite to bradyzoite conversion. *Parasitology Research, 107*(2), 253-260. doi:10.1007/s00436-010-1899-6

Suzuki, Y. (1998). Critical role for IFN-γ in prevention of toxoplasmic encephalitis in mice genetically resistant to development of the disease. *Parasitology International, 47*, 115. doi:10.1016/s1383-5769(98)80247-x

Tomavo, S. (1996). The major surface proteins of Toxoplasma gondii: structures and functions. *Toxoplasma Gondii,* 45-54. doi:10.1007/978-3-642-51014-4_4