4-14-2017

RNA Directed Gene Regulation in Toxoplasma gondii

Anna Crater-Potter
University of Windsor

RNA Directed Gene Regulation in *Toxoplasma gondii*

By

Anna Crater-Potter

A Dissertation
Submitted to the Faculty of Graduate Studies
through the Department of Chemistry and Biochemistry
in Partial Fulfillment of the Requirements for
the Degree of Doctor of Philosophy
at the University of Windsor

Windsor, Ontario, Canada

2017

"RNA Directed Gene Regulation in *Toxoplasma gondii"*

by

Anna Crater-Potter

APPROVED BY:

_____

M. Bitzer, External Examiner
University of Michigan

_____

W. Crosby
Department of Biological Sciences

_____

M. Boffa
Department of Chemistry & Biochemistry

_____

L. Lee
Department of Chemistry & Biochemistry

_____

S. Ananvoranich, Advisor
Department of Chemistry & Biochemistry

January 13, 2017

DECLARATION OF CO-AUTHORSHIP/PREVIOUS PUBLICATIONS

I. Declaration of Co-Authorship

I hereby declare that this dissertation incorporates material that is the result of joint research as follows:

Chapter 2: This work incorporates data generated by Ahmed Cherry, Michael Holmes and Dema Kadri under the supervision of Dr. Sirinart Ananvoranich. Ahmed Cherry, Michael Holmes and Dema Kadri contributed to data collection and analysis.

Chapter 3: This work includes the generation of a plasmid to incorporate miR60a binding sites to a gene of interest by Emad Manni under the supervision of Dr. Sirinart Ananvoranich.

Chapter 4: This review was written in collaboration with Scott Roscoe, Mikayla Roberts and Dr. Sirinart Ananvoranich.

I am aware of the University of Windsor Senate Policy on Authorship and I certify that I have properly acknowledged the contribution of other researchers to my thesis, and have obtained written permission from each of the co-author(s) to include the above material(s) in my thesis.

I certify that, with the above qualification, this thesis, and the research to which it refers, is the product of my own work.

II. Declaration of Previous Publications

This thesis includes three original papers that have been previously published in peer-reviewed journals, as follows:

Chapter	Publication	Status
2	Crater AK, Cherry A, Holmes M, Kadri D, Ananvoranich S. (2012) Evaluation of the ability of short and long double-stranded RNAs to induce homologous gene silencing in the protozoan parasite, *Toxoplasma gondii*. American Journal of Biomedical Sciences. 4(1):1-13.	Published

3	Crater AK, Manni E, Ananvoranich S. (2015) Utilization of inherent miRNAs in functional analyses of *Toxoplasma gondii* genes. Journal of Microbiological Methods. 108:92-102.	Published
4	Crater AK, Roscoe S, Roberts M, Ananvoranich S. (2015). Antisense technologies in the studying of *Toxoplasma gondii*. Review. Journal of Microbiological Methods.	Published

I certify that I have obtained a written permission from the copyright owner(s) to include the above-published material(s) in my thesis. I certify that the above material describes work completed during my registration as a graduate student at the University of Windsor.

I declare that, to the best of my knowledge, my dissertation does not infringe upon anyone's copyright nor violate any proprietary rights that any ideas, techniques, quotations, or any other material from work of other people included in my thesis, published or otherwise, are fully acknowledged in accordance with the standard referencing practices. Furthermore, to the extent that I have included copyrighted material that surpasses the bounds of fair dealing within the meaning of the Canada Copyright Act, I certify that I have obtained a written permission from the copyright owner(s) to include such material(s) in my thesis.

I declare that this is a true copy of my thesis, including any final revisions, as approved by my thesis committee and the Graduate Studies office, and that this thesis has not been submitted for a higher degree to any other University or Institution.

ABSTRACT

Toxoplasma gondii is an intracellular obligate parasite of phylum Apicomplexa. *Toxoplasma* is able to infect any nucleated cell including up to one third of the world's population. Within its hosts, the asexual life cycle of *Toxoplasma* consists of two distinct forms; a rapidly growing form called tachyzoite and a latent cyst encapsulated form called bradyzoite. Although tachyzoites can be removed by the host's immune system, parasites can convert to bradyzoites thereby evading the host's immune response. If the host's immune system becomes weakened, bradyzoites are able to reconvert to tachyzoites. Taken together, these observations suggest *Toxoplasma* contains intricate gene regulation mechanisms that could be shared by other intracellular parasites. MicroRNAs (miRNAs) are crucial genetic effectors involved in numerous gene regulation mechanisms in eukaryotes. Although a post-transcriptional gene silencing can be observed in *Toxoplasma* the roles and functions of Tg-miRNAs are still elusive. In this work an engineered dual luciferase reporter system was used to examine and standardize the ability of long and short double-stranded RNA to control gene expression in *Toxoplasma*. The effects of endogenous Tg-miRNAs were also evaluated based on (i) their abundance and (ii) the number of binding sites within a transcript. Tg-miRNA effectiveness can be altered by use of miRNA mimics and inhibitors. Also a genetic knockdown system that exploits and directs Tg-miRNAs for loss-of-function analysis of essential and non-essential genes was developed. To further our understanding of miRNA induced gene silencing on parasite cell biology, a model gene target was selected for the most abundant Tg-miR-60a. The

locus of the Ubiquitin-like protease 1 (Ulp1) contains the highest number of predicted Tg-miR-60a binding sites. The antisense RNA expresses a likely precursor of Tg-miR-60a suggesting TgUlp1 expression could be self-regulating. Ulp1 activity is required for the first and last step in the SUMOylation pathway. TgUlp1 is able to compliment the function of yeast Ulp1 and exhibits cleavage activity *in vitro*. Knockout of TgUlp1 is detrimental to parasite egress and survival. Since a miR-60a inhibitor or antisense RNA can alter TgUlp1 expression and activity, a TgUlp1-SF line was created to study the effects of miR-60a and its effects on total protein SUMOylation to demonstrate the importance of proper gene regulation by miRNAs. Furthermore, *Toxoplasma* Argonaute, the core protein of RNA-induced gene silencing that is required for a gene silencing effect, is expressed as two isoforms from the same locus, suggesting alternative translational starts sites. Overall, this dissertation reveals the importance of proper gene regulation directed by non-coding RNAs in *Toxoplasma* that could lead to the development of anti-*Toxoplasma* therapeutic.

ACKNOWLEDGEMENTS

I would like to start by acknowledging my supervisor, Dr. Sirinart Anavoranich for her guidance and support. I am truly grateful for the opportunity to be a part of her lab and learn from her vast range of knowledge. I will always value her time, effort and our conversations about my research, ideas and techniques. I would also like to extend my many thanks to committee member, Dr. Lee, for her greatly appreciated financial support and for the many academic and motivational conversations. I will always value our mutual love for the Detroit Red Wings and "Tea with Lee" times. It will never be forgotten. I would also like to thank committee members whose suggestions and guidance are truly appreciated, Dr. Boffa and Dr. Crosby.

I would also like to extend my thanks to the current members of the Ananvoranich lab: Scott, Farzana, Amran and my undergraduate minions, Mikayla, Nadya, Karlee and Ambreen as well as past members Mike, Dema and Ahmed. I would also like to thank the Department of Chemistry and Biochemistry family, especially, Beth, Marlene and Cathy.

Thank you to my friends, whose support and encouragement I will always cherish, Norah, Ahmed, Corey, Dennis, Pam, Tanya and Besa. Particularly, I want to thank Norah, practicality my sister, who was always ready to help, lend an ear or have a glass of wine at the ready.

Finally, I would like to thank my family, my wonderful and understanding husband, Brad and my encouraging parents, Ron and Cathy. To my son Joshua, mommy is truly grateful you're an amazing child and my little prodigy. Also, my

friends who are close enough to be family, Todd, Dawn, Kyle and Sara. I love you all. Thank you.

TABLE OF CONTENTS

CHAPTER 4: Antisense technologies in the studying of *Toxoplasma gondii*

CHAPTER 5: Ubiquitin-like protease 1 (Ulp1) as a model for the study of miRNA-induced gene regulation in *Toxoplasma gondii*

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

LIST OF ABBREVIATIONS

Ago	Argonaute
AIDS	Acquired Immune Deficiency Syndrome
Bz	Bradyzoite
cDNA	Complimentary DNA
ChFP	Cherry Fluorescent Protein
cis-nat siRNA	Cis-natural siRNA
CP	Coat Proteins
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
D-MEM	Dulbecco's Modified Eagle Media
DCL1	Dicer-Like 1
dd	destabilization domain
DHFR	DiHydroFolate Reductase
DLR	Dual Luciferase Reporter
DNA	Deoxyribonucleic Acid
dsRNA	Double-stranded RNA
endo-siRNA	Endogenous siRNA
exo-siRNA	Exogenous siRNA
Ffluc	Firefly luciferase
FRET	Fluorescence Resonance Energy Transfer
GFP	Green Fluorescent Protein
HEN1	Hua-Enhancer 1
HFF	Human Foreskin Fibroblast
HoDI	Homolog of Dozi
HST	Hasty
HXGPRT	HypoXanthine Guanine PhosphoRibosylTransferase
Hz	Hoechst
IB	ImmunoBlot
kDa	kiloDalton
KO	KnockOut
LDH1	Lactate DeHydrogenases 1
lncRNA	long non-coding RNA
MEM	Minimum Essential Media
miRNA	microRNA
MRE	MiRNA Responsive Element
mRNA	messenger RNA
mRNP	messenger RiboNucleoProtein
ncRNA	non-coding RNA
NHEJ	Non-Homologous End Joining
P-bodies	Processing bodies
PAZ	Piwi-Argonaute-Zwille

PCR	Polymerase Chain Reaction
piRNA	Piwi interacting RNA
Piwi	P-element induced wimpy testis
Pol	Polymerase
pre-miRNA	Premature miRNA
pre-mRNA	Premature mRNA
pri-miRNA	Primary miRNA
PRMT1	ARg Methyl Transferase 1
PV	Parasitophorous Vacuole
qPCR	Quantitative PCR
RdRp	RNA-dependent RNA polymerase
RISC	RNA Induced Silencing Complex
RITS	RNA Induced Transcriptional Silencing
RLM-RACE	RNA Ligase-Mediated Rapid Amplification of cDNA
RNA	RiboNucleic Acid
RNAi	RNA interference
Rnluc	Renilla luciferase
rNTP	riboNucleoside Tri-Phosphate
RT	Reverse Transcriptase
SAGE	Serial Analysis of Gene Expression
SBP	Streptavidin Binding Protein
SF	Strep-Flag
siRNA	small interfering RNA
snoRNA	small nucleolar RNA
snRNA	small nuclear RNA
SUMO	small ubiquitin-related modifier
T. gondii	*Toxoplasma gondii*
Tg-miRNA	*Toxoplasma gondii* miRNA
TSN	Tudor Staphylococcal Nuclease
Tub	Tubulin
Tz	Tachyzoite
Ulp1	Ubiquitin-like protease 1
UPRT	Uracil PhosphoRibosyl Transferase
UTR	UnTranslated Region
WAGO	Worm Argonaute
WT	WildType

CHAPTER 1

General Introduction

1.1. *Toxoplasma gondii*, the one and only

Over 100 years ago in 1908, one of the most successful parasitic protozoans was discovered in the tissue of a rodent, the gundi, at the Pasteur Institute in Tunis by Charles Nicolle and Louis Manceaux, and in that same year in the tissue of a rabbit in Sao Paulo, Brazil by Alfonso Splendore (Splendore 1908). It was originally mis-identified as *Leishmania*, another unicellular eukaryotic parasite by both laboratories (Duby 2008). However, Nicolle later realized that he discovered a new organism and named it *Toxoplasma gondii*, based on the parasite's morphology; toxo "bow" and plasma "life" and the host.

Later *Toxoplasma* was isolated from birds by Sabin and Olitsky in 1973 and was discovered to be identical to the human isolates. This lead to further investigations of *Toxoplasma* isolates from humans and animals (Sabin and Olitsky 1973). Furthermore, an in-depth study of the genetic differences among 275 isolates worldwide resulted in the grouping of three dominant clonal lineages within North American called type I, II and III. Type I parasites are highly virulent, while type II and III are most commonly found within humans and animals, respectively (Dubey 2002, Montoya and Liesenfeld 2004). *Toxoplasma* infection occurs worldwide with the largest range of hosts despite the fact that *Toxoplasma gondii* is the only one species of its genus (Duby 2008). Although great advances have been made in the last 50 years to prevent *Toxoplasma* infection, today one third of the world's population has been estimated to become infected by *Toxoplasma* (Flint 2005). Typically infection does not produce any clinical symptoms, however,

severe disease can be observed in congenitally infected children and immunosuppressed individuals, including patients undergoing chemotherapy or by acquired immune deficiency syndrome (AIDS, Dubey 1996).

1.2. The life cycle of *Toxoplasma gondii*

As an intracellular parasite, *Toxoplasma* spends most of its life cycle within the cells of its host. The life cycle proceeds in two stages; sexual reproduction within the gut of felines and asexual reproduction within its host. In the asexual stage, the parasite can alternate between two forms called tachyzoite and bradyzoite (Figure 1.1.).

1.2.1. Sexual reproduction

In 1965, William Hutchison pursued the theory of a fecal transmission route for *Toxoplasma via* cats after researchers reported no evidence of the parasite within canine fecal samples ten years prior. Hutchison not only discovered the oocyst form of the parasite but also its definitive host (Hutchison 1965). Later it was discovered that *Toxoplasma* undergoes sexual reproduction and genetic crossing within the epithelial cells of the feline small intestine (Frenkel et al., 1970, Hutchison et al., 1970, Ferguson et al., 1974). The oocysts are shed for up to 21 days in the feces of infected felines and able to survive in the environment for up to 18 months (Frenkel 1975). Sporulation usually occurs between 1-21 days, and the oocyst-containing sprorzoites have been found to be the most significant source of infection for the intermediate hosts (Figure 1.1., Montoya and Liesenfeld 2004). Despite the parasite's promiscuous ability to infect any nucleated cells including mammal and bird, felines remain the only definitive host for sexual reproduction.

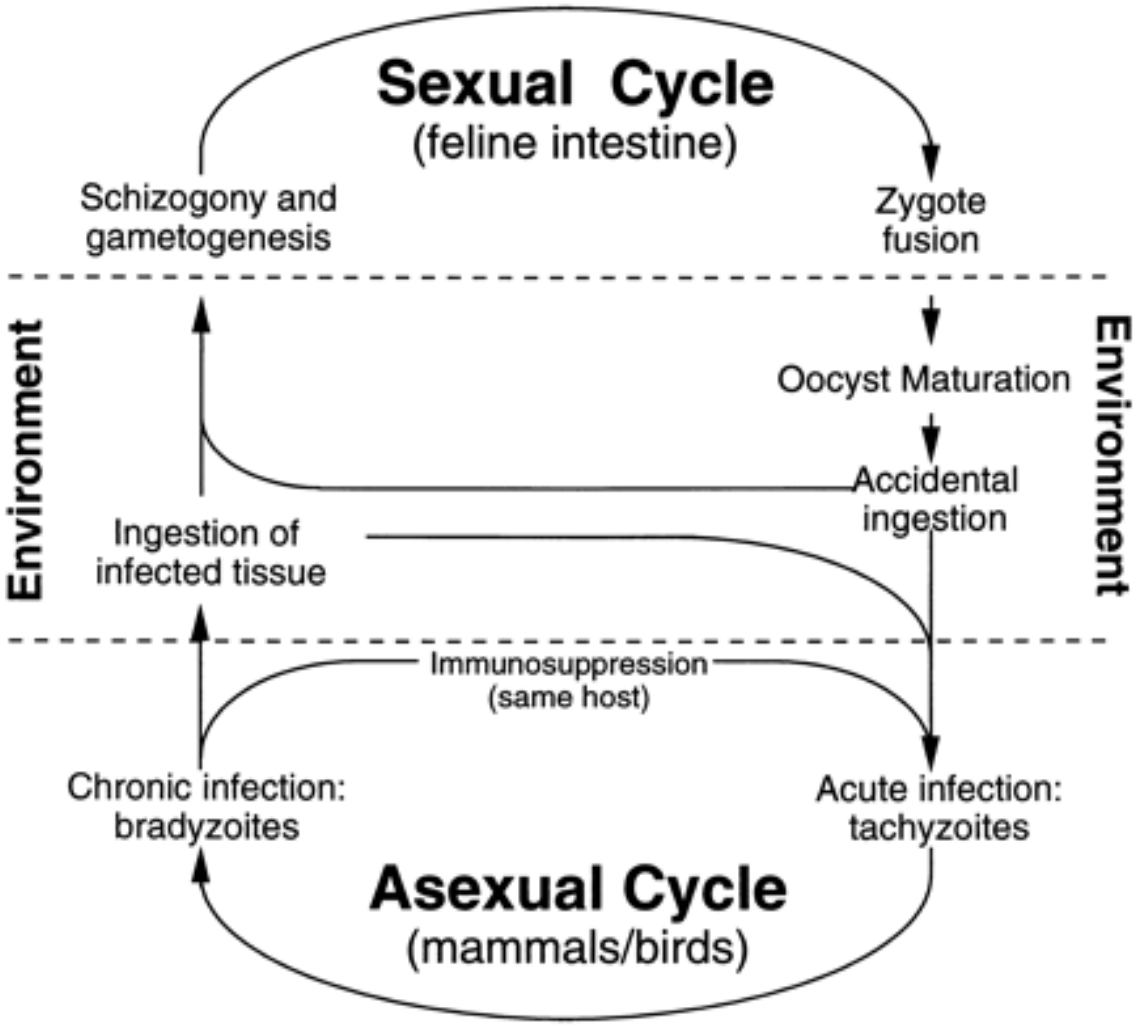


Figure 1.1. *Toxoplasma gondii* life cycle. Sexual reproduction takes place within the gut of felines. Cats then shed oocysts that can be taken up by any warm-blooded animal. Acute infection occurs by the rapidly growing tachyzoites. An attack by the host's immune system induces tachyzoite convert to a slow growing cyst-encapsulated bradyzoite. Immune suppression allows the bradyzoite to reconvert to tachyzoites. If a cat ingests bradyzoites cyst infected tissue, the parasites re-enter the sexual life cycle. (Adapted from Black and Boothroyd, 2000)

1.2.2. Asexual reproduction: A tale of two forms

The proliferative form of *Toxoplasma* called the tachyzoite, was the stage observed by Nicolle and Manceaux in the gundi. Tachyzoites are crescent shaped approximately 2 by 6 μm with a pointed anterior and a rounded posterior end. This small "banana-like" structure is the rapidly growing form of *Toxoplasma.* Tachyzoites are able to divide rapidly by a process called endodyogeny, where two daughter parasites form within the parent parasite (Glodman et al., 1958). Rapid parasite growth causes a strong inflammatory response by the host's immune system, and the tachyzoites can easily be cleared (Figure 1.1). However, *Toxoplasma* is able to avoid an immune system attack by converting to an encysted slow-growing, dormant form called bradyzoite. This form does not cause harm to the host or elicit a host immune response. The cyst wall also makes the parasite less susceptible to destruction. Although the parasite in the bradyzoite form continues to grow, the growth is slower than that of the tachyzoite form that the parasite remain intracellular and dormant (Montoya and Liesenfeld 2004). In individuals with a weakened immune system, bradyzoites could be released from their cysts and convert back to the rapidly growing tachyzoite form. When bradyzoite tissue cysts from infected animals are eaten, the parasites are released within a new host and re-enter the asexual life cycle (Figure 1.1., Dubey et al., 1998, Black and Boothroyd, 2000).

1.3. Lytic cycle

As an intracellular parasite, *Toxoplasma* must invade a host cell to survive. Tachyzoites move through the extracellular environment by gliding motion until attachment to a host cell (Figure 1.2.). Once attached, the parasite enters by penetrating

through the host cell membrane or by phagocytosis (Dubey et al., 1998). Once inside, the parasite is surrounded by a parasitophorous vacuole (PV) derived from parasite and host cell lipids and membrane proteins (Dubey et al., 1998). The parasite can then undergo multiple rounds of replication while remaining within the PV (Figure 1.2.). Rapid growth is achieved because *Toxoplasma* is auxotrophic and sequesters materials from the host cell such as amino acids, nucleotides, sugars, cholesterol and lipids which requires less energy (Laliberte and Carruthers, 2008). Growth rates have been shown to be directly related to parasite virulence. Type I parasites are highly virulent where types II and III are avirulent in mice (Weiss and Kim 2007). After multiple rounds of replication, the PV fills with parasites and causes the host cell to lyse (Figure 1.2.), and allow the parasite to initiate the next round of infection, the lytic cycle. The destruction of the infected host cell is the serious consequence of toxoplasmosis (Blader et al., 2015).

1.4. *Toxoplasma* infection

Toxoplasma infection occurs through the ingestion of tissue cysts or sporulated oocysts from infected meat or from the environment such as contaminated water or soil (Dubey and Jones, 2008). *Toxoplasma* has been found in animals worldwide, regardless of environmental or habitat (Dubey and Jones, 2008). This is especially a concern for human health and consumption of contaminated animals and products. Tachyzoites can spread throughout the hosts system and can be transmitted to a pregnant host's offspring (Montoya and Liesenfeld, 2004). Congenital toxoplasmosis can cause severe complications to fetal development including mental retardation, blindness, and brain lesions and in the most severe cases, can cause abortion or stillbirth (Montoya and Liesenfeld, 2004). Initial *Toxoplasma* infection mimics flu-like symptoms. Persisting

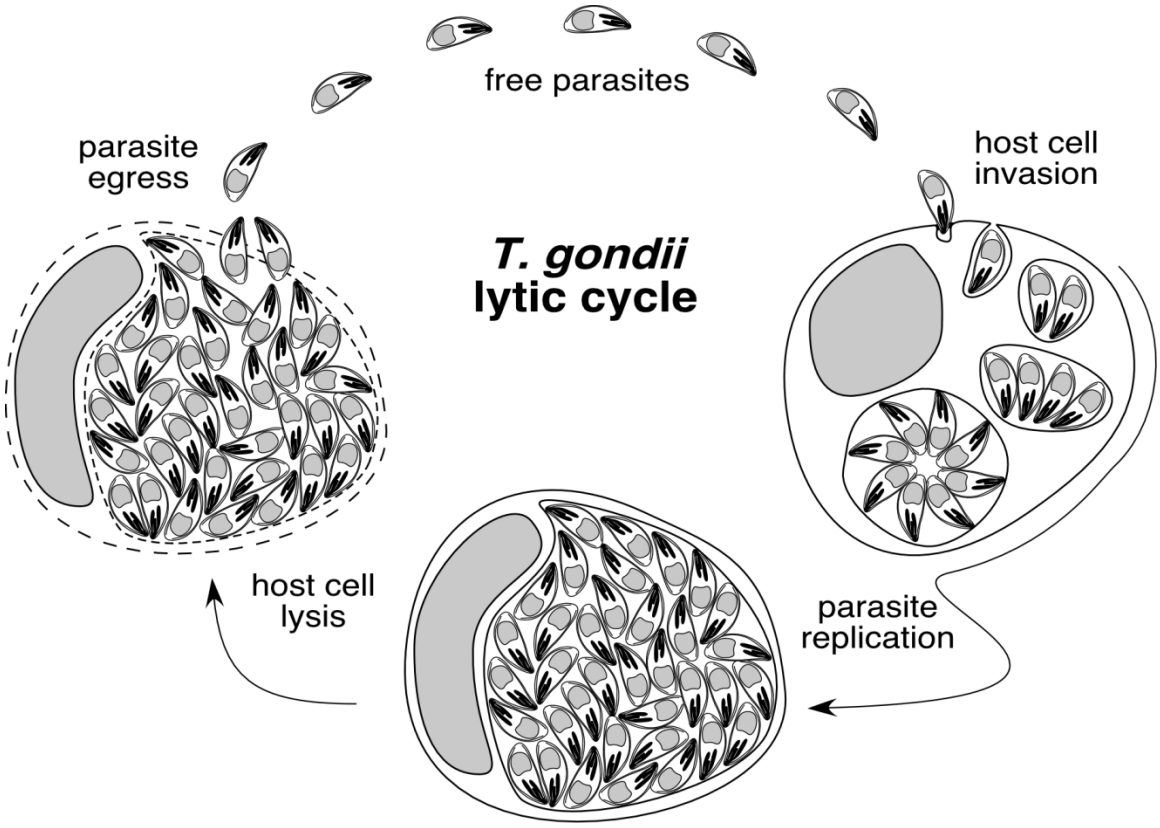


Figure 1.2. *Toxoplasma gondii* lytic cycle. Extracellular parasites attach and invade a new host cell. The intracellular parasites become surrounded by a parasitophorous vacuole where they undergo multiple rounds of replication. Once the host cell fills with parasites it ruptures releasing free tachyzoites that are able to invade a new host cell. (Adapted from Gubbles, 2014)

Toxoplasma infection can cause neurological and behavioral changes, and has been linked to psychological disorders such as depression, schizophrenia and suicidal behavior (Flegr, 2013, Dubey, 1996). The most interesting behavioral change has been found in rodents, where infected rats display an attraction to the scent of feline urine thereby leading them to be consumed by the parasite's definitive host, completing its life cycle (Berdoy et al., 2000, Dubey 1996, Montoya and Liesenfeld, 2004).

1.5. The model Apicomplexan

Toxoplasma is a member of the phylum Apicomplexa that includes other important pathogenic genus such as *Plasmodium*, the cause of malaria, *Cryptosporidium*, a water borne parasite and *Eimeria*, *Neospora*, and *Theleria* all of which are important zoonotic parasites (Montoya and Liesenfeld 2004). Among intracellular parasites, *Toxoplasma* is the most experimentally tractable organism and is amenable to genetic manipulation with reliable protocols for both forward and reverse genetics. These include high transient transfection efficacy and expression of epitope tags, reporter constructs and heterologous proteins. The genomic sequences for type I, II, and III parasite strains are available (toxodb.org) with regular updates (Gajria et al., 2008). Furthermore, *Toxoplasma* can be used as an expression system to study biological functions of other apicomplexan parasites that cannot be expressed in other organisms (Soete et al., 1999, Weiss and Kim 2007). *Toxoplasma* has also been widely used as a model to study host-parasite interactions, basic biology of host immune response, and mouse models (Weiss and Kim 2007). *Toxoplasma* is a unique model organism for intracellular parasitic research.

1.6. Gene regulation

The regulation of gene expression is vital to the survival of all organisms. The levels of gene products, both functional proteins and RNAs, can be modulated at the level of transcription, post-transcription, translation and post-translation. Transcription regulation is the control of DNA expression orchestrated *via* chromatin modifications, transcription factors such as enhancers and repressors, and promoter activity (Bell et al., 2011, Bruce et al., 2007). Post-transcriptional regulation is the control of gene expression at the RNA level *via* RNA editing, intron splicing and addition of 5' cap and poly(A) tail; interactions with RNA binding proteins for nuclear export and processing; and non-coding RNAs for mRNA degradation or translational repression (Bruce et al., 2007). Within this complex network of controlled gene expression, my work focuses on post-transcriptional gene regulation *via* non-coding RNAs, which is another layer of the control of gene regulation. Thus this section is focused upon the current knowledge of post-transcriptional gene regulation.

1.6.1. Post-transcriptional gene regulation by RNA interference

In 1998 Fire and Mello published one of the most influential discoveries in science that described the RNA interference (RNAi) pathway in *Caenorhabditis elegans*. The study began with an investigation of antisense RNA function in modulating the expression level of endogenous genes at the mRNA level. Double-stranded RNA (dsRNA) molecules were found to be more potent and specific at interfering with genes than antisense RNA (Fire at al., 1998). English and Jones (1998) described a gene silencing effect caused by integration of T-DNA transposons in tobacco plants that

caused heritable epigenetic changes to promote drug resistance. Shortly thereafter, Chuang and Meyerowitz (2000) reported that in *Arabidopsis thaliana*, RNA duplexes were able to induce significant gene-specific down-regulation when compared to single-stranded sense or antisense alone. The study confirmed the dsRNA-mediated interference (Chuang and Meyerowitz, 2000). RNA interference (RNAi) is thus described as a biological process by which the expression of a gene can be altered by non-coding RNA molecules. The pathway has been found in many single-cell and multicellular eukaryotes. Not all eukaryotes contain the required enzymes or effector RNA for a functional RNAi mechanism. For example, some protozoa such as *Leishmania major*, *Trypanasoma cruzi* and *Saccharomyces cerevisiae* lack the RNAi pathway, while other protozoa such as *Trypanasoma brucei*, *Tetrahymena thermophile*, and *Saccharomyces castelli* contain the canonical RNAi components (Kolev et al., 2011, Aravind et al., 2000). The RNAi pathway is believed to be similar to an immune system in cells that can defend against viral infection and transposons for chromatin integrity and development (Matzke and Matzke, 2004). It remains unclear for what purpose the presence or absence of a functional RNAi pathway is among these single-celled organisms. It has been hypothesized that select organisms might lose the need for an RNAi mechanism thus the necessary genes for a functional pathway (Kolev et al., 2011).

The RNAi mechanism is initiated by the presence of long dsRNA. In the cytosol, dsRNA molecules are cleaved into smaller 20-25 nucleotide duplexes with a 2-nucleotide 3' overhang by a member of the RNAse III family called Dicer (Siomi and Siomi, 2009, Bernstein et al., 2001). The RNA product is called either microRNA (miRNA) or small interfering RNA (siRNA, Figure 1.3.). The RNA duplex is then transferred to an

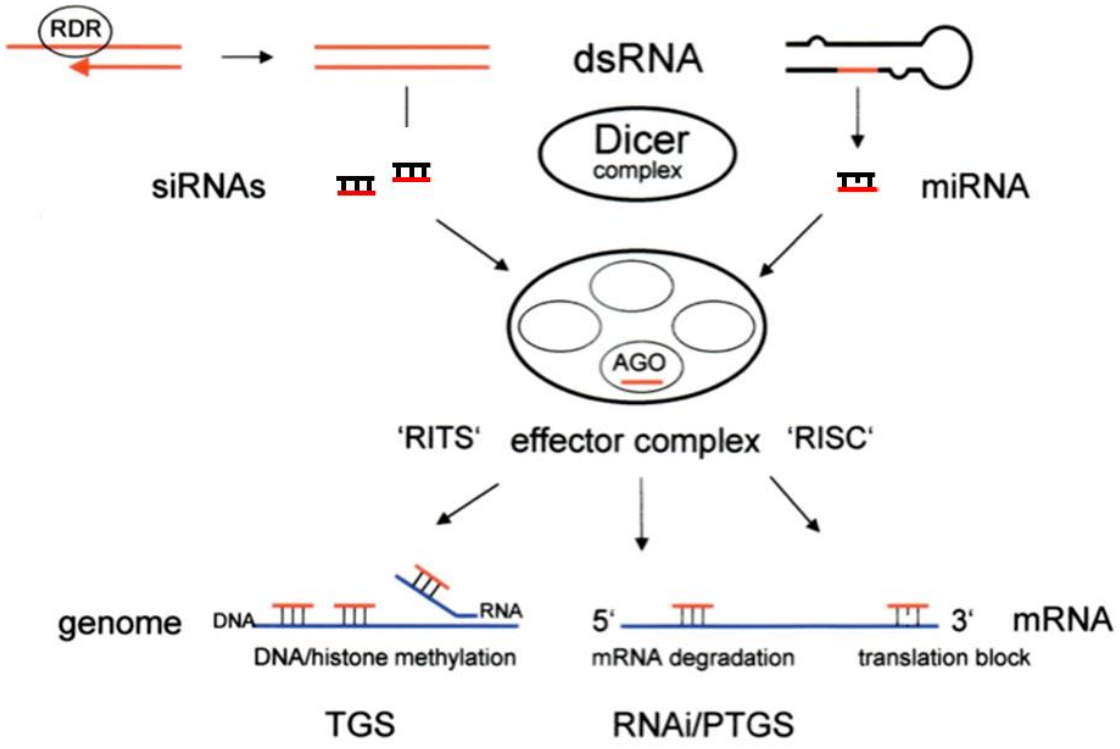


Figure 1.3. RNA interference pathway. RNAi is initiated by the presence of long dsRNA that is cleaved by Dicer. Within RISC, small RNAs are bound by an Argonaute protein and guided to their target mRNA through either a perfect match or mismatch complementarity (guide strand in red). In the nucleus, the RITS complex controls transcription through histone modifications or small RNA interactions with DNA or RNA. An enhanced silencing effect can be observed by amplification of effector RNA by an RdRp. (Adapted from Matzke and Matzke, 2004)

Argonaute protein in the RNA-induced silencing complex (RISC). Within RISC, the antisense strand of the RNA duplex guides the RISC complex to its target mRNA, which gives its name as the guide strand. The sense strand, or passenger strand, is removed and degraded (Preall et al., 2006, Okamura et al., 2004). Generally the outcome of gene regulation occurs based on the complementarity of the miRNA or siRNA to the mRNA target. Perfect match interactions result in mRNA degradation, while mismatch interactions result in translational repression (Figure 1.3.). MiRNAs generally function in translational repression, as they rarely contain perfect match complementarity to mRNA targets (Okamura et al., 2004). MiRNA binding sites can be found throughout target mRNA structures (i.e. CDS, 5' and 3'-UTR), but most often located within the 3' UTR of the mRNA (Lim et al., 2004, Friedman et al., 2009). The effects of these small non-coding RNAs can be amplified through the synthesis of secondary siRNAs produced by an RNA-dependent RNA polymerase (RdRp). The RNAi mechanism also allows for the maintenance and organization of the genome by transcriptional regulation *via* the RNA induced transcriptional silencing (RITS) complex (Figure 1.3., Holmquist and Ashley, 2006).

Since its discovery, RNAi has quickly become the preferred methodology to precisely, efficiently, and stably down-regulate any gene of interest for functional studies (Saurabh et al., 2014). DsRNA can be easily generated and introduced into a cell or organism. Since synthetic RNAs can be recognized and utilized by the enzymes involved in the RNAi pathway, functional genomic applications such as mapping and annotation have been possible in organisms where traditional engineering methodologies are difficult (Saurabh et al., 2014, Kupferschmidt, 2013). RNAi has also been used to

genetically engineer plants to produce lower levels of toxins, improve stress tolerance and act as an insecticide (Grimm et al., 2006, Kupferschmidt, 2013).

1.6.2. Argonaute protein

Argonaute proteins are central to the RNAi mechanisms, where they direct small non-coding RNAs to induce gene silencing. In eukaryotes, gene silencing can occur at the transcriptional, post-transcriptional and chromatin remodeling levels by (i) degradation of mRNA target, (ii) translational repression or (iii) by causing chromatin modifications. Argonaute proteins were first discovered in plants and named after a phenotype of *A. thaliana* mutant that resembled the animal *Argonauta argo* (Bohmert et al., 1998). Among eukaryotes, the total number of Ago proteins can vary among organisms. *Drosophila melongaster*, *Homo sapiens*, and *C. elegans*, contain five, eight and 27 paralogs, respectively, while *Schizosaccharomyces pombe* and *S. castellii* contain only one (Jinek and Doudna, 2009). When an organism contains more than one ortholog of Ago, each paralog may have a different function from one another despite their sequence similarity. For example, one paralog of Ago could be found within the nucleus while another in the cytosol to facilitate transcriptional or post-transcriptional gene silencing, respectively. Organisms that contain a single ortholog of Ago, the protein can be found within the nucleus and cytosol and participate in both transcriptional and post-transcriptional gene regulation.

Based on sequence homology and phylogenetic analysis, Ago proteins are categorized in three subfamilies; (i) the Ago-like subfamily, whose members resemble *A. thaliana* Ago1. Members of the Ago-like subfamily contain two conserved domains

called PAZ (Piwi-Argonaute-Zwille) and Piwi (P-element induced wimpy testis). They are typically ~100 kDa and interact with miRNAs and siRNAs. Their Piwi domain also contains the canonical catalytic triad, DDH/E. (ii) The Piwi-like subfamily, whose members resemble *D. melanogaster* Piwi protein. Piwi-like subfamily members contain only the Piwi domain, are generally ~40 kDa and interact solely with PIWI-interacting RNAs (piRNAs). (iii) The WAGO subfamily, whose members are only found in *C. elegans*. For the purpose of this thesis I will be focusing on members of the Ago-like subfamily.

1.6.2.1. The function of Argonaute in the mechanism of gene silencing

Ago proteins bind small non-coding RNAs and play important part in the mechanism of gene silencing. Ago-bound small RNAs alone cannot catalyze the reaction without first forming a fully functional RISC. The formation of RISC involves (i) loading and (ii) subsequent unwinding of the small RNA duplex by Ago. Theoretically either strand of the small RNA duplex could act as the guide strand, but thermodynamically the termini of small RNA determine which strand is loaded or discarded in Ago proteins. Generally, the less stably paired 5'-end of the RNA duplex is loaded to Ago, and is known as the guide strand (Khvorova et al., 2003, Schwarz et al., 2003). This reaction requires ATP hydrolysis to mediate a conformational change of Ago to accommodate the small RNA. After the RNA duplex is bound, Ago reverts back to its closed state, which drives the removal and degradation of the passenger strand (Kawamata and Tomari, 2010). The exact mechanism of unwinding and discarding of the passenger strand remains unclear. The complex between Ago and the single-strand guide RNA, referred to as the miRISC, recognizes its target mRNA through

complimentary base pairing. Although many software applications aimed to predict miRNA targets have been developed and refined, the success in the identification of mRNA targets is still limited. This, in part, is due to the lack of consensus rules describing the exact base pairing or structure of small RNA to be loaded onto miRISC. Single-molecule fluorescence resonance energy transfer (FRET) analysis showed that human HsAgo2 initially scans for its target by using 2-4 nucleotides of the guide strand sequence. The stability of the interaction increases with base pairing of the seed sequence 2-8 nucleotides from the 5' end of the guide strand. Perfect-matched complementarity promotes mRNA cleavage, while mismatches in the central region lead to translational repression of the target mRNA (Chandradoss et al., 2015). Translational repression could still lead to mRNA degradation due to deadenylation and decapping of the target mRNA (Nishihara et al., 2013, Arribas-Layton et al., 2013).

Under stress conditions Ago proteins have been shown to localize to the nucleus to facilitate transcriptional regulation by associating with chromatin modifiers and splicing factors (Ameyar-Zazoua et al., 2012). For example, the catalytically inactive HsAgo1 and catalytically active HsAgo2 have both been shown to promote methylation of the H3K9 histone to slow RNA polymerase II elongation rate (Wu et al., 2006).

1.7. Non-coding effector RNAs

Regulatory non-coding RNAs (ncRNAs) are single stranded or double-stranded RNA molecules transcribed by either RNA polymerase II or III to produce a functional RNA molecule (Morey and Avner, 2004). NcRNAs originate from intergenic regions as antisense RNA or from overlapping protein-coding genes. NcRNAs are classified by

their function as a housekeeping or regulatory RNA. Examples of housekeeping ncRNAs include transfer RNAs and ribosomal RNAs while examples of regulatory RNAs include microRNAs (miRNAs) and small interfering RNAs (Morey and Avner, 2004, Brosius, 2005, Palazzo and Lee, 2015).

1.7.1. Long non-coding RNAs

Long non-coding RNAs (lncRNAs) are grouped based on their length of 200-2,000 nucleotides. LncRNAs function to control gene expression by (i) sequence complementarity for gene silencing, (ii) secondary structure formation of scaffolds to assemble complexes, (iii) or act as a decoy for transcription factors (Yoon et al., 2013).

1.7.1.1. Antisense RNA

Long antisense RNAs are single-stranded RNAs that can be transcribed by either RNA polymerase II or III or the product of spliced introns of pre-mRNA. Antisense RNAs function by binding to the complementary nucleotides in mRNA gene targets. The first studies of antisense RNA as an effector RNA is the hok/sok system of *E. coli.* The R1 plasmid, which encodes for the hok gene (host killing), also encodes an antisense RNA, sok. When the sok antisense is expressed, the translation of complementary hox mRNA is inhibited by base pair complementation, causing an obstruction to the translation machinery (Gerdes et al., 1985). Transcription of long antisense RNA is often associated with the expression of its protein-coding gene (Engstrom et al., 2006, Korneev et al., 2015). Xist and Tsix are perhaps the most well-known non-coding single-stranded RNAs, involved in random X inactivation in female mammalians. Xist ncRNA is

expressed from one X chromosomes, and is regulated by the expression of an antisense ncRNA Tsix to induce transcriptional silencing (Stavropoulos et al., 2001).

1.7.1.2. Long double-stranded RNA

In eukaryotes long dsRNA is formed by the transcription of complementary sense and antisense RNA molecules or by transposable elements (Wicker at al., 2007). The presence of transposable elements can cause truncated coding and non-coding transcripts in both sense and antisense directions. DsRNAs are shown to involve in chromatin remodeling, RNA editing, processing and nuclear localization (Meister and Tuschl, 2004). DsRNA can be acquired as a result of viral infection. RNA viruses use RNA rather than DNA as its genetic material which can be present as long dsRNA that initiates the RNAi mechanism (Blevins et al., 2006).

1.7.2. Small double-stranded RNAs

MiRNAs and siRNAs are the most pertinent small RNAs to be discussed here. They are derived from long dsRNAs which are (i) exogenous dsRNA such as viral RNA, (ii) endogenously transcribed as RNA-encoding genes, or (iii) spliced introns, called mirtrons (Rother and Meisster, 2011). These small dsRNAs are central to the post-transcriptional gene regulation mechanism which alters gene expression by causing mRNA destruction or translational repression. Conversely, small dsRNAs can also bind mRNA targets to increase their activity *via* recruitment of transcriptional or translational initiation machinery (Rother and Meisster, 2011).

1.7.2.1. MicroRNAs

Lee and co-workers at Dartmouth College described the first single-stranded non-protein-coding regulator RNA, *lin*-4 in *C. elegans* (Lee et al., 1993). The *lin*-4 locus encodes for small RNA molecules with partial complimentary to several sequences located on the 3' untranslated region of *lin*-14 mRNA, which is important for larval development. *lin*-4 regulates the expression of *lin*-14 through an antisense RNA mechanism (Lee et al., 1993, Wightman et al., 1993). *let*-7 is another excellent example of functional small RNA. *let*-7 represses the expression of *lin*-41 which is important in developmental transitions of *C. elegans* (Reinhart et al., 2000). MiRNAs are well conserved in both plants and animals. However, miRNA biogenesis are different among organism kingdoms (Axtell et al., 2011).

1.7.2.2. Mammalian miRNAs

Mammalian miRNAs are able to recognize and bind their target mRNA via Watson-Crick base pairing at positions 2 to 8 from the 5' end of the miRNA, called the seed sequence (Bartel, 2004). Thus, miRNA families are generally grouped based on their seed sequence. By using transcriptome-wide mapping of miRNA binding sites (Ago HITS-CLIP), Soek et al (2016) showed that miRNA target binding occurred through non-canonical sites and interactions. Canonical miRNA target interactions would be the 2-8 nt seed sequence of the guide RNA. These non-canonical interactions may include a near perfect match (1 nt mismatch), a 1 nt bulge, or a wobble (G:U non-Watson-Crick base-pairing). Another contributing factor is the central nucleotides termed seed-like motifs. The interaction with the seed-like motif negates the need for

perfect match binding within the seed region (Soek et al., 2016). These above described interactions induce translational repression, which may eventually lead to mRNA degradation rather than direct cleavage of target mRNA. Noteworthy, a miRNA family could have hundreds of different mRNA targets, and a target could be regulated by multiple miRNAs (Bartel, 2004).

Mammalian miRNAs can be transcribed by either RNA polymerase II or III from intergenic regions or from spliced introns (Lee at al., 2004, Faller et al., 2008). Multiple miRNAs can be produced as ploycistronic units by one long dsRNA primary miRNA (pri-miRNA). A single pri-miRNA, ~80 nucleotides, may contain one up to six miRNAs. Pri-miRNAs produced by pol II result in a transcript that is capped at the 5' end and polyadenylated. The resulting pri-miRNA folds into a secondary double-stranded structure (Figure 1.4, Lee et al., 2004, Altuvia et al., 2005). RNA polymerase III can also transcribe pri-miRNAs, but these RNAs are not capped or polyadenylated (Faller at al., 2008). A member of the RNaseIII family cleaves the pri-miRNA to form a single hairpin structure, referred to as pre-miRNA (Figure 1.4., Murchison et al., 2004). Pre-miRNAs produced from spliced introns are also called mirtrons. Pre-miRNAs are exported from the nucleus by the nucleocytoplasmic Exportin-5. Once exported, Dicer, which is another RNAseIII family member, cleaves the pre-miRNA to produces a 20-25-nucleotide miRNA duplex (Figure 1.4. B, Lund et al., 2006).

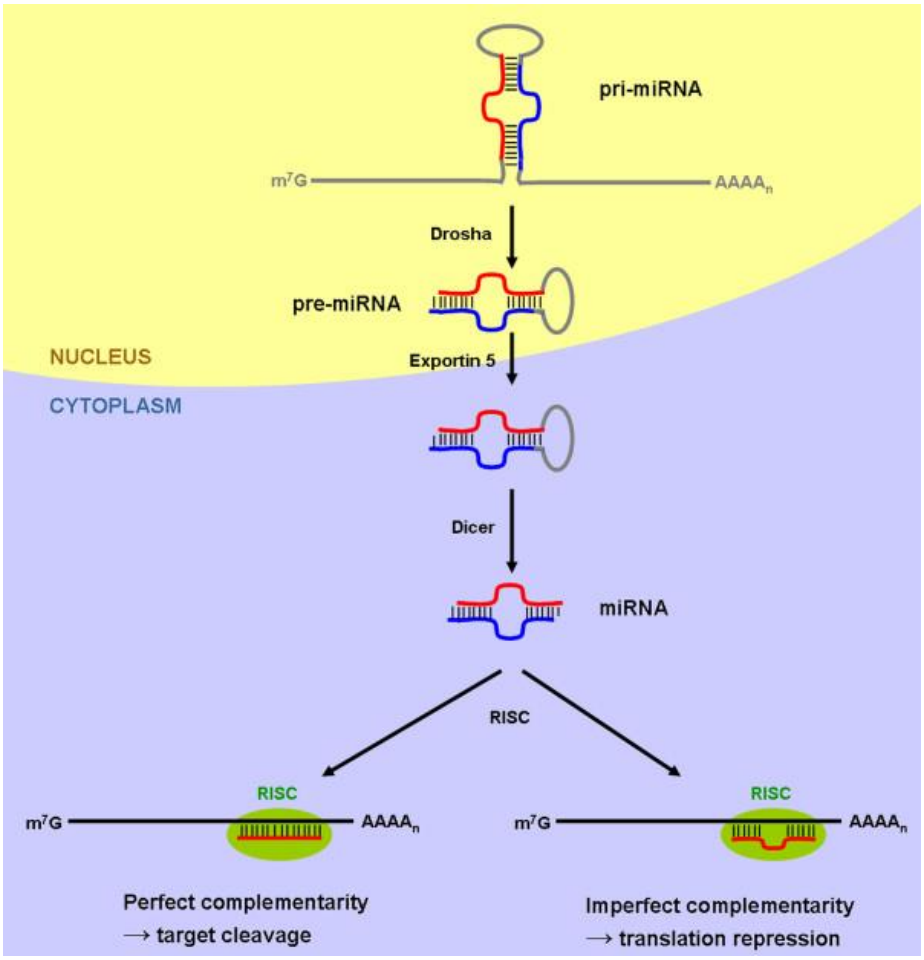


Figure 1.4. MicroRNA biogenesis. Pri-miRNAs produced by pol II result in a transcript that is capped at the 5' end and polyadenylated. The resulting pri-miRNA folds into a secondary double-stranded structure. Drosha cleaves pri-miRNAs to form pre-miRNAs that are exported from the nucleus by Exportin-5. Dicer cleaves the pre-miRNAs to produces a 20-25-nucleotide miRNA duplex with a 2-nucleotide overhang on the 3' end. MiRNAs and their target mRNA interaction cause either mRNA degradation or translational repression (Adapted from Hoffmann, et al., 2012).

Dicer cleavage creates a 2-nucletoide overhang on the 3' end of each miRNA strand (Lund et al., 2006, Murchison et al., 2004).

1.7.2.3. Plant miRNAs

Unlike mammalian miRNAs, plant miRNAs usually have a near-perfect base pairing interaction with their target mRNAs. Plant pri-miRNAs are generally transcribed by pol II from independent non-coding intergenic regions, by inverted duplication of protein-coding or non-protein-coding regions. This accounts for the extensive target complementarity required by plant gene regulation mechanisms. The majority of pri-miRNAs encode a single mature miRNA. Therefore, pri-miRNAs derived from intron splicing are rare (Axtell et al., 2011). Processing of pri-miRNAs to a mature miRNA duplex occurs in the nucleus by an RNaseIII Dicer homolog, Dicer-like 1 (DCL1). The miRNA duplex is also methylated at the 3' overhang by an RNA methyltransferase protein called Hua-Enhancer 1 (HEN1); which has not been observed in mammalian miRNAs. The duplex is then transported from the nucleus by an Exportin-5 homolog called Hasty (HST, Axtell et al., 2011, Lelandais-Briere et al., 2010).

1.7.2.4. Small interfering RNAs

In 1999, Baulcombe and Hamiliton discovered small interfering RNAs (siRNAs) while studying viral infection in plants (Hamilton and Baulcombe, 1999). However, siRNAs can also be endogenous (endo-siRNA) generated by three mechanisms; (i) a perfect match hairpin derived from inverted repeats, (ii) a cis-natural siRNAs (cis-nat siRNAs) transcribed from the same genomic locus as their target but from the opposite DNA strand or (iii) trans-natural siRNAs (trans- nat-siRNAs) transcribed at different loci

that encode for a gene and a pseudogene (Figure 1.5A, Elbashir et al., 2001). Endo-siRNAs are generated by a similar mechanism as miRNAs mentioned above (1.7.2.2.). However, exo-siRNA precursors originate from viral infection generated by an RNA-dependent RNA polymerase and processed by a cytosolic RNAseIII member (Figure 1.5, D.). On the other hand, endo-siRNAs are first transcribed and processed within the nucleus by an RNAseIII member then transported to the cytosol by an Exportin-5 homolog, as observed in mammalian miRNA biogenesis (Figure 1.5.). The resultant mature siRNAs duplexes contain a phosphorylated 5' end and a hydroxylated 3' end with a two-nucleotide overhang (Bernstein et al., 2001, Elbashir et al., 2001).

1.8. MiRNA and siRNA structure leads to function

When Dicer processes long dsRNA RNA molecules, the PAZ domain binds to the 3' end of the RNA substrate to aid in stabilization of the dsRNA for cleaving. Upon binding, dsRNA interacts with the intermolecular dimer of two RNAseIII domains of Dicer. Each RNAseIII domain cleaves one of the two-dsRNA strands and creates the canonical 3' 2 nucleotide overhang (Tomari and Zamore, 2005). This interaction and subsequent enzymatic reaction yield products of 20-25 nucleotides long. Argonaute proteins preferentially bind and keep the stand with the less stable paired 5' end of the small dsRNA. The passenger strand is degraded from the complex. Thermodynamically, the unwinding of A:U bonds are preferred to G:C bonds because of the fewer hydrogen bonds necessary to disrupt, hence, A:U rich 5' ends are preferred. Binding of the small RNA to Argonaute requires ATP for conformational changes to accommodate the small RNA (Kim, 2005). The "slicer" activity of Argonaute takes place within the Piwi domain where the phosphodiester linkage between nucleotides of the target mRNA are cleaved.

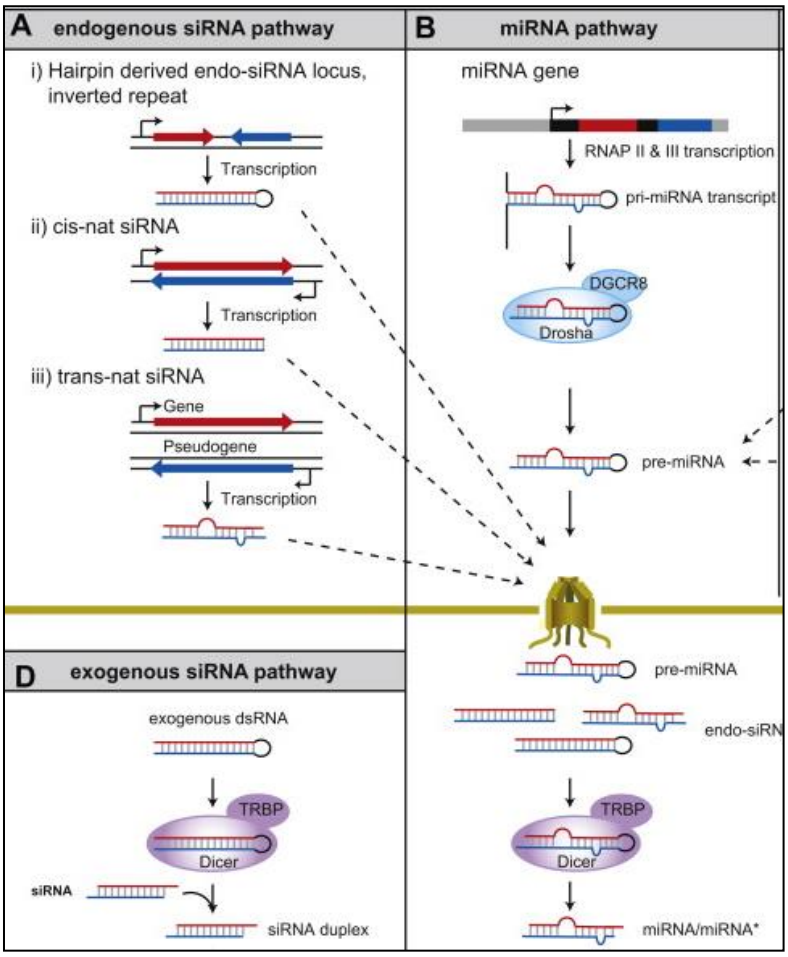


Figure 1.5. Origins of small non-coding RNAs. A. Endo-siRNAs originate from hairpin derived from inverted repeats, cis-natural siRNAs, or trans-natural. B. MiRNAs biogenesis. D. Exo-siRNA precursors originate from viral infection. Dicer processes small RNAs in the cytosol to produce effector RNAs. (Adapted from Rother and Meisster, 2011)

Once the cleaved mRNA is removed from the complex the guide strand can remain bound to Argonaute to engage another target mRNA (Tomari and Zamore, 2005).

1.9. Post-translation modifications by SUMOylation

In 1997, Melchior discovered a mysterious protein mobility shift that was due to the addition of a novel protein modification that she would later name SUMO (small ubiquitin-related modifier). This later lead to the discovery of the post-translational modification pathway called SUMOylation described as the reversible addition of a SUMO (small ubiquitin-like modifier) peptide (90-110 amino acids) to a protein substrate (Mahajan et al., 1997). Unlike ubiquitination, where proteins are polyubiquitinylated and sent to lysosomes for degradation, the addition of a SUMO peptide aids in protein localization, stabilization and activity (Hay, 2005, Komander and Rape, 2012). The addition of SUMO has been observed in proteins involved in DNA repair, transcription, RNA processing, nuclear-cytoplasmic transport, stress response and progression through the cell cycle (Hay, 2005, Liu and Shuai, 2008).

The pathway is initiated through the cleavage of a SUMO precursor by a protease called Ulp1 (ubiquitin-like protease 1). Ulp1 is a cysteine peptidase that is highly specific for the SUMO peptide as it recognizes the tertiary structure of SUMO rather than an amino acid sequence (Mossessova and Lima, 2000). To form a mature SUMO, Ulp1 cleaves amino acid residues from the C-terminal end of the SUMO precursor to reveal a di-glycine motif that is required for attachment to target proteins (Figure 1.6, A.). Ulp1 cleaves peptide bonds *via* hydrolysis that involves a thiol group of a cysteine residue as a nucleophile, a histidine and an aspartic acid, which forms the Ulp1 catalytic triad. The

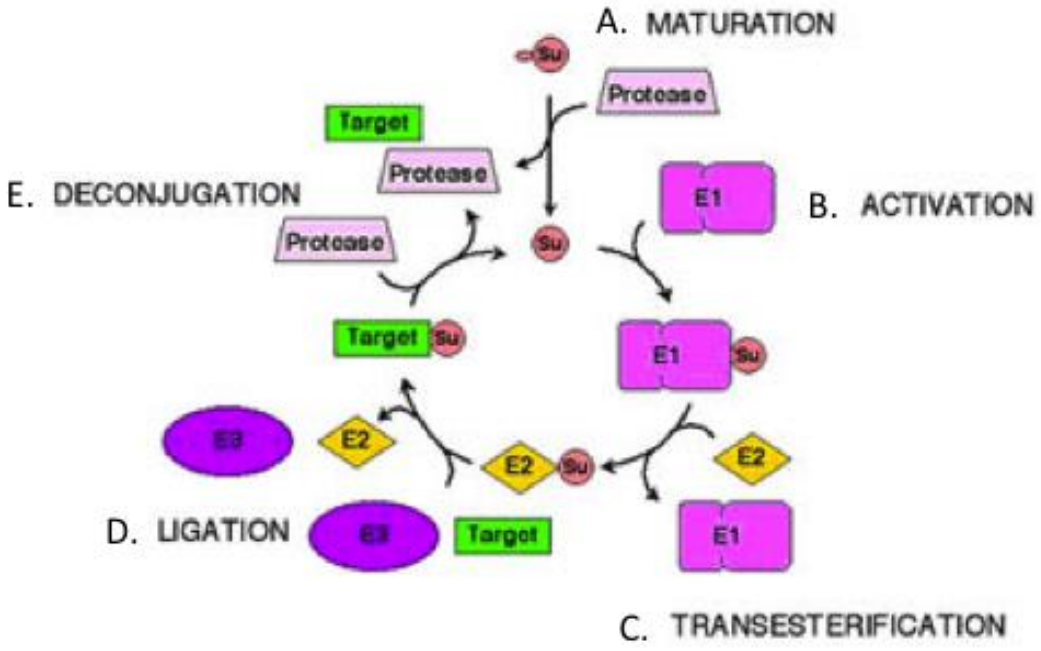


Figure 1.6. The SUMOylation pathway. The SUMOylation pathway is initiated by the presence of a SUMO precursor that is cleaved from the C-terminal end to reveal a di-glycine motif by Ulp1 (A). The mature SUMO is transferred to an E1 then E2 complex (B, C). E3 ligates SUMO to a protein substrate (D), which recognizes the ψKxE consensus sequence (where ψ is a large hydrophobic residue and x is any residue). Ulp1 also removes SUMO from conjugated protein substrates (E). (Adapted from Hay, 2013)

mature form of SUMO with its revealed di-glycine motif is transferred to the catalytic cysteine of an E1 activating enzyme by the formation of a thioester bond (Figure 1.6, B.). Next, SUMO is transferred to the catalytic cysteine residue in the E2 conjugating enzyme which directly binds to a protein substrate. SUMO is then conjugated to the protein substrate by an E3 ligase that recognizes a ψKxE consensus sequence (where ψ is a large hydrophobic residue and x is any residue) on the protein substrate (Figure 1.6. C and D.). The carboxyl group of the di-glycine motif is ligated to the ε–amino group of a lysine residue on the substrate. Ulp1 is also responsible for the removal of SUMO from SUMOylated proteins (Mossessova and Lima, 2000, Watts 2013).

1.10. Research Objectives

Toxoplasma contains all the canonical components of a post-transcriptional gene silencing mechanism including RNAi genes Argonaute and Dicer as well as effector RNAs. Although an RNAi-like gene silencing can be observed in *Toxoplasma* with long dsRNA, the functions of the underlying mechanism are still unknown. The following chapters describe my investigation towards the better understanding of RNA directed gene regulation in *Toxoplasma*.

The first goal (Chapter 2) describes the development of a reporter system allowing for the study of gene silencing in *Toxoplasma*. Using our dual luciferase reporter system we investigated the ability of endogenous *Toxoplasma* miRNAs to induce a gene silencing effect (Chapter 3). The *Renilla* luciferase transcript was engineered to contain binding sites for the two most abundant miRNA families, miR-60 and miR-4. The silencing effect caused by endogenous miRNAs could be altered by the use of miRNA mimics and inhibitors. As the most abundant miRNA family, miR-60 was

directed to down-regulate a gene of interest for functional studies. This new methodology proved extremely useful for fast and efficient down-regulation of an essential and non-essential gene. We compiled a comprehensive review on antisense RNA technologies and their current uses in *Toxoplasma* in Chapter 4. This review outlines the applications, advantages and disadvantages of antisense RNA as a genetic modifier. Although *Toxoplasma* miRNAs are functional, their underlying importance in gene regulation induced gene silencing remained unclear. Chapter 5 describes my work on the identification and characterization of a model gene target of *Toxoplasma* miRNAs. This endeavor uncovered an interesting self-regulating gene (TgUlp1) by the miR60 family whose expression and activity could greatly impact overall parasite cell biology. Disregulation of TgUlp1 expression created a cascade of events that can lead to parasite termination. Chapter 6 describes my work on *Toxoplasma* Argonuate. Even though the expression of TgAgo has been identified to a single locus, two forms of TgAgo are expressed. Further investigation is still required in this area however these are the first studies to suggest that multiple forms of Argonaute could arise from alternative translational start sites.

Overall, this dissertation reveals the importance of proper gene regulation directed by non-coding RNAs in *Toxoplasma*. These studies have led to the development of a methodology that can be used as a genetic tool to quickly and efficiently down-regulate a gene for functional analysis and to study its importance to parasite cell biology. Thereby, this tool could also lead to quicker and more effective drug development studies. Although the use of synthetic siRNAs in biomedical research to induced gene silencing as a drug therapy is promising (Schiffelers et al., 2004), and Tg-miRNA mimics and

inhibitors could be directed as an anti-*Toxoplasma* therapy, these studies were first aimed at understanding the post-transcriptional gene silencing mechanism by miRNAs.

1.11. References

Altuvia Y, Landgraf P, Litthwick G, Elefant N, Pfeffer S, Aravin A, Brownstein MJ, Tuschl T, Margalit. (2005) Clustering and conservation patterns of human microRNAs. Nucleic Acids Res. 12;33(8):2697-706.

Aravind L, Watanabe H, Lipman DJ, Koonin EV 2000. Lineage-specific loss and divergence of functionally linked genes in eukaryotes. Proc Natl Acad Sci 97: 11319–11324

Axtell MJ, Westholm JO, Lai EC. Vive la lectropor: biogenesis and evolution of microRNAs in plants and animals. Genome Biol. 2011;12(4):221.

Bartel DP, Chen CZ. Micromanagers of gene expression: the potentially widespread influence of metazoan microRNAs. Nat Rev Genet. 2004 May;5(5):396-400.

Bell JT, Pai AA, Pickrell JK, Gaffney DJ, Pique-Regi R, Degner JF, Gilad Y, Pritchard JK. DNA methylation patterns associate with genetic and gene expression variation in HapMap cell lines. Genome Biol. 2011;12(1):R10. doi:10.1186/gb-2011-12-1-r10. Erratum in: Genome Biol. 2011;12(6):405.

Berdoy M, Webster JP, Macdonald DW. Fatal attraction in rats infected with *Toxoplasma gondii*. Proc Biol Sci. 2000 Aug 7;267(1452):1591-4.

Bernstein E, Caudy AA, Hammond SM, Hannon GJ. Role for a bidentate ribonuclease in the initiation step of RNA interference. Nature. 2001;409:363–366.

Black MW, Boothroyd JC. Lytic cycle of *Toxoplasma gondii*. Microbiol Mol Biol Rev. 2000 Sep;64(3):607-23.

Blader IJ, Coleman BI, Chen CT, Gubbels MJ. Lytic Cycle of *Toxoplasma gondii*: 15 Years Later. Annu Rev Microbiol. 2015;69:463-85.

Blevins T, Rajeswaran R, Shivaprasad PV, Beknazariants D, Si-Ammour A, Park HS, Vazquez F, Robertson D, Meins F Jr, Hohn T, Pooggin MM. Four plant Dicers mediate viral small RNA biogenesis and DNA virus induced silencing. Nucleic Acids Res. 2006;34(21):6233-46.

Brkljacic J, Grotewold E. Combinatorial control of plant gene expression. Biochim Biophys Acta. 2016 Jul 15. Pii: S1874-9399(16)30137-7.

Brosius J. Echoes from the past–are we still in an RNP world? Cytogenet Genome Res. 2005;110(1-4):8-24.

Bruce Alberts; Alexander Johnson; Julian Lewis; Martin Raff; Keith Roberts; Peter Walter (2007). Molecular Biology of the Cell (Fifth ed.). Garland Science.

Carthew RW, Sontheimer EJ. Origins and Mechanisms of miRNAs and siRNAs. Cell. 2009 Feb 20;136(4):642-55.

Chuang CF, Meyerowitz EM. Specific and heritable genetic interference by double-stranded RNA in *Arabidopsis thaliana*. Proc Natl Acad Sci U S A. 2000 Apr 25;97(9):4985-90.

Dubey JP. *Toxoplasma Gondii*. In: Baron S, editor. Medical Microbiology. 4th edition. Galveston (TX): University of Texas Medical Branch at Galveston; 1996. Chapter 84.

Dubey JP. Tachyzoite-induced life cycle of *Toxoplasma gondii* in cats. J Parasitol. 2002 Aug;88(4):713-7.

Dubey JP, Jones JL. *Toxoplasma gondii* infection in humans and animals in the United States. Int J Parasitol. 2008 Sep;38(11):1257-78.

Dubey JP. *Toxoplasma gondii* oocyst survival under defined temperatures. J Parasitol. 1998 Aug;84(4):862-5.

Dubey JP. *Toxoplasma Gondii*. In: Baron S, editor. Medical Microbiology. 4th edition. Galveston (TX): University of Texas Medical Branch at Galveston; 1996. Chapter 84.

Dubey JP. The history of *Toxoplasma gondii*–the first 100 years. J Eukaryot Microbiol. 2008 Nov-Dec;55(6):467-75.

Elbashir SM, Lendeckel W, Tuschl T. RNA interference is mediated by 21- and 22-nucleotide RNAs. Genes Dev. 2001 Jan 15;15(2):188-200.

English JJ, Jones JD. Epigenetic instability and trans-silencing interactions associated with an SPT::Ac T-DNA locus in tobacco. Genetics. 1998 Jan;148(1):457-69.

Engstrom PG, Suzuki H, Ninomiya N, Akalin A, Sessa L, Lavorgna G, Brozzi A, Luzi L, Tan SL, Yang L, et al. 2006. Complex loci in human and mouse genomes. PloS Genet 2: e47.

Faller M, Matsunaga M, Yin S, Loo JA, Guo F 2007. Heme is involved in microRNA processing. Nat Struct Mol Biol 14: 23–29

Ferguson DJ, Hutchison WM, Siim JC. Proceedings: Comparison of the methods of asexual multiplication undergone by *Toxoplasma gondii*. Parasitology. 1974 Oct;69(2):xiii.

Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. Nature. 1998 Feb 19;391(6669):806-11.

Flegr J. How and why *Toxoplasma* makes us crazy. Trends Parasitol. 2013 Apr;29(4):156-63.

Flint JA. Van Duynhoven YT. Angulo FJ. DeLong SM. Braun P. Kirk M. Scallan E. Fitzgerald M. Adak GK. Sockett P. Ellis A. Hall G. Gargouri N. Walke H. Braam P. Estimating the burden of acute gastroenteritis, foodborne disease, and pathogens commonly transmitted by food: An international review. Clin Infect Dis. 2005;41:698–704.

Frenkel JK, Ruiz A, Chinchilla M. Soil survival of *Toxoplasma* oocysts in Kansas and Costa Rica. Am J Trop Med Hyg. 1975 May;24(3):439-43.

Frenkel JK. Pursuing *Toxoplasma*. J Infect Dis. 1970 Dec;122(6):553-9.

Friedman JM, Jones PA. MicroRNAs: critical mediators of differentiation, development and disease. Swiss Med Wkly. 2009 Aug 22;139(33-34):466-72. Doi: smw-12794.

Gerdes K, Larsen JE, Molin S. Stable inheritance of plasmid R1 requires two different loci. J Bacteriol. 1985 Jan;161(1):292-8.

Gajria B, Bahl A, Brestelli J, Dommer J, Fischer S, Gao X, Heiges M, Iodice J, Kissinger JC, Mackey AJ, Pinney DF, Roos DS, Stoeckert CJ Jr, Wang H, Brunk BP. ToxoDB: an integrated *Toxoplasma gondii* database resource. Nucleic Acids Res. 2008 Jan;36(Database issue):D553-6.

Goldman M, Carver RK, Sulzer AJ. Reproduction of *Toxoplasma gondii* by internal budding. J Parasitol. 1958 Apr;44(2):161-71.

Grimm D, Streetz KL, Jopling CL, Storm TA, Pandey K, Davis CR, Marion P, Salazar F, Kay MA. Fatality in mice due to oversaturation of cellular microRNA/short hairpin RNA pathways. Nature. 2006 May 25;441(7092):537-41.

Gubbels MJ, Morlon-Guyot J, Berry L, Chen CT, Lebrun M, Daher W. The *Toxoplasma gondii* calcium-dependent protein kinase 7 is involved in early steps of parasite division and is crucial for parasite survival. Cell Microbiol. 2014 Jan;16(1):95-114.

Hamilton AJ, Baulcombe DC. A species of small antisense RNA in posttranscriptional gene silencing in plants. Science. 1999 Oct 29;286(5441):950-2.

Hay RT. SUMO: a history of modification. Molecular Cell. 2005, 1;18(1):1-12

Hay RT. Decoding the SUMO signal. Biochem Soc Trans. 2013 Apr;41(2):463-73.

Holmes M, Liwak U, Pricop I, Wang X, Tomavo S, Ananvoranich S. Silencing of tachyzoite enolase 2 alters nuclear targeting of bradyzoite enolase 1 in *Toxoplasma gondii*. Microbes Infect. 2010 Jan;12(1):19-27.

Holmquist GP, Ashley T. Chromosome organization and chromatin modification: influence on genome function and evolution. Cytogenet Genome Res. 2006;114(2):96-125.

Hutchison WM. Experimental transmission of *Toxoplasma gondii*. Nature. 1965 May 29;206(987):961-2.

Hutchison WM, Dunachie JF, Siim JC, Work K. Coccidian-like nature of *Toxoplasma gondii*. Br Med J. 1970 Jan 17;1(5689):142-4.

Hüttenhofer A, Schattner P, Polacek N. Non-coding RNAs: hope or hype? Trends Genet. 2005 May;21(5):289-97.

Khan A, Jordan C, Muccioli C, Vallochi AL, Rizzo LV, Belfort R Jr, Vitor RW, Silveira C, Sibley LD. Genetic divergence of *Toxoplasma gondii* strains associated with ocular toxoplasmosis, Brazil. Emerg Infect Dis. 2006 Jun;12(6):942-9.
Kim, K., Weiss, L.M., 2004. *Toxoplasma gondii*: the model apicomplexan. Int. J. Parasitol. 34 (3), 423–432.

Kim VN. MicroRNA biogenesis: coordinated cropping and dicing. Nat. Rev. Mol. Cell Biol. 2005;6:376–385.

Kolev NG, Tschudi C, Ullu E. RNA interference in protozoan parasites: achievements and challenges. Eukaryot Cell. 2011 Sep;10(9):1156-63.

Komander D, Rape M. (2012) The ubiquitin code. Annual Review of Biochemistry. 81:203-29.

Korneev SA, Maconochie M, Naskar S, Korneeva EI, Richardson GP, O'Shea M. A novel long non-coding natural antisense RNA is a negative regulator of Nos1 gene expression. Sci Rep. 2015 Jul 8;5:11815.

Kupferschmidt K. A lethal dose of RNA. Science. 2013 Aug 16;341(6147):732-3.

LaCount DJ, Bruse S, Hill KL, Donelson JE. Double-stranded RNA interference in *Trypanosoma brucei* using head-to-head promoters. Mol Biochem Parasitol. 2000 Nov;111(1):67-76.

Laliberté J, Carruthers VB. Host cell manipulation by the human pathogen *Toxoplasma gondii*. Cell Mol Life Sci. 2008 Jun;65(12):1900-15. Doi: 10.1007/s00018-008-7556-x.

Lee Y, Kim M, Han J, Yeom KH, Lee S, Baek SH, Kim VN. MicroRNA genes are transcribed by RNA polymerase II. EMBO J. 2004 Oct 13;23(20):4051-60.

Lee RC, Feinbaum RL, Ambros V. The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell. 1993 Dec 3;75(5):843-54.

Lelandais-Brière C, Sorin C, Declerck M, Benslimane A, Crespi M, Hartmann C. Small RNA diversity in plants and its impact in development. Curr Genomics. 2010 Mar;11(1):14-23.

Liu B, Shuai K. Regulation of the sumoylation system in gene expression. Curr Opin Cell Biol. 2008 Jun;20(3):288-93. Doi: 10.1016/j.ceb.2008.03.014. Epub 2008 May 28.

Lund E, Dahlberg JE. Substrate selectivity of exportin 5 and Dicer in the biogenesis of microRNAs. Cold Spring Harb Symp Quant Biol. 2006;71:59-66.

Mahajan R, Delphin C, Guan T, Gerace L, Melchior F. A small ubiquitin-related polypeptide involved in targeting RanGAP1 to nuclear pore complex protein RanBP2. Cell. 1997 Jan 10;88(1):97-107.

Matzke MA, Matzke AJ. Planting the seeds of a new paradigm. PloS Biol. 2004 May;2(5):E133. Epub 2004 May 11.

Meister G, Tuschl T. Mechanisms of gene silencing by double-stranded RNA. Nature. 2004 Sep 16;431(7006):343-9.

Mendler L, Braun T, Müller S. The Ubiquitin-Like SUMO System and Heart Function: From Development to Disease. Circ Res. 2016 Jan 8;118(1):132-44.

Montoya, J.G.; Liesenfeld, O. Toxoplasmosis. Lancet, 2004, 363(9425), 1965-1976.

Morey C, Avner P. Employment opportunities for non-coding RNAs. FEBS Lett. 2004 Jun 1;567(1):27-34.

Mossessova E, Lima CD. (2000) Ulp1-SUMO crytal structure and genetic anaylsis reveal conserved interactions and a regulatory element essential for cell growth in yeast. Molecular Cell. 5(5):865-76.

Murchison EP, Hannon GJ. miRNAs on the move: miRNA biogenesis and the RNAi machinery. Curr Opin Cell Biol. 2004 Jun;16(3):223-9.

Musilova K, Mraz M. MicroRNAs in B-cell lymphomas: how a complex biology gets more complex. Leukemia. 2015 May;29(5):1004-17.

Okamura K, Ishizuka A, Siomi H, Siomi MC. Distinct roles for Argonaute proteins in small RNA-directed RNA cleavage pathways. Genes Dev. 2004 Jul 15;18(14):1655-66.

Palazzo AF, Lee ES. Non-coding RNA: what is functional and what is junk? Front Genet. 2015 Jan 26;6:2.

Preall JB, He Z, Gorra JM, Sontheimer EJ. Short interfering RNA strand selection is independent of dsRNA processing polarity during RNAi in *Drosophila*. Curr Biol. 2006 Mar 7;16(5):530-5.

Reinhart BJ, Slack FJ, Basson M, Pasquinelli AE, Bettinger JC, Rougvie AE, Horvitz HR, Ruvkun G. The 21-nucleotide let-7 RNA regulates developmental timing in *Caenorhabditis elegans*. Nature. 2000 Feb 24;403(6772):901-6.

Röther S, Meister G. Small RNAs derived from longer non-coding RNAs. Biochimie. 2011 Nov;93(11):1905-15.

Sabin AB, Tarro G. Herpes simplex and herpes genitalis viruses in etiology of some human cancers. Proc Natl Acad Sci U S A. 1973 Nov;70(11):3225-9.

Sarge KD, Park-Sarge OK. SUMO and its role in human diseases. Int Rev Cell Mol Biol. 2011;288:167-83.

Saurabh S, Vidyarthi AS, Prasad D. RNA interference: concept to reality in crop improvement. Planta. 2014 Mar;239(3):543-64.

Siomi H, Siomi MC. On the road to reading the RNA-interference code. Nature. 2009 Jan 22;457(7228):396-404.

Soete M, Hettman C, Soldati D. The importance of reverse genetics in determining gene function in apicomplexan parasites. Parasitology. 1999;118 Suppl:S53-61.

Splendore A. Un nuovo protozoa parassita de' conigli. Incontrato nelle lectr anatomiche d'une malattia che ricorda in molti punti il Kala-azar dell' uomo. Nota preliminare pel. Rev Soc Scient Sao Paulo.

Stavropoulos N, Lu N, Lee JT. A functional role for Tsix transcription in blocking Xist RNA accumulation but not in X-chromosome choice. Proc Natl Acad Sci U S A. 2001 Aug 28;98(18):10232-7. Epub 2001 Jul 31.

Tomari Y, Zamore PD. Perspective: machines for RNAi. Genes Dev. 2005 Mar 1;19(5):517-29. Review.

Verdel A, Jia S, Gerber S, Sugiyama T, Gygi S, Grewal SI, Moazed D. RNAi-mediated targeting of heterochromatin by the RITS complex. Science. 2004 Jan 30;303(5658):672-6.

Wapinski O, Chang HY. Long noncoding RNAs and human disease. Trends Cell Biol. 2011 Jun;21(6):354-61.

Watts FZ. (2013) Starting and stopping SUMOylation. What regulates the regulator? Chromosoma. 122(6):451-63.

Wicker T, Sabot F, Hua-Van A, Bennetzen JL, Capy P, Chalhoub B, Flavell A, Leroy P, Morgante M, Panaud O, Paux E, SanMiguel P, Schulman AH. A unified classification system for eukaryotic transposable elements. Nat Rev Genet. 2007 Dec;8(12):973-82.

Wightman B, Ha I, Ruvkun G. Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in *C. elegans*. Cell. 1993 Dec 3;75(5):855-62.

Yoon J-H, Abdelmohsen K, Gorospe M. 2013. Posttranscriptional gene regulation by long noncoding RNA. J Mol Biol 425: 3723–3730

CHAPTER 2

Evaluation of the ability of short and long double-stranded RNAs to induce homologous gene silencing in the protozoan parasite, *Toxoplasma gondii*

2.1. Introduction

Since its discovery (Elbashir et al., 2001 and Fire et al., 1998), RNA interference (RNAi) has become one of the most effective tools used in biological research for altering gene expression patterns for studying functional genetics. The RNAi methodology has expanded to serve many applications and purposes in both diagnostic and medical sciences (Grimm, 2009 and Kim and Rossi, 2008). Fundamentally, RNAi is a naturally occurring, evolutionary conserved, highly efficient and specific pathway by which short double-stranded RNAs (dsRNA) trigger the inhibition of gene expression (Carthew et al., 2009 and Meister et al., 2004). In metazoans, dsRNAs are expressed in the form of long primary transcripts, called pri-miRNAs. Pri-miRNAs are single, fold-back molecules that form dsRNA structures by intramolecular interactions, which are then processed by one or more RNAseIII members (i.e. Drosha and Dicer) into short and functional silencing RNAs, called miRNAs (Carthew et al., 2009 and Grimm, 2009). One single strand of the miRNA, called the mature guide or antisense strand, is then loaded onto the RNA-induced silencing complex (RISC), which allows it to bind its mRNA target to induce gene silencing (Brodersen et al., 2009). In protozoans, such a mechanism has been elucidated in *Trypanosoma brucei*, a parasitic and pathogenic kinetoplast (Ngo et al., 1998), which has a more complex silencing mechanism than *Tetrahymena*, a ciliated paramecium (Schramke et al., 2004).

The discovery of an RNAi mechanism in *T. brucei* (Ngo et al., 1998) has sparked considerable interest in exploiting the mechanism for the study of pathogenesis in a

variety of pathogenic protozoan parasites, including *T. cruzi* (Chagas disease), *Leishmania spp*. (leishmaniasis), *Plasmodium spp*. (malaria) and *T. gondii* (toxoplasmosis). Limited success has been reported (Kolev et al., 2011), and its possible causes could lay in their evolutionary divergence (Ullu et al., 2004). The genome sequence revealed the absence of RNAi genes in *T. cruzi* (El-Sayed et al., 2005) and old-world *L. major* and *L. donovani* (Ivens et al., 2005), thus explaining their lack of response to dsRNA. Enigmatically, *P. falciparum*, whose genome lacks conserved RNAi-genes, was shown to exhibit homologous gene silencing, akin to RNAi, where dsRNA could trigger down-regulation of gene expression (McRobert et al., 2002). Based on their evolutionary divergence, only some species have conventional RNAi genes and functions; while others retain only some homologs and/or functions. More investigations are required to unravel the mechanism of homologous gene silencing and to render a fruitful genetic tool for these protozoan parasites.

Among these evolutionarily distant protozoan organisms, *T. gondii* harbors an ortholog of Dicer, Argonaute and RNA-dependent RNA polymerase (Al Riyahi et al., 2006, Braun et al., 2010 and Kolev et al., 2011). *T. gondii*, an obligate intracellular parasite of the phylum Apicomplexa, is one of the most successful parasitic pathogens that has infected an estimated one third of the population and is a major threat in immunocompromised individuals (Montoya and Liesenfeld, 2004). To gain a better understanding of the biology and pathogenesis of *T. gondii*, a homologous gene silencing *via* RNAi methodology offers an excellent genetic tool. Although the canonical RNAi mechanism has not yet been demonstrated in *T. gondii*, it was shown that miRNAs in *T. gondii* (*Tg*-miRNAs) are expressed in a high abundance in some lineage strains such as,

RH (type I), PRU (type II) and CTG (type III) (Braun et al., 2010). Some of these annotated *Tg*-miRNAs were shown to associate with translation machinery, suggesting their functions in the regulation of gene expression similar to other eukaryotes. Moreover the possible targets of *Tg*-miRNA were predicted to interact by perfect or partial base-pair interaction similar to those of siRNA and miRNA of other eukaryotes (Braun et al., 2010).

Despite a lack of a mechanistic characterization, our group is one of few laboratories which successfully showed that homologous gene silencing, similar to RNAi, is functional and can be used in the study of *T. gondii* biology (Al-Anouti and Ananvoranich, 2002, Al-Anouti et al., 2003, Al-Anouti et al., 2004, Adams et al., 2005, Ananvoranich et al., 2006, Yu et al., 2008 and Yu et al., 2009, Holmes et al., 2010). However, these results have proved difficult to reproduce in other laboratories (Kolev et al., 2011). Here, in an attempt to standardize homologous gene silencing in *T. gondii*, a dual luciferase reporter system was employed so that the expression of *Photinus* (firefly, Ffluc) and *Renilla* (Rnluc) luciferases, served as an internal control and a test target, respectively. Such a reporter system would allow any interested laboratories to set up comparable experiments to ensure reproducibility of the RNAi activity. A similar reporter system was successfully used in the study of an RNAi mechanism in mammalian cell cultures, including Hela, Cos-7 and Huh-7 cells (Choe et al., 2010 and Elbashir et al., 2001). The reporter system would thus allow for an unbiased analysis of the silencing capability and effectiveness found in *T. gondii*, in comparison to those of other organisms. In addition, we have demonstrated that both long and short dsRNA are capable of invoking homologous gene silencing in *T. gondii* with different silencing

abilities. Here we referred to the short dsRNA with perfect-matched central bases as "siRNA", and that with mismatched central bases as "miRNA". This study was the first to show the ability of miRNA to induce gene silencing in *T. gondii*.

2.2. Material and Methods

2.2.1. Cell and parasite cultures

Human foreskin fibroblasts (HFF; ATCC, #1041) were maintained using Dulbecco's Modified Eagle Media (D-MEM) with high D-glucose and L-glutamine (Invitrogen, #12100046) supplemented with 10% cosmic calf serum (ThermoFisher Scientific, Hyclone, #H3008704N) and diluted (0.5x) Antibiotic-Antimycotic (Invitrogen, #15240-062) at 37 °C in 5% CO_2. RHΔHX, *T. gondii* type I strain (NIH AIDS Research and Reference Reagent Program, #2857) was cultured in confluent HFF using Minimum Essential Media (MEM, Invitrogen, #61100061) supplemented with 1% dialyzed fetal bovine serum (ThermoFisher Scientific, Hyclone, #SH3007903) and diluted (0.5x) Antibiotic-Antimycotic.

2.2.2. Reporter plasmids

Expression plasmids, pTubRn and pTubFf, were derived from pRL-3xBulge and pFL-Con (Pillai et al., 2005) and pTub_eGFP_mCherry_CAT which were gifts from W. Filipowicz (Friedrich Miescher Institute for Biomedical Research, Switzerland) and J. Murray (University of Pennsylvania, USA), respectively. In separate reactions, the DNA fragments containing the coding sequence of Rnluc and miRNA-testing target sites and the coding sequence of Ffluc were excised from the plasmids pRL-3xBulge and pFL-Con

following (i) Not I digestion, (ii) Klenow Fragment treatment, (iii) NheI digestion and (iv) gel extraction. The resulting DNA fragment containing either designed Rnluc-silencing target or Ffluc was ligated to prepared pTub_eGFP_mCherry_CAT. In the preparation of pTub_eGFP_mCherry_CAT, the coding sequence of eGFP_mCherry fusion was removed following (i) Afl II digestion, (ii) Klenow Fragment treatment, (iii) Nhe I digestion and (iv) gel extraction to give the cloning vector carrying the *T. gondii*-Tubulin (TgTub) promoter and the polyadenylation signal, contained within the 3'-UTR of *T. gondii* dihydrofolate reductase (*Tg*DHFR), flanking the Rnluc or Ffluc coding sequence, and chloramphenicol acetyltransferase cassette. Resultant pTubRn and pTubFf plasmids were subjected to restriction endonuclease analyses and nucleotide sequencing reactions to confirm their identity and correct open reading frame.

2.2.3. *In vitro* transcription of long and short dsRNAs

In vitro transcription reactions were carried out using either 2 μg DNA amplicons or 2 μM annealed primers as templates in the presence of T7 RNA polymerase (~5 units), 80 mM HEPES-KOH (pH 7.5) 24 mM $NaCl_2$, 2 mM spermidine, 40 mM DTT, 10 mM rNTPs, and ~1 unit pyrophosphatase in 50 μL reaction mixtures at 37° C for 16 hours, similar to previously described conditions (Al-Anouti and Ananvoranich, 2002). Resultant RNAs were extracted with one volume of phenol-chloroform mixture (1:1), subsequently precipitated and quantified using a spectrophotometer (Thermo Scientific NanoDrop2000). The DNA amplicons were prepared using oligonucleotide primers containing a gene specific sequence and the sequence of the T7 promoter, as listed in Table 2.1, under standard conditions for Taq DNA polymerase (New England Biolabs).

Table 2.1. Oligonucleotide primers used in the study. The sequence of T7 RNA polymerase promoter and its complementary bases are in bold and underlined, respectively.

Names	Nucleotide sequences
FW_RnLuc_1	**TAATACGACTCACTATA**GGTGTACGACCCCGAGCAAC
RV_RnLuc_1	**TAATACGACTCACTATA**GGTAGGCAGCGAACTCCTCA
FW_RnLuc_2	**TAATACGACTCACTATA**GGGACGAGTGGCCTGACATC
RV-RnLuc_2	**TAATACGACTCACTATA**GGGCAGCTTCTGTTTACTTA
FW_UP	**TAATACGACTCACTATA**GGTTCATGAGCACCACATTGGG
RV_UP	**TAATACGACTCACTATA**GGGTTCCAAAGTACCGGT
T7promoterGG	**TAATACGACTCACTATA**GG
Sense_Tpl	AATGAGGTAGTTCAATAGGCTGTGCC<u>TATATGAGTCGTATTA</u>
Antisense_match_Tpl	AACACAGCCTATTGAACTACCTCACC<u>TATATGAGTCGTATTA</u>
Antisense_mismatch_Tpl	AACACAGCCTACCTGGCTACCTCACC<u>TATATGAGTCGTATTA</u>
Sense_CTRL	AAGGGATCGACAAGGAGCTTTCC<u>TATAGTGAGTCGTATTAC</u>
Antisense_CTRL	AAGGAAAGCTCCTTGTCGATCCC<u>TATAGTGAGTCGTATTAC</u>

For *Rnluc*-dsRNAs, pTubRnluc was used as templates for PCR amplification using two sets of oligonucleotide primers: FW_RnLuc_1 and RV_RnLuc_1 to yield *Rnluc*-dsRNA_1, or FW_RnLuc_2 and RV-RnLuc_2, to yield *Rnluc*-dsRNA_2. The amplicon used in the generation of control dsRNA was amplified from a plasmid containing *T. gondii* uracil phosphoribosyl transferase (TgUPRT) and a pair of oligonucleotide primers, called FW_UP and RV_UP.

For short dsRNA *in vitro* transcription, annealed oligonucleotide primers were used as templates. The oligonucleotide primers were designed to contain a complementary sequence of the T7 RNA polymerase promoter and the target sequences to independently generate sense and antisense strands of the RNA. The templates for the sense RNA strand synthesis were formed by heating a mixture of 2 μM T7promoterGG and Sense_Tpl for 5 minutes at 70 °C and slowly cooling to room temperature. The mixture was then used in the *in vitro* transcription reaction. Similar procedures were carried out to create the templates for the antisense strands carrying a perfect or mismatch complementarity by using T7promoterGG to anneal with Antisense_match_Tpl or Antisense_mismatch_Tpl respectively. Following the *in vitro* transcription, resultant RNAs were quantified. Equal amounts of sense and antisense RNAs were mixed and annealed by heating the mixture to 70 °C for 5 minutes and cooled to room temperature to generate double-stranded silencing RNA with perfect-matched central bases (siRNA) or with mismatched central bases (miRNA) (Figure 2.1.). The sense and antisense strands of the control RNA were generated similarly using T7promoterGG to anneal with Sense_CTRL or Antisense_CTRL, respectively, to create templates for *in vitro* transcription reactions.

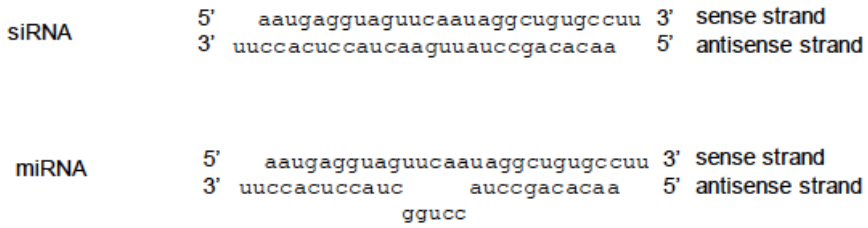


Figure 2.1. Sequences of short dsRNAs.

2.2.4. Dual luciferase assay following transfection by electroporation

Freshly lysed parasites were harvested and counted. Approximately $2x10^6$ parasites were used for each transfection *via* electroporation using a BTX ECM 630 (1500 volts, 25 Ω, and 25 μF). For each transfection, reporter plasmids with or without tested RNAs were mixed with harvested parasites in an electroporation mixture (400 μL) containing 120 mM KCl, 0.15 mM $CaCl_2$, 10 mM K_2HPO_4/KH_2PO_4 (pH 7.6) 2 mM EDTA, 5 mM $MgCl_2$, 2 mM ATP, 5 mM glutathione in a 4 mm-gap cuvette and electroporated (Roos et al., 1994). Following electroporation, the parasites were cultured in confluent HFF monolayers grown in T25 flasks using MEM media with L-glutamine without phenol red (Sigma Aldrich, #M3024) and supplemented with 1% dialyzed fetal bovine serum and diluted (0.5x) Antibiotic-Antimycotic at 37 °C in 5% CO_2 for selected periods.

For dual luciferase assays, the infected monolayers were harvested and lysed with 100 μL of 1x Passive Lysis Buffer (Promega, #E1531) and incubated for 10 minutes at room temperature. Lysates were cleared of debris by centrifugation (10,000 x g for 1 minute), and the supernatant was kept on ice and used in the dual luciferase assay, which was carried out in a 2-step fashion. For a Ffluc assay, 20 μL of lysate was added to a freshly made reaction mixture (100 μL) containing 200 μM D-luciferin (Sigma Aldrich, #L9504), 20 mM Tricine, 10 mM $MgSO_4$, 5 mM DTT, 250 μM ATP and 250 μM Coenzyme A. The mixture was incubated for 10 seconds at room temperature. For a Rnluc assay, 20 μL of lysate was added to a freshly made reaction mixture (100 μL) containing 0.1 μM Coelenterazine (Santa Cruz Biotechnology Inc, #sc-205908), 100 mM K_2HPO_4/KH_2PO_4 (pH 7.67.6) 500 mM NaCl, 1 mM EDTA, and 0.02% BSA. The mixture was incubated for 10 seconds at room temperature. The luminescence signals

from both reactions were measured with a 20/20n Luminometer (Turner Biosystems). Rnluc signals were normalized to Ffluc signals for a direct comparison across independent experiments. For each experiment, the reporter plasmid electroporation without tested RNA, called mock electroporation, was carried out to allow for the measurement of the dual luciferase expression under independently tested conditions. At least two independent experiments were performed.

2.3. Results and Discussion

2.3.1. Strategy and dual luciferase reporter system

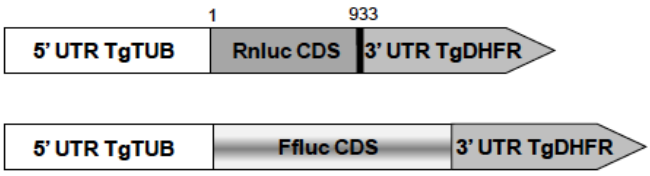
In the design of the dual luciferase reporter system for the gene silencing study, we aimed to have an equal transcript production of both transcription units. Therefore, the reporter system was engineered so that the expression of both Ffluc and Rnluc are under the control of a strong and constitutive *T. gondii* promoter, called *Tg*Tubulin, which also provides the necessary 5'-UTR for the protein's expression. Furthermore, to ensure similar levels of transcript stability and protein synthesis from both constructs, the 3'-UTR of *Tg*DHFR (dihydrofolate reductase) were placed downstream from the coding sequence (Figure 2.2a.). The transcript derived from Rnluc was used as the target for the homologous gene silencing by three different RNA modulators, including long dsRNA, miRNA and siRNA, while the Ffluc transcript served as an internal control for transfection, expression, and detection.

Due to the parasitic nature of *T. gondii*, in a gene expression and silencing analysis, the reporter system and silencing RNAs would be introduced into the parasite prior to infection of a host monolayer. We viewed that the success of the subsequent

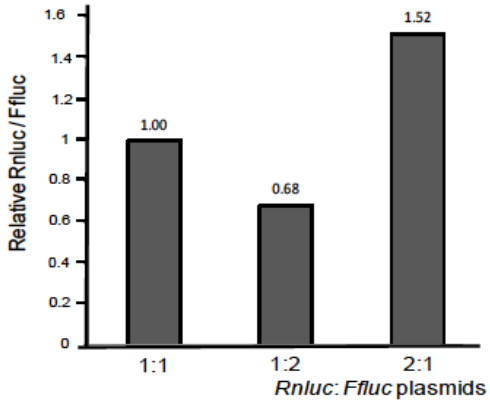
analysis would rely greatly on the viability of the post-electroporated parasite and the efficiency of transfection. The viability of the post-electroporated parasite was evaluated using the ability of the viable parasite to exclude trypan blue and propidium idodide (Jones et al. 1985). From three independent experiments we detected that 75 ± 10% were viable and able to multiply when up to 60 μg of nucleic acids (RNA or DNA) were used in the electroporation conditions (see Material and Methods). The efficiency of transfection was evaluated by determining the amounts of remaining and internalized nucleic acids following electroporation. Nucleic acids (RNA or DNA) were randomly labeled with fluorescein during *in vitro* transcription or PCR amplification, and used in the electroporation. Approximately 52% of labeled nucleic acids were internalized (Al-Anouti & Ananvoranich, unpublished data). We thus concluded that under similar controlled conditions, the electroporation of nucleic acids (60 μg, DNA plasmids and/or RNAs) would allow sufficient numbers of viable parasites for gene silencing assays.

Next we assessed the expression of our newly engineered dual luciferase system to allow for a sensitive detection of a broad range of different silencing activities, while remaining within a linear range for a direct comparison. Based on a previous report (Behnke et al., 2008) using a similar reporter system in *T. gondii*, when a promoter of interest and a well-characterized promoter were placed to control the expression of Rnluc and Ffluc, respectively, the transfection was carried out using 60 μg of the Rnluc and Ffluc reporter plasmids (1:2 ratio). Keeping the total amount of DNA used at 60 μg, we varied the ratio of Rnluc:Ffluc at either 1:1, 1:2 or 2:1. All ratios tested yielded reproducible data (Figure 2.2b). Consequently, the plasmid ratio at 1:1 was kept for

A.



B.



C.

Rnluc/Ffluc Total amount (µg)	Rnluc activity (light units)	Ffluc activity (light units)
0.2	1974 ± 7.1	2851 ± 3.2
2	2972 ± 5.8	3975 ± 2.1
10	4434 ± 10.2	10319 ± 4.8
20	34086 ± 5.4	38675 ± 4.4
40	41980 ± 3.9	50044 ± 3.6

Figure 2.2. Dual luciferase reporter system. (A) Schematic representation of mRNA reporters used in this study. Numbers indicate the first nucleotide of the start codon and the last nucleotide of coding sequence. The transcript encoding Rnluc was used as target for the gene silencing study and from which silencing RNAs were designed (see Table 2.1 and Materials and Methods). (B) Dual luciferase expression was assayed using three different ratios of Rnluc and Ffluc plasmids, while the total amount of plasmids were kept at 60 µg for the transfection. The graph is a representative of compiled experimental results. (C) Dose-response of dual luciferase expression was assessed when the ratio of Rnluc:Ffluc plasmids were kept at 1:1 ratio. The expression of the dual luciferase is shown as relative Rnluc/Fluc light forming units. The assays were carried out 24 hours post-electroporation. Values were obtained from at least two independent experiments.

simplicity. Other experimental parameters, including the total plasmid concentration, the number of extracellular parasites, and incubation time (post-transfection and conditions luciferase assays), were evaluated in different combinations to establish optimization for the study. When varied numbers of harvested parasites were used in the transfection, ~10^5-10^6 freshly released parasites were the most suitable for the study because the parasites remained intracellular at 24, 48 or 72 hours post-electroporation under the conditions tested (see Materials and Methods). These data indicated that both transcription units were efficiently expressed and maintained during the testing period (3 days). Because lower amounts of the reporter system were used in other eukaryotic systems (Choe et al., 2010 and Elbashir et al., 2001), we thus progressively lowered the total amounts of plasmids from 40 to 0.2 μg (Ffluc:Rnluc = 1:1) so that the gene silencing data could be directly compared. It was detected that within the amounts of plasmids tested; both transcription units were still sufficiently expressed and allowed for detection within a linear range (Figure 2.2.). In all succeeding silencing analyses, 1 μg of each reporter plasmid (2 μg total) was used.

2.3.2. Homologous gene silencing by long double-stranded RNA

Long dsRNAs (*ca.* 300-800 nts) have previously been shown to induce homologous gene silencing in *T. gondii* (Al-Anouti and Ananvoranich 2002, Al-Anouti et al., 2003, Al-Anouti et al., 2004, Ananvoranich et al., 2006, Holmes et al., 2010). To test our newly developed reporter system, two species of long dsRNAs were designed to have homologous sequences to the *Rnluc* mRNA. *Rnluc*-dsRNA_1 (596 nts) was homologous to nucleotides 12 to 607 of the *Rnluc* coding sequence, and *Rnluc*-dsRNA_2 (589 nts) was derived from the Rnluc coding sequence at nucleotides 458-933 and the proximal

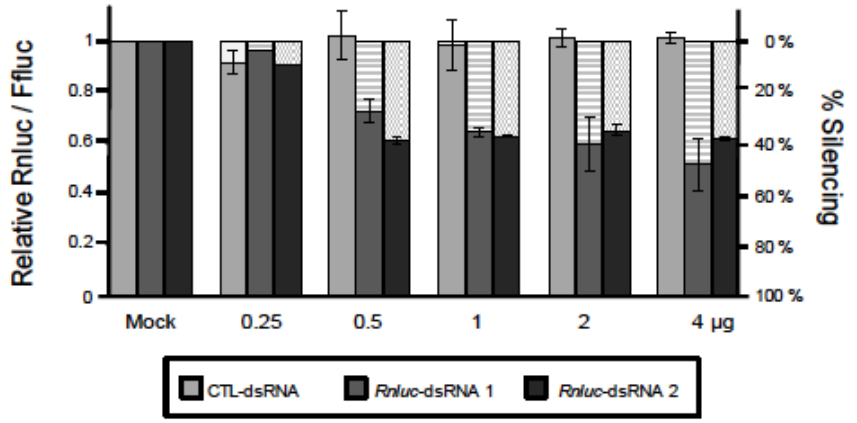
region (113 nts) of its 3'-UTR (Table 2.2.). These long dsRNAs have an overlapping region (149 nts) spanning the nucleotides 458 to 607 of the Rnluc coding sequence. Using these dsRNAs, we aimed to evaluate whether dsRNAs derived from these regions exhibited similar silencing inducing ability. This in turn would reflect inter-molecular interactions formed between the mRNA target and different dsRNAs, given potential structural preferences.

To determine the extent of dsRNA ability to induce gene silencing, varied amounts of tested dsRNAs (0.25 to 4 μg) were co-transfected along with 2 μg reporter plasmids (1:1) by electroporation. When 0.25 μg dsRNA was used, a negligible reduction (<10%) of the relative Rnluc/Ffluc level was detected at 24 hours post-electroporation (Figure 2.3, A). A clear silencing effect (~30-40%) was observed when 0.5 μg of *Rnluc*-dsRNA_1 and 2 were used and able to reduce the relative Rnluc/Ffluc level to 73 ± 4 % and 61 ± 1%, in comparison with a mock silencing control (without dsRNA). The silencing effects were slightly increased when higher amounts of both dsRNAs were used. The maximal silencing ability was observed when 2-4 μg dsRNA caused the relative Rnluc/Ffluc reduction to 52 ± 10% and 62 ± 1% for *Rnluc*-dsRNA_1 and dsRNA_2, respectively at 24 hours post-electroporation (Figure 2.2a). Similar levels of silencing were observed when the electroporated parasites were allowed to incubate longer, and the dual luciferase assays were carried out at 48 hours post-electroporation (Supplementary Figure 2.1.). The findings suggested that the transfection of long dsRNAs at 2 μg or ~5.1 pmoles, despite their targeted regions, have saturated the silencing machinery and reached their maximal ability to cause a silencing effect under conditions tested.

Table 2.2. List of silencing RNAs used in the study. Numbers of the nucleotides are corresponding to the Rnluc transcript depicted in Figure 2.2a.

Silencing RNAs	**Homologous to regions**	**Homologous to *Rnluc* transcript (Nucleotides)**
Rnluc dsRNA_1	CDS	12 - 607
Rnluc dsRNA_2	CDS & 3'-UTR	458 - 1045
siRNA	3'-UTR	943-965, 972-994, 1000-1022
miRNA	3'-UTR	943-965, 972-994, 1000-1022

A.



B.

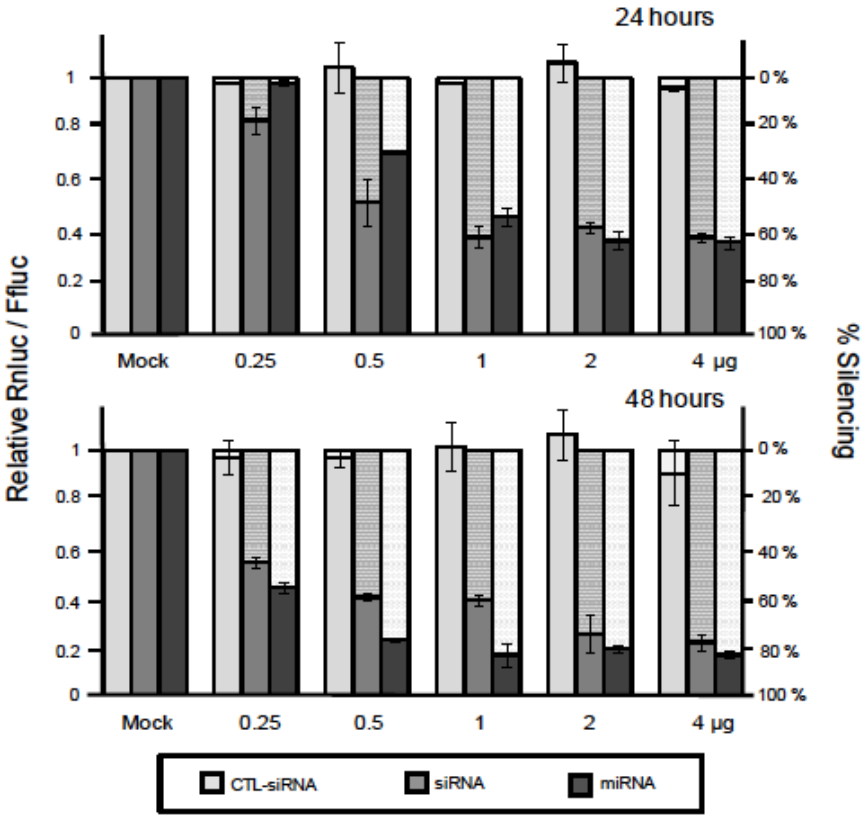


Figure 2.3. Gene silencing effect caused by (A) long dsRNA was detected at 24 hours post-electroporation (B) and siRNA and miRNA at 24 hours and 48 hours post- plasmids (1 µg of each Rnluc and Ffluc plasmid) in the absence of tested silencing RNAs.

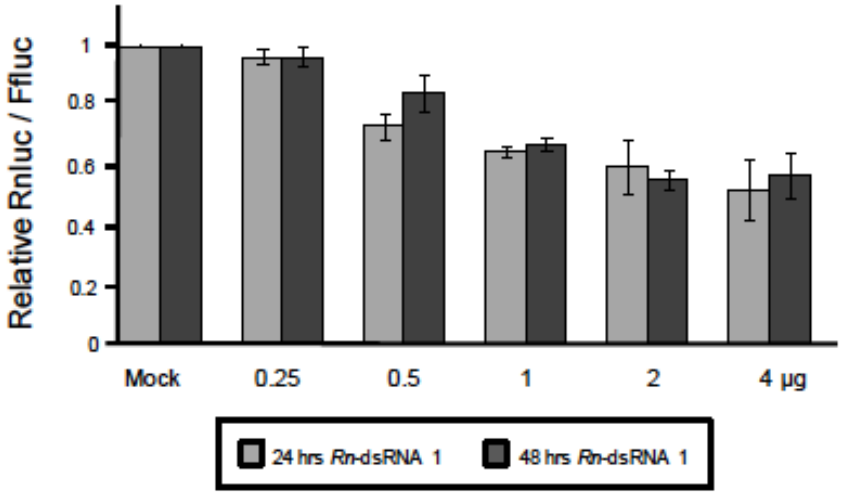


Figure Supplementary 2.1. Gene silencing effect caused by Rnluc-dsRNA_1 was detected at 24 and 48 hours post-electroporation.

To evaluate the specificity of the homologous gene silencing, an unrelated dsRNA whose nucleotide sequence is different from the intended target (the *Rnluc* transgene) was tested under similar conditions. *TgUPRT* dsRNA (629 nts) did not cause any silencing effect on Rnluc expression (Figure 2.3a, the lightest shaded bars), indicating that the silencing effect is sequence specific. The presence of unrelated dsRNA (as high as 4 μg or 10.3 pmoles) did not affect the activity of Rnluc or Ffluc. It thus indicated that in a unicellular organism, such as *T. gondii*, the presence of long dsRNA in a large quantity (4 μg) did not cause a general cellular response affecting the transcription processes and/or the degradation of gene products over the tested conditions.

In summary, we have confirmed that a homologous long dsRNA has the ability to induce gene silencing in *T. gondii*. DsRNA exhibited the ability to induce *Rnluc* silencing and to drop the relative Rnluc/Ffluc to 60% in comparison to a mock electroporation. This silencing effect (~40% reduction of gene expression) was considered a mild effect in comparison to when miRNA was used in silencing (~80% reduction) the expression of Rnluc in mammalian cell cultures (Choe et al., 2010 and Elbashir et al., 2001). However, such a mild silencing ability could be useful, when it is sufficient to produce a detectable phenotype (i.e. the change in an enzymatic activity). However, when *Tg*UPRT or *Toxoplasma* lactate dehydrogenases (*Tg*LDHs) were silencing targets, their homologous dsRNAs were able to reduce their expression by ~90%. In those experiments, Al-Anouti et al., (2003) detected the silencing outcomes by radioactive uracil uptake assays for *Tg*UPRT study and by *in vitro* and *in vivo* growth and differentiation assays for *Tg*LDH1 and *Tg*LDH2 experiments (Al-Anouti et al., 2004). It is highly likely that the level or extent of silencing effect is gene-dependent, since here the silencing effect towards Rnluc

was observed at ~40%, similar to that of *Toxoplasma* enolase experiments (Holmes et al., 2010).

Despite its mild silencing ability, a homologous dsRNA-induced gene silencing procedure offers a very easy-to-use method. One would require little information about the gene target, such as the sequence of the transcripts or the exons, prior to initiating a silencing protocol. Moreover, long dsRNAs can quickly be synthesized in a very cost-effective way *via in vitro* transcription. One drawback of the system would be due to its inherited long target region(s). When more than one species of transcripts are produced from the same exon(s), the long homologous dsRNA would not be able to discern one transcript over another. Considering its *pros* and *cons*, this homologous dsRNA-induced gene silencing methodology is a valuable method-of-choice. Keeping in mind, with its potentially mild silencing ability, the lack of a desirable phenotype might not indicate the failure of the silencing procedure, but could be an indication of the critical level of gene expression required for maintaining a phenotype.

2.3.3. Homologous gene silencing by short double-stranded RNAs

Due to their availability and reliability to induce gene silencing, gene-specific siRNAs and miRNAs have become powerful reagents used in biological research. Recent genomic and biochemical research has pointed that the *T. gondii* genome harbors both the RNAi genes and miRNAs (Al Riyahi et al., 2006, Braun et al., 2010, Kolev et al., 2011). It was thus crucial to investigate whether difference in complementarity could induce gene silencing of designated gene target. Based on predicted secondary structures and known constraints of short dsRNAs (Kim et al., 2005), we used the mfold nucleic acid folding and hybridization prediction software (Zuker 2003) and SciTool RNAi design

software (Integrated DNA Technologies, Inc., Iowa) to predict the most probable target sites and to design short dsRNAs specific to the *Rnluc* transcripts. Twenty possible targets and siRNA candidates were predicted and ranked from the most to the least probable location for the interactions. We chose the three most probable sites located between the *Rnluc* stop codon and the proximal region (103 nts) of its 3'-UTR (Table 2.2). Two silencing RNAs were designed and synthesized for the study. One contains nucleotides perfectly complementary to its target, representing a siRNA interaction; while the other, with five-central nucleotide mismatch, represents a miRNA interaction (Figure 2.1). These silencing RNAs have three possible recognition sites located within the 3' UTR between the nucleotides 943 and 1022 of the *Rnluc* transcript (Table 2.2).

Varied amounts of siRNA and miRNA (0.25 to 4 μg, or 15 to 246 pmoles) were independently tested for their ability to induce gene silencing using conditions similar to those tested for the long dsRNA-induced gene silencing. At 0.25 or 0.5 μg, siRNA was better than long dsRNA or miRNA (at the same amounts) to lower the *Rnluc* expression (Figure 2.3b upper panel) at 24 hours post-electroporation. Both siRNA and miRNA at 1 to 4 μg exhibited similar ability to silence the expression of *Rnluc* (~60% reduction). Interestingly, when the silencing assays were carried out at 48 hours post-electroporation, the effect was more pronounced for all amounts of siRNA and miRNA tested (Figure 2.3b, lower panel). Both siRNA and miRNA (at 0.25 μg) exhibited similar silencing ability (54 ± 2% reduction), which was higher than those at 24 hours. At higher amounts used, miRNA seems to show slightly better ability than siRNA to induce gene silencing. This silencing effect was not detected with an unrelated small RNA at either 24 or 48

hours (Figure 2.3, B, lightly shade bars), confirming the homologous nucleotide specificity.

An enhanced silencing effect found at a longer incubation was not observed in dsRNA-induced gene silencing (Figure S2.1.). It is possible that siRNA- or miRNA-engaged machinery could be more stable than, or different from, that of a dsRNA-system. Considering the nature of transient expression of the reporter plasmids, the more stable the silencing machinery is, the more robust the machinery is to compete against the expression of reporters. Another possibility could be due to the presence of an RNA-dependent RNA polymerase, to which *T. gondii* carries a homologue, which would allow for the amplification of short dsRNAs using the transcripts as templates. Thus, their silencing ability would remain longer and could be enhanced as long as the transcript is present.

In summary, our findings clearly indicated that small silencing RNAs (siRNA and miRNA) are able to specifically induce gene silencing in *T. gondii*. This study is the first to demonstrate the activity of miRNA in gene silencing in the Apicomplexan protozoan parasites. More importantly, we showed that the silencing effect was within the range (~80% reduction siRNA and miRNA tested (Figure 2.3b, lower panel). Both siRNA and miRNA (at 0.25 μg) exhibited similar silencing ability (54 ± 2% reduction), which was higher than those at 24 hours. At higher amounts used, miRNA seems to show slightly better ability than siRNA to induce gene silencing. This silencing effect was not detected with an unrelated small RNA at either 24 or 48 hours (Figure 2.3b, lightly shade bars), confirming the homologous nucleotide sequence was required.

2.3.4. Which silencing RNA (long dsRNA, siRNA and miRNA) is the best to induce homologous gene silencing?

Using the dual luciferase reporter system, we demonstrated that silencing RNAs, including long dsRNA, siRNA (perfect-match bases) and miRNA (central-mismatch bases), were capable of inducing gene-specific silencing. Long dsRNA exhibited the least silencing effect (~40% reduction of gene expression). The mild silencing effect was maintained for 48 hours, which could be due to (i) additional processing steps required to process the long dsRNA into small silencing RNAs, or (ii) a different silencing mechanism being engaged. As stated earlier, long dsRNA-induced gene silencing could be a useful tool when a quick assessment is required for evaluating a putatively essential gene. In a shorter tested period (24 hours), siRNA and miRNA were equally effective where the silencing effect was ~50%. When a prolonged and high silencing effect is required, miRNA offers the best tool among these tested silencing RNAs. Tested miRNA affected the Rnluc expression the most (80% reduction), which is equivalent to those tested in different mammalian cell cultures (Choe et al., 2010 and Elbashir et al., 2001). Therefore, we concluded that miRNA, with central-mismatch bases, is the best silencing RNA to use when a homologous gene silencing is required for the study of a gene of interest.

2.4. Acknowledgements

This work was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC, Canada)

2.5. References

Adams, B.; Musiyenko, A.; Kumar, R.; Barik, S. A novel class of dual-family immunophilins. Journal of Biological Chemistry. 2005, 280(26), 24308-24314.

Al-Anouti, F.; Ananvoranich, S. Comparative analysis of antisense RNA, double-stranded RNA, and delta ribozyme-mediated gene regulation in *Toxoplasma gondii*. Antisense and Nucleic Acid Drug Development. 2002, 12(4), 275-281.

Al-Anouti, F.; Quach, T.; Ananvoranich, S. Double-stranded RNA can mediate the suppression of uracil phosphoribosyltransferase expression in *Toxoplasma gondii*. Biochemical and Biophysical Research Communications. 2003, 302(2), 316-323.

Al-Anouti, F.; Tomavo, S.; Parmley, S.; Ananvoranich, S. The expression of lactate dehydrogenase is important for the cell cycle of *Toxoplasma gondii*. Journal of Biological Chemistry. 2004, 279(50), 52300-52311.

Al Riyahi, A.; Al-Anouti, F.; Al-Rayes, M.; Ananvoranich, S. Single argonaute protein from *Toxoplasma gondii* is involved in the double-stranded RNA induced gene silencing. International Journal for Parasitology, 2006, 36(9), 1003-1014.

Ananvoranich, S.; Al Rayes, M.; Al Riyahi, A.; Wang, X. RNA silencing of glycolysis pathway in *Toxoplasma gondii*. Journal of Eukaryotic Microbiology, 2006, 53(1), S162-163

Braun, L.; Cannella, D.; Ortet, P.; Barakat, M.; Sautel, C.F.; Kieffer, S.; Garin, J.; Bastien, O.; Voinnet, O.; Hakimi, M.A. A complex small RNA repertoire is generated by a plant/fungal-like machinery and effected by a metazoan-like Argonaute in the single-cell human parasite *Toxoplasma gondii*. PloS Pathogens, 2010, 6(5), e1000920.

Behnke, M.S.; Radke, J.B.; Smith, A.T.; Sullivan, W.J.; White, M.W. The transcription of bradyzoite genes in *Toxoplasma gondii* is controlled by autonomous promoter elements. Mol. Microbiol. 2008, 68(6), 1502-18.

Brodersen, P.; Voinnet, O. Revisiting the principles of microRNA target recognition and mode of action, Nature Reviews Molecular Cell Biology, 2009, 10, 141–148.

Carthew, R.W.; Sontheimer,E.J. Origins and mechanisms of miRNAs and siRNAs, Cell, 2009, 136, 642–655.

Choe, J.; Cho, H.; Lee, H.C.; Kim, Y.K. MicroRNA/Argonaute 2 regulates nonsense -mediated messenger RNA decay. *EMBO reports*, 2010, 11, 380-386.

Dykxhoorn, D.M.; Lieberman, J. The silent revolution: RNA interference as basic biology, research tool, and therapeutic, Annual Review of Medicine. 2005, 56, 401–423.

Elbashir, S.M.; Harborth, J.; Lendeckel, J.W.; Yalcin, A.; Weber, K.; Tuschl, T. Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cell. Nature, 2001, 411, 494-498.

El-Sayed, N. M.; Myler, P. J.; Bartholomeu, D. C.; Nilsson, D.; Aggarwal, G.; Tran, A. N.; et al. The genome sequence of *Trypanosoma cruzi*, etiologic agent of Chagas disease. Science, 2005, 309, 409-415.

Fire, A.; Xu, S.; Montgomery, M. K.; Kostas, S. A.; Driver, S. E.; Mello, C. C. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. Nature, 1998, 391, 806-811.

Grimm, D. Small silencing RNAs: state-of-the-art. Advanced Drug Delivery Review, 2009, 61(9), 672-703.

Holmes, M.; Liwak, U.; Pricop, I.; Wang, X.; Tomavo, S.; Ananvoranich, S. Silencing of tachyzoite enolase 2 alters nuclear targeting of bradyzoite enolase 1 in *Toxoplasma gondii* Microbes and Infection. 2010, 12(1), 19-27.

Ivens, A. C.; Peacock, C. S.; Worthey, E. A.; Murphy, L.; Aggarwal, G.; Berriman, M.; et al. The genome of the kinetoplastid parasite, *Leishmania major*. Science, 2005, 309, 436-442.

Jones, K.H.; Senft, J.A. An improved method to determine cell viability by simultaneous staining with fluorescein diacetate-propidium iodide. Journal of Histochemistry and Cytochemistry. 1985, 33(1), 77-79.

Kim, D.; Rossi, J. RNAi mechanisms and applications. Biotechniques, 2008, 44(5), 613-616.

Kim, D.H.; Behlke, M.A.; Rose, S.D.; Chang, M.S.; Choi, S.; Rossi, J.J. Synthetic dsRNA Dicer substrates enhance RNAi potency and efficacy. Nature Biotechnology, 2005, 23, 222–226.

Kolev, N.G.; Tschudi, C.; Ullu, E. RNA Interference in Protozoan Parasites: Achievements and Challenges. Eukaryotic Cell, 2011.

McRobert, L.; McConkey, G. A. RNA interference (RNAi) inhibits growth of *Plasmodium falciparum*. Molecular and Biochemical Parasitology, 2002, 119, 273-278.

Meister, G.; Tuschl, T. Mechanisms of gene silencing by double-stranded RNA, Nature, 2004, 431, 343–349.

Montoya, J.G.; Liesenfeld, O. Toxoplasmosis. Lancet, 2004, 363(9425), 1965-1976.

Ngo, H.; Tschudi,C.; Gull, K.; Ullu, E. Double-stranded RNA induces mRNA degradation in *Trypanosoma brucei*. Proceeding of the National Academy of Sciences of USA, 1998, 95, 14687-14692.

Pillai, R.S.; Bhattacharyya, S.N.; Artus, C.G.; Zoller, T.; Cougot, N.; Basyuk, E.; Bertrand, E.; Filipowicz, W. Inhibition of translational initiation by Let-7 MicroRNA in human cells. Science. 2005, 309(5740), 1573-1576.

Roos, D.S.; Donald, R.G.; Morrissette, N.S.; Moulton, A.L. Molecular tools for genetic dissection of the protozoan parasite *Toxoplasma gondii.* Methods in Cell Biology. 1994, 45, 27-63.

Schramke, V.; Allshire, R. Those interfering little RNAs! Silencing and eliminating chromatin. Current Opinion in Genetics & Development, 2004, 14(2), 174-180.

Ullu, E.; Tschudi, C.; Chakraborty, T. RNA interference in protozoan parasites. Cellular Microbiology, 2004, 6(6), 509-519.

Yu, L.; Gao, Y.F.; Qiao, Z.P.; Li, C.L.; Li, X.; Shen, J.L. *Toxoplasma gondii:* siRNA can mediate the suppression of adenosine kinase expression. Experimental Parasitology. 2008, 118(1), 96-102.

Yu, L.; Gao, Y.F.; Li, X.; Qiao, Z.P.; Shen, J.L. Double-stranded RNA specific to adenosine kinase and hypoxanthine-xanthine-guanine-phosphoribosyltransferase retards growth of *Toxoplasma gondii*. Parasitology Research, 2009, 104(2), 377-383.

Zuker, M. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Research. 2003, 31(13), 3406-3415.

CHAPTER 3

Utilization of inherent miRNAs in functional analysis of *Toxoplasma gondii* genes

3.1. Introduction

Toxoplasma gondii, a member of the phylum Apicomplexa, is one of the most successful intracellular parasitic pathogens. It was reported that up to one third of the world's population has been infected, and could develop a disease called toxoplasmosis (Montoya and Liesenfeld, 2004). In general *T. gondii* infection does not cause clinical illness because the parasite in its rapidly growing tachyzoite form could easily be cleared by the host's immune system during the initial course of infection. If not cleared, residual tachyzoites could differentiate into slowly growing bradyzoites, which could be very resistant to an immune response and persist throughout the lifetime of the infected human host. When the dormant bradyzoites reconvert to tachyzoites, they could cause a relapse leading to fatality in individuals with weaken immunity (Cenci-Goga et al., 2011).

Throughout the life cycle of *T. gondii*, a tight regulation of gene expression is required to ensure and sustain its viability (i) within its human and other non-definitive hosts during its asexual cycle including tachyzoite –bradyzoite interconversion, and (ii) within the gut of feline definitive host during its sexual reproduction to yield zygote-containing cysts or oocysts (Montoya and Liesenfeld, 2004). To date, various molecular and genetic tools have been generated and employed to modulate the expression of the gene of interest at various stages of parasite development and aimed at gaining a better understanding of its function(s) in the host-parasite interaction and pathogenesis (Roos et al., 1994 and Boothroyd et al., 1997). An effective molecular manipulation of the parasites could easily be achieved following a transfection via electroporation of the

freshly released tachyzoites (Soldati and Boothroyd, 1993, Roos et al., 1994 and Boothroyd et al., 1997). In 2007, Meissner and co-workers reviewed and cataloged molecular tools available for *T. gondii* to include reliable non-homologous and homologous recombination, as well as an elusive RNAi-targeting technology. *T. gondii* tool box has continuingly been expanded to include a tetracycline-inducible gene expression system (Meissner et al., 2001, van Poppel et al., 2006 and Jammallo et al., 2011), a site-specific cre/loxP-based recombination system (Heaslip et al., 2010), and a ligand-induced protein stability system (Herm-Götz et al., 2007). Moreover, homologous recombination events could be enhanced with the use of a permissive parasite strain (Huynh and Carruthers, 2009) and well-designed plasmid systems (Upadhya et al., 2011). The Miessner group later reviewed and cataloged advantages and disadvantages of currently available controllable gene expression systems acting at either transcriptional and/or post-transcriptional levels (Jiménez-Ruiz et al., 2014). In June 2014, two reports showed that CRISPR (clustered regularly interspaced short palindromic repeats) technology can successfully be applied to engineer the genome of *T. gondii* and to allow for deletion, replacement, site-directed mutagenesis and epitope tagging (Shen et al., 2014 and Sidik et al., 2014). Collectively these methodologies have enabled a selective and specific down-regulation of a gene of interest and proven to be very valuable.

Although many sophisticated and varieties of genetic tools have become available, RNA-based silencing techniques have not yet been explored and improved to provide consistent outcomes. One contributing factor could stem from the lack of understanding of how endogenous RNA effectors (miRNAs and/or siRNAs) and their engaged machineries operate at molecular levels in *T. gondii*. The presence of hallmark

RNAi genes have been repeatedly confirmed in *T. gondii* to include Argonaute (*Tg*Ago), Dicer, and an RNA-dependent RNA polymerase (Ullu et al., 2004, Al Riyahi et al., 2006, Braun et al., 2010 and Kolev et al., 2011). *Tg*-miRNAs have been detected via deep-sequencing methodologies (Braun et al., 2010, Wang et al., 2012 and Xu et al., 2013). It was also shown that *T. gondii* harbors a chimeric RNAi mechanism containing plant- and fungal-like effector RNAs, and metazoan-like RNAi genes (Braun et al., 2010). Musiyenko et al. (2012) further demonstrated that *Tg*Ago displays a weak target RNA cleavage activity linked to its lack of canonical DDE/H catalytic triad. A greater target RNA cleavage activity was observed upon the subsequent association of a Tudor staphylococcal nuclease (TSN) and a protein Arg methyl transferase, PRMT1. The study strongly suggested that the RNA-induced silencing complex of *T. gondii* utilizes the slicer activity of TSN as for partially mismatched miRNA target pairs, and that of Ago for a perfectly matched miRNA target pairs. With its unique silencing machinery and/or mechanism, the RNA-directed gene silencing can effectively be achieved in *T. gondii*. Taken together these findings, a consistent RNA silencing methodology for *T. gondii* should be within reach.

The study here was aimed to gain a better understanding of how *T. gondii* utilizes its miRNAs in the RNA silencing mechanism to give rise to eventual outcomes of down regulation of gene expression. To do so, we used a dual luciferase reporter system to evaluate (i) the extent of their silencing effect with the two most abundant *Tg*-miRNAs, namely *Tg*-miR-60a and *Tg*-miR-4a, (ii) the nature of their interactions with target transcripts, and (iii) the effect of *Tg*-miR-60a miRNA mimics and inhibitors on their silencing effect. Furthermore, we explored the utility and versatility of *Tg*-miR-60a as a

genetic tool in modulating the expression of a gene of interest. As a proof of concept, the highly expressed *Tg*-miR-60 family was directed to control the expression of (i) DEAD-box RNA helicase (*Tg*HoDI), and (ii) a metabolically important lactate dehydrogenase isoform 1 (*Tg*LDH1). We showed that when the binding sites of *Tg*-miR-60a were introduced into the parasite's transcripts via homologous recombination at either locus, the expression levels of the chosen genes can significantly be altered. It was thus proven that *Tg*-miRNAs could be directed and controlled to assist in loss-of-function analyses.

3.2. Materials and Methods

3.2.1. Mammalian cell and parasite cultures

Human foreskin fibroblasts (HFF, ATCC-1041) were grown using Dulbecco's Modified Eagle Medium (DMEM) with high D-glucose and L-glutamine (Invitrogen, #12100046) supplemented with 10% v/v cosmic calf serum (ThermoFisher Scientific, Hyclone, # H3008704N) and diluted (0.5x) Antibiotic-Antimycotic (Invitrogen, # 15240-062) and maintained at 37 °C in 5% CO_2. RHΔHX, *T. gondii* type 1 strain, was obtained from NIH AIDS Research and Reference Reagent Program (#2857); ΔKu80 strain which is a RHΔHX-derived strain, was obtained from Dr. Vern B. Carruthers (University of Michigan). Unless stated otherwise, both strains were maintained in HFF using Minimum Essential Medium (MEM, Invitrogen, #61100061) supplemented with 1% v/v dialyzed fetal bovine serum (ThermoFisher Scientific, Hyclone, # SH3007903) and diluted (0.5x) Antibiotic-Antimycotic.

3.2.2. Plasmid construction

The nucleotide sequences of primers used in the study are shown in Table 3.1. The dual luciferase plasmid system, which was previously described in Crater et al.

(2011), was adopted for the study. The system consists of two plasmids referred to as pTubFf and pTubRnLet7a3x. The former was unmodified and served as an internal control for transfection efficiency and expression assays. The latter was used as the parental plasmid of all construction in the study to create multiple *Rnluc* expression plasmids carrying one or three binding sites for any *Tg*-miRNAs. To replace the three consecutive *Hs*-let-7a binding sites in the parental plasmid, Phusion High-Fidelity DNA polymerase reaction (New England Biolabs, #F-530) was used according to the manufacturer's protocol with specific primer sets to reconstitute the desired binding site(s). In addition, each primer of the set was designed to carry a partial sequence of one chosen restriction endonuclease to allow for future restriction analyses of resultant plasmids when they were reconstituted following construction. Once PCR reactions were completed, the amplicons were purified, treated with T4 polynucleotide kinase (Fermentas, #EK0031) and re-circularized using T4 DNA ligase (Fermentas, #EK0016). pTubRn construct, which lacks the binding site for *Tg*-miRNAs, was generated using FwRn and RvRn primers. For the construction of pTubRnMIR-60a3x plasmid to express *Rnluc* transcripts carrying three binding sites of *Tg*-miR-60a, the primers named miR-60aFw and miR-60aRv were used, and the newly generated EcoRI site was introduced for its endonuclease analysis. Similarly, three primer pairs, including miR-4aAFw/miR-4aARv, miR-4aBFw/miR-4aBFw, and miR-4aCFw/miR-4aCRv, were used in the construction of plasmids to express *Rnluc* transcripts carrying three binding sites of *Tg*-miR-4a with different complementary interactions, respectively. When the plasmids were engineered to carry only one binding site, FwRn or RvRn primer was paired with another primer with sequence encoding one or two binding site for *Tg*-miR-60a or *Tg*-miR-4a.

Table 3.1. Oligonucleotide primers used in the study. The nucleotide sequence coding for Tg-miRNA binding sites are in bold. Nucleotides corresponding to restriction endonuclease recognition sites are in italics.

Names	Sequences
Construction of expression plasmids	
FwRn	GCGTTAAGTAAACAGAAGCTGCC
RvRn	TCTAGATTACTGCTCGTTCTTCAG
miR-60aFW (*EcoRI*)	*TT*CAGTATGGATTTCGTACCGACTGTGTGGCCTAGTATGGATTTCGTACCGACTGTGTGCGTTAAGTAAACAGAAGCTGCCC
miR-60aRV (*EcoRI*)	*TT*CACACAGTCGGTACGAAATCCATACTTCTAGATTACTGCTCGTTCTTCAG
miR-4aAFw (*EcoRI*)	*TT*CAATGACTACAGCTTCCAAGCAAACATGGCCTAATGACTACAGCTTCCAAGCAAACATGCGTTAAGTAAACAGAAGCTGCCC
miR-4aARv (*EcoRI*)	*TT*CATGTTTGCTTGGAAGCTGTAGTCATTTCTAGATTACTGCTCGTTCTTCAG
miR-4aBFw (*HindIII*)	*CTT*AGTGACCCCAGCTTCCAAGCAAACATGCGTTAAGTAAACAGAAGCTGCCC
miR-4aBFw (*HindIII*)	*CTT*ATGTTTGCTTGGAAGCTGGGGTCACTTCCGAGATGTTTGCTTGGAAGCTGGGGTCACTTCTAGATTACTGCTCGTTCTTCAG
miR-4aCFw (*XhoI*)	*GA*GAATGAATACAGCTTCCTAGCGAACATGCGTTAAGTAAACAGAAGCTGCCC
miR-4aCRv (*XhoI*)	*GA*GATGTTCGCTAGGAAGCTGTATTCATTCCGAGATGTTCGCTAGGAAGCTGTATTCATTTCTAGATTACTGCTCGTTCTTCAG
In vitro transcription reactions	
miR-60a_upper	ACACAGTCGGTACGAAATCCATACTCC*TATAGTGAGTCGTATTA*
miR-60a_lower	AGTATGGATTTCGTACCGACTGTGTCC*TATAGTGAGTCGTATTA*
hs-let7a_upper	AATGAGGTAGTTCAATAGGCTGTGCC*TATAGTGAGTCGTATTA*
hs-let7a_lower	AACACAGCCTATTGAACTACCTCACC*TATAGTGAGTCGTATTA*
T7promoterGG	TAATACGACTCACTATAGG
RT-qPCR	
qFF_Fw	CGCCTGAAGTCTCTGATTAAGTACAA
qFF_Rv	TCCGTGCTCCAAAACAACAA
qRn_Fw	CGCAACTACAACGCCTACCTT
qRn_Rv	GCGTCCTCCTGGCTGAAGT
qHoDI_Fw	CAAAGGACAACCGCGTGAAA
qHoDI_Rv	TCGAAGATGCCCATCAACAATT
qGAPDH_Fw	GGTGTTCCGTGCTGCGAT
qGAPDH_Rv	GCCTTTCCGCCGACAAT
Generation of transgenic parasites	
miR-60aSFFw (*EcoRI*)	*TT*CAGTATGGATTTCGTACCGACTGTGTGGCCTAGTATGGATTTCGTACCGACTGTGTGCGCCCAGCCCACAGAAGCTGCCCGT
miR-60aSFRv (*EcoRI*)	*TT*CACACAGTCGGTACGAAATCCATACTGCCTTATTTATCGTCGTCGTCTTTGT
LDH1_Fw	TACTTCCAATCCAATTTAATGCAGGTCGTACCGTACCCATTTG
LDH1_Rv	TCCTCCACTTCCAATTTTAGCCTGAAGAGCAGCAACCGCCTTG
HoDI_Fw	TACTTCCAATCCAATTTAATGCTACGACGTGAATTGTTGATGG
HoDI-RV	TCCTCCACTTCCAATTTTAGCGTAAATTGCTTCGTCAATCTGG

For example, FwRn was paired with miR-60aRv to generate a plasmid encoding for one binding site of *Tg*-miR-60a. Following successful confirmation by restriction analyses, a large scale plasmid preparation of reporter plasmids were prepared for a transfection by electroporation, which the reporter plasmids (1 μg of Ffluc and 1 μg of Rnluc) were introduced into approximately 2 x 10^6 freshly lysed parasites via electroporation using a BTX ECM 630 (1500 volts, 25 Ω, and 25 μF) in a mixture (400 μL) containing 120 mM KCl, 0.15 mM $CaCl_2$, 10 mM K_2HPO_4/KH_2PO_4. pH 7.6, 2 mM EDTA, 5 mM $MgCl_2$, 2 mM ATP, 5 mM glutathione in a 4 mm-gap cuvette.

3.2.3. *In vitro* transcription of miRNA mimics and inhibitors

Using partial DNA duplexes as templates, *in vitro* transcription reactions were carried out in 50 μL-reaction mixtures in the presence of T7 RNA polymerase (~ 5 units), 80 mM Hepes-KOH, pH 7.5, 24 mM $NaCl_2$, 2 mM spermidine, 40 mM DTT, 10 mM rNTPs, and ~ 1 unit pyrophosphatase at 37 °C for 16 hours. Reactions were subsequently extracted with one volume of phenol-chloroform mixture (1:1), and resultant RNAs were precipitated and quantified (Thermo Scientific NanoDraop2000). As shown in Table 3.1, oligonucleotide templates were designed to carry two parts of complementary sequence (i) to encode for desirable resultant transcripts of the upper or lower strand of *Tg*-miR-60a or *Hs*-let-7a, and (ii) to T7 RNA polymerase promoter (underlined, Table 3.1.). The individual oligonucleotide template (2 μM) was annealed with an equimolar of T7promoterGG following mixing and heating the mixture for 5 minutes at 70 °C and slowly cooled to room temperature to yield partial DNA duplexes. One fifth of the annealing reaction (or 400 nM of the partial DNA duplex) was used as templates for the RNA synthesis. Resultant single-stranded RNA products, which were generated using the

oligonucleotide template named miR-60a_upper or hs-let7a_upper, were referred to as *Tg*-miR-60a inhibitor and *Hs*-let-7a inhibitor, respectively. When an equal amount of *Tg*-miR-60a inhibitor was mixed and annealed with the single-stranded RNA products obtained from the reaction using miR-60a_lower oligonucleotide template, the double-stranded RNA products was referred to as *Tg*-miR-60a mimic. *Hs*-let-7a mimic was similarly generated from RNA products obtained from hs-let7a_upper and hs-let7a_lower oligonucleotide templates. Quality of RNA products were verified by gel electrophoresis to confirm their expected sizes prior to electroporation into freshly released parasites.

3.2.4. RNA isolation and RT-qPCR

Forty-eight hours following electroporation, total RNA was isolated from freshly lysed parasites (approximately 3×10^6 parasites) using TRI reagent (Molecular Research Center, Inc., #TR 118) according to the manufacturer's protocol, and treated with RQ1 Rnase-Free Dnase (Promega, #M6101) for 60 minutes at 37 °C. RNA samples were electrophoresed on a 1% agarose gel to verify for the presence of distinct rRNAs and mRNA smear bands, and the absence of DNA templates. Samples were kept for future use at −20 °C. Isolated total RNA (~500 ng) was annealed with 250 ng of an oligo-(dT) following heating at 65 °C for 5 minutes and quickly cooled on ice. The reactions were then divided in half. A half of each annealed reaction was used in a SuperScript III reverse transcriptase (RT) reaction (Invitrogen #18080-093), while the other aliquot was used in a control reaction in the absence of SuperScript III RT. The reaction mixtures were diluted 1/10, and diluted samples were subjected to real-time qPCR using Kapa Sybr Fast qPCR Kit Master Mix (2x) ABI Prism (Kapa Biosystems #KR0390) and OneStepPlus thermocycler (Applied Biosystems). All test samples were analyzed to give

technical triplicates. The expression level derived from the dual reporter system was measured and calculated, so that the Rnluc expression from tested expression plasmid was normalized by Ffluc expression using ΔΔCt protocol carried out by the StepOne Software v 2.1 (Applied Biosystem). To generate a data set which allowed for comparison across all technical and biological triplicates, the relative expression obtained from the dual reporter system where Rnluc transcripts lack miRNA binding sites was measured for each independent experiment set, and used for normalizing the effect of *Tg*-miRNAs on the dual reporter expression.

3.2.5. Western blot analyses

Cell-free lysates were prepared by addition of an appropriate amount of RIPA buffer to collected parasites, followed by sonication for 5 seconds. Protein concentration was measured using the RC DC Protein Assay according to manufacturer's instructions (Bio-Rad Laboratories, #500- 0119). Approximately 10 μg was used for the analysis, of which the lysates were denatured using 6x SDS protein loading dye and resolved on a 10% SDS-PAGE. The primary antibodies used were α-Flag mouse antisera (1:5,000, Applied Biological Materials Inc., Richmond, B.C, Canada, #G191), α-tub mouse antisera (12G10, 1: 1000, Developmental Studies Hybridoma Bank, University of Iowa, USA), and α-TgLDH1 rabbit antisera (1:2,000, Yang and Parmley, 1997). The signals were revealed using secondary antibodies: horseradish peroxidase (HRP) conjugated goat anti-mouse antisera (Rockland Immunochemicals Inc., #610-1319) or HRP conjugated goat anti-rabbit antisera (Rockland Immunochemicals Inc., #611-1322) at a 1:10,000 dilution, and a Chemiluminescent HRP Substrate Kit (Millipore, #WBKLS0500).

3.2.6. Immunocytochemistry

HFF monolayers were cultured on coverslips until confluent, infected with ~1 x 10^6 freshly lysed parasites and incubated at 37 °C and 5% CO_2. Infected monolayers on the cover slips were rinsed three times with 1x phosphate buffer saline (PBS), fixed with 3% paraformaldehyde for 5 minutes and permeabilized with 0.25% Triton X-100 in PBS for 10 minutes. Coverslips were then washed in 1x PBS and subsequently incubated in a blocking solution (5% bovine serum albumin, BSA in PBS) with gentle agitation for 1 hour at room temperature. Coverslips were then incubated for 1 hour in a humidity chamber with a specific primary antibody diluted in 2% BSA. The primary antibodies used were α-FLAG mouse antisera at 1:500 dilutions and α-TgLDH1 rabbit antisera at 1: 5,000 dilutions. Coverslips were washed in 1× PBS (5 minutes, twice) and incubated for 1 hour with secondary antibody diluted in 2% BSA. Corresponding secondary antibodies were used at 1:1000 dilutions. Cells were then incubated with Hoechst staining solution (10 μM) at room temperature for 5 minutes, washed twice with 1x PBS and mounted on glass slides. For the study carried out using extracellular parasites, the parasites were collected and filtered through a 3-μm filtering set. Approximately 1×10^6 parasites were transferred onto coverslips and further incubated for an hour to allow for adhering onto the surface of coverslips. The coverslips were rinsed in 1x PBS, fixed with 3% paraformaldehyde for 5 minutes and permeabilized with 0.25% Triton X-100 for 10 minutes. Subsequent blocking and antibody interaction steps were as same as those of intracellular parasites. Image acquisition was carried out using a Leica DMI 6000B inverted fluorescent microscope using HCX PL Apo 100×/1.40–0.70 objective and a

Leica DFC 360FX camera with Leica Advanced Fluorescence Application Software V.2.3.0. Post-acquisition image processing was performed with Adobe Photoshop 7.0.

3.2.7. Utilization of *Tg*-miR-60a family for loss-of-function analysis

3.2.7.1. Rationale

Despite their possible redundancy, most genes are tightly regulated to allow for their necessary functions to display in a spatial- and temporal-specific manner. In order to accurately assess the functionality of a gene, the pattern of spatial and temporal requirements should be taken into consideration. Particularly when the specific gene silencing was implemented, a natural expression profile of the gene product should be used as a point of reference.

3.2.7.2. Experimental design

To standardize the silencing effects with a gene's natural expression profile, the binding site for the *Tg*-miR-60a family was introduced into the genomic locus at the position immediate after the protein coding sequence of the gene of interest via homologous recombination (Figure 3.5a). Transforming plasmids were constructed using pSF-TAP-LIC-HX plasmid (obtained from Dr. Vern B. Carruthers, University of Michigan). First the parental plasmid was modified by site directed mutagenesis to integrate three binding sites of Tg-miR-60a using miR-60aFWSF and miR-60aRVSF primers (Table 3.1), and the resultant plasmid was named pSF-TAP MIR603x_LIC-HX. Both pSF-TAP-LICHX and pSF-TAP MIR603x_LIC-HX were linearized by PacI digestion prior to the treatment with T4 DNA polymerase to create compatible ends in the presence of dGTP for a ligation-independent cloning procedure as described by Huynh and Carruthers (2009). Amplicons corresponding to either of TgLDH1 (1073 bps) or

TgHoDI (2767 bps) were prepared using gene specific primer pair of LDH1_Fw/LDH1_Rv and HoDI_Fw/HoDI-RV respectively (Table 3.1), and purified prior to the treatment with T4 DNA polymerase to create compatible ends in the presence of dCTP. Equimolar concentrations of the prepared plasmid and insert were allowed to anneal and subsequently used in bacterial transformations. The resultant plasmids were analyzed to confirm their sequences via restriction endonuclease digestions. The plasmids were then used in the transformation of *T. gondii*. Transgenic parasites were selected and grown in culturing media containing mycophenolic acid (25 μg/mL) and xanthine (50 μg/mL).

3.2.8. Loss-of-functional analysis

Functional analyses were carried out using parasites generated from test or control plasmids which were cultured and treated under similar conditions. Assessment of growth profile was performed after the parasites were allowed to infect HFF monolayers for 24 hours. The numbers of parasites per vacuole were counted and calculated to obtain population distribution. Analyses were performed in triplicate (n = 3), and at least 100 vacuoles were counted for each test.

3.3. Results and discussion

3.3.1. Evaluating the silencing effect of *Toxoplasma gondii* miRNAs

A large collection of *Tg*-miRNAs were identified via a deep sequencing methodology, from which 14 families were grouped according to their seed sequences (6 nucleotides) and their expression levels (Braun et al., 2010). Two of the highest expressed *Tg*-miRNA families, named *Tg*-miR-60a and *Tg*-miR-4a, were selected for the study. The guide strands (26–27 nucleotides) and seed sequences of *Tg*-miR-60a and *Tg*-miR-4a are shown Table 3.2. A dual reporter system was designed and engineered so that *Renilla* luciferase (Rnluc) transcripts carry the complementary nucleotide sequence to the guide strand, referred to here as a *Tg*-miRNA binding site(s). Also as part of the reporter system, the transcripts of *Firefly* luciferase (Ffluc) were used as an internal control for transfection efficiency among various experiments (Figure 3.1a). Both transcripts have the same controlling elements of *Tg*-Tubulin-5′ UTR and *Tg*-DHFR-3′ UTR, and expressed under the control of a TgTubulin promoter to ensure equal expression and stability of transcripts (Crater et al., 2011). The engineered reporter system was utilized in the first part of the study, which was aimed for gauging the extent of the silencing effect of *Tg*-miR-60a and *Tg*-miR-4a toward reporter transcripts, in corresponding to (i) their expression levels and (ii) the nature of their complementarity to their target transcripts.

3.3.2. Extent of *Tg*-miRNAs to induce a silencing effect

Between the two selected *Tg*-miRNA families, *Tg*-miR-60a was found to be about 2.4 times more abundant than *Tg*-miR-4a (Braun et al., 2010). The Rnluc transcript was

Table 3.2. Chosen Tg-miRNAs and their predicted gene targets and interactions. The sequences of TgmiRNAs are in lower case, while their predicted targets are in capital letters. Underline indicates their seed sequence.

Tg-miRNA (name and sequence)	Predicted gene target and interaction (Gene ID and sequence)
Tg-miR-60a acacaqucgguacgaaauccauacu	5′ AGUAUGGAAUUCGUACCGACUGUGU 3′ 3′ ucauaccuaaagcauggcugacaca 5′
Tg-miR-4a auguuugcuuggaagcuguagucauu	TGGT1_203470 (A) 5′ AAUGACUACAGCUUCCAAGCAAACAU 3′ 3′ uuacugaugucgaagguucguuugua 5′
	TGGT1_246940 (B) CC 5′ AGUGAC CAGCUUCCAAGCAAACAU 3′ 3′ uuacug gucgaagguucguuugua 5′ au
	TGGT1_268390 (C) A U 5′ AAUGA UACAGCUUCC AGCAAACAU 3′ 3′ uuacu augucgaagg ucguuugua 5′ g u

designed for the assessment of the extent of *Tg*-miRNA silencing ability in respect to two factors: (i) the abundance of *Tg*-miRNA species, and (ii) the number of binding sites found on target transcripts. The resultant Rnluc transcripts carried either one or three binding sites for *Tg*-miR-60a or *Tg*-miR-4a. Forty-eight hours following transfection of the reporter system (1 μg of each construct) into freshly released *T. gondii* (10^6 parasites), the parasites were collected, and their cell-free extracts prepared and used for the measurement of relative expression of Rnluc and Ffluc luciferases. At least three independent experiments were performed for each construct, and the data were normalized using the reporter system without any binding sites for *Tg*-miRNAs, for each independent test and to allow for cross referencing among various sets of experiments. The presence of one binding site for individual *Tg*-miR-60a and *Tg*-miR-4a in the Rnluc transcripts caused the relative expression to reduce to 20 ± 2.7% and 60 ± 5.3%, reflecting 80% and 40% silencing effect, respectively (Figure 3.1b). The presence of three binding sites for individual *Tg*-miR-60a and *Tg*-miR-4a in the Rnluc transcripts decreased the relative expression to 17 ± 2.4% and 39 ± 2.3%, reflecting 83% and 61% silencing effect, respectively. Notably, the higher abundance of *Tg*-miR-60a conferred the higher silencing effect, as detected in the constructs carrying one binding site of *Tg*-miR- 60a. An increasing in the number of binding sites for *Tg*-miR-60a did not result in a drastic increase in the silencing effect, suggesting a saturation of the silencing machinery employing *Tg*-miR-60a as its effector. On the other hand, when the saturation of the silencing machinery was not reached by an excess amount of *Tg*-miRNA, such as those utilizing a lower abundant *Tg*-miR-4a, an increase in the number of binding sites, from one to three, could augment the silencing effect. When the Rnluc transcript carrying three

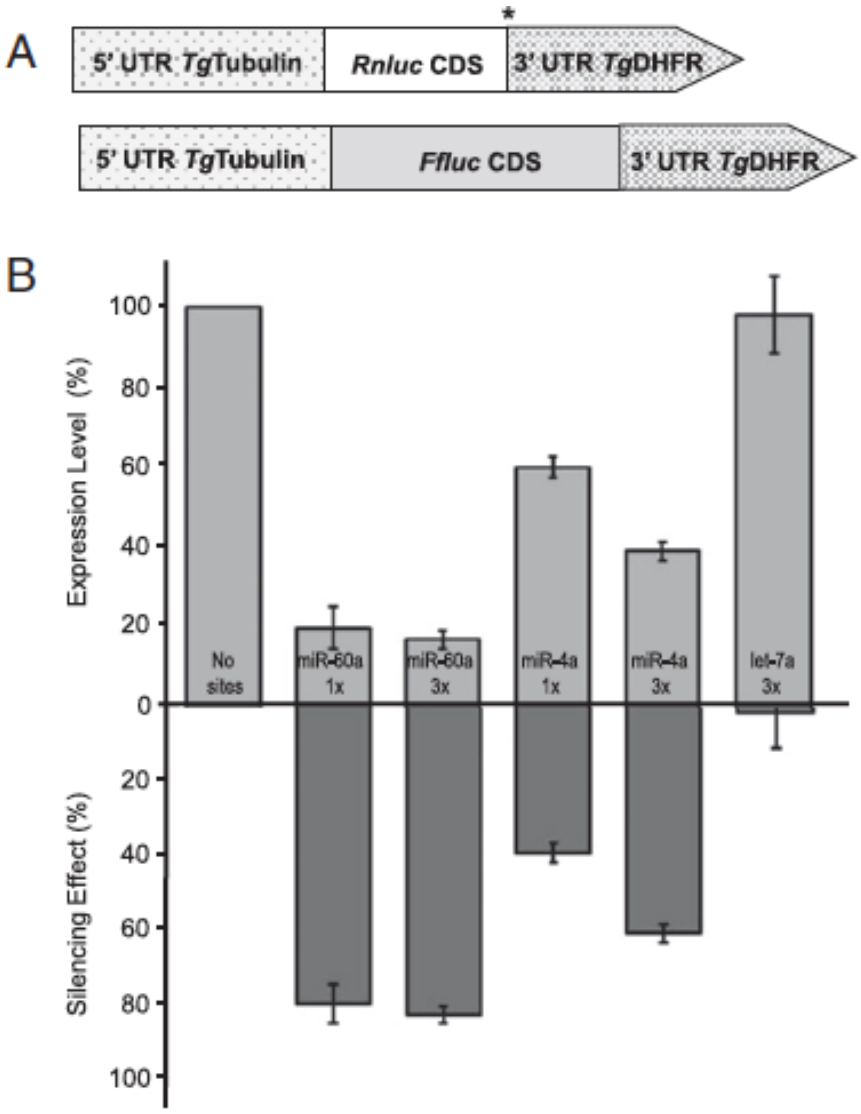

A
*
5' UTR TgTubulin
Rnluc CDS
3' UTR TgDHFR
5' UTR TgTubulin
Ffluc CDS
3' UTR TgDHFR
B
Expression Level (%)
Silencing Effect (%)
100
80
60
40
20
0
20
40
60
80
100
No sites
miR-60a 1x
miR-60a 3x
miR-4a 1x
miR-4a 3x
let-7a 3x

Figure 3.1. Gene silencing effect induced by endogenous Tg-miRNAs. A. Graphical representation of transcripts generated from the dual luciferase reporter system used in the study. Renilla luciferase (Rnluc) transcripts carry either one (1×) or three (3×) binding sites for Tg-miRNAs, whose location is indicated by asterisk (*). Firefly luciferase (Ffluc) transcripts were generated under the control of same promoter and controlling elements as those of Rnluc transcripts, and were used as an internal control for the transfection and expression analyses. B. The expression levels of the dual luciferase reporter system whose Rnluc transcripts carrying either one (1x) or three (3x) binding site(s) for *Tg*-miR60a, *Tg*-miR4a or *Hs*-let-7a were calculated to obtain relative expression levels when compared to those without a *Tg*-miRNA binding site (no sites). The extent of *Tg*-miRNAs to induce gene silencing, which was derived from the reduction of relative gene expression, was plotted to show relative silencing effect (%). Data was obtained from three independent trials, and plotted as means ± standard deviations.

binding sites for *Hs*-let-7a miRNA was used, and no changes in the relative expression level and silencing effect were observed. It strongly suggested that *T. gondii* does not express any miRNAs closely related to *Hs*-let-7a miRNA, and/or that the level of Hs-let-7a-related *Tg*-miRNA is much lower than those of *Tg*-miR-60 and *Tg*-miR-4 families. Moreover the findings indicated that the silencing effect occurs only (i) when the transcripts have the binding site and (ii) when the miRNA is present.

3.3.3. The effect of nucleotide sequence complementarity on *Tg*-miRNA-induced gene silencing

It has been demonstrated that miRNAs bind their target transcripts using perfect or imperfect complementary base-pair interactions (Pillai et al., 2005, Wakiyama et al., 2007). While no predicted targets for *Tg*-miR-60a was reported, three potential gene targets of *Tg*-miR- 4a were predicted to include three independent loci in the genome of *T. gondii* (Braun et al., 2010). These loci, namely (A) TGGT1_203470, (B) TGGT1_246940, and (C) TGGT1_268390, encode for three hypothetical proteins (Gajria et al., 2008). The transcripts from these loci were predicted to interact with *Tg*-miR-4a in three different complementarity formats; via a perfect match (A), one bulge with two mismatch bases (B), and two bulges with one mismatch base (C), respectively (Table 3.2). Three binding sites corresponding to these individual interactions were placed into the Rnluc transcripts. In addition, a similar design with a perfect match was made for *Tg*-miR-60a to allow for an equivalent comparison between the effect of *Tg*-miR-60a and *Tg*-miR-4a. As shown in Figure 3.2a, with a perfect complementarity, we observed a 20% higher silencing effect when Rnluc transcripts carry *Tg*-miR-60a than those of *Tg*-miR-4a (format A). The levels of silencing effect caused by *Tg*-miR-4a were not

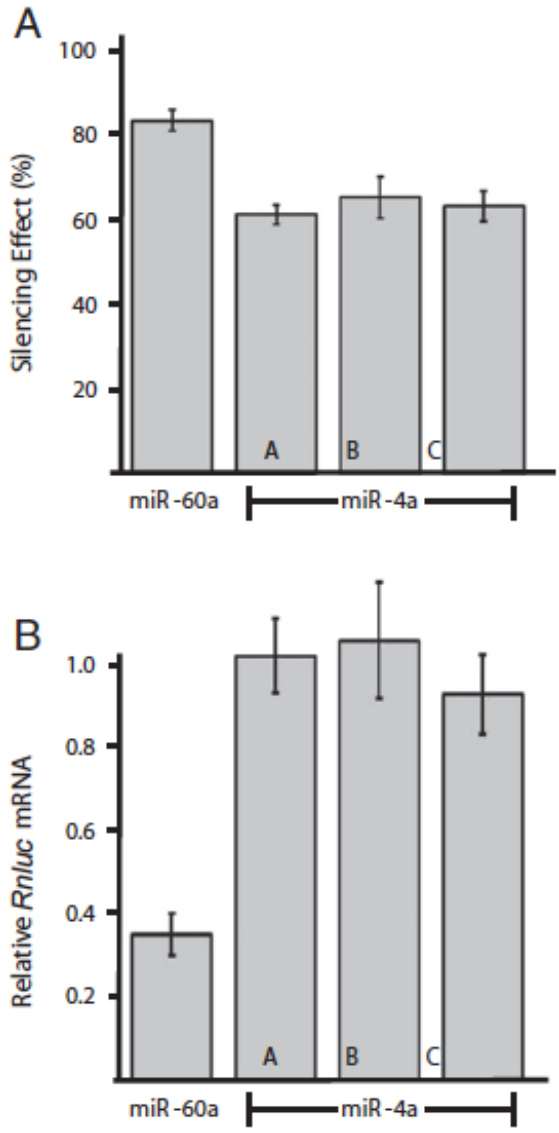

A
100
80
60
40
20
Silencing Effect (%)
A
B
C
miR-60a
miR-4a
B
1.0
0.8
0.6
0.4
0.2
Relative Rnluc mRNA
A
B
C
miR-60a
miR-4a

Figure 3.2. Effect of nucleotide complementarity. Rnluc transcripts of the dual luciferase reporter system were engineered to have three binding sites for *Tg*-miR60a with a perfect match, and for *Tg*-miR-4a mimicking the different base-pair interactions of its three predicted target genes. Their sequences and interactions are shown in Table 3.2. A. Relative silencing effect was calculated as compared to those without a binding site. B. Relative Rnluc mRNA levels were determined for each construct using the (ΔΔCt) method, where the obtained mRNA levels were normalized and compared to the construct containing no binding sites. Data was obtained from at least three independent experiments for each constructed tested, and plotted as means ± standard deviations.

significantly different with perfect or imperfect complementarity at ~60% silencing effect. Using RT-qPCR analysis to assess the levels of *Rnluc* and *Ffluc* mRNAs in the evaluation of the different formats of target pairing interactions, it was detected that perfect and imperfect matches of *Tg*-miR-4a target pairs gave no significant difference in the levels of *Rnluc* mRNA in comparison to those of *Ffluc* mRNA (Figure 3.2b). The data indicated that the silencing effect resulted from translational repression following *Tg*-miR-4a and target interactions. According to Musiyenko et al. (2012), the imperfect matched pair of miRNA and target would engage a stronger silencing activity of TSN than that of levels the perfect match of miRNA to engage Ago slicer activity, both leading to a directed mRNA degradation. Moreover, we detected a drastic difference in the relative Rnluc mRNA between the perfect matched pairing by *Tg*-miR-60a and *Tg*-miR-4a (Figure 3.2b). The reduction of *Rnluc* target mRNA by *Tg*-miR-60a strongly suggested that transcript degradation was a mode of *Tg*-miR-60a action, while the unchanged level of *Rnluc* target mRNA of *Tg*-miR-4a was indicative of translational suppression. The findings suggested that *T. gondii* RISC could be directed by *Tg*-miRNAs and resulted in mRNA degradation as well as translational suppression. The finding was not surprising because TgAgo, via pull-down and MS analysis, was shown to associate with various initiation and elongation factors of translational machinery (Braun et al., 2010). However, further investigation is needed to address how the *Tg*-miR-4a-induced silencing machinery could silence its target genes without discriminating among various forms of base-pair interactions. Most importantly the silencing effect was not a result of the degradation of the transcripts. Until we could identify authentic targets of

Tg-miR-60a, here it was assumed that the perfect match of *Tg*-miR-60a caused transcript degradation.

3.3.4. The effect of exogenous small RNAs on *Tg*-miRNA-induced gene silencing

Despite the significance of the underlying mechanism of *Tg*-miRNA-induced gene silencing, no indication from genome analysis could currently suggest how different species of *Tg*-miRNAs could engage different silencing machineries. Here we focused our effort in making use of the abundant *Tg*-miRNA species to direct functional analyses of *T. gondii* genes. Therefore it was of importance to examine whether the extent of *Tg*-miR-60a-induced gene silencing effect could be altered by exogenous small RNAs. We generated two species of small RNAs carrying a sequence pertinent to *Tg*-miR-60a, including (i) one with exact sequence and structure as *Tg*-miR-60a, referred to as miR-60a mimic, and (ii) the other with only the one strand of *Tg*-miR-60a, referred to as miR-60a inhibitor. Using the Rnluc construct carrying three *Tg*-miR-60a binding sites, we monitored the silencing effect following the transfection of various amounts of small RNAs along with the reporter constructs where the ratio of Rnluc to Ffluc (2 μg Rnluc, 1 μg Ffluc) was altered to ensure the detection of silencing alteration. In the absence of small RNAs, the reporter expression was reduced by the abundant *Tg*-miR-60a with the silencing effect detected at 65 ± 2.2% (Figure 3.3a), confirming (i) a higher amount of transcripts was generated, and (ii) a larger portion of the transcripts escaped the *Tg*-miR-60a-induced silencing machinery to produce final gene products than when 1 μg of Rnluc was used (Figure 3.2a, 83 ± 2.5%). Subsequent experiments were carried out with the increased amount of Rnluc (2 μg) construct to allow for the comparison of *Tg*-miR-60a mimic and inhibitor under similar testing conditions. Exogenous supplies of *Tg*-miR-60a

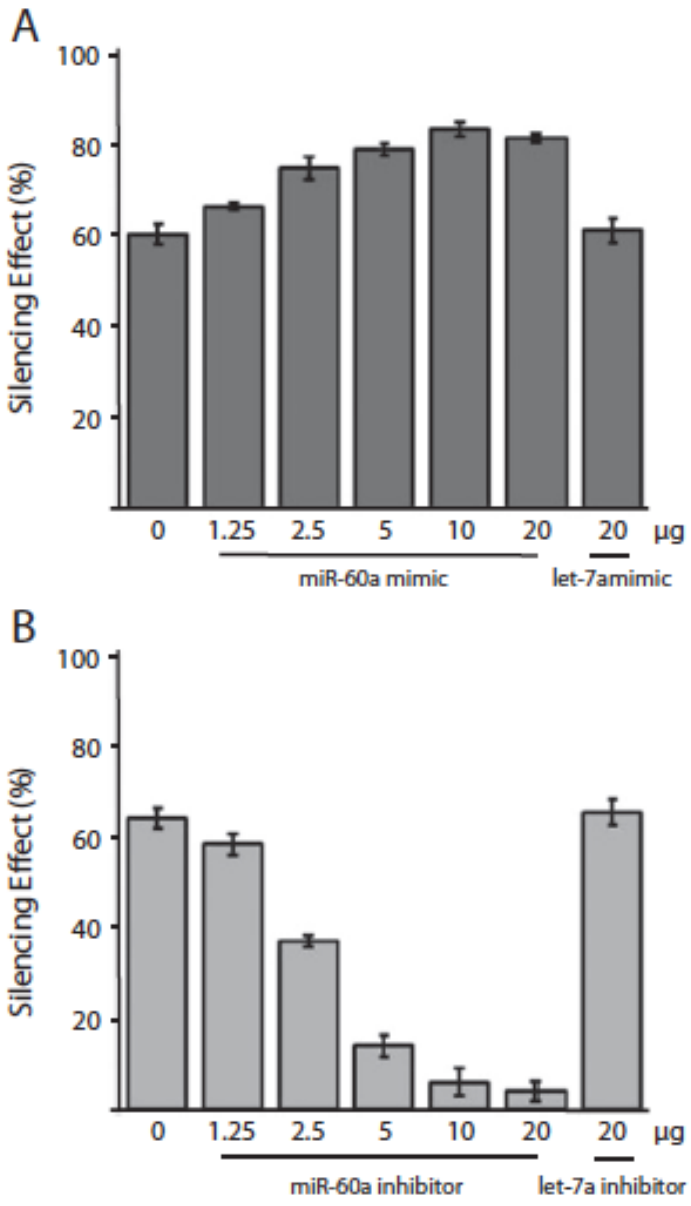


Figure 3.3. Effect of miRNA mimics and inhibitors. The engineered dual reporter system, whose resultant Rnluc transcripts carry three *Tg*-miR60a binding sites, was electroporated into the parasites along with various amounts of small RNAs (µg): A. Double-stranded *Tg*-miR60a mimic or B. Single-stranded *Tg*-miR60a inhibitor. *Hs*-let-7a mimic and inhibitor were used as control. Three independent experiments were performed for each test, and the data were plotted as means ± standard deviations.

mimic gave rise to an increase in the silencing effects from 65 ± 2.2% in the absence of small RNAs, to 83 ± 2.5% in the presence of 10 or 20 μg of Tg-miR-60a mimic (Figure 3.3a). The presence of 20 μg *Hs*-let-7a mimic did not cause any silencing effect. The results suggested that miRNA mimics could be engaged in the same silencing machinery as endogenous miRNAs. The saturation of *Tg*-miR-60a-induced silencing machinery gave rise to a maximum of ~80% silencing effect, under conditions used with the engineered reporter system. When an increase amount of *Tg*-miR-60a inhibitor was supplied, the silencing effect was diminished in a dose-dependent manner (Figure 3.3b). The presence of *Tg*-miR-60a inhibitor as little as 1 μg could perturb the silencing effect caused by *Tg*-miR-60a-induced silencing machinery. A maximal disturbance was observed with 10 - 20 μg of *Tg*-miR-60a inhibitor, where the silencing effect was abolished to ~6.2%. When miRNA inhibitors with an unrelated nucleotide sequence, namely *Hs*-let-7a-inhibitor (20 μg) was supplied, the relative expression and silencing effect was unaffected. These findings suggested that small RNAs with nucleotide sequences identical to one or both strands of miRNA could be exploited as genetic tools for the moderation of miRNA-directed gene regulation in *T. gondii*, to gain a better understanding of the function of gene of interest following a negative regulation at the post-transcription level.

3.3.5. Effective gene silencing caused by directed *Tg*-miR-60a

As a proof of principle that the highly expressed *Tg*-miR-60 family can be directed to control the expression of any genes of interest, we chose two unrelated target genes to evaluate (i) the effect of directed gene silencing and (ii) the practicality and usefulness of this genetic tool in functional analysis. The first chosen target is a less

studied gene, named DEAD-box RNA helicase (TgHoDI, TGGT1_089230; Cherry and Ananvoranich, 2014), whose ortholog is dispensable for growth in yeast (Tseng-Rogenski et al. 2003). The second is a metabolically crucial gene, lactate dehydrogenase (TgLDH1, TGGT1_232350). Due to (i) the relatively high abundance of *Tg*-miRNAs and (ii) the lack of information on authentic targets of *Tg*-miRNAs, it was possible that any given transcript species might already be under the control of a *Tg*-miRNA species. To ensure that the effect observed here was the result of our *Tg*-miR-60a-directed silencing machinery and not from other *Tg*-miRNA effects, the entire nucleotide sequences of *Tg*HoDI and *Tg*LDH1 transcripts (pre-mRNA and mature mRNA) were searched for a possible *Tg*-miR-60a binding site to partially (at least 50%) or fully interact with the seed or entire sequence of *Tg*-miR-60a (Table 3.2, underlined sequence). No potential binding sites of *Tg*-miR-60a were found on either transcript.

3.3.5.1. Silencing of DEAD-box RNA helicase (*Tg*HoDI) expression by *Tg*-miR60a caused a subtle change, but sufficiently demonstrated its physiological importance during the re-invasion

*Tg*HoDI is a member of DEAD-box RNA helicase superfamily, which contains highly conserved proteins and implicated in almost all aspects of RNA metabolism in organisms ranging from yeast to complex eukaryotes (Hooper and Hilliker, 2013; Carroll et al., 2011). *Tg*HoDI was reported to be involved in the formation of mRNA-ribonucleoprotein (mRNP) complexes and granules during stress response (Cherry and Ananvoranich, 2014). Tseng-Rogenski et al. (2003) showed that a null mutant of DEAD-box RNA helicase (DHH1) in yeast is not lethal but exhibits a temperature sensitive phenotype. Using this Δdhh1 strain, Cherry and Ananvoranich (2014) showed that the expression of *Tg*HoDI cDNA in the yeast strain under DHH1 promoter could

complement and partially restore its phenotype. In an attempt to study the physiological significance of *Tg*HoDI in *T. gondii*, a ligand-controlled destabilization domain (ddFKBP) version of *Tg*HoDI was generated. Although stable parasite lines expressing TgHoDI-ddFKBP fusion protein were obtained, the ligand-induced protein degradation was not achieved (unpublished data). Indirectly this suggested (i) that the C-terminus of TgHoDI was not solvent-exposed, and (ii) that *Tg*HoDI could be involved in the rapid formation of mRNPs similar to other DEAD-box RNA helicase members (Hooper and Hilliker, 2013). Due to its possibly tight regulation and physiological importance, TgHoDI was chosen as a target for *Tg*- miR60a-directed gene silencing.

Stable parasite lines expressing either *Tg*HoDI-SF (control) or *Tg*HoDI-SF-miR (test) were generated and maintained for the study. Under normal growth conditions the test strain expressing *Tg*HoDI-SF-miR expressed approximately 50% and 30% less *Tg*HoDI protein and mRNA, respectively, in comparison to the control strain (−, Figure 3.4, A and B). The introduction of *Tg*-miR- 60a mimic (+, Figure 3.4, A and B) enhanced the extent of *Tg*-miR60adirected silencing and decreased the levels of *Tg*HoDI protein and mRNA to approximately 25% and 10%, but did not affect the expression of *Tg*HoDI in the control strain. *Tg*HoDI was previously shown as a protein marker of mRNP granules called stress granules (Cherry and Ananvoranich, 2014). The accumulation of cytoplasmic mRNA granules, including those of stress granules, in extracellular tachyzoites was showed to be important to the ability of extracellular parasites to survive and re-establish their infectivity. The higher numbers and larger sizes of the stress granules formed during the stressful conditions translated to a higher success in reestablishment of infection (Lirussi and Matrajt, 2011). Using these stable parasite lines,

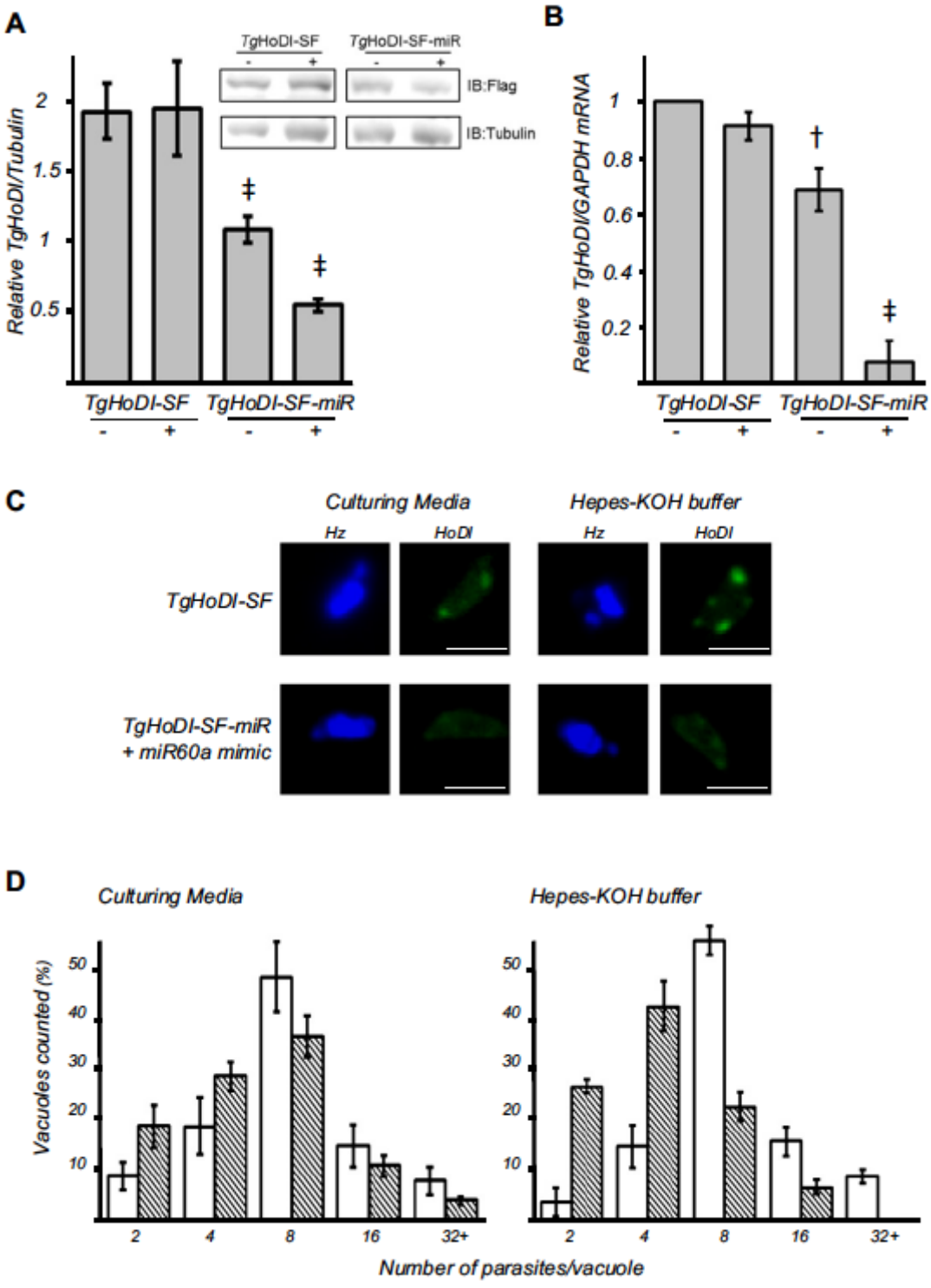

A
TgHoDI-SF
TgHoDI-SF-miR
IB:Flag
IB:Tubulin
Relative TgHoDI/Tubulin
TgHoDI-SF
TgHoDI-SF-miR
B
Relative TgHoDI/GAPDH mRNA
C
Culturing Media
Hepes-KOH buffer
Hz
HoDI
TgHoDI-SF
TgHoDI-SF-miR
+ miR60a mimic
D
Culturing Media
Hepes-KOH buffer
Vacuoles counted (%)
Number of parasites/vacuole

Figure 3.4. TgHoDI expressions under the control of *Tg*-miR60a-induced gene silencing. Freshly released transgenic parasites, namely TgHoDI-SF and TgHoDI-SF-miR were electroporated with Tg-miR60a mimic (20 μg, +; 0 μg, −). A. Thirty-six hours following electroporation, the parasites were released and harvested for the cell-free lysates which were analyzed for the levels of TgHoDI-SF fusion protein using α-Flag antisera. The levels of tubulin detected using α-tubulin antisera were used as a loading control. Relative signals of TgHoDI-SF fusion protein was normalized by those of tubulin. The data were obtained from three independent tests. B. At a similar time point, total RNAs were isolated and subjected to RT-qPCR analyses. Daggers (†) represent p-value ≤ 0.05, while double daggers (‡) represent p-value ≤ 0.01. C. Representative images of extracellular parasites following incubation in either culturing medium or Hepes-KOH solution (60-minute treatment), the accumulation of TgHoDI were assessed using α-Flag antisera in conjunction with α-mouse FITC conjugated antisera. Micrometer bar represents 5 μm. D. A parallel analyses of those shown in C. An aliquot of treated parasites were transferred onto HFF monolayers to allow for a reinvasion. Their ability to reinvade and recover from the treatment was assessed 24 hours later. The number of parasites per vacuole was counted. Data was collected and calculated from at least three independent trials, where clear bars represent those from TgHoDI-SF, and hatched bars, TgHoDI-SF-miR.

we aimed to validate the function of *Tg*HoDI in re-establishing their infectivity. To establish a stress condition where the accumulation of *Tg*HoDI and mRNP granules could clearly be monitored, the extracellular parasites of *Tg*HoDI-SF control strain were released from infected monolayers and subjected to stressful conditions. We utilized a Hepes-KOH solution treatment. Use of a buffering solution containing 27 times higher KCl, and 22 times less NaCl than those of the culturing media was previously shown to induce a stress response in *T. gondii* (Lirussi and Matrajt, 2011). The Hepes-KOH solution (25 mM Hepes-KOH pH 7.2, 142 mM KCl, 5 mM NaCl, 2 mM EGTA, 5 mM $MgCl_2$, and 1 mg/mL BSA) mimicked the composition of the host cell cytosol ionic conditions, which the parasites come in contact with during egress. It was observed that the incubation of extracellular parasites in this Hepes-KOH solution for 60 minutes at 37 °C caused maximal formation of mRNP granules, as observed by the accumulation of *Tg*HoDI aggregates throughout the cytosol (Figure 3.4c, upper right panel). Subsequently we silenced the *Tg*HoDI expression to the lowest level possible by introducing *Tg*-miR-60a mimic into freshly released *Tg*HoDI-SF-miR parasites via electroporation. The parasites were allowed to infect and to grow in HFF monolayers for 48 hours to ensure that the level of *Tg*HoDI protein was approximately 75% lower than those of the control untreated parasites (as shown in Figure 3.4a). Extracellular parasites were then released and subjected to the solution treatment for 60 minutes at 37 °C. It was observed that the numbers and sizes of *Tg*HoDI aggregates were diminished as compared those of the control conditions (Figure 3.4c, lower right panel). In parallel, these treated parasites were allowed to re-invade HFF monolayers after the stressful conditions and were monitored for their reinvasion ability by microscopic examination 24 hours following an

infection of HFF monolayers. When the extracellular parasites were kept in their culturing media for 60 minutes at 37 °C (without the solution treatment), both parasite strains were able to retain a similar infectivity, namely reinvasion and multiplication (Figure 3.4d left histogram, clear bars for *Tg*HoDI-SF, control; hatched bars for *Tg*HoDI-SF-miR, test). The majority (50%) of the intracellular vacuoles contained 8 parasites per vacuole at 24 hours post-infection. Both stable parasite lines exhibited a similar growth profile with ~ 6-8 hours doubling time, indicating that the changes in the level of *Tg*HoDI gene products did not significantly affect the parasite ability to multiply under non-stress conditions. On the other hand, following the solution treatment, the extracellular parasites of the test strain with electroporated *Tg*-miR-60a mimic were slower in progressing through reinvasion and multiplicity. The majority of the formed vacuoles of the test strain contained 4 parasites per vacuoles (hatched bars), in comparison to 8 parasites per vacuoles in those of control strain (clear bars) when subjected to a similar treatment. The decrease of *Tg*HoDI protein (four times less than the control) was concomitantly detected along with the slowing down of parasite multiplicity (Figure 3.4d, right histogram). The changes in growth profiles were not observed in mock or *Hs*-let-7a- mimic electroporation, nor were the detectable changes in *Tg*HoDI accumulation and growth observed in other electroporation reactions in either strain (data not shown). The findings thus confirmed that the accumulation of stress granules in extracellular tachyzoites is important to the ability of extracellular parasites to re-establish their infectivity.

3.3.5.2. Expression of lactate dehydrogenase isoform 1 is critical to the parasite cell biology

Being an obligate intracellular parasite, *T. gondii* obtains its nutrients from the host cells. Under anaerobic conditions, the parasite relies on *Tg*LDH which is the last

committed enzyme of the glycolytic pathway to interconvert pyruvate and lactate, as well as NAD+ and NADH. Because of the necessity of the NAD+-NADH conversion under anaerobic conditions, *Tg*LDH is thus considered one of potential drug targets. Two active *Tg*LDH isoforms, namely *Tg*LDH1 and *Tg*LDH2, are produced in a stage specific manner. *Tg*LDH1 is detected in tachyzoites, while *Tg*LDH2 in bradyzoites (Yang and Parmley, 1997). To date, the mechanism underlining the requirement of each specific isoform is still under investigation. The presence of stage specific *Tg*LDH1 protein is commonly used as a marker in displaying the tachyzoite development stage Yang and Parmley, 1997). While no documentation on a *Tg*LDH1 null mutant, one successful *Tg*LDH1 knockdown was reported using long double-stranded RNA expression generated from a head-to-head promoter construct (Al-Anouti et al., 2004). The *Tg*LDH1 knockdown strain was documented to grow slower and to convert into bradyzoites in a lesser degree than the parental strain, indicating its physiological importance. As part of the proof of principle, the metabolic important *Tg*LDH1 gene was chosen as a second target to further gauge the level of *Tg*-miR-60a-induced silencing system toward a critical and essential gene. Similar to the *Tg*HoDI study, two stably transgenic lines, *Tg*LDH1-SF and *Tg*LDH1-SF-miR were desirable for the study. While *Tg*LDH1-SF parasite strain was easily obtained and maintained, we were unable to obtain or maintain the *Tg*LDH1-SF-miR strain, despite a number of control conditions tested. It thus indirectly indicated that the level of *Tg*LDH1 expression was effectively silenced below a threshold level to allow the parasite to maintain a metabolic requirement for its viability. To further investigate the cause(s) of our difficulties in the generation and/or maintenance of *Tg*LDH1-SF-miR line, the transfection reactions were performed to include those with

pTgLDH1-SF, or pTgLDH1-SF-miR (Figure 3.5a). The electroporated parasites were allowed to infect HFF monolayers for 36 hours using selective media. Due to the nature of transfection reactions, the bulk of parasites contained both transformed and non-transformed parasites. The level of *Tg*LDH1-SF fusion protein was thus quantified in comparison to the total level of *Tg*LDH1 and tubulin background from three independent experiments by using immunoblot-densitometric analyses (Figure 3.5b). The level of *Tg*LDH1 protein was drastically decreased in the *Tg*-miR- 60-induced silencing strain (T1), whose expression could be restored by the introduction of *Tg*-miR60a inhibitor (T2). The silencing effect was also detected by immunocytochemistry (Figure 3.5c). Using α-Flag antisera, the epitope tagged *Tg*LDH1 protein was detected in CTL, and T2, respectively. However, no α-Flag signals were detected in T1. Altogether, it indicated that transgenic *Tg*LDH1-SF-miR parasites can be generated, but the stable clonal line cannot be maintained for further study. The findings thus reiterated the importance of *Tg*LDH1 functionality in parasite viability. In addition, the findings also confirmed that the abundant *Tg*-miR60a can be directed to silence a target gene, and that *Tg*-miR60a inhibitor can be used to restore the parasites from detrimental effect of gene silencing.

3.3.5.3. Versatility and possible limitation of the methodology

The miRNA-directed gene silencing methodology reported here offers an alternative approach to currently available functional genetic procedures. Due to the abundance and variety of *Tg*-miRNAs, necessary reagents could be tailored to the appropriate gene of interest. The plasmids could easily be modified to introduce the binding site of another species of *Tg*-miRNAs, when *Tg*-miR-60a is not appropriate for the gene of interest. For example, when the gene of interest is already under the control of

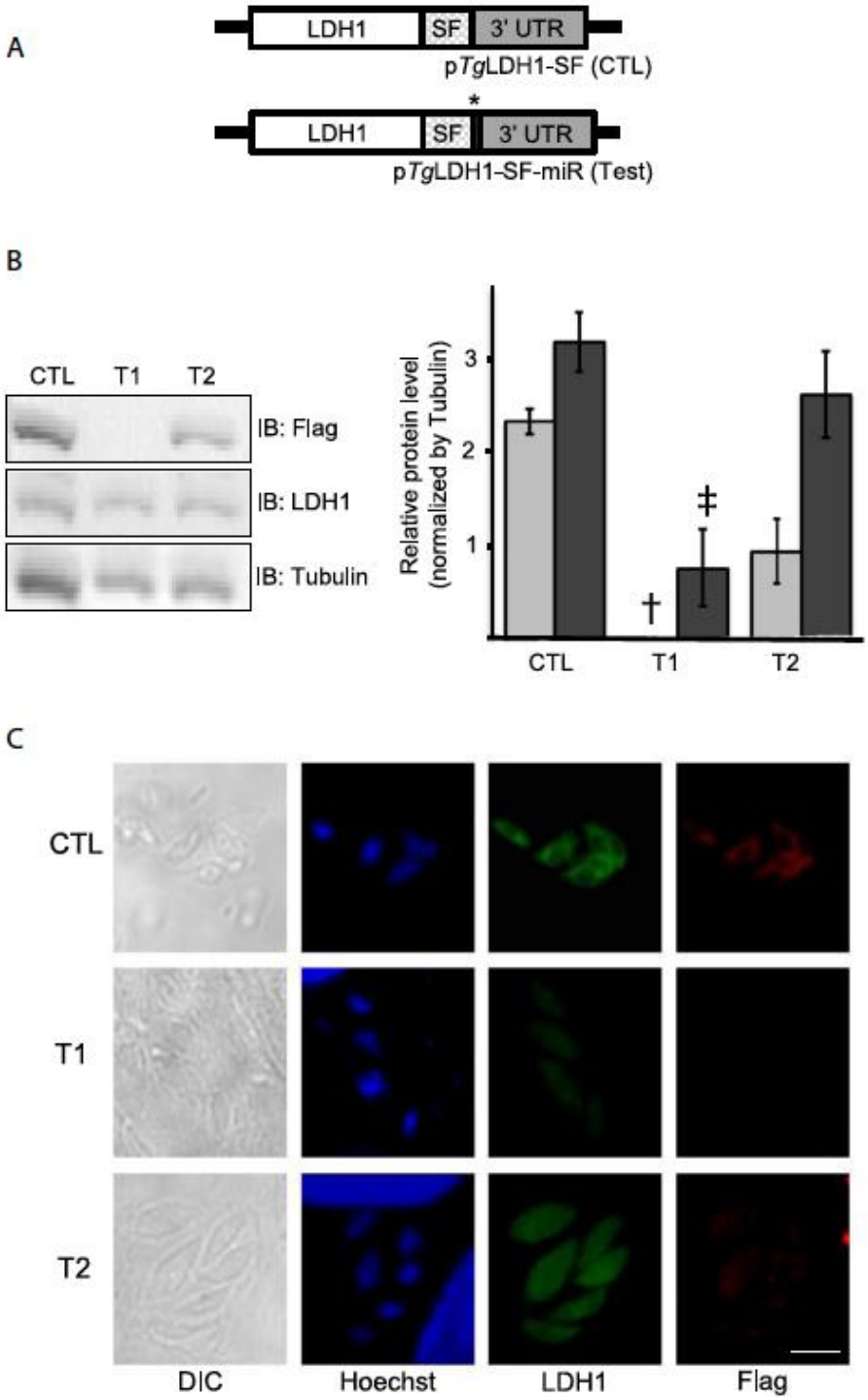

A
LDH1
SF
3' UTR
pTgLDH1-SF (CTL)
*
LDH1
SF
3' UTR
pTgLDH1-SF-miR (Test)
B
CTL
T1
T2
IB: Flag
IB: LDH1
IB: Tubulin
Relative protein level (normalized by Tubulin)
3
2
1
†
‡
CTL
T1
T2
C
CTL
T1
T2
DIC
Hoechst
LDH1
Flag

Figure 3.5. TgLDH1 loss-of-function analyses. Freshly released Δku80 tachyzoites were electroporated with indicated plasmid construct (20 μg), which was engineered to allow for a homologous recombination at the TgLDH1 locus. A. Representation of plasmid constructs aimed for a homologous recombination to incorporate the nucleotide sequence coding epitope tags of Strep and Flag (SF) to be fused to the coding sequence of the gene (pTgLDH1-SF, CTL), and for the similar construct with the presence of three binding sites for Tg-miR60a immediate after the stop codon (as indicated by asterisk (*), pTgLDH1-SF-miR; test). B. Thirty-six hours following electroporation, the parasites were released and harvested for the cell-free lysate preparation. The lysates were prepared from electroporation reactions containing CTL, test construct (T1) and test construct in the presence of Tg-miR60a inhibitor (20 μg, T2). The levels of TgLDH1 and TgLDH1-SF fusion proteins were detected using α-TgLDH1 and α-Flag antisera. A-Tubulin antiserum was used as a loading control. Notably, immunoblots probed with α-Flag antisera were over-exposed to allow for the visualization of Flag tagged signals from T2 samples. Densitometric data were obtained from three independent immunoblot analyses of the samples, and plotted as means ± standard deviations. The light bars represent those of TgLDH1-SF and the dark bars of TgLDH1. Both were normalized the signals obtained from TgTubulin. Daggers (†) represent p-value ≤ 0.05, while double daggers (‡) represent p-value ≤ 0.01. C. The expression of TgLDH1 and TgLDH1-SF proteins was monitored by immunocytochemistry analyses thirty-six hours following the electroporation. Micrometer bar represents 5 μm.

Tg-miR-60a, the introduction of more *Tg*-miR-60a binding sites will not yield any significant down regulation effect for the analysis. This possible limitation, however, could easily be remedied by utilizing or directing other *Tg*-miRNAs, such as *Tg*-miR-4a to the target gene products. It is also possible that the expression profile of *Tg*-miR-60a is not suitable for that of the gene of interest or the parasite archetype strain. When such conditions occur, the reagents could be modified so that the expression profile and strain are suitable and compatible.

3.4. Conclusion

We showed that *Tg*-miRNAs, at least the two most abundant *Tg*-miRNAs, namely *Tg*-miR-60a and *Tg*-miR-4a, are functional effectors to alter the expression of a reporter gene, Rnluc of the dual luciferase reporter system. The extent of their silencing effect was influenced by (i) the number of binding sites and (ii) the abundance of *Tg*-miRNA species. A higher silencing effect was observed (i) when the transcripts carry multiple binding sites, and (ii) when the levels of *Tg*-miRNA effector are present in a higher concentration. The nature of the interactions between target transcripts and the *Tg*-miRNA species directed the mode of gene silencing, either mRNA degradation or translational suppression. It remains to be investigated how the complementarity and specific *Tg*-miRNA species dictate the mode of reactions, as unexpectedly we observed that perfect matches of *Tg*-miR-60a and *Tg*-miR-4a had a different mode of silencing despite their similar end point. Moreover, we could take advantage of *Tg*-miR-60a miRNA mimics and inhibitors to modulate the silencing effect. As a proof of principle, we successfully demonstrated that the endogenously abundant miRNAs can be directed to down regulate the expression of two model targets, *Tg*HoDI and *Tg*LDH1. The extent

of gene silencing can be further enhanced by miRNA mimics or diminished by miRNA inhibitors, as needed. In conjunction with recent and powerful CRISPR technology, the binding site(s) of *Tg*-miRNAs can easily be introduced into any ectopic sites to cause specific gene silencing, when a knockdown, but not knockout, of the target gene is needed. Nonetheless, the most important consideration should be placed on developing of an effective readout of the gene silencing. Naturally to maintain cellular activities, various regulatory systems work in concert. Gene and functional redundancy would allow the expression of genes to be temporarily compromised. While *Tg*HoDI knockdown was effective, it required an effective measurement to detect the subtle changes. In the case of an essential gene, such as *Tg*LDH1, a gene knockout would be lethal and difficult for gaining further information. By having a gene knockdown protocol, the information could be extracted and deduced.

3.5. Acknowledgement

This work was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC); University of Windsor Doctoral Entrance Scholarship to ACP.

3.6. References

Al-Anouti, F., Tomavo, S., Parmley, S., Ananvoranich, S., 2004. The expression of lactate dehydrogenase is important for the cell cycle of *Toxoplasma gondii*. J. Biol. Chem. 279, 52300–52311.

Boothroyd, J.C., Black, M., Bonnefoy, S., Hehl, A., Knoll, L.J., Manger, I.D., Ortega-Barria, E., Tomavo, S., 1997. Genetic and biochemical analysis of development in *Toxoplasma gondii*. Philos. Trans. R. Soc. Lond. B Biol. Sci. 352, 1347–1354.

Braun, L., Cannella, D., Ortet, P., Barakat, M., Sautel, C.F., Kieffer, S., Garin, J., Bastien, O., Voinnet, O., Hakimi, M.A., 2010. A complex small RNA repertoire is

generated by a plant/fungal-like machinery and effected by a metazoan-like Argonaute in the single-cell human parasite *Toxoplasma gondii*. PloS Pathog. 6, e1000920.

Carroll, J.S., Munchel, S.E., Weis, K., 2011. The DexD/H box ATPase Dhh1 functions in translational repression, mRNA decay, and processing body dynamics. J. Cell Biol. 194, 527–537.

Cenci-Goga, B.T., Rossitto, P.V., Sechi, P., McCrindle, C.M., Cullor, J.S., 2011. *Toxoplasma* in animals, food, and humans: an old parasite of new concern. Foodborne Pathog. Dis. 8, 751–762.

Cherry, A.A., Ananvoranich, S., 2014. Characterization of a homolog of DEAD-box RNAhelicases in *Toxoplasma gondii* as a marker of cytoplasmic mRNP stress granules. Gene 543, 34–44.

Crater, A.K., Cherry, A., Holmes, M., Kadri, D., Ananvoranich, S., 2011. Evaluation of the ability of short and long double-stranded RNAs to induce homologous gene silencing in the protozoan parasite, *Toxoplasma gondii*. Am. J. Biomed. Sci. 4, 1–13.

Gajria, B., Bahl, A., Brestelli, J., Dommer, J., Fischer, S., Gao, X., Heiges, M., Iodice, J., Kissinger, J.C., Mackey, A.J., Pinney, D.F., Roos, D.S., Stoeckert Jr., C.J., Wang, H., Brunk, B.P., 2008. ToxoDB: an integrated *Toxoplasma gondii* database resource. Nucleic Acids Res. 36, D553–D556.

Heaslip, A.T., Dzierszinski, F., Stein, B., Hu, K., 2010. TgMORN1 is a key organizer for the basal complex of *Toxoplasma gondii*. PloS Pathog. 6, e1000754.

Herm-Götz, A., Agop-Nersesian, C., Münter, S., Grimley, J.S., Wandless, T.J., Frischknecht, F., Meissner, M., 2007. Rapid control of protein level in the apicomplexan *Toxoplasma gondii*. Nat. Methods 4, 1003–1005.

Hooper, C., Hilliker, A., 2013. Packing them up and dusting them off: RNA helicases and mRNA storage. Biochim. Biophys. Acta 1829, 824–834.

Huynh, M.H., Carruthers, V.B., 2009. Tagging of endogenous genes in a *Toxoplasma gondii* strain lacking Ku80. Eukaryot. Cell 8, 530–539.

Jammallo, L., Eidell, K., Davis, P.H., Dufort, F.J., Cronin, C., Thirugnanam, S., Chiles, T.C., Roos, D.S., Gubbels, M.J., 2011. An insertional trap for conditional gene expression in *Toxoplasma gondii*: identification of TAF250 as an essential gene. Mol. Biochem. Parasitol. 175, 133–143.

Jiménez-Ruiz, E., Wong, E.H., Pall, G.S., Meissner, M., 2014. Advantages and disadvantages of conditional systems for characterization of essential genes in *Toxoplasma gondii*. Parasitology 13, 1–9.

Kolev, N.G., Tschudi, C., Ullu, E., 2011. RNA interference in protozoan parasites: achievements and challenges. Eukaryot. Cell 10, 1156–1163.

Lirussi, D., Matrajt, M., 2011. RNA granules present only in extracellular *Toxoplasma gondii* increase parasite viability. Int. J. Biol. Sci. 7, 960–967.

Meissner, M., Brecht, S., Bujard, H., Soldati, D., 2001. Modulation of myosin A expression by a newly established tetracycline repressor-based inducible system in *Toxoplasma gondii*. Nucleic Acids Res. 15, 29.

Meissner, M., Breinich, M.S., Gilson, P.R., Crabb, B.S., 2007. Molecular genetic tools in *Toxoplasma* and *Plasmodium*: achievements and future needs. Curr. Opin. Microbiol. 10, 349–356.

Montoya, G., Liesenfeld, O., 2004. Toxoplasmosis. Lancet 363, 1965–1976.

Musiyenko, A., Majumdar, T., Andrews, J., Adams, B., Barik, S., 2012. PRMT1 methylates the single Argonaute of *Toxoplasma gondii* and is important for the recruitment of Tudor nuclease for target RNA cleavage by antisense guide RNA. Cell. Microbiol. 14, 882–901.

Pillai, R.S., Bhattacharyya, S.N., Artus, C.G., Zoller, T., Cougot, N., Basyuk, E., Bertrand, E., Filipowicz, W., 2005. Inhibition of translational initiation by Let-7 MicroRNA in human cells. Science 309, 1573–1576.

Roos, D.S., Donald, R.G., Morrissette, N.S., Moulton, A.L., 1994. Molecular tools for genetic dissection of the protozoan parasite *Toxoplasma gondii*. Methods Cell Biol. 45, 27–63.

Shen, B., Brown, K.M., Lee, T.D., Sibley, L.D., 2014. Efficient gene disruption in diverse strains of *Toxoplasma gondii* using CRISPR/CAS9. mBio 5 (6), e02021-14.

Sidik, S.M., Hackett, C.G., Tran, F., Westwood, N.J., Lourido, S., 2014. Efficient genome engineering of *Toxoplasma gondii* using CRISPR/Cas9. PloS One 9, e100450.

Soldati, D., Boothroyd, J.C., 1993. Transient transfection and expression in the obligate intracellular parasite *Toxoplasma gondii*. Science 260, 349–352.

Tseng-Rogenski, S.S., Chong, J.L., Thomas, C.B., Enomoto, S., Berman, J., Chang, T.H., 2003. Functional conservation of Dhh1p, a cytoplasmic DexD/H-box protein present in large complexes. Nucleic Acids Res. 31, 4995–5002.

Ullu, E., Tschudi, C., Chakraborty, T., 2004. RNA interference in protozoan parasites. Cell. Microbiol. 6, 509–519.

Upadhya, R., Kim, K., Hogue-Angeletti, R., Weiss, L.M., 2011. Improved techniques for endogenous epitope tagging and gene deletion in *Toxoplasma gondii*. J. Microbiol. Methods 85, 103–113.

Van Poppel, N.F., Welagen, J., Duisters, R.F., Vermeulen, A.N., Schaap, D., 2006. Tight control of transcription in *Toxoplasma gondii* using an alternative tet repressor. Int. J. Parasitol. 36, 443–452.

Wakiyama, M., Takimoto, K., Ohara, O., Yokoyama, S., 2007. Let-7 microRNA-mediated mRNA deadenylation and translational repression in a mammalian cell-free system. Erratum in Genes Dev. 21:1857–1862. Genes Dev. 21, 2509.

Wang, J., Liu, X., Jia, B., Lu, H., Peng, S., Piao, X., Hou, N., Cai, P., Yin, J., Jiang, N., Chen, Q., 2012. A comparative study of small RNAs in *Toxoplasma gondii* of distinct genotypes. Parasites Vectors 5, 186.

Xu, M.J., Zhou, D.H., Huang, S.Y., Zhao, F.R., Nisbet, A.J., Lin, R.Q., Song, H.Q., Zhu, X.Q., 2013. Comparative characterization of microRNA profiles of different genotypes of *Toxoplasma gondii*. Parasitology 140, 1111–1118.

Yang, S., Parmley, S.F., 1997. *Toxoplasma gondii* expresses two distinct lactate dehydrogenase homologous genes during its life cycle in intermediate hosts. Gene 184, 1–12.

CHAPTER 4

Antisense technologies in the studying of *Toxoplasma gondii*

4.1. *Toxoplasma gondii*, a model organism

T. gondii is one of the most successful intracellular parasites which is reported to infect any warm-blooded animals, and an estimated one third of the world's population (Kim and Weiss, 2008). As a member of the phylum Apicomplexa, *Toxoplasma* has a unique non-photosynthetic apicoplast organelle, similar to that found in other important and pathogenic microorganism members, including *Plasmodium* which is the cause of malaria, Cryptosporidium a waterborne parasite, as well as *Eimeria*, *Neospora*, and *Theleria* all can cause disease in animals. These pathogenic microorganisms share a remarkable evolutionarily relationship. Using *Toxoplasma* as a study model, we review various types of non-coding RNAs which have emerged and could be effectors in the regulation of gene expression (Radke et al., 2005; Hassan et al., 2012; Ramaprasad et al., 2015). Recent research shows that *Toxoplasma* contains a "mix-and-match" of plant- and animal-like machinery used to direct RNA for gene regulation (Braun et al., 2010; Musiyenko et al., 2012). *Toxoplasma* can be easily cultured using any nucleated cell cultures and has proven previously to be a very useful model to develop a number of molecular and genetic tools (Roos et al., 1994; Kim and Weiss, 2004). Taken together, *Toxoplasma* is an important model organism, and a fascinating system for exploring the progress in antisense technologies.

4.2. History of antisense RNAs and other regulatory RNAs

The term "antisense RNA" was coined in the early 1970s, and first used in describing "RNA repressors" in prokaryotes. Spiegelman et al. (1972) showed that

antisense RNA molecules can hybridize with complementary sense RNA to render it inactive. This experiment was among the first to define the function of regulatory RNA as a highly specific repressor or inhibitor of gene expression. Coleman et al. (1985) showed that antisense RNA could be used to target foreign genes in an experiment, in which resistance to bacteriophage SP infection could be conferred to Escherichia coli. They achieved this by creating plasmids coding for RNA sequences which targeted genes coding for bacteriophage coat protein and/or replicase. By the time that Green et al. (1986) published a review on antisense RNA technologies, it was widely accepted that antisense RNA can perform a repressor function, and that the 5′ end of the messenger RNA (mRNA) was the key target sequence. Antisense RNA which did not include sequences complementarity to the 5′ end of the target mRNA required much higher concentrations to achieve an effective repression. Additionally, some genes remained resistant to repression by antisense RNA. It was suggested that mRNA with a high degree of secondary structure was resistant to hybridization with antisense RNA. Pioneer work on antisense RNAs in prokaryotes has provided us with fundamental knowledge on various aspects of RNA directed gene regulation, both inhibition and activation of replication, transcription and translation.

The applications of antisense technology in eukaryotic systems began as early as 1978. Zamecnik and Stephenson (1978) were able to temporarily inhibit the production of Rous sarcoma virus (RSV) in chick embryo fibroblast tissue cultures. This study targeted the intermediate provirus DNA, and as such used antisense DNA rather than RNA. As with antisense RNA, hybridization between the provirus DNA and the injected antisense DNA was believed to prevent the normal function of the oligonucleotide, in this

case the necessary circularization and/or incorporation of the viral genome. By 1984 artificial antisense RNA was successfully used to regulate the expression of a eukaryotic gene, named thymidine kinase in murine cells (Izant and Weintraub, 1984). Bunch and Goldstein (1989) developed a system for generating antisense RNA *in vivo* with cultured Drosophila melanogaster cells. The methodology used by Bunch and Goldstein (1989) offered an alternative to the commonly used microinjection as a method by which artificial antisense RNA was introduced into the eukaryotic cells. The study also suggested that hybridization activity between sense and antisense RNA in eukaryotes occurred primarily in the nucleus with un-spliced mRNA. However, it was later shown by Vanhee-Brossollet and Vaquero (1998) that antisense RNA molecules have regulatory function both in the nucleus and cytoplasm, and at any step of gene expression. It has more recently been found that this multi-level expression control is possible due to several different types of antisense RNA.

4.3. Application of antisense technologies and their progress in *Toxoplasma*

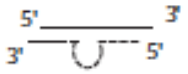
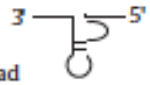
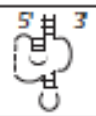
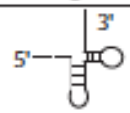
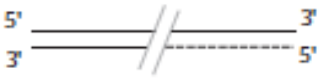
After decades of the discovery of antisense RNAs and other regulatory RNAs, underlining mechanisms and functions are still being investigated and appear to vary greatly among organisms and various types of regulatory RNAs. These RNAs are non-protein-coding transcripts and referred to as non-coding RNAs (ncRNAs). Due to the varied structures and mechanisms, regulatory RNAs or ncRNAs are generally categorized based on their sizes: (i) small ncRNAs, whose lengths range from ~20 nucleotides to ~200 nucleotides, include siRNA, miRNA, and piwiRNA. (ii) Long ncRNAs are varied from over 200 nucleotides to a few thousands nucleotides and include natural antisense transcripts (NAT), intra- and inter-genic ncRNAs, and intronic RNAs. Both small and

large regulatory ncRNAs could inhibit gene expression and be exploited as target specific genetic tools (Wahlestedt, 2013; Krishnan and Mishra, 2014; Miller, 2014). It is clear that the modes of inhibition are different depending on the type of ncRNAs. First, siRNAs and miRNAs affect mRNA stability mostly through the involvement of the RNA-induced silencing complex (RISC) in the cytoplasm and cause mRNA degradation and/or translational repression, while long ncRNAs primarily affect chromatin regulators, which suppress transcription, and thus act in the nucleus. Their spatial requirement could be a limiting factor, if we aim to perturb the natural function(s) of these regulatory RNAs. Second, small ncRNAs, miRNAs in particular, can affect the stability of multiple mRNA targets, while long ncRNAs, particularly cis-acting NAT, are locus specific. When target specificity is critical, the use of miRNAs and their derivatives require proper controls to discern between targeted-inhibition and an off-target effect. Thirdly, the abundance of small ncRNAs is likely to be higher than that of large ncRNAs. This could be an important factor in considering for their dose-to-efficacy property. In the following sections, we summarize genetic silencing tools and technologies used in *Toxoplasma* according to their capability related to naturally found regulatory RNAs. Table 4.1 shows the representative structures, relative sizes, and targets of antisense technologies summarized here.

4.3.1. Small non-coding RNAs

The *Toxoplasma* transcriptome contains a large repertoire of small ncRNAs, including the most prominent group of miRNAs (Braun et al., 2010; Wang et al., 2012; Xu et al., 2013). Two of the most abundant miRNAs families, namely miR60a and miR4a of parasite strain GT1 (Braun et al., 2010), were tested using engineered dual luciferase

Table 4.1. Representative structures and relative sizes of regulatory ncRNAs used in *Toxoplasma gondii*. The dash lines represent the antisense strands which will pair with their targets.

Effector RNAs	Structures	Targets
siRNA (20–25 nt)		mRNA
miRNA (20–25 nt)		mRNA
miRNA inhibitor (20–30nt)		siRNA/miRNA
Antagomir (20–30 nt)		miRNA
Morpholino (~25 nt)		mRNA
Ribozyme (7–20 nt)	Hammerhead 	RNA
	Delta 	
	RNAseP 	
Guide RNA (20 nt–NGG)		DNA/RNA
Long dsRNA (200–600 nt, up to 2 kb)		mRNA
Antisense RNA (~2,000 nt)		mRNA and DNA

reporter system (Crater et al., 2015). The expression of engineered transcripts harboring binding sites of either miR60a or miR4a was repressed in a sequence specific manner, clearly indicating that *Toxoplasma* harbors miRNA-inducing silencing machinery (Crater et al., 2015). Also the silencing effect detected using the reporter system indicated that the extent of silencing effect is dependent on the abundance of miRNA present and the number of miRNA binding sites on a target transcript. It was noted in the same study that the nature of silencing effect could depend on the miRNA species. While miR60 was detected to induce mRNA degradation, miR4a was found to induce translational repression. The findings indicated that *Toxoplasma* silencing machinery could differ from those of other organisms. Musiyenko et al. (2012) showed that *Toxoplasma*-Argonaute (Ago) displays a weak RNA cleavage activity and requires an association with a Tudor staphylococcal nuclease (TSN) and Arginine methyl transferase for its activity. The study also strongly suggested that the RNA- induced silencing complex of *Toxoplasma* utilizes the slicer activity of TSN for partially mismatched miRNA target pairs, and that of Ago for perfect match targets.

4.3.1.1. siRNA and miRNA

It has been demonstrated that exogenous supplies of miRNAs or siRNA could enhance or induce an inhibition of gene expression. We here referred to the short dsRNA (20–25 nt) with perfect-matched central bases as “siRNA”, and that with mismatched central bases as “miRNA” (Table 4.1). Yu et al. (2008) successfully showed that commercially designed and synthesized Stealth™ RNAi duplexes (25-mer dsRNA, Invitrogen) against *Toxoplasma* adenosine kinase can inhibit the gene expression. When individual siRNAs of the Stealth™ RNAi duplexes were introduced into the parasite by

transfection, it was detected that each siRNA exhibited varied efficacy depending on the target locations, and the dosage used. The findings are in accordance with other RNA-targeted gene silencing methodologies in that not all target regions of an mRNA are equally accessible. Wahlestedt (2013) stated that only 10–15% of designed antisense oligonucleotides will reduce the expression of the target. To enhance the inhibition effect, a common practice is to make a cocktail of siRNA, as suggested and used with the commercial Stealth™ RNAi duplexes (Invitrogen).

The capability of siRNA and miRNA to inhibit gene expression was systemically compared using engineered dual luciferase reporter system (Crater et al., 2012). By *in vitro* transcription, but not by chemical synthesized as those of Stealth™ RNAi system, siRNAs and miRNAs were made, and three binding sites for either siRNA or miRNA were placed onto the 3′-UTR of Renilla luciferase transcripts. Both siRNA and miRNA were shown to exhibit a similar silencing ability, and their effect appeared to be enhanced with longer incubation time (24 h vs. 48 h). In the same study, the effect of these small regulatory RNAs was compared to that of long double-stranded RNAs (~600 bp) which are specifically targeted Renilla luciferase transcripts. The silencing effect of long dsRNAs is lower than those of siRNA and miRNA. It is suggested that an additional processing step for the long dsRNA might be necessary, although we cannot exclude the possibility that it might be due to the different site of action of long and short RNA regulators.

The above mentioned studies clearly indicate that small regulatory RNA, in the form of siRNA and miRNA can be designed and effectively used as a target-specific silencing tool. With the aid of reliable software's algorithms, effective siRNA and/or

miRNA cocktails can be chemically or enzymatically synthesized. Most companies that commercialize the production of siRNA or miRNA provide designing services. Although we cannot advocate for a particular company or software to be superior to others, the best practice is to use multiple sites to confirm and re-confirm the design of siRNA and miRNA whenever possible.

4.3.1.2. miRNA inhibitors

miRNA inhibitors are short single stranded RNA (~20–30 nucleotides) with sequence complementary to intercept the interaction of miRNA and its target mRNA or gene. While it is a challenge to identify the target gene of a miRNA species in an organism, due to the nature of perfect or impartial base-pairing of these small regulatory RNAs to their targets, Braun et al. (2010) identified a number of potential targets by allowing up to 3 mismatches between their identified miRNA and their targets. It was stated that more than 80 putative target transcripts were identified for miR-60a alone, and that an average of 25 cellular targets could be retrieved for each of the other identified *Toxoplasma* miRNAs. The effect of miR-60a inhibitor was investigated using an engineered dual luciferase reporter system (Crater et al., 2015). It was reported that the silencing effect of miR60a on the reporter system was diminished in a dose–response manner, when an increasing amount (1–20 μg) of miR-60a inhibitor was used. To evaluate the global effect of miR60a inhibitor on *Toxoplasma* infectivity, miR-60a inhibitor (20 μg) was introduced into freshly released extracellular parasites. The parasites electroporated with miR-60a inhibitor exhibited 80% less ability to form plaques than those electroporated with unrelated single stranded RNA of a similar length (Figure 4.1, A and B). The data indicates that the early process of *Toxoplasma* infection

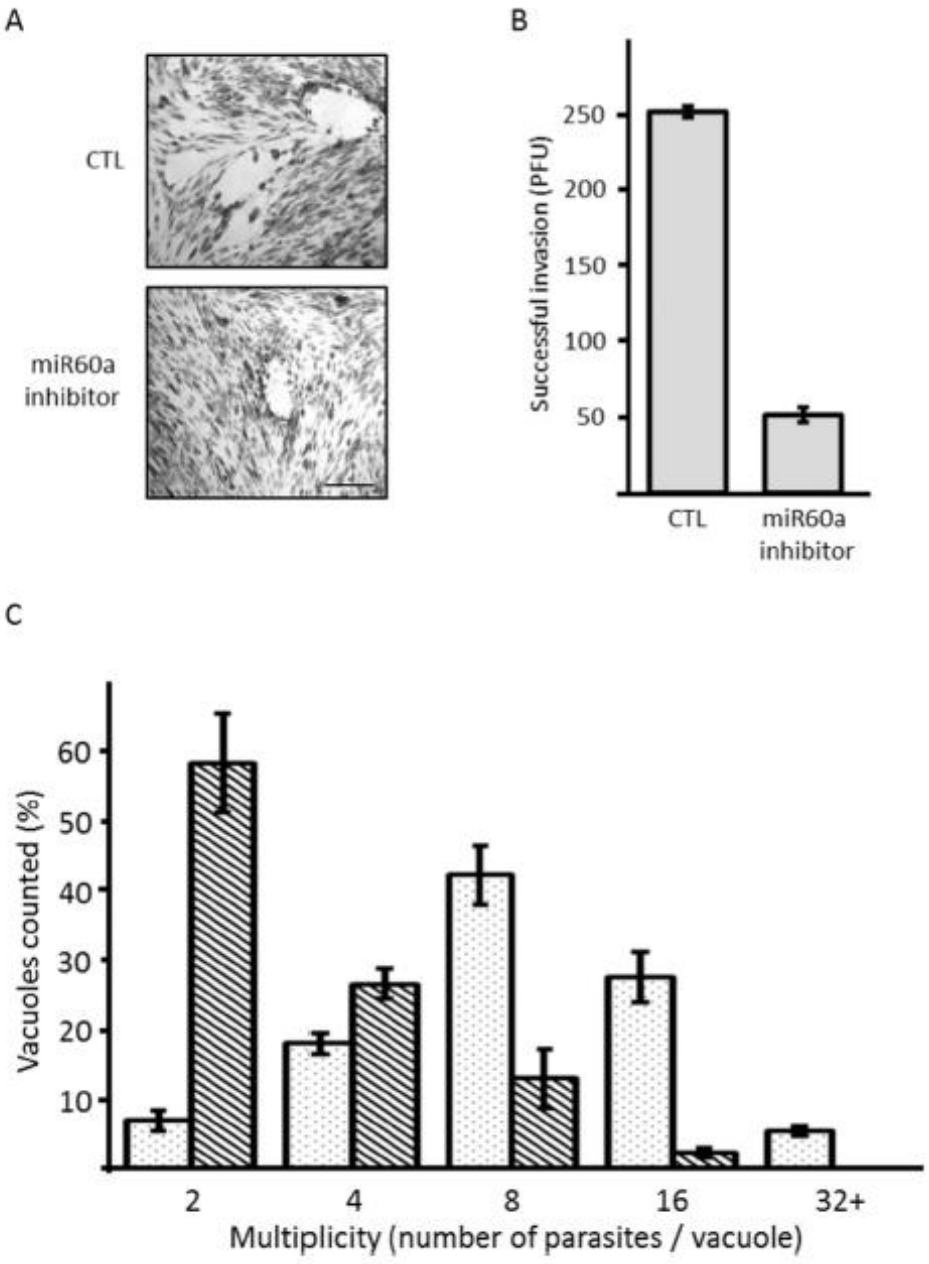


Figure 4.1. Effects of miR60a inhibitor. Freshly lysed parasites (~1 × 10^6) were resuspended in electroporation buffer (400 µl of 120 mM KCl, 0.15 mM $CaCl_2$, 10 mM K_2HPO_4/KH_2PO_4. pH 7.6, 2 mM EDTA, 5 mM $MgCl_2$, 2 mM ATP, 5 mM glutathione in a 4 mm-gap cuvette (Roos et al., 1994). Twenty micrograms of single-stranded RNAs, either miR60a inhibitor or an unrelated RNA with similar size and structure (CTL), were

in vitro synthesized, purified and added to the parasite suspension as described in Crater et al. (2015). The mixtures were then pulsed with 1500 V, 25 Ω, and 25 μF in a BTX ECM 630 (Roos et al., 1994). To allow for the antisense effect to take place, the transfected parasites were grown on confluent monolayers of human fibroblasts (60 × 15 mm) for 48 h. Transfected parasites were released, and approximately 1000 parasites were used to re-infect a new plate (35 × 10 mm) and incubated for 72 h. Plates were then fixed with cold methanol and stained with 0.5% crystal violet blue and examined for the formation of plaques, as an indicator of successful invasion and infection. (A) Representative images of CTL and miR60a inhibitor plaque assays. Images were acquired using a Leica DMIRB with C PLAN 10×/0.22 LMC objective, and Northern Eclipse software. Micrometer bar represents 40 μm. (B) The number of plaques formed were counted from 50 random fields of view, and calculated from three independent experiments. This data indicates that miR60a inhibitor interfered with the infection process. In addition, 200 released parasites were transferred onto confluent monolayers grown on coverslips (22 × 22 mm) and incubated for 24 h. The coverslips were then fixed by 3% paraformaldehyde and stained with Hoechst (10 μM). (C) From three independent experiments, the numbers of parasites in 100 random vacuoles were counted to monitor multiplicity. Hatched bars represent those electroporated with miR60a inhibitor, and dotted bars are for control experiments. This indicates that miR60a inhibitor can also affect the parasites ability to multiply once entering.

was impaired. This impairment could be due to direct or indirect disturbance effect of miR-60a inhibitor on attachment and/or invasion. When the electroporated parasites were allowed to multiply for a limited period (24 h) without reaching their lytic phase, it was detected that the parasite multiplied at a slower rate than those of the control electroporation. Majority (60%) of vacuoles of parasites electroporated with miR-60a inhibitor contained 2 parasites/vacuole (hatched bars, Figure 4.1c), where those of control (40%) carried 8 parasites/vacuole (dotted bars, Figure 4.1c). Considering an average doubling time of the parasite is 6 h (CTL), the presence of miR-60a inhibitor clearly increased the doubling time. This finding confirms the global effect of miR60a inhibitor, which could be direct effects on the predicted 80 target genes (Braun et al., 2010) and/or indirect effects of these genes partaking in various steps of parasite pathogenesis.

4.3.1.3. Antagomir

The term ‘antagomirs’ is used for a specific group of antisense single-stranded RNAs that have a phosphothioate backbone with 3′- cholesterol-conjugation and nucleobases complementary to miRNAs. Antagomirs first appeared and were successfully used by Krützfeldt et al. (2005) to inhibit the effect of miR-122-directed gene regulation in mice. With the success and relatively high stability of antagomirs, these antisense molecules have been extensively exploited for therapeutic uses in various diseases whose pathogenesis and/or prognosis involve miRNAs, such as cancers and cardiovascular diseases (Kaboli et al., 2015; Romaine et al., 2015). Although at the time of this manuscript preparation we did not find any reports on antagomirs in *Toxoplasma*, we present and include it here because antagomirs and miRNA inhibitors are closely

related both structurally and functionally. In addition we anticipate greater use of antagomirs in molecular microbiological and parasitological studies.

4.3.1.4. Morpholino oligomer

Morpholino oligomer is a common name for phosphorodiamidate morpholino oligomer (PMO) which is used in an antisense silencing of a gene expression at translation step. Although the name "PMO" is telling more about its structure, which is composed of nucleobases bound to morpholine rings (instead of deoxyribose rings) linked through phosphorodiamidate groups (instead of phosphodiester linkage), morpholino oligomers is a more widely used name. Because of their chemical constituents, morpholino oligomers are less negatively charged than natural oligonucleotides, while still having high hydrophilicity and ability to form canonical base-pairing with target mRNA. Generally, morpholino oligomers are ~25 nt in length, which is similar to that of other small regulatory RNAs described above, but their binding location on the mRNA is preferably near the translational start site spanning the 5′-UTR. Most morpholino oligomers are used as research tools for reverse genetics by knocking down gene function. It was noted that the ability of this antisense technology to cross multiple cellular membranes is drastically increased when the 3′-terminus of morpholino oligomers is conjugated to transductive peptides (Lai et al., 2012). In the same study, the transductive peptide conjugated morpholino oligomers were shown to successfully block the translation of three genes encoding for *Toxoplasma's* dihydrofolate reductase, enoyl-ACP reductase and apetela 2 domain transcription factor XI-3. Witola et al. (2014) used a similar design of peptide-conjugated morpholino oligomer, but with nucleobases specific to interact with *Toxoplasma* mRNA encoding a major secretory protein, named GRA10

and showed that this antisense technology caused a reduction in growth and development of intracellular *Toxoplasma*.

4.3.1.5. Ribozymes

Ribozymes are catalytically active RNA molecules. Thus far, naturally discovered ribozymes catalyze either the transfer of peptidyl or phosphoryl bond for nucleotide cleaving (Zhang et al., 2010). Those with catalytic ability to perform sequence-specific destruction of phosphoryl bonds of mRNA have been the most exploited for their potential as disease-specific therapeutics tools, and stretched further into our possible personalized medicine. Similarly, in the development of genetic tools and technologies various natural types of ribozymes (i.e. hammerhead, delta ribozymes), and artificial types (i.e. RNA ligase, RNA kinase) have been tested for gene sequence specific activity. Based on their mechanistic reactions, most ribozymes employ a step-wise reaction, whose initial step involves a nucleotide base-pairing or an "antisense" interaction to recognize their substrate whose final step yields destruction of their targets, for the purpose of knock-down or silencing of gene expression. The antisense pairing or substrate recognition often requires a short span of nucleobases ranging from 7 to ~20 nucleotides (Table 4.1.). On this basis, we summarize the application of ribozymes in *Toxoplasma* under antisense technology related to small ncRNAs. Sheng et al. (2004) successfully used engineered delta ribozymes (antisense or substrate recognition sequence of 7 nt in length) to simultaneously down regulate the expression of uracil phosphoribosyltransferase (UPRT) and hypoxanthine guanine phosphoribosyltransferase (HXGPRT). These trans-acting ribozymes were engineered to express independently from the target gene, and multiple targets can be silenced simultaneously. More

elegantly, a self-cleaving hammerhead ribozyme, called Sm1, was engineered to perform specific nucleolytic activity in an inducible manner (Agop-Nersesian et al., 2008). In the study, toyocamycin-responsible hammerhead ribozymes were shown to control the expression of adenosine kinase, when the sequence of the ribozymes was placed onto the upstream sequence within the transcriptional units, called “cis-acting”. The self-cleaving or cis-acting riboyzmes, such as those of toyocamycin-responsive hammerhead ribozymes, have a possible drawback in that the sequence encoding the ribozyme has to be integrated into transcription unit of the gene of interest. As a consequence, it requires effective knock-in genetic manipulation. However, it also offers the advantage of maintaining the level of natural expression as the transcriptional unit is expressed by its own or endogenous promoter.

It is important to note that although the genome of *Toxoplasma* contains no sequence encoding hammerhead-like or delta-like ribozymes, both ribozymes can effectively execute their desirable action of nucleolytic activity in a sequence specific manner. Thus it is of importance to point out that *Toxoplasma* actually has a naturally found ribozyme, called RnaseP (Mack et al., 1994). RnaseP has been used in other system and organisms to perform effective nucleolytic activity (Kim and Liu, 2007). The presence of natural RnaseP activity in *Toxoplasma* should be explored for the development of genetic technology.

4.3.1.6. Small guide RNAs

In recent years, one of the most talked about regulatory RNA-derived technologies is that of the guide RNA of the CRISPR (clustered regulatory interspaced short palindromic repeats) system, originally identified as an adaptive immune system

found in bacteria. Since numerous reviews and commercial sites discuss the use of the CRISPR system, it is thus more appropriate here to limit our scope to those used in *Toxoplasma* to serve our aims in comparing various antisense technologies. A simplified version of the system involves an engineered plasmid expressing (i) a nuclease-directing guide RNA, and (ii) Cas9 nuclease (Sidik et al., 2014). The nuclease-directing guide RNA contains nuclease bases (i) acting as an antisense recognition sequence to pair with genomic DNA target, and (ii) directing the Cas9 nuclease to the desired location, which subsequently causes double stranded DNA breaks. Sidik et al. (2014) showed that the CRISPR/Cas9 system can be used to induce homologous recombination in the parasite to, for example, introduce point mutations without the need for antibiotic selection. The study also indicated that a non-homologous end-joining (NHEJ) pathway is required for the prevention of chromosomal instability and to achieve the best efficiency of CRISPR/Cas9 genetic modulation system. The system was shown to cause 100-fold increase in gene disruption, in comparison to traditional methods used via homologous recombination (Sidik et al., 2014). A similar study performed by Shen et al. (2014), showed that CRISPR/Cas9 not only disrupts genes by insertion but is also capable of deleting full genomic regions using homologous recombination. Shen et al. (2014) noted that the CRISPR/Cas9 system can be used to insert selectable markers without the need for homology arms. These reports show that gene disruption is most effective in wild-type strains of *Toxoplasma* because of the functional NHEJ; however, genome editing in processes such as epitope tagging has superior success in ΔKU80 strains that lack NHEJ capability and therefore rely on homologous recombination (Sidik et al., 2014). The CRISPR/Cas9 system has many promising applications in *Toxoplasma* as well as a wide

variety of other organisms (Shen et al., 2014); however, there is still room for improvement in efficiency and off targeting (Sander and Joung, 2014) which will likely be addressed as it continues to come to the forefront of genome editing research.

4.3.2. Long non-coding RNAs

Not only because of their sizes, are lncRNAs constituted as an interesting class of regulatory RNAs. lncRNAs, not mainly depending on sequence or structure, can function both by sequence complementarity to bind other nucleic acids and/or by structure to form molecular scaffolds for assembly of macromolecular complexes to result in either repression or activation of gene expression (Yoon et al., 2013; Krishnan and Mishra, 2014). When the genome of *Toxoplasma* was surveyed using SAGE methodologies, more than 50% of the *Toxoplasma* genome was shown to transcribe during various stages of parasite development (Radke et al., 2005). Among polyadenylated transcripts analyzed by SAGE tags, 21.5% of total SAGE tags observed were annotated as antisense polyadenylated transcripts. Based on the average density of SAGE tags detected throughout 14 chromosomes of *Toxoplasma*, an average antisense polyadenylated transcripts is ~4000 nucleotides and derived from the opposite strands of the transcripts. It is interesting to note that the expression of antisense transcripts are inversely proportional to those of the sense transcripts and there are a number of intergenic regions (~ 2000 base-pairs in length) to be explored (Radke et al., 2005). Hassan et al. (2012) using high throughput RNA-sequencing data analysis showed that the *Toxoplasma* genome has a number of ncRNAs. Ramaprasad et al. (2015) used a manual evaluation of strand-specific RNA sequencing and shotgun proteomics to identify cis-natural antisense

transcripts and long intergenic ncRNAs. Also, in this study a set of long antisense polyadenylated transcripts and sense coding transcript were reported, and the reverse correlation between the levels of the antisense noncoding transcripts and coding (sense) RNAs reported by Radke et al. (2005) was confirmed. However the length of long antisense polyadenylated transcripts is reported to be ~1300 nucleotides (Ramaprasad et al., 2015), instead of ~2000. The presence of cis-natural antisense RNA and other forms of long ncRNAs are indicative of the existence of a possible mode of gene regulation via long ncRNAs in *Toxoplasma*.

4.3.2.1. Long antisense RNAs

The first report on utilizing long antisense RNA to regulate gene expression in *Toxoplasma* was by Nakaar et al. (1999). In the study, antisense RNA (~2 kb) corresponding to the coding sequence of nucleoside triphosphate hydrolase (NTPase) was expressed under the control of its own promoter to yield antisense transcript, whose 3′-UTR was tailored by a cis-acting hammerhead ribozyme to ensure its nuclear retention. Resulting antisense NTPase transcripts subsequently pair with the mRNA with perfect matches to cause mRNA degradation in the nucleus. The silencing of NTPase was detected by immunoblotting to show the diminished level of target gene product. The same strategy was used to down regulate selectable markers, called HXGRPT and CAT (Nakaar et al., 2000). Both studies not only showed the silencing effect of nuclear antisense RNA, but also showed that catalytically active hammerhead ribozymes are able to silence the target without relying on endogenous machinery. Without a compartment-directed expression, the over-expression of antisense RNA transcripts against the gene

encoding uracil phosphoribosyltransferase (UPRT) was shown to effectively silence the gene expression (Al-Anouti and Ananvoranich, 2002). Although these studies did not offer an insight into the silencing mechanism, the reported silencing effect strongly suggested that the basic mechanism of antisense-directed gene silencing must be conserved between *Toxoplasma* and other eukaryotic organisms.

4.3.2.2. Long double-stranded RNA

Since the discovery of RNA interference mechanism, it is very well accepted that siRNA and long double-stranded RNA can be used in silencing gene expression. LaCount et al. (2000) showed that the overexpression of a transcript with promoters arranged in a head-to-head fashion causes a gene-specific knockdown of gene expression in *Trypanosoma brucei*. Al-Anouti and Ananvoranich (2002) adopted a similar strategy for *Toxoplasma*. Despite the lack of mechanistic evidence to show the presence of functional RNA interference in *Toxoplasma*, a similar expression system was used to down regulate a number of genes, including marker, metabolic and growth-related essential genes (Al-Anouti and Ananvoranich, 2002; Al-Anouti et al., 2003, 2004; Adams et al., 2005; Al Riyahi et al., 2006; Ananvoranich et al., 2006; Yu et al., 2009 and Holmes et al., 2010). It was noted that the silencing effect of double-stranded RNAs is better than that of antisense RNAs (Al-Anouti and Ananvoranich, 2002). The superior effect of double-stranded RNA could be due to (i) a higher stability of double stranded RNA and/or (ii) fewer processes involved in the double stranded RNA induced gene silencing.

4.4. Future of antisense technologies

This review summarizes a limited set of our ever-expanding toolbox for *Toxoplasma*, based on regulatory ncRNAs. With that in mind, using conventional genetic tools and the various mix-and-match technologies, antisense technologies could yield an array of valuable information. Two examples should be mentioned to encourage future development of genetic tools via mix-and-match of known methods. By taking advantage of homologous recombination system offered by ΔKU80 strains that lack NHEJ capability, the binding sites of miR60a can be placed onto the 3'UTR of the gene of interest in *Toxoplasma* (Crater et al., 2015). The resultant system was used to direct the endogenous miR-60a-directed machinery to silence *Toxoplasma* genes encoding for DEAD-box RNA helicase, and lactate dehydrogenase. In a more complex system, endogenous tagging with DiCre-mediated DNA recombination system can be used to insert U1 recognition sites at the last exon of the gene of interest to induce pre-mRNA degradation (Pieperhoff et al., 2015). As a concluding remark, it is definitely an exciting time to witness the expansion of genetic tools and technologies. More effective technologies will be developed and realized using a combination of approaches and methodologies.

4.5. Acknowledgments

This work was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) .

4.6. References

Adams, B., Musiyenko, A., Kumar, R., Barik, S., 2005. A novel class of dual-family immunophilins. J. Biol. Chem. 280 (26), 24308–24314.

Agop-Nersesian, C., Pfahler, J., Lanzer, M., Meissner, M., 2008. Functional expression of ribozymes in Apicomplexa: towards exogenous control of gene expression by inducible RNA-cleavage. Int. J. Parasitol. 38 (6), 673–681.

Al Riyahi, A., Al-Anouti, F., Al-Rayes, M., Ananvoranich, S., 2006. Single argonaute protein from *Toxoplasma gondii* is involved in the double-stranded RNA induced gene silencing. Int. J. Parasitol. 36 (9), 1003–1014.

Al-Anouti, F., Ananvoranich, S., 2002. Comparative analysis of antisense RNA, doublestranded RNA, and delta ribozyme-mediated gene regulation in *Toxoplasma gondii*. Antisense Nucleic Acid Drug Dev. 12 (4), 275–281.

Al-Anouti, F., Quach, T., Ananvoranich, S., 2003. Double-stranded RNA can mediate the suppression of uracil phosphoribosyltransferase expression in *Toxoplasma gondii*. Biochem. Biophys. Res. Commun. 302 (2), 316–323.

Al-Anouti, F., Tomavo, S., Parmley, S., Ananvoranich, S., 2004. The expression of lactate dehydrogenase is important for the cell cycle of *Toxoplasma gondii*. J. Biol. Chem. 279 (50), 52300–52311.

Ananvoranich, S., Al Rayes, M., Al Riyahi, A., Wang, X., 2006. RNA silencing of glycolysis pathway in *Toxoplasma gondii*. J. Eukaryot. Microbiol. 53 (Suppl. 1), S162–S163.

Braun, L., Cannella, D., Ortet, P., Barakat, M., Sautel, C.F., Kieffer, S., Garin, J., Bastien, O., Voinnet, O., Hakimi, M.A., 2010. A complex small RNA repertoire is generated by a plant/fungal-like machinery and effected by a metazoan-like Argonaute in the single-cell human parasite *Toxoplasma gondii*. PloS Pathog. 6 (5), e1000920.

Bunch, T.A., Goldstein, L.S.B., 1989. The conditional inhibition of gene expression in cultured *Drosophila* cells by antisense RNA. Nucleic Acids Res. 17, 9761–9782.

Coleman, J., Hirashima, A., Inokuchi, Y., Green, P.J., Inouy, M., 1985. A novel immune system against bacteriophage infection using complementary RNA (micRNA). Nature 315, 601–603.

Crater, A.K., Cherry, A., Holmes, M., Kadri, D., Ananvoranich, S., 2012. Evaluation of the ability of short and long double-stranded RNAs to induce homologous gene silencing in the protozoan parasite, *Toxoplasma gondii*. Am. J. Biomed. Sci. 4, 1–13.

Crater, A.K., Manni, E., Ananvoranich, S., 2015. Utilization of inherent miRNAs in functional analyses of *Toxoplasma gondii* genes. J. Microbiol. Methods 108, 92–102.

Green, P.J., Pines, O., Inouye, M., 1986. The role of antisense RNA in gene regulation. Annu. Rev. Biochem. 55, 569–597.

Hassan, M.A., Melo, M.B., Haas, B., Jensen, K.D., Saeij, J.P., 2012. De novo reconstruction of the *Toxoplasma gondii* transcriptome improves on the current genome annotation and reveals alternatively spliced transcripts and putative long non-coding RNAs. BMC Genomics 13, 696.

Holmes, M., Liwak, U., Pricop, I., Wang, X., Tomavo, S., Ananvoranich, S., 2010. Silencing of tachyzoite enolase 2 alters nuclear targeting of bradyzoite enolase 1 in *Toxoplasma gondii*. Microbes Infect. 12 (1), 19–27.

Izant, J.G., Weintraub, H., 1984. Inhibition of thymidine kinase gene expression by anti-sense RNA: a molecular approach to genetic analysis. Cell 36, 1007–1015.

Kaboli, P.J., Rahmat, A., Ismail, P., Ling, K.H., 2015. MicroRNA-based therapy and breast cancer: a comprehensive review of novel therapeutic strategies from diagnosis to treatment. Pharmacol. Res. 97, 104–121.

Kim, K., Liu, F., 2007. Inhibition of gene expression in human cells using RnaseP-derived ribozymes and external guide sequences. Biochim. Biophys. Acta 1769 (11–12), 603–612.

Kim, K., Weiss, L.M., 2004. *Toxoplasma gondii*: the model apicomplexan. Int. J. Parasitol. 34 (3), 423–432.

Kim, K., Weiss, L.M., 2008. *Toxoplasma*: the next 100 years. Microbes Infect. 10 (9), 978–984. Krishnan, J., Mishra, R.K., 2014. Emerging trends of long non-coding RNAs in gene activation. FEBS J. 281 (1), 34–45.

Krützfeldt, J., Rajewsky, N., Braich, R., Rajeev, K.G., Tuschl, T., Manoharan, M., Stoffel, M., 2005. Silencing of microRNAs in vivo with ‘antagomirs’. Nature 438 (7068), 685–689.

LaCount, D.J., Bruse, S., Hill, K.L., Donelson, J.E., 2000. Double-stranded RNA interference in *Trypanosoma brucei* using head-to-head promoters. Mol. Biochem. Parasitol. 111 (1), 67–76.

Lai, B.S., Witola, W.H., El Bissati, K., Zhou, Y., Mui, E., Fomovska, A., McLeod, R., 2012. Molecular target validation, antimicrobial delivery, and potential treatment of *Toxoplasma gondii* infections. Proc. Natl. Acad. Sci. U. S. A. 109 (35), 14182–14187.

Mack, D., McLeod, R., Stark, B., 1994. Characterization of ribonuclease P from *Toxoplasma gondii*. J. Eukaryot. Microbiol. 41 (5), 13S. Miller, A.D., 2014. Delivering the promise of small ncRNA therapeutics. Ther. Deliv. 5 (5), 569–589.

Musiyenko, A., Majumdar, T., Andrews, J., Adams, B., Barik, S., 2012. PRMT1 methylates the single Argonaute of *Toxoplasma gondii* and is important for the recruitment of Tudor nuclease for target RNA cleavage by antisense guide RNA. Cell. Microbiol. 14 (6), 882–901.

Nakaar, V., Samuel, B.U., Ngo, E.O., Joiner, K.A., 1999. Targeted reduction of nucleoside triphosphate hydrolase by antisense RNA inhibits *Toxoplasma gondii* proliferation. J. Biol. Chem. 274 (8), 5083–5087.

Nakaar, V., Ngo, E.O., Joiner, K.A., 2000. Selection based on the expression of antisense hypoxanthine–xanthine–guanine-phosphoribosyltransferase RNA in *Toxoplasma gondii*. Mol. Biochem. Parasitol. 110 (1), 43–51.

Pieperhoff, M.S., Pall, G.S., Jiménez-Ruiz, E., Das, S., Melatti, C., Gow, M., Wong, E.H., Heng, J., Müller, S., Blackman, M.J., Meissner, M., 2015. Conditional U1 gene silencing in *Toxoplasma gondii*. PloS One 10 (6), e0130356.

Radke, J.R., Behnke, M.S., Mackey, A.J., Radke, J.B., Roos, D.S., White, M.W., 2005. The transcriptome of *Toxoplasma gondii*. BMC Biol. 3, 26.

Ramaprasad, A., Mourier, T., Naeem, R., Malas, T.B., Moussa, E., Panigrahi, A., Vermont, S.J., Otto, T.D., Wastling, J., Pain, A., 2015. Comprehensive evaluation of *Toxoplasma gondii* VEG and Neospora caninum LIV genomes with tachyzoite stage transcriptome and proteome defines novel transcript features. PloS One 10 (4), e0124473.

Romaine, S.P., Tomaszewski, M., Condorelli, G., Samani, N.J., 2015. MicroRNAs in cardiovascular disease: an introduction for clinicians. Heart 101 (12), 921–928.

Roos, D.S., Donald, R.G., Morrissette, N.S., Moulton, A.L., 1994. Molecular tools for genetic dissection of the protozoan parasite *Toxoplasma gondii*. Methods Cell Biol. 45, 27–63.

Sander, J., Joung, J., 2014. CRISPR-Cas systems for editing, regulating and targeting genomes. Nat. Biotechnol. 32 (4), 347–355.

Shen, B., Brown, K.M., Lee, T.D., Sibley, L.D., 2014. Efficient gene disruption in diverse strains of *Toxoplasma gondii* using CRISPR/CAS9. Mbio 5 (3), e01114.

Sheng, J., Al-Anouti, F., Ananvoranich, S., 2004. Engineered delta ribozymes can simultaneously knock down the expression of the genes encoding uracil

phosphoribosyltransferase and hypoxanthine–xanthine–guanine phosphoribosyltransferase in *Toxoplasma gondii*. Int. J. Parasitol. 34 (3), 253–263.

Sidik, S.M., Hackett, C.G., Tran, F., Westwood, N.J., Lourido, S., 2014. Efficient genome engineering of *Toxoplasma gondii* using CRISPR/Cas9. PloS One 9 (6), e100450.

Spiegelman, W.G., Reichardt, L.F., Yaniv, M., Heinemann, S.F., Kaiser, A.D., Eisen, H., 1972. Bidirectional transcription and the regulation of phage λ repressor synthesis. Proc. Natl. Acad. Sci. U. S. A. 69, 3156–3160.

Vanhee-Brossollet, C., Vaquero, C., 1998. Do natural antisense transcripts make sense in eukaryotes? Gene 211, 1–9.

Wahlestedt, C., 2013. Targeting long non-coding RNA to therapeutically upregulate gene expression. Nat. Rev. Drug Discov. 12 (6), 433–446.

Wang, J., Liu, X., Jia, B., Lu, H., Peng, S., Piao, X., Hou, N., Cai, P., Yin, J., Jiang, N., Chen, Q., 2012. A comparative study of small RNAs in *Toxoplasma gondii* of distinct genotypes. Parasites Vectors 5, 186.

Witola, W.H., Bauman, B., McHugh, M., Matthews, K., 2014. Silencing of GRA10 protein expression inhibits *Toxoplasma gondii* intracellular growth and development. Parasitol. Int. 63 (5), 651–658.

Xu, M.J., Zhou, D.H., Huang, S.Y., Zhao, F.R., Nisbet, A.J., Lin, R.Q., Song, H.Q., Zhu, X.Q., 2013. Comparative characterization of microRNA profiles of different genotypes of *Toxoplasma gondii*. Parasitology 140 (9), 1111–1118.

Yoon, J.H., Abdelmohsen, K., Gorospe, M., 2013. Post-transcriptional gene regulation by long noncoding RNA. J. Mol. Biol. 425 (19), 3723–3730.

Yu, L., Gao, Y.F., Qiao, Z.P., Li, C.L., Li, X., Shen, J.L., 2008. *Toxoplasma gondii*: siRNA can mediate the suppression of adenosine kinase expression. Exp. Parasitol. 118 (1), 96–102.

Yu, L., Gao, Y.F., Li, X., Qiao, Z.P., Shen, J.L., 2009. Double-stranded RNA specific to adenosine kinase and hypoxanthine–xanthine–guanine-phosphoribosyltransferase retards growth of *Toxoplasma gondii*. Parasitol. Res. 104 (2), 377–383.

Zamecnik, P.C., Stephenson, M.L., 1978. Inhibition of Rous sarcoma virus replication and cell transformation by a specific oligodeoxynucleotide. Proc. Natl. Acad. Sci. U. S. A. 75 (1), 280–284.

Zhang, J., Lau, M.W., Ferré-D'Amaré, A.R., 2010. Ribozymes and riboswitches: modulation of RNA function by small molecules. Biochemistry 49 (43), 9123–9131.

CHAPTER 5

Ubiquitin-like protease 1 (Ulp1) as a model for the study of miRNA-induced gene regulation in *Toxoplasma gondii*

5.1. Introduction

Toxoplasma gondii is an intracellular parasitic protozoan and a member of the phylum Apicomplexa whose members also include important pathogens such as *Plasmodium*, the cause of malaria; *Cryptosporidium*, a water borne parasite; and *Neospora* and *Eimeria*, important zoological parasites (Kim and Weiss, 2004). *T. gondii* is able to infect any warm-blooded mammal, and has been reported to infect up to one third of the world's population. The life cycle of *Toxoplasma* consists of two stages; a sexual reproduction stage that occurs only in the gut of its feline definitive hosts, and an asexual reproduction stage in other warm-blooded hosts. The asexual stage contains two distinct forms: a rapidly growing form called tachyzoite, and a latent cyst encapsulated form called bradyzoite. Although the tachyzoite form is easily removed by the host's immune system, the parasites are able to convert to the bradyzoite form and evade the host's immune response. When the host's immune system becomes weakened, the bradyzoites are able to reconvert to their rapidly growing tachyzoite form.

One of our research focus is to study a mechanism of post-transcriptional gene regulation directed in *Toxoplasma* by small interfering RNAs, including siRNAs and miRNAs. We showed that the abundance of Tg-miRNAs and the number of binding sites within a transcript can influence the levels of silencing effects (Crater et al., 2015). The presence of miRNA mimics and inhibitors altered the silencing effects. To gain a better understanding of miRNA roles in the regulation of gene expression and parasite cellular

biology, we endeavored to establish a model gene to facilitate the study. It was shown previously that *Toxoplasma* small interfering RNAs were generated by plant/fungal-like machinery (Braun et al, 2010). In *Toxoplasma* type I strain, the most abundant miRNAs are members of Tg-miR-60 family, and predicted to have approximately 80 putative target genes. Using this information as our starting point, we used the seed sequence (**CACAGT**, 6 nts) of the Tg-miR-60a guide strand (A**CACAGT**CGGTACGAAATCCATACT, 25 nts) as a query in a blast search. Seventy-five genes were identified to have at least one location with six consecutive nucleotides to base-pair with the seed sequence. To ensure strong interactions between Tg-miR-60a and the transcripts, we filtered out any genes with less than 18 nucleotides complementarity, including those to the seed sequence. The transcripts of eight genes were found to have at least one location with at least 18 nucleotides complementary to Tg-miR-60a guide strand. Among these genes, three genes were annotated to have a putative function in the *Toxoplasma* database (Gajria et al, 2008). Sorting based on the number of probable Tg-miR-60a binding sites, we selected the locus containing a gene with the highest number of probable Tg-miR-60a binding sites. Seven probable sites were identified in a locus encoding ubiquitin-like protease 1 (TgUlp1, TGGT1_214470); four sites locate on the non-coding strand, and three sites on the coding strand of the gene encoding TgUlp1 (Figure 2a, Table 5.1.). This identification prompted us to investigate and explore whether TgUlp1 could offer a model for the study of miRNA-induced gene silencing in *Toxoplasma*.

Prototypic Ulp1, was first identified in baker's yeast by Li and Hochstrasser (1999) where it was found to have an essential role in the G2/M phase progression of the

Table 5.1. Predicted miR-60a binding sites on TgUlp1-pre-mRNA based on plant miRNA binding. Predicted base pairing is indicated by *. Seed sequence is highlighted in orange and target sequence is red.

Predicted base pairing	Location on pre-mRNA
5′ GCGAACGAAACGAAAAUGACUGUGA 3′ pre-mRNA * ** * ******** 3′ TCATACCTAAAGCATGGCTGACACA 5′ miR60a	Intron 2
5′ UUCGCCUGCCUCUUUGCUACUGUGG 3′ pre-mRNA * * ****** 3′ TCATACCTAAAGCATGGCTGACACA 5′ miR60a	Intron 2
5′ UGUCUCGACUUCGUACCGACUGUGU 3′ pre-mRNA ** * ** **** ********* 3′ TCATACCTAAAGCATGGCTGACACA 5′ miR60a	Intron 7

cell cycle (Li and Hochstrasser, 1999). *Ulp1* is an essential gene in eukaryotes, which encodes an enzyme involved in a post-translational modification pathway called SUMOylation. The pathway of SUMOylation is involved in the reversible addition of a small ubiquitin-like modifier (SUMO) peptide (90-110 amino acids) to a protein substrate. Unlike ubiquitination, where target proteins are polyubiquitinylated and sent to lysosomes for degradation, SUMOylation of target proteins aids in the protein's localization, stabilization and activity (Hay, 2005; Komander and Rape, 2012). SUMOylation plays an important role in DNA repair, transcription, RNA processing, nuclear-cytoplasmic transport, stress response and progression through the cell cycle (Hay, 2005). The pathway is initiated by Ulp1 in the activation of a SUMO precursor. Ulp1 is a cysteine peptidase that is highly specific for the SUMO peptide, as it recognizes the tertiary structure of SUMO rather than an amino acid sequence (Mossessova and Lima, 2000). Ulp1 cleaves amino acid residues from the C-terminal end of the SUMO precursor to reveal a di-glycine motif that is required for attachment to target proteins. The hydrolysis of a peptide bond by Ulp1 involves the thiol group of a cysteine residue as a nucleophile, a histidine, and an aspartic acid to assist the reaction. The mature form of SUMO with its revealed di-glycine is subsequently transferred by E1 and E2 enzymes and ligated to its protein substrate by an E3 enzyme which recognizes the ψKxE consensus sequence on the substrate (where ψ is a large hydrophobic residue and x is any residue). The catalytic activity of Ulp1 is also responsible for the removal of SUMO peptide from SUMOylated proteins (Watts 2013).

In *Toxoplasma*, a study was undertaken to explore the biological significance of SUMOylation (Braun et al., 2009). The cDNA coding for the SUMO peptide was

identified (TGME49_266460). Since a knockout or over-expression of TgSUMO was found to be lethal, a N-terminal HA-Flag tagged version was generated under the control of its endogenous promoter. Affinity purification of HA-Flag-TgSUMO expressed in tachyzoites coupled to mass spectrometry identified 120 putative SUMOylated proteins. The majority of identified SUMOylated proteins (~27 %) were found to be involved in protein translation and folding. Other identified proteins were involved in (i) metabolism and post-translational modification and (ii) degradation, which makes up 19 and 11 % of the total identified proteins, respectively (Braun et al., 2009). Using an anti-HA antisera, HA-Flag-TgSUMO was found to localize to the nucleus of transgenic parasites. They next developed a polyclonal antibody raised against a recombinant TgSUMO protein expressed from bacteria. To confirm the specificity of the new antibody for TgSUMO, the HA antisera signal was compared to the TgSUMO antisera signal for localization and western blot analysis. Since the anti-TgSUMO antisera could recognizes free SUMO and SUMOylated proteins, localization during the parasite lytic cycle was analyzed. Upon contact with a host cell during invasion, TgSUMO was found on the parasite membrane at the point of contact. As the parasite entered the host cell, TgSUMO staining migrated down the length of the parasite at the moving junction. Once the parasite successfully invaded, the TgSUMO signal was found within the parasite's nucleus. Under bradyzoite conditions a TgSUMO signal was found on the parasitophorous vacuole membrane (PVM). These studies suggest SUMOylation in *T. gondii* is a dynamic processes required for parasite invasion and bradyzoite conversion. SUMOylation orthologs, Ulp1, E1, E2 and E3 have been identified yet their functions remain unknown (Braun et al., 2009).

5.2. Materials and Methods

5.2.1. Cell cultures

5.2.1.1. Mammalian cell cultures

Human foreskin fibroblasts (HFF; ATCC, #1041) were grown in Dulbecco's Modified Eagle Media with high D-glucose and L-glutamine (DMEM; Invitrogen, #12100046) supplemented with 10% v/v Cosmic Calf Serum (Hyclone, #H3008704N) and 0.5x Antibiotic–Antimycotic (Invitrogen, #15240-062). Cultures were maintained at 37 °C in 5% CO_2.

5.2.1.2. *Toxoplasma gondii* culture

RHΔHX, T. *gondii* type I strain, was obtained from NIH AIDS Research and Reference Reagent Program (#2857); Δku80 strain, which is an RHΔHX-derived strain, was obtained from Dr. Vern B. Carruthers (University of Michigan). Unless stated otherwise, both strains were maintained in HFF using Minimum Essential Medium (MEM, Invitrogen, #61100061) supplemented with 1% v/v dialyzed fetal bovine serum (ThermoFisher Scientific, Hyclone, #SH3007903) and diluted (0.5x) Antibiotic-Antimycotic and maintained at 37 °C in 5% CO_2. Selection of transgenic parasites was grown in culture media containing either, mycophenolic acid (25 μg/mL) and xanthine (50 μg/mL) or pyrimethamine (1μM), for hypoxanthine-xanthine-guanine phosphoribosyltransferase (HXGPRT) or dihydrofolate reductase (DHFR) resistance, respectively.

5.2.1.3. *Saccharomyces cerevisiae*

Wild type strain, MHY500 (MATa his3-Δ200 leu2-3, 112 ura3-52, lys2-80, 1trp1-1, gal2) and the Δulp1 shuffle strain, YOK 40 (ulp1::HIS3, ULP1/URA3) gifts from Dr.

Oliver Kerscher, (The College of William and Mary, Williamsburg, Virginia) were grown under their nutrient requirements on YEPD plates.

5.2.2. Plasmid construction

TgUlp1-SF transforming plasmid was constructed using pSF-TAP-LIC-HX (obtained from Dr. Vern B. Carruthers, University of Michigan). The plasmid was linearized by PacI digestion prior to the treatment with T4 DNA polymerase to create compatible ends in the presence of dGTP for a ligation-independent cloning procedure as described by Huynh and Carruthers (2009). TgUlp1 (1298 bp) amplicon was prepared using gene specific primer pair of TgUlp1LICFw/TgUlp1LICRv (Table 5.2.), and purified prior to the treatment with T4 DNA polymerase to create compatible ends in the presence of dCTP. Equimolar concentrations of the prepared plasmid and insert were allowed to anneal and subsequently used in bacterial transformations. The resultant plasmid, pUlp1-SF-TAP-LIC-Hx was analyzed to confirm its sequence via restriction endonuclease digestions. The plasmid was then used in the transformation of *T. gondii* (ΔKu80). Transgenic parasites were selected and grown in culturing media containing mycophenolic acid (25 μg/mL) and xanthine (50 μg/mL).

CRISPR/Cas9 constructs for TgUlp1 knock out were generated by PCR reaction using the pU6-universal parental plasmid (obtained from S. Lourido, Whitehead Institute for Biomedical Research, Cambridge, MA) as a template. Opposing primers, gRNAUlp1NsiI and U3upstrem and gRNAUlp1XhoI and U3upstrem were used to generate gRNAUlp1_1 and gRNAUlp1_2 plasmids (Table 5.1). PCR products were gel purified using QiaexII Gel Extraction kit (Qiagen #20021) according to the manufacturer protocol. Purified PCR products were then phosphorylated using T4 PNK (New

Table 5.2. Oligo nucleotide primers used in the study.

Primer name	Sequence
Plasmid Construction and verification	
TgSUMO-GFP	
XbaIRBS5Sumo	AATCTAGAAGGAGATATACCATGTCGGACGACAAGAAG
EcoRI3Sumo	AAGAATTCTGCGCCCCCCGTCTGTT
EcoRISer5eGFP	AAGAATTCTCCATGGTGAGCAAGGGCGAG
HindIII3eGFP	AAAAGCTTCGAACTCCAGCAGGACCA
LIC constructs	
TgUlp1LICFw (b)	TACTTCCAATCCAATTTAATgcacgaacaaaacctggttc
TgUlp1LICRv (b')	TCCTCCACTTCCAATTTTAGCgtcgagttcgccgccgacga
FlagRv (e)	TGCGGATGAGACCAAGAACC
GST-TgUlp1	
GST_Ulp1_FwCf	GATCTGGTTCCGCGTggatccATGATGTATGTTGACCTC
Ulp1EcoR1_RvCf	CTCGAGTCGACCCGGgaattcTCAGGGGTCGAGTTCGCCGC
Ulp1RvRT	TCAGGGGTCGAGTTCGCCGC
TgUlp1 yeast construction	
ScUlp1Fw	GGTACCGGGCCCCCCCTCGAGACACATATATTACACAAGTA
ScUlp1Rv	GTCAACATACATCATATGCATTATTATGATATCCAAAGGTC
TgUlp1Rv	TCCTTTGCTAGCCATGCGGCCGCGGTCGAGTTCGCCGCCGA
TgUlp1Fw	TTGGATATCATAATAATGCATATGATGTATGTTGACCTCCC
CRISPR	
gRNAUlp1NsiI	ggacagctcgagtgtctgtaGTTTTAGAGCTAGAAATAGCAAG
gRNAUlp1XhoI	ggacagctcgagtgtctgtaGTTTTAGAGCTAGAAATAGCAAG
TgUlp1DHFRFw	gcgggacagctcgagtgtctGCTTCGCCAGGCTGTAAATC
TgUlp1DHFRRv	aaaatcaaagcttcgtccgtCAGGAATTCATCCTGCAAGTG
U3upstrem	aacttgacatccccatttacc
RT-PCR	
qGAPDH_Fw	GGTGTTCCGTGCTGCGAT
qGAPDH_Rv	GCCTTTCCGCCGACAAT
qUlp1_Fw (a)	ACGCGCTCTGGGACAAAA
qUlp1_Rv (a')	GAACAGAGGCGAAGTCGTAGGT
Exon6Fw (c)	AGCTCGAGTGTCTGTACGGCT
UPRTFw	GGTCACTGCTGCTGTTGACATC
UPRTRv	TTTTCCAGTCCGCGATTC
***In vitro* transcription**	
T7promoterGG	TAATACGACTCATATAGG
antiMIR60a_tpl	ACACAGTCGGTACGAAATCCATACTccTATATGAGTCGTATTA
antiMIR4a_tpl	ATGTTTGCTTGGAAGCTGTAGTCATTccTATATGAGTCGTATTA
Ulp1intron7FwT7	TAATACGACTCACTATAGGacgcgctctgggacaaaa

England BioLabs #B0201S) according to the manufacturer protocol. Subsequent reactions were then ligated by a T4 DNA ligase (Fermentas #02076201) and analyzed by NsiI and XhoI restriction endonuclease reactions, respectively. A DHFR (dihydrofolate reductase) selection cassette was prepared by a PCR reaction with primers TgUlp1DHFRFw and TgUlp1DHFRRv that contained 20-nucleotide homology to TgUlp1 genomic sequence where Cas9 cleavage would occur. The resultant PCR product was extracted with 1 volume of phonol:chloroform then ethanol precipitated. Parasites were transformed with 50 μg of each CRISPR/Cas9 plasmid, gRNAUlp1_1 and gRNAUlp1_2, and 20 μg of purified DHFR selection insert. Parasites were cultured in Ed1 media containing 1 μM pyrimethamine.

For expression of GST-Ulp1 in bacteria, the pGEX-4T-1 plasmid was prepared by restriction endonuclease reaction with BamHI and EcoR1 restriction endonucleases. For generation of the 2355 bp TgUlp1 amplicon, cDNA which was generated by SuperScriptIII reverse transcriptase (Invitrogen #18080-093) with 250 ng of total RNA and 2 pmol gene specific primer Ulp1RvRT, was used as a template. Amplicons were produced with primers, GST_Ulp1_FwCf and Ulp1EcoR1_RvCF in a PCR reaction. Each primer carries an 18-nucleotide homology to the linearized plasmid. Purified vector and insert were subjected to a Cold Fusion cloning reaction (System Biosciences #MC010A-1) according the manufacture's protocol at a 1:1 ratio. Resultant plasmid was confirmed by EcoRI restriction endonuclease.

Expression of the TgSUMO-eGFP recombinant protein substrate was generated by linearization of pET28b plasmid by a XbaI and HindIII restriction endonuclease reaction. To generate a TgSUMO (TGGT1_266460) insert, a PCR reaction with primers

XbaIRBS5Sumo and EcoRI3Sumo with cDNA was used. The resultant PCR product was digestion with XbaI and EcoR1 restriction endonuclease reaction and gel purified. For generation of the eGFP insert, a PCR reaction with primers EcoR1ser5eGFP and HindIII3eGFP were used with a ptub-mcherryFP-EGFP plasmid (from Dr. Vern B. Carruthers, University of Michigan) as a template. Resultant amplicon was digested with EcoRI and HindIII and gel purified. Purified linear plasmid and TgSUMO and eGFP inserts were ligated at a 2:1:1 ratio using T4 DNA ligase. Resultant plasmid was confirmed by restriction endonuclease reactions.

For yeast complementation assays, the parental plasmid pScUlp1-GFP/LEU2 (gift from Dr. Oliver Kerscher, The College of William and Mary, Williamsburg, Virginia) was prepared by digested with XhoI and NotI endonucleases to remove the ScUlp1 promoter and coding region. Primers ScUlp1Fw and ScUlp1Rv were used with the parental plasmid as a template to generate the ScUlp1 promoter insert. The resultant product was subjected to XhoI endonuclease reaction to create a complimentary end between the 5' end of the promoter insert and the linearized plasmid used for ligation. TgUlp1 coding region was generated in a PCR reaction with primers TgUlp1Fw and TgUlp1R with the pGEX-TgUlp1 plasmid as a template for TgUlp1. The PCR product was then digested with NotI endonuclease to create complimentary 3' end used for ligation. Purified components were ligated with T4 DNA ligase at a 2:1:1 ratio. Resultant plasmid was confirmed by XhoI restriction endonuclease reaction.

5.2.3. Yeast complementation assay

Standard lithium acetate transformation protocol for yeast cultures was used (Gietz and Schiestl, 2007). YOK 40 strain (referred to as Δulp1 shuffle strain) was grown

overnight at 30 °C and prepared for transformation. While a parental pAA3 plasmid was used as a negative control (empty vector), a plasmid expressing Ulp1 from its endogenous promoter (pUlp1-GFP/LEU2, obtained from Dr. Oliver Kerscher, The College of William and Mary, Williamsburg, Virginia) served as a positive control (Elmore et al., 2011). The construct of TgUlp1 cDNA under the endogenous yeast Ulp1 promoter was used in the transformation to monitor the level of functional complementation. All yeast strains and their transformed cultures were grown to a mid-log phase of growth and then subsequently diluted to ~0.5 (OD_{600}). Serial dilutions were performed using sterile 1x TE. Each dilution was spotted (3 μL) onto YEPD plates containing a final concentration of 1 mg/mL of 5-Fluoroorotic acid (5-FOA, Bio Basic #703-95-7) and 20 μg/mL of uracil (Bio Basic #UD0564) at pH 4.5 for induction of Ulp1 knockout. Plates were incubated at 30 °C for 72 hours. Growth analyses were performed using densitometry (ImageJ 1.44p) for three independent experiments.

5.2.4. RNA isolation and qRT-PCR

Total RNA was isolated from freshly lysed parasites (approximately 10^6 parasites) using TRI reagent (Molecular Research Center, Inc., #TR 118) according to the manufacturer's protocol, and treated with RQ1 Rnase-Free Dnase (Promega, #M6101) for 30 minutes at 37 °C. Samples were kept at −20 °C until use. Isolated total RNA (~500 ng) was annealed with 250 ng of an oligo-(dT) following heating at 65 °C for 5 minutes and slowly cooled. The reactions were then divided in half. A half of each annealed reaction was used in SuperScript III reverse transcriptase (RT) reaction (Invitrogen #18080-093), while the other aliquot was used as a negative control reaction in the absence of SuperScript III RT. For real-time qPCR, the reaction mixtures were

diluted 1/10, and analyzed by using Kapa Sybr Fast qPCR Kit Master Mix (2x) ABI Prism (Kapa Biosystems #KR0390) and OneStepPlus thermocycler (Applied Biosystems). Samples were analyzed in technical triplicates among at least biological tests. Relative TgUlp1 expression levels in comparison to TgGAPDH expression levels were measured and calculated using **ΔΔ**Ct protocol carried out by the StepOne Software v 2.1 (Applied Biosystem). To generate a data set, relative TgUlp1 expression from no addition of RNA was used for normalization of Tg-miR-60a and Tg-miR-4a (control) inhibitor TgUlp1 mRNA levels. PCR was performed using Hot Start *Taq* DNA Polymerase (New England Biolabs, Inc., #M0495S) according to the manufacturer's protocol.

5.2.5. *In vitro* transcription of effector RNAs

Using partial DNA duplexes as templates, *in vitro* transcription reactions were carried out in 50 μL-reaction mixtures in the presence of T7 RNA polymerase (~5 units), 80 mM Hepes-KOH, pH 7.5, 24 mM NaCl, 2 mM spermidine, 40 mM DTT, 10 mM rNTPs, and ~1 unit pyrophosphatase at 37 °C for 16 hours. Reactions were subsequently extracted with one volume of phenol-chloroform mixture (1:1), and resultant RNAs were precipitated and quantified (Thermo Scientific NanoDraop2000). The oligonucleotide template (2 μM) was annealed with an equimolar of T7promoterGG following mixing and heating the mixture for 5 minutes at 70 °C and slowly cooled to room temperature to yield partial DNA duplexes. One fifth of the annealing reaction (or 400 nM of the partial DNA duplex) was used as templates for the RNA synthesis. Resultant single-stranded RNA products, which were produced using the oligonucleotides named antimiR-60a_tpl and antimiR4a_tpl were referred to as miR-60a inhibitor and miR-4a inhibitor,

respectively. For the production of pri-miR-60, DNA amplicons were prepared using an oligonucleotide primers containing a gene specific sequence and the sequence of the T7 promoter only on the forward primer, as listed in Table 5.1, under standard conditions for Hot Start *Taq* DNA polymerase (New England Biolabs). Isolated genomic DNA from type I parasites was used as a template with primers: TgUlp1intron7T7 and qUlp1Rv to produce a 710 bp product and primers: TgUlp1intron7T7 and Ulp1LICRv to produce a 2308 bp product. *In vitro* transcription reactions were carried out in similar conditions as described above using 1 μg of DNA amplicon. Quality of RNA products were verified by gel electrophoresis to confirm their expected sizes prior to electroporation into freshly released parasites.

5.2.6. Dual luciferase assay

For dual luciferase assays, 48 hours post-transfection, freshly lysed parasites were harvested and lysed with 100 μL of 1x Passive Lysis Buffer (Promega, #E1531) and incubated for 10 minutes at room temperature. Lysates were cleared of debris by centrifugation (10,000 xg for 1 minute), and the supernatant was kept on ice and used in the dual luciferase assay, which was carried out in a 2-step fashion as previously described by Crater et al. (2015). Luminescence signals were measured with a 20/20n Luminometer (Turner Biosystems). Rnluc signals were normalized to Ffluc signals for a direct comparison among different independent experiments. For each experiment, the reporter plasmid electroporation without tested *in vitro* synthesized RNA, was carried out to allow for the measurement of the dual luciferase expression under at least three independently tested conditions.

5.2.7. Western blot analysis

Cell-free lysates were prepared by the addition of RIPA buffer to collected parasites and sonicated for 30 seconds (x3). Protein concentration was measured using the RC DC Protein Assay according to manufacturer's instructions (Bio-Rad Laboratories, #500-0119). Approximately 20-25 μg was used for the analysis, of which the lysates were denatured using 6x SDS protein loading buffer and resolved on a 10% SDS-PAGE. The primary antibodies used were α-Flag mouse antisera (1:5,000, Applied Biological Materials Inc., Richmond, B.C, Canada, #G191), α-tub mouse antisera (12G10, 1:1,000, Developmental Studies Hybridoma Bank, University of Iowa, USA), and α-GFP goat antisera (1:2,500, Rockland Immunochemicals Inc., #600-101-215). The signals were revealed using secondary antibodies: horseradish peroxidase (HRP) conjugated goat anti-mouse antisera (Rockland Immunochemicals Inc., #610-1319) or rabbit anti-goat antisera (Rockland Immunochemicals Inc., #605-4302) at a 1:10,000 dilution, and a Chemiluminescent HRP Substrate Kit (Millipore, #WBKLS0500).

5.2.8. Recombinant protein purification

Both GST-TgUlp1 and TgSUMO-GFP fusion proteins were expressed in *E. coli*, BL21-codonPlus (DE3)-RIL cells (Agilent #230245), by the addition of 1 mM IPTG and grown overnight at room temperature. GST-TgUlp1 was purified on glutathione-agarose beads (Novagen #7044-3) and eluted with 10 mM glutathione and 50mM Tris-HCL, pH 8.0. Freshly purified protein was then used for cleavage assays. TgSUMO-GFP was purified on Ni-NTA agarose beads (Invitrogen #537735A) and eluted with 250 mM imidazole, then dialyzed with 50 mM Tris-HCL, pH 8.0, 150 mM NaCl, and 0.2 % Tween 20 and stored at -80 °C for future use.

5.2.9. TgUlp1 cleavage assays

5.2.9.1. TgUlp1 activity within parasite lysates

For TgUlp1 activity within type I parasites, 48 hours post-transfection with the indicated concentration of miR-60a inhibitor, parasites were lysed in reaction buffer containing 50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 0.2 % Tween 20 and 1 mM dithiothreitol, and no protease inhibitors, and sonicated for 30 seconds. Total protein was quantified and 25 μg of parasite lysates were combined with 4 μM of substrate and incubated at 37 °C for two hours. Aliquots were removed every 30 minutes and reactions were stopped by the addition of SDS loading buffer and boiled for 10 minutes. Samples were then resolved by 10 % SDS-PAGE and subjected to western blot analysis. TgSUMO-GFP amounts were quantified by densitometry for at least three experimental trials.

5.2.9.2. *In vitro* TgUlp1 activity

Equal amounts of GST-TgUlp1 and TgSUMO-GFP (4 μM) were combined in reaction buffer containing 50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 0.2 % Tween 20 and 1 mM dithiothreitol, and no protease inhibitors, and incubated at 37 °C for 30 minutes. Reactions were terminated by the addition of SDS loading buffer and boiled for 10 minutes. Samples were then resolved by SDS-PAGE and visualized via coomassie staining. TgSUMO-GFP amounts were quantified by densitometry for at least three experimental trials.

5.2.10. Bioinformatics

Multiple sequence alignments were performed with the Clustal Omega tool (Sievers et al., 2011) while sequence similarities were performed by GGSEARCH for global and by SSEARCH for local protein analysis (Pearson 1988).

5.2.11. Statistical analysis

Statistical analysis was completed by the paired-wise student t-test methodology assuming a null hypothesis. All values were calculated from at least three independent tests.

5.3. Results

5.3.1. TgUlp1 is the most probable target of Tg-miR-60a

To verify if TgUlp1 expression is controlled by miR-60a, we used a miRNA inhibitor to alter its expression. The mRNA levels of TgUlp1 were determined by qRT-PCR. To confirm the effects of the miR-60a inhibitor on protein levels, we used our engineered dual luciferase reporter system where our *Renilla* luciferase transcript (Rnluc) contains three perfect match sites to miR-60a in the 3' UTR as previously described (Crater et al. 2015). Upon the introduction of increasing concentrations of miR-60a inhibitor (5, 10 and 20 μg), relative TgUlp1 mRNA levels increased to 2.4 ± 0.23, 5.15 ± 0.35 and 10.15 ± 0.60, respectively (Figure 5.1a). When an unrelated miRNA inhibitor for Tg-miR-4a was used, relative TgUlp1 mRNA levels were 1.28 ± 0.22. To confirm the sequence specific effects of the miR-60a inhibitor, two unrelated genes (TgGAPDH and TgUPRT) showed no change in relative mRNA levels (Figure 5.1b). The effects of the miR-60a inhibitor (5, 10, and 20 μg) on relative Rnluc activity increased to 46.6 ±

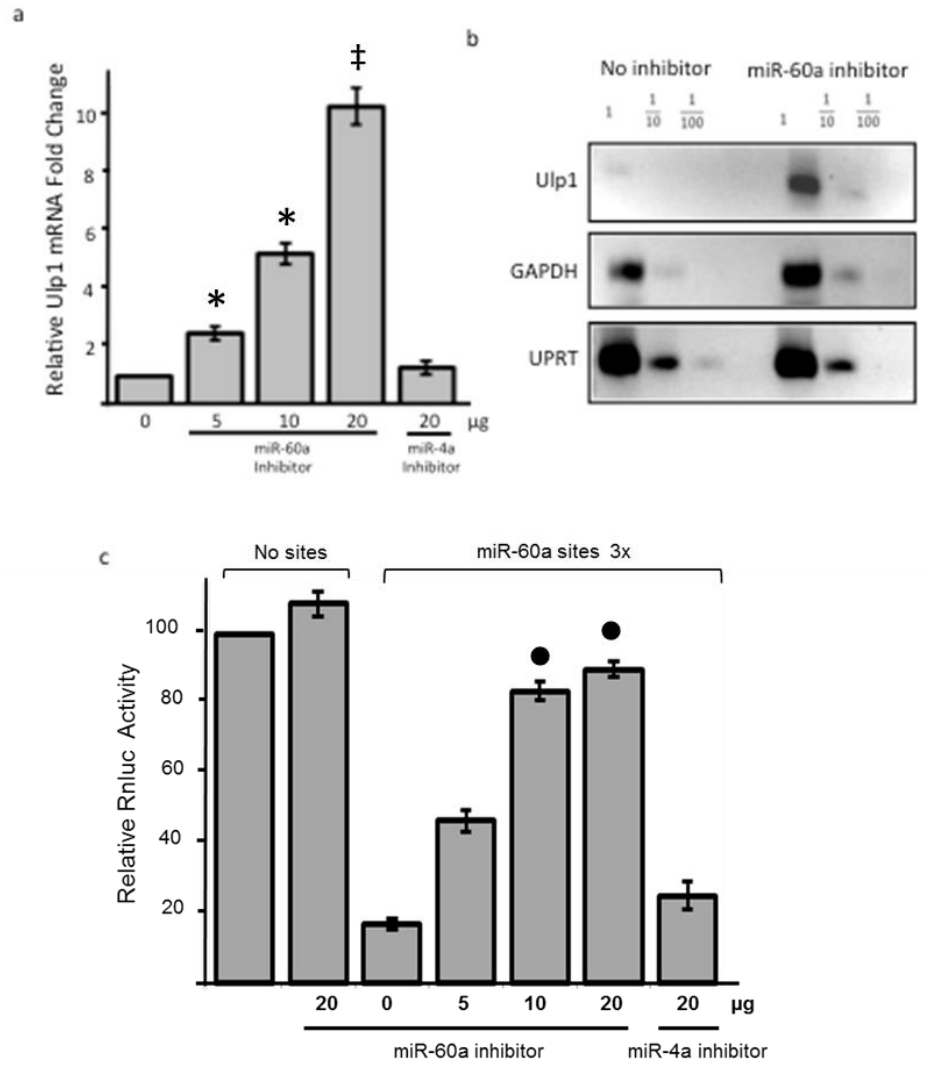


Figure 5.1. TgUlp1 expression is affect by a miR-60a inhibitor. A. qRT-PCR indicates that the presence of a miR-60a inhibitor can increased relative TgUlp1 mRNA levels in a sequence specific manor. B. RT-PCR indicates an overall low expression of TgUlp1 when compared to housekeeping genes. TgUlp1 mRNA levels increase in the presence of a miR-60a inhibitor while transcripts with no miR-60a binding sites remain unchanged. C. Rnluc activity was also compared in subsequent parasite lysates. (P-values from at least three independent trials are, ‡ $p \leq 0.001$, * $p \leq 0.005$ and ● $p \geq 0.01$)

3.2, 83.8 ± 2.7 and 90 ± 2.2 %, respectively, when compared to no inhibitor at 17 ± 1.6 (Figure 5.1c). The unrelated miR-4a inhibitor caused no change in relative Rnluc activity, 25.0 ± 4.0 %. As a control, the maximal amount of miR-60a inhibitor (20 μg) was also tested with the Rnluc construct with no miR-60a binding sites. Relative Rnluc activity remained unchanged indicating the sequence specificity of the miR-60a inhibitor (Figure 5.1c).

5.3.2. TgUlp1 genomic arrangement and miR-60a binding sites

Since miR-60a has a near perfect complementarity to intron 7 of TgUlp1 pre-mRNA we sought to examine the genomic arrangement of this region in relationship to the predicted miR-60a binding sites. We hypothesized that if miR-60a is able to bind intron 7, it may cause intron retention. Total RNA was isolated and reverse transcribed using a poly(T) oligonucleotide (polyA RT). For polyA generated cDNA, the PCR reaction using primer set a-a', which spans exons 7-8, produced two transcript products at 110 and 710 bps. The 110 bp product corresponds to joined exons 7 and 8 of TgUlp1 mRNA, and the 710 bp product would include exon 7, intron 7, and exon 8. The lack of a product in the no RT reaction suggests the 710 bp amplicon is not the result of genomic DNA contamination and could be the effect of intron 7 retention (Figure 5.2b). PCR reactions for primer sets b-b', a-b' and c-a' produced one transcript size each that corresponds to TgUlp1 mRNA at 281, 538 and 198 bp respectively (Figure 5.2b). As a control, we also generated cDNA by a gene specific primer for TgUlp1 (primer b' RT). Unexpectedly, primer set a-a' only produced a PCR product at 110 bp. Subsequent reactions (c-a'and a-b') produced products of 198 and 281 bp, respectively, which correspond to TgUlp1 mRNA.

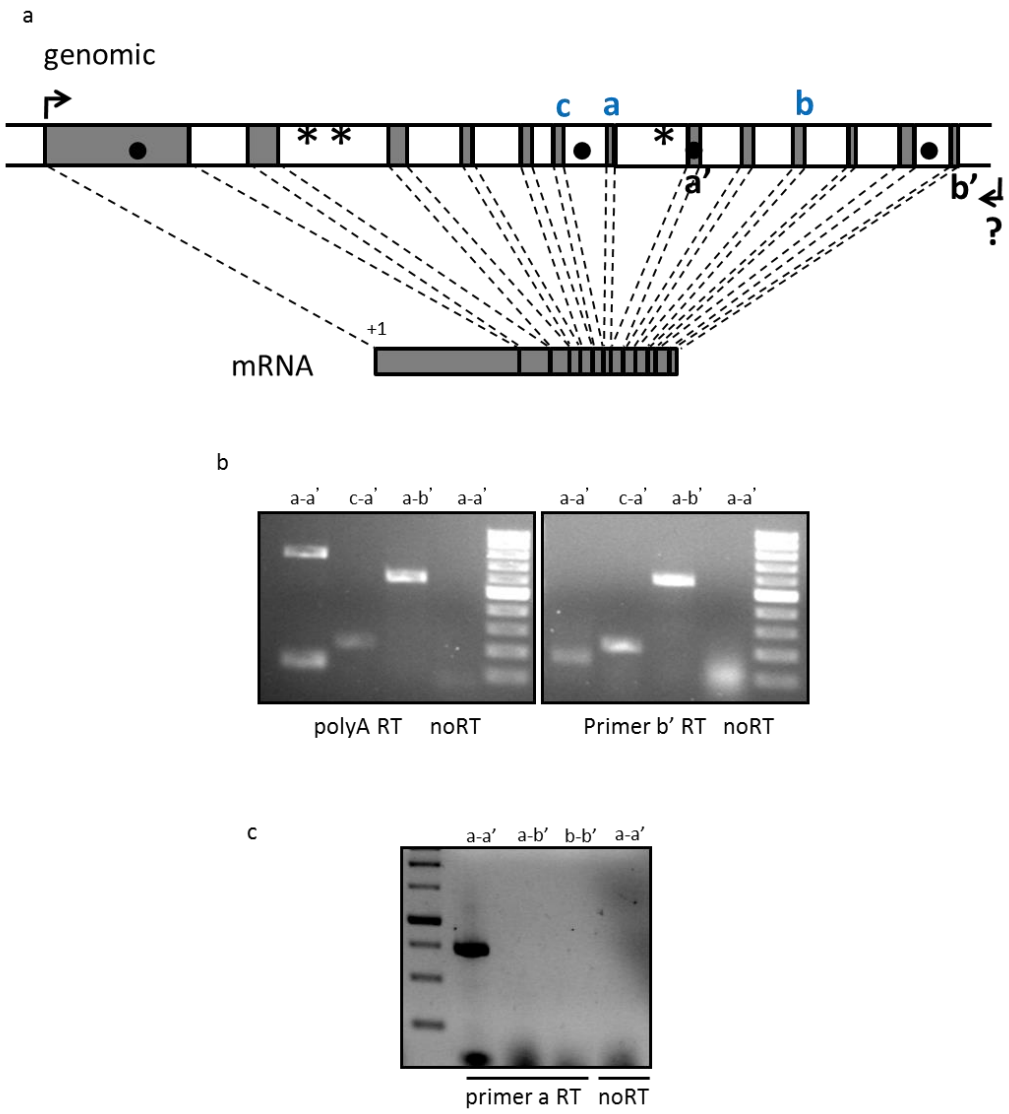


Figure 5.2. TgUlp1 genomic arrangement. A. Predicted miR-60a binding sites on TgUlp1 locus are indicated by * and ● on the coding and non-coding strand, respectively. B. Primer set a-a' spans intron 7 where a near perfect match binding to miR-60a is located. Polyadenylated generated cDNA produced two PCR bands at ~110 and ~710 bp. The 710 bp product is not present in gene specific generated cDNA (by primer b'), only the 110 bp from mRNA. C. The 710 bp product is only detected in antisense generated cDNA by primer a.

Complementary DNA for the antisense RNA was also generated and tested (primer a RT). To our surprise, primer set a-a' produced a product at 710 bp. The lack of amplicons in the subsequent reactions with antisense cDNA (a-b', b-b', and No RT) could indicate the location for the promoter of the antisense RNA (Figure 5.2c). These data suggests the 710 bp product is not the effect of intron retention but of a polyadenylated antisense RNA.

The antisense RNA contains a sequence that resembles the sequence of miR-60a guide strand. To investigate the ability of the identified 710 nt antisense RNA to confer a gene silencing effect, we used our dual luciferase reporter assay. The 710 nt nucleotide sequence of antisense RNA was *in vitro* transcribed and tested. In the presence of *in vitro* synthesized 710 RNA, Rnluc activity decreased when compared to no *in vitro* synthesized RNA (Figure 5.3.). When the 710 *in vitro* synthesized RNA was RNaseIII treated prior to transfection, Rnluc activity decreased further to 24.8 ± 2.3 % (Figure 5.3.). A longer antisense RNA molecule (2803 nt) from the same locus was used as a control. In the presence of the 2803 nt RNA and RNAseIII treated RNA Rnluc activity was 43.6 ± 0.36 and 36.9 ± 1.1 %, respectively, when compared to no RNA added at 47.2 ± 3.7 % (Figure 5.3.). These data suggest that the expressed antisense RNA could confer gene silencing and a possible locus is pri-miR-60a.

Taken altogether, our findings suggest that TgUlp1 could be identified as the first self-regulating gene in *T. gondii* by miRNAs. A self-regulating gene could be an indication of an essential gene. Therefore we sought to determine the function and importance of TgUlp1 to parasite cell biology and the effects of miRNA-induced gene regulation of TgUlp1 in the SUMOylation pathway.

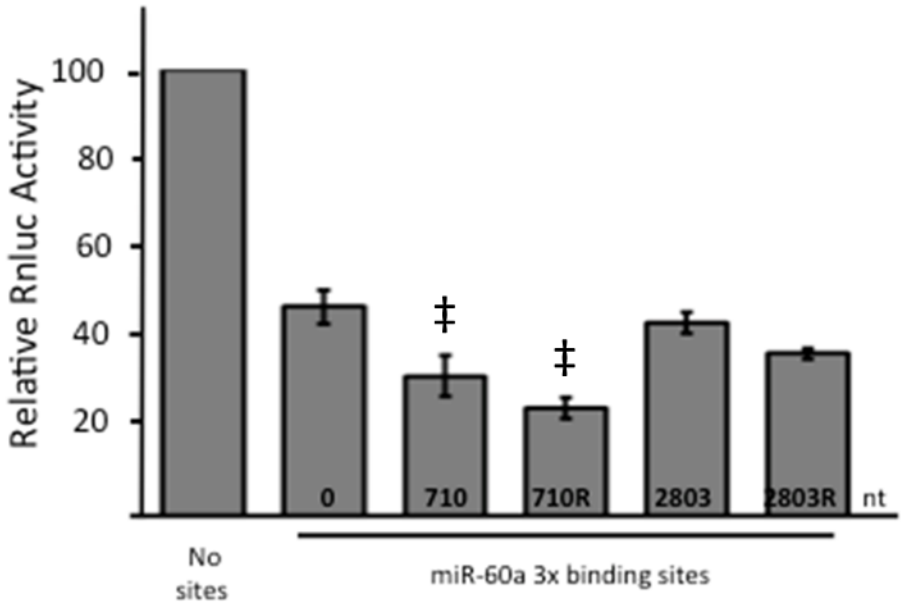


Figure 5.3. Antisense RNA acts as effector RNA. *In vitro* transcribed antisense RNA (710) is able to decrease Rnluc activity when the transcript contains miR-60a binding sites. RNAseIII treatment of antisense RNA (710R) can decrease Rnluc activity further. To confirm the silencing effect, a 2803 nt RNA corresponding to the antisense region to the 3' coding region of TgUlp1 was tested (2803). No alteration in Rnluc activity was observed even when the RNA was RNAseIII pre-treated (2803R). (P-values compared to Rnluc activity in the presence of miR-60a binding sites, but no additional RNA from at least three independent trials are ‡ $p \leq 0.001$)

5.3.3. TgUlp1 function by yeast complementation

Due to the conserved catalytic triad and high sequence similarity of the functional domain of TgUlp1 (TGGT1_214470) with yeast Ulp1 (GenBank:KZV07496.1), we anticipated TgUlp1 to display similar activities as observed with yeast Ulp1. To determine if TgUlp1 could function similarly to ScUlp1, we used a conventional yeast complementation analysis. A Ulp1 shuffle strain (ulp1::HIS3, Ulp1/URA3), referred to as ΔUlp1, which is a conditional knock-out in the presence of 5-fluoroorotic acid (5-FOA), was used (Li et al., 2003, Elmore et al., 2011). To determine a background level of growth, the ΔUlp1 strain was grown on plates containing 5-FOA and uracil at 30 °C for 72 hours and growth was determined via densitometry. Growth of the ΔUlp1 strain was 20.6 ± 2.6 % when compared to the wild-type (WT) (Figure 5.4.). Transfection of ΔUlp1 with the pAA3 (LEU2) empty vector displayed a similar growth phenotype to 19.6 ± 2.5 % as compared to WT. When ΔUlp1 was transformed with a pAA3 plasmid containing the yeast wild-type Ulp1 under the control of its own promoter (ScUlp1, Elmore et al., 2011), growth was restored to 87.4 ± 5.3 % of the wild type phenotype. When pAA3 plasmid engineered to carry TgUlp1 (cDNA 2316 bp) under the control of the yeast Ulp1 promoter (TgUlp1) was used, complementation by TgUlp1 was achieved. Growth was restored to 48.4 ± 3.7 % (1x) and 80.0 ± 4.4 % (2x) as compared to the WT (Figure 5.4.).

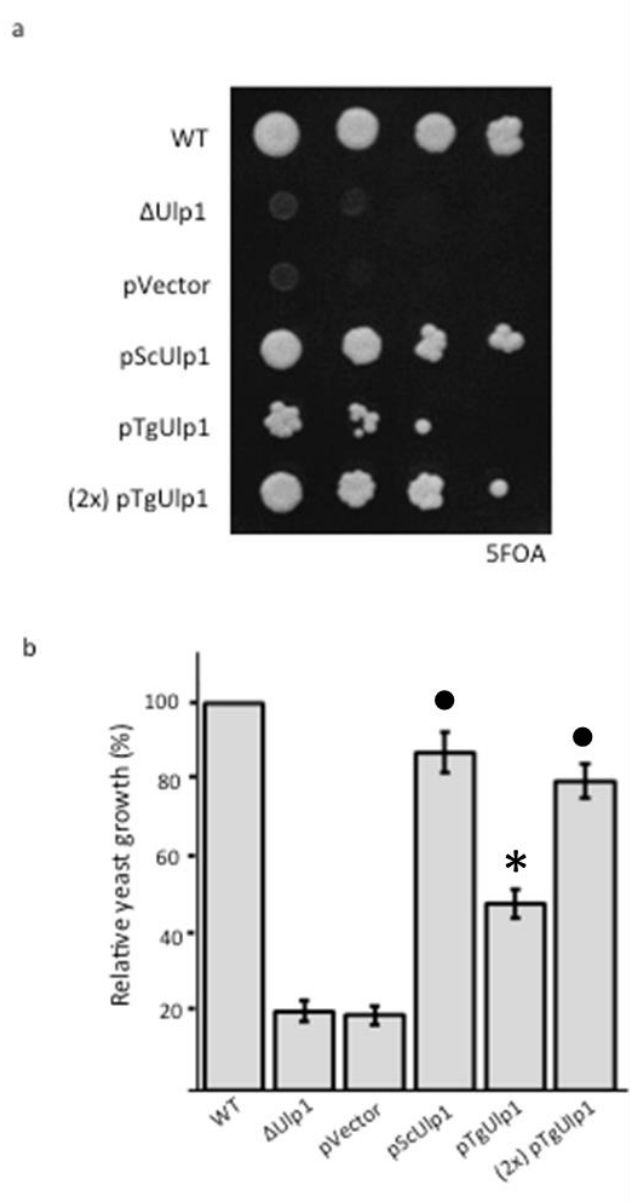


Figure 5.4. Yeast complementation analysis in ΔUlp1 strain. The ability of yeast to utilize the *T. gondii* ortholog of Ulp1 was tested by growth analysis in the conditional ScUlp1 KO yeast strain. Under the ScUlp1 KO conditions, phenotypic growth was only restored upon transfection of a Ulp1 expressing plasmid. (P-values compared to WT yeast growth by densitometry from at least three independent trials are ● $p \geq 0.01$, * $p \leq 0.01$)

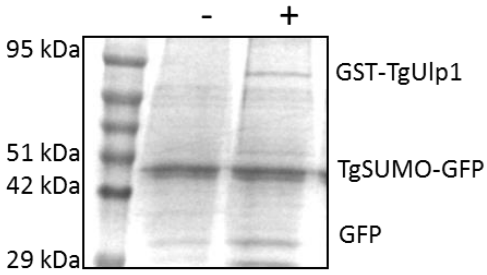
5.3.4. TgUlp1 peptidase activity

To examine the peptidase activity of TgUlp1, we expressed and purified a recombinant protein and a SUMOylated recombinant protein substrate for in vitro assays. TgUlp1 (85 kDa) was N-terminally tagged with a glutathione S-transferase (GST, 26 kDa) that would produce a 111 kDa protein. The peptide of TgSUMO (TGGT1_266460, 10.3 kDa) was joined to the N-terminus of GFP that carries a C-terminal His tag (6x) to give a 45 kDa protein substrate (Figure 5.5a). Equal amounts (4 μM) of purified GST-TgUlp1 enzyme and TgSUMO-GFP substrate were combined and incubated at 37 °C for 30 minutes. In the presence of GST-TgUlp1 (+), the amount of TgSUMO-GFP substrate decreased to 57 ± 6.9 % when compared to no enzyme (-) (Figure 5.5a). A GFP product at 34.7 kDa was also detected (Figure 5.5a). This indicates TgUlp1 is able to recognize and hydrolyze the peptide bond from a SUMO conjugated protein.

Separate parasite lysates (~25 μg) were prepared 48-hours post-transfection without or with the miR-60a inhibitor and combined with purified TgSUMO-GFP substrate. First, mock (no inhibitor) parasite lysates displayed a substrate cleavage rate of 0.21 $nM \cdot min^{-1} \cdot \mu g^{-1}$ (Figure 5.5b). Second, the introduction of 5, 10, and 20 μg of inhibitor increased substrate cleavage rates to 0.58, 0.77 and 1.10 $nM \cdot min^{-1} \cdot \mu g^{-1}$, respectively (Figure 5.5b). These data suggest the presence of the miR-60a inhibitor can increase TgUlp1 activity in parasite lysates.

A


N-terminus
TgSUMO
GFP
His 6x
C-terminus
N-terminus
GFP
His 6x
C-terminus


-
+
95 kDa
51 kDa
42 kDa
29 kDa
GST-TgUlp1
TgSUMO-GFP
GFP

B

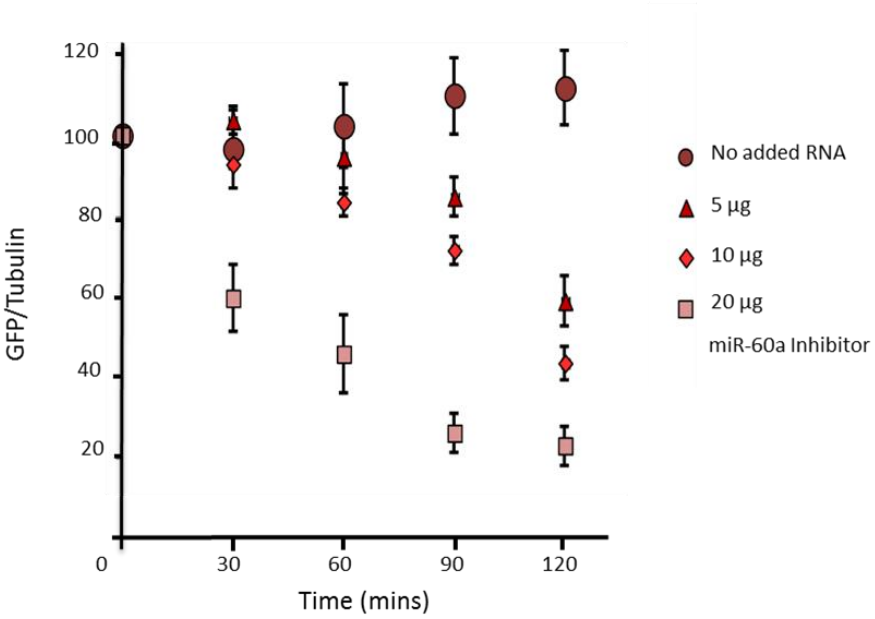

GFP/Tubulin
120
100
80
60
40
20
0
30
60
90
120
Time (mins)
No added RNA
5 µg
10 µg
20 µg
miR-60a Inhibitor

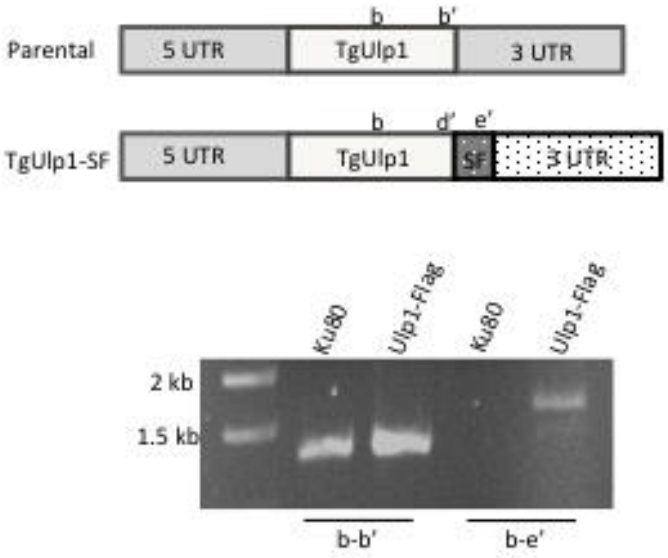
Figure 5.5. TgUlp1 peptidase activity. A. Purified GST-TgUlp1 can remove TgSUMO from a recombinant protein substrate, TgSUMO-GFP. B. Total TgSUMO-GFP levels were analyzed by western blot and an anti-GFP antisera within parasite lysates. Parasites were transfected with the miR-670a inhibitor and allowed to grow for 48 hours. Extracellular parasites were lysed in reaction buffer (see section 5.2.9.1.) and incubated with purified TgSUMO-GFP. Activity was calculated by the total amount of TgSUMO-GFP substrate present at each time point. (A, p-value compared to GFP signal without GST-TgUlp1 is $p \leq 0.005$, B, p-value for all substrate cleavage rates are $p \leq 0.0001$ calculated by densitometry from at least three independent trials)

5.3.5. TgUlp1 localization

To monitor TgUlp1 at the protein level, we generated a tagged version of TgUlp1 by a genetic knock-in strategy. The nucleotides encoding for a Strep/Flag (SF) tag were introduced immediately before the stop codon for TgUlp1 cDNA to create a ~93 kDa protein (85 kDa TgUlp1, ~8 kDa SF) (Huynh and Carruthers, 2009). To ensure proper integration of the SF tag to the TgUlp1 locus, PCR analysis was performed (Figure 5.6a). Primer set b-b' spans 1298 bps in exons 10-13 of TgUlp1 and was present in both the parental (Ku80) and transgenic TgUlp1-SF parasites (Figure 5.6a). Primer set b-e', spans 1798 bps in exon 10 to the Flag tag sequence and was only present in the TgUlp1-SF parasites (Figure 5.6a).

SUMOylation could be important for parasite invasion and cyst formation, as TgSUMO localizes to specific areas during these events (Braun et al., 2010). Localization of TgUlp1 could suggest its involvement in similar events. Within intracellular tachyzoites, TgUlp1 localization is mainly dispersed throughout the cytoplasm, however, in extracellular tachyzoites, TgUlp1 localizes to distinct structures in the apical and basal ends of the parasite (Figure 5.6b). Under bradyzoite conditions (*Dolichos* stain for cyst walls not shown), the distinct structures do not appear to change (Figure 5.6b).

a


b
b'
Parental
5 UTR
TgUlp1
3 UTR
b
d'
e'
TgUlp1-SF
5 UTR
TgUlp1
SF
3 UTR
Ku80
Ulp1-Flag
Ku80
Ulp1-Flag
2 kb
1.5 kb
b-b'
b-e'

b

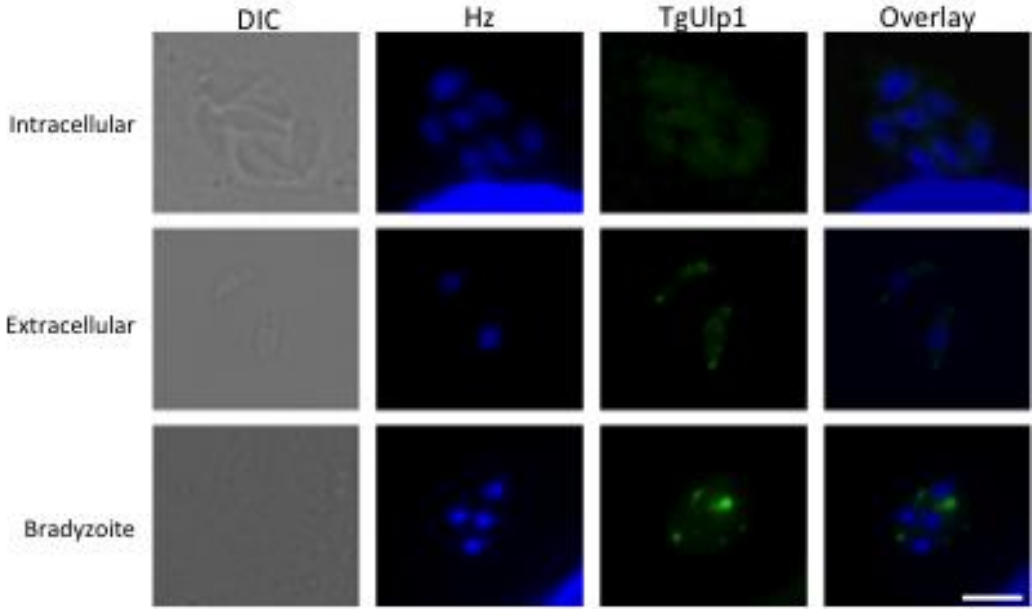

DIC
Hz
TgUlp1
Overlay
Intracellular
Extracellular
Bradyzoite

Figure 5.6. Endogenously tagged TgUlp1 localization. A. PCR confirmation of generation of TgUlp1-SF transgenic parasites by genetic knock-in. B. TgUlp1 localization changes throughout the different stages of the lytic cycle. Within intracellular parasites, TgUlp1 localization is found within the cytoplasm, which changes to distinct structures in the apical and basal ends of the parasite in extracellular parasites. Under bradyzoite conditions the distinct structures become more prevalent. (Hz, Hoechst, micrometer represents 5 μm.)

5.3.6. Physiological function of TgUlp1

To determine the function of TgUlp1, we attempted to generate a knockout parasite strain by using the CRISPR/Cas9 mythology in type I parasites. Two guide RNAs were designed to target and cleave in the genomic sequence of TgUlp1 at nucleotide 4162 in exon 6 and nucleotide 7023 in exon 12 to remove a 2880 bp region that encodes for the catalytic domain of TgUlp1. Using homologous recombination, a 3640 bp DHFR (dihydrofolate reductase) selectable marker was inserted at the cut sites. Upon selection pressure, transfected parasites would begin to grow normally with an average of 8 parasites per vacuole at 24-hours post-invasion (data not shown). However, we noticed at 48-hours post-invasion, vacuole morphology would appear somewhat abnormal and the parasites remained intracellular (Figure 5.7., vacuole indicated by red triangle). Generally under our culturing conditions at 48-hours post- invasion, parasites would require a fresh monolayer of host cells to invade. As the parasites were allowed to growth further (up to 5 days), vacuole morphology continued to look abnormal and worsened in appearance and size (Figure 5.7.). A lack in host cell destruction or debris due to parasite egress, or exit, of the host cell was also observed. The abnormal vacuole morphology was noted to be similar in parasites with an inability to egress (Heaslip et al., 2011). The inability to create a stable TgUlp1 KO parasite strain may indicate TgUlp1 is as an essential gene and may be involved in parasite egress.

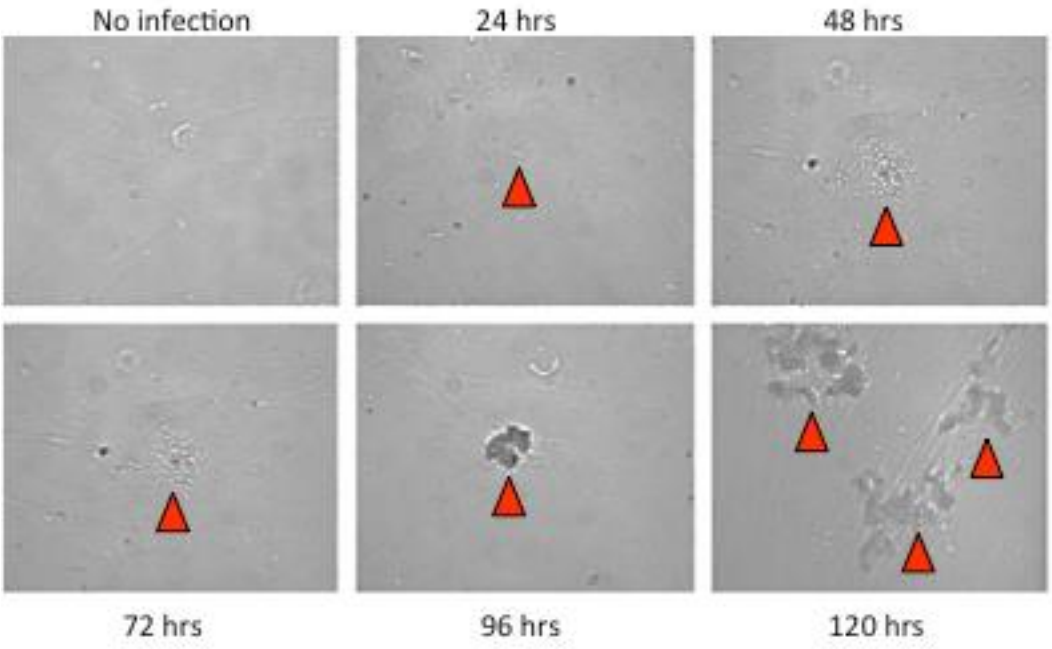


Figure 5.7. Effects of TgUlp1 knockout by CRISPR/Cas9 methodology. After 48-hours post-infection, TgUlp1KO parasites display abnormal vacuole morphology (indicated by red triangle). By 72 hours post-infection, parasites display an inability to egress and remained intracellular.

5.3.7. Effects of TgUlp1 expression on SUMOylation

The effects of TgUlp1 on total SUMOylated substrates in parasite lysates was tested by alteration of TgUlp1 expression accomplished by using the miR-60a inhibitor and *in vitro* synthesized 710 nt antisense RNA (710as). Extracellular TgUlp1-SF parasites were transfected with 20 μg of miR-60a inhibitor, 710 antisense RNA (710as), or a mixture of miR-60a inhibitor and 710as. TgUlp1-SF expression was detected by an anti-Flag antisera. Densitometry values were calculated based on the expression of an alpha-tubulin loading control. At 48 hours post-transfection the presence of miR-60a inhibitor increased Ulp1-SF expression to 232 ± 17 % when compared to WT (Figure 5.8.). Introduction of 20 μg of the 710as RNA decreased TgUlp1 expression to 65 ± 13%. When equal amounts of miR-60a inhibitor and 710as RNA were mixed and introduced to the parasites, TgUlp1 expression was 129 ± 10% (Figure 5.8.).

Subsequent lysates were also analyzed for total SUMOylated substrates with an anti-SUMO antisera. The SUMO antisera would recognizes free SUMO and SUMOylated protein substrates. The presences of the miR-60a inhibitor increased total SUMOylated proteins levels to 266 ± 13 % when compared to WT (Figure 5.8.). Total SUMOylated proteins levels decreased to 47 ± 4.1 % in the presences of 710as RNA. Combined miR-60a inhibitor and 710as RNA restored total SUMO levels to 129 ± 17 %. Selected individual SUMOylated proteins (A, B, C, and D) were also monitored for an alteration of expression in the presence of the effector RNAs (Figure 5.8).

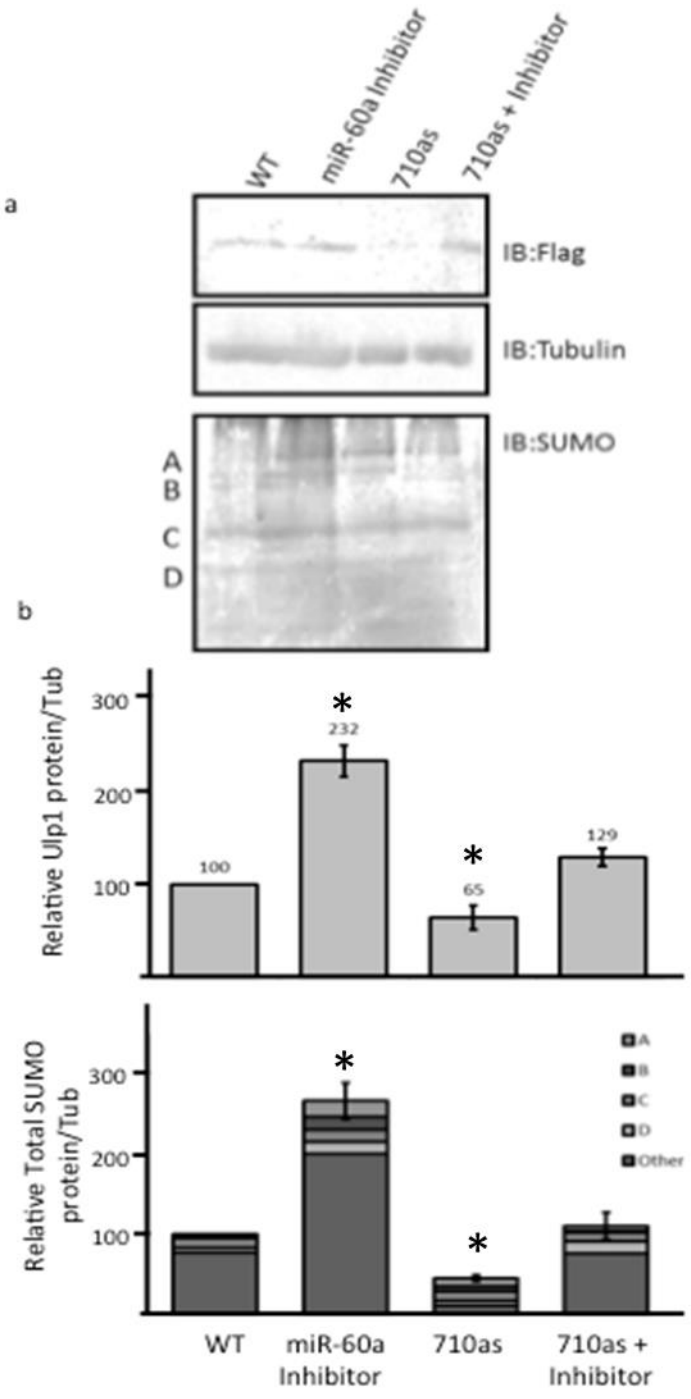

a
WT
miR-60a Inhibitor
710as
710as + Inhibitor
IB:Flag
IB:Tubulin
IB:SUMO
A
B
C
D
b
Relative Ulp1 protein/Tub
300
200
100
100
232
65
129
Relative Total SUMO protein/Tub
300
200
100
A
B
C
D
Other
WT
miR-60a Inhibitor
710as
710as + Inhibitor

Figure 5.8. Alteration of TgUlp1 expression affects total SUMOylation levels. A. Representation of immunoblot for analysis at 48-hours post-transfection with or without effector RNA molecules. B. TgUlp1-SF parasites were transfected with an effector RNA or combined effector RNAs and allowed to grow for 48 hours. Changes in protein expression of TgUlp1-SF and total SUMOylated proteins were observed in the presence of the miR-60a inhibitor and 710 antisense RNA (710as) when tested independently. (p-values were calculated by densitometry from at least three independent trials compared to WT expression relative to Tubulin expression,* $p \leq 0.01$)

5.4. Discussion

For the identification of the most probable target for miR-60a, we used miRNA target binding criteria based on plant miRNAs, which contain a perfect to near perfect complementary binding. The identified near perfect miR-60a binding site is located in intron 7 of the pre-mRNA. Two other putative miR-60a binding sites with only a seed match complementarity (6 nt) were found in intron 2 (Table 5.1., Figure 5.2a). Since these three sites were identified in the introns of TgUlp1 pre-mRNA, we sought to identify possible miR-60a binding sites on the mRNA based on mammalian miRNA binding criteria. Mammalian miRNAs need as few as 6-nt complementary to bind target mRNA and can contain a 1 nt mismatch in the miRNA seed sequence (Seok et al., 2016). Either strand of a miRNA can be used as the "guide" strand therefore we sought to analyze both the lower and upper strand of miR-60a to target TgUlp1 mRNA. Only the guide strand sequence of miR-60a has been reported, therefore we used two predicted upper strands from miR-60b as reported by Braun et al., (2010). These miRNAs come from the same miRNA family (miR-60), contain the same seed sequence and should target the same mRNAs. Using these criteria, we identified ten putative binding sites for Tg-miR-60. Table 5.3. indicates the sequence of the miR-60 strand (lower or upper +) and their predicted base pair interactions (indicated by *) and location on the TgUlp1 mRNA. The predicted seed sequence base pairing is highlighted. Using the lower strand of miR-60a, eight putative binding sites were identified. Two of the binding sites are located in the predicted 5' UTR, four in the coding region, and one in the predicted 3' UTR. The miR-60b upper 1 strand has a binding site in the predicted 5' UTR and exon 6 while miR-60b upper 2 would bind once in the predicted 5' UTR.

Table 5.3. Predicted miR-60a binding site on TgUlp1 mRNA based on mammalian miRNA binding. Seed sequence is highlighted in orange and target sequence is red. Predicted base pairing is indicated by * (+ indicates upper strand of pridicited miR60).

Predicted base pairing	Location on mRNA (nt)
5′ UCCCCCCGUUGGAAUUUUCCCUUCU 3′ mRNA * * ******* * 3′ GUACAUCGUGUAUAUCAUGGGAAGA 5′ miR60b+	Predicted 5' UTR -364 to -339
5′ CCGCUGUUCGUGGUGUCUAAUCGAG 3′ mRNA * * ***** * 3′ CUGAUUUCAGUCGUAUUAGUUAA 5′ miR60b+	Predicted 5' UTR -322 to -297
5′ AAUCGAGGGCGACUUCGCGUCUGUU 3′ mRNA ** * * ** **** 3′ UCAUACCUAAAGCAUGGCUGACACA 5′ miR60a	Predicted 5' UTR -304 to -279
5′ GAGCUUCUGUACACCCACGAGUGUA 3′ mRNA ** * * *** *** 3′ UCAUACCUAAAGCAUGGCUGACACA 5′ miR60a	Predicted 5' UTR -68 to -43
5′ GAGGAAGGAUUCGAGAGACGGUGUC 3′ mRNA * * * *** **** 3′ TCATACCTAAAGCATGGCTGACACA 5′ miR60a	Exon 1 681 to 705
5′ CGUCCUGGACACAAGGACGACUAUCU 3′ mRNA **** ***** * * 3′ UCAUACCUAAAGCAUGGCUGACACA 5′ miR60a	Exon 2 1165 to 1189
5′ CAGGCCCGUCAAGCGACGAUUGUGC 3′ mRNA * * *** **** 3′ UCAUACCUAAAGCAUGGCUGACACA 5′ miR60a	Exon 2 1348 to 1373
5′ CUGGCUGAACGACGAAGUUAUCAAC 3′ mRNA * ** * * * ***** 3′ CUGAUUUCAGUCGUAUUAGUUAA 5′ miR60b+	Exon 6 1710 to 1735
5′ AGUGGUGCAUCCCCGACGACUUUGC 3′ mRNA *** * * * ***** ** 3′ UCAUACCUAAAGCAUGGCUGACACA 5′ miR60a	Exon 11 2120 to 2145
5′ UGCCAUGCAUGGAAAGAAAUGUGUG 3′ mRNA * * * * ***** 3 UCAUACCUAAAGCAUGGCUGACACA 5′ miR60a	Predicted 3' UTR 2700 to 2725

TgUlp1 was identified as the most probable target for miR-60a. The expression of TgUlp1 mRNA and protein levels can be effected by the introduction of a miR-60a inhibitor indicating miR-60a acts upon TgUlp1. Antisense RNA produced from the TgUlp1 locus, which contains a possible miR-60a guide sequence, is able to down regulated Rnluc activity when the transcript carries miR-60a binding sites and decrease TgUlp1 protein expression. Based on this data, the locus of TgUlp1 has other regions that would require investigation. The pre-mRNA of TgUlp1 could be a source of miR-60a mirtrons. Upon splicing, introns 2, 7, and 12 might produce a structure that resembles a mirtron with sequences similar to miR-60a. These regions would also need to be tested for their ability to induce gene silencing.

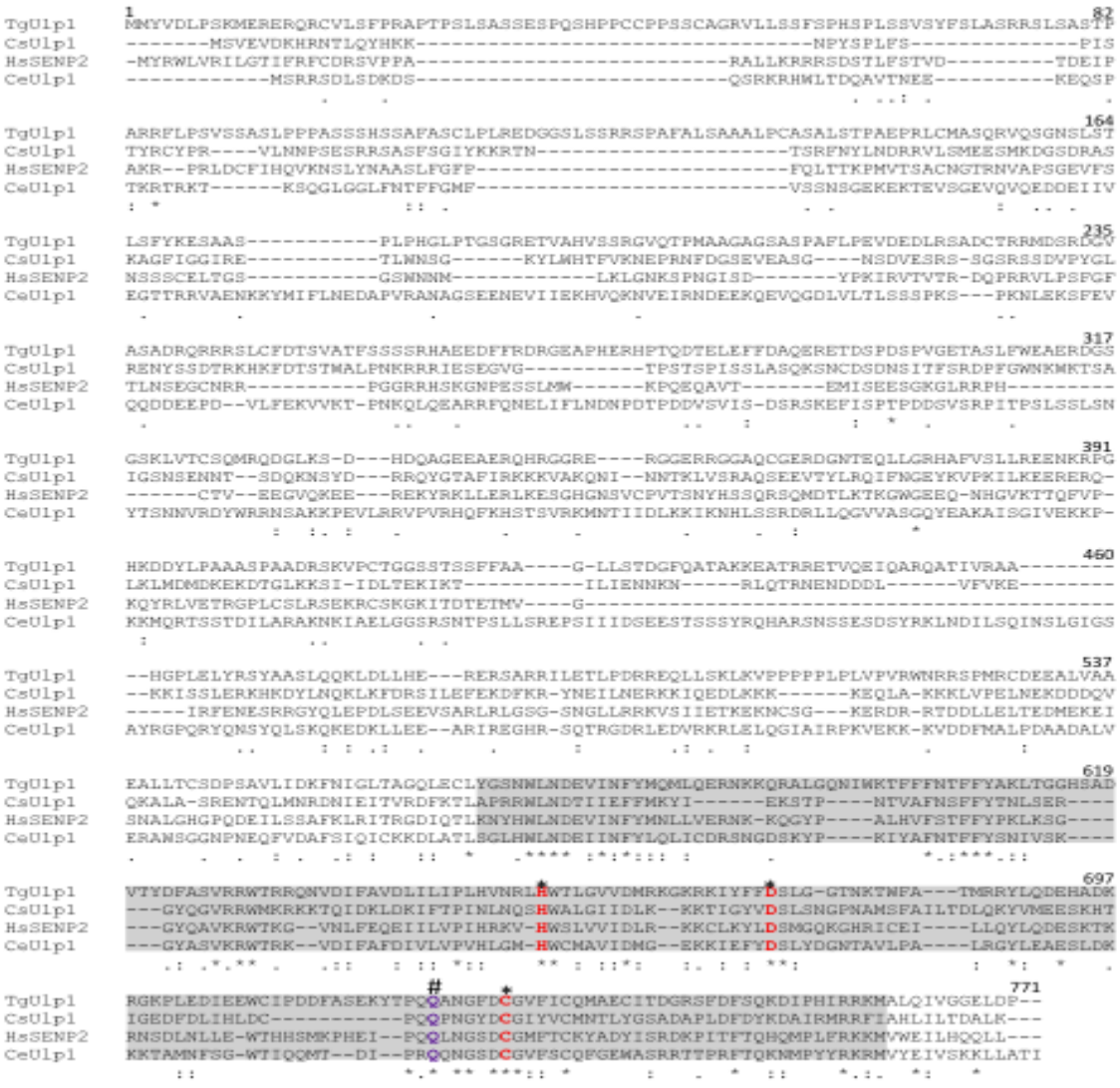
TgUlp1 locus is on chromosome X and composed of 13 exons and 12 introns that span 7432 bp. The coding sequence is 2316 nucleotides (Gajria et al., 2008). Pairwise sequence alignment of TgUlp1 from three archetype strains revealed 8 SNPs (Type I, GT1, Type II, ME49 and type III, Veg) in the N-terminal variable region. Each SNP results in an amino acid change, but are conservative substitutes. Reported gene expression profile of TgUlp1 indicates an overall low expression (40^{th} percentile) in the three strains (Bahl, et al., 2010). Overall, this indicates that TgUlp1 would have the same function and activity among the parasite archetypes (type I, II, and III).

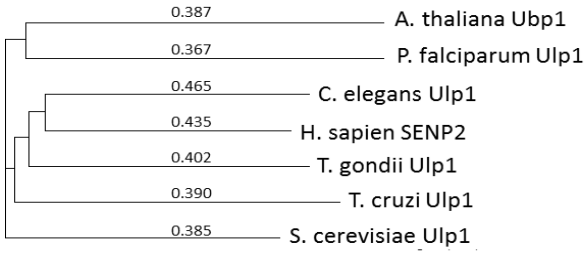
The amino acid sequence alignment of TgUlp1 (TGGT1_214470) with reported homologs, showed a C-terminal cysteine peptidase domain that contains a canonical catalytic triad, His-Asp-Cys. TgUlp1 sequence also contains a conserved Gln that is important for stabilizing the acyl intermediate that forms during catalysis (Supplemental Figure 5.1a). The global sequence similarity among TgUlp1 (771 amino acids), *H.*

sapiens SENP2 (589 amino acids), *C. elegan* Ulp1 (697 amino acids), and *S. cerevisiae* Ulp1 (621 amino acids) is, 31.9, 35.5, and 30.2 %, respectively. The catalytic domains alone have a 57, 50.6, and 47.4% sequence similarity respectively (accession; SENP2, NP_067640.2, *C. elegan* Ulp1, NP_498095.3, yeast Ulp1, GenBank:KZV07496.1). Global phylogenic tree analysis revealed TgUlp1 to be more closely related to human SENP2 and *C. elegan* Ulp1 rather than *P. falciparum* Ulp1, another Apicomplexan (Supplemental Figure 5.1b). Therefore, we would expect TgUlp1 to have similar function and activity to these identified homologs.

TgUlp1 was only able to complement the growth of yeast to ~50 % (Figure 5.4.). The level of rescue observed could be due to the longer N-terminal domain of TgUlp1 compared to yeast Ulp1 (~150 amino acids) or the lower sequence similarity between the functional peptidase domains (47.4 %). The N-terminal domain of yeast Ulp1 has been shown to be important for nuclear localization (Li et al., 2003), therefore, the differences in length, tertiary structure and amino acids may have an impact on the ability of TgUlp1 to localize and function in the yeast nucleus. Furthermore, these differences may also hinder the ability of TgUlp1 to interact with and cleave yeast SUMO-1 and SUMOylated products.

A recent study used the CRISPR/Cas9 system to perform a genome-wide screening of *T. gondii* genes by knockout (Sidik et al., 2016). After three lytic cycles, parasites were analyzed for their ability to grow and given a score based on their fitness. Any gene with a negative score was considered to contribute to parasite fitness and identified as an indispensable gene. Under these criteria, TgUlp1 KO was found to have




0.387
A. thaliana Ubp1
0.367
P. falciparum Ulp1
0.465
C. elegans Ulp1
0.435
H. sapien SENP2
0.402
T. gondii Ulp1
0.390
T. cruzi Ulp1
0.385
S. cerevisiae Ulp1

Supplemental Figure 5.1. Amino acid analysis and phylogenic tree for TgUlp1. A. Sequence alignment of *Toxoplasma* (TgUlp1) with *C. elegan* (CsUlp1), *H. sapiens* (HsSENP2), and *S. cerevisiae* (CsUlp1). Catalytic domain in gray while catalytic triad, H-D-C is highlighted in red (*), conserved Q in purple (#). B. Phylogenic tress analysis of full length Ulp1 orthologs.

a score of -2.73 (Sidik et al., 2016). Combining our inability to produce a stable TgUlp1 KO parasite indicates that TgUlp1 is an essential gene.

TgUlp1 has been identified as the first gene target of miR-60a and could be self-regulating gene. TgUlp1 KO and alteration of expression by the miR-60a inhibitor suggests a tight regulation of TgUlp1 expression is necessary for parasite cell biology. Overall, these data show the importance of an RNA induced gene silencing mechanism in *T. gondii.*

5.5. Acknowledgements

This work was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC). Also, ACP received an additional financial support, which was kindly allocated from NSERC to Dr. Lana Lee and from University of Windsor.

5.6. References

Braun, L.; Cannella, D.; Ortet, P.; Barakat, M.; Sautel, C.F.; Kieffer, S.; Garin, J.; Bastien, O.; Voinnet, O.; Hakimi, M.A. A complex small RNA repertoire is generated by a plant/fungal-like machinery and effected by a metazoan-like Argonaute in the single-cell human parasite *Toxoplasma gondii.* PloS Pathogens, 2010, 6(5), e1000920.

Crater AK, Manni E, Ananvoranich S. (2015) Utilization of inherent miRNAs in functional analyses of *Toxoplasma gondii* genes. Journal of Microbiological Methods. 108:92-102.

Elmore ZC, Donaher M, Matson BC, Murphy H, Westerbeck JW, Kerscher O. Sumo-dependent substrate targeting of the SUMO protease Ulp1. BMC Biol. 2011 Oct 28;9:74.

Gajria B, Bahl A, Brestelli J, Dommer J, Fischer S, Gao X, Heiges M, Iodice J, Kissinger JC, Mackey AJ, Pinney DF, Roos DS, Stoeckert CJ Jr, Wang H, Brunk BP. ToxoDB: an integrated *Toxoplasma gondii* database resource. Nucleic Acids Res. 2008 Jan;36(Database issue):D553-6.

Hay RT. (2005) SUMO: a history of modification. Molecular Cell. 1;18(1):1-12

Heaslip AT, Nishi M, Stein B, Hu K. The motility of a human parasite, *Toxoplasma gondii*, is regulated by a novel lysine methyltransferase. PLoS Pathog. 2011 Sep;7(9):e1002201.

Huynh MH, Carruthers VB. Tagging of endogenous genes in a *Toxoplasma gondii* strain lacking Ku80. Eukaryot Cell. 2009 Apr;8(4):530-9.

Kim, K., Weiss, L.M., 2004. *Toxoplasma gondii:* the model apicomplexan. Int. J. Parasitol. 34 (3), 423–432.

Komander D, Rape M. (2012) The ubiquitin code. Annual Review of Biochemistry. 81:203-29.

Li SJ, Hochstrasser M. The Ulp1 SUMO isopeptidase: distinct domains required for viability, nuclear envelope localization, and substrate specificity. J Cell Biol. 2003 Mar 31;160(7):1069-81.

Montoya, J.G.; Liesenfeld, O. Toxoplasmosis. Lancet, 2004, 363(9425), 1965-1976.

Mossessova E, Lima CD. (2000) Ulp1-SUMO crytal structure and genetic anaylsis reveal conserved interactions and a regulatory element essential for cell growth in yeast. Molecular Cell. 5(5):865-76.

Pearson WR, Lipman DJ. Improved tools for biological sequence comparison. Proc Natl Acad Sci U S A. 1988 Apr;85(8):2444-8.

Seok H, Ham J, Jang ES, Chi SW. MicroRNA Target Recognition: Insights from Transcriptome-Wide Non-Canonical Interactions. Mol Cells. 2016 May 31;39(5):375-81.

Sidik SM, Huet D, Ganesan SM, Huynh MH, Wang T, Nasamu AS, Thiru P, Saeij JP, Carruthers VB, Niles JC, Lourido S. A Genome-wide CRISPR Screen in *Toxoplasma* Identifies Essential Apicomplexan Genes. Cell. 2016 Sep 8;166(6):1423-1435.e12.

Watts FZ. (2013) Starting and stopping SUMOylation. What regulates the regulator? Chromosoma. 122(6):451-63.

CHAPTER 6

Toxoplasma gondii Argonaute: A core protein of RNA-induced silencing complex

6.1. Introduction

RNA-directed gene silencing mechanisms play a critical role in the control of gene expression at the transcriptional, post-transcriptional and chromatin remodeling levels in eukaryotes. Argonaute (Ago) is the core protein found in the RNA-induced silencing complex (RISC) (Fire et al., 1998, Shabalian and Koonin, 2008, Carthew and Sontheimer, 2009, Siomo and Siomi, 2009). A member of the Argonaute family is responsible for guiding small RNAs (siRNAs and miRNAs) to their regulatory target mRNAs. Subsequent RISC assembly leads to (i) degradation of mRNA target, (ii) translational repression or (iii) causing chromatin modifications. The Ago protein family was first identified in plants (Bohmert at al., 1998). Ago proteins were named after a phenotype of a knockout in *Arabidopsis thaliana*, as it resembles the tentacles of the octopus, *Argonaute argo* (Bohmert at al., 1998). Originally, Ago was called "Slicer" based on its nuclease activity to hydrolyze phosphodiester bonds of messenger RNA (mRNA). Based on sequence homology, Ago proteins have been divided into three subfamilies; (i) Ago subfamily, which resembles *A. thaliana* Ago1, (ii) Piwi subfamily (P-element *i*nduced wimpy testis) that is related to *Drosophila melanogaster* Piwi protein and (iii) WAGO which are specific to *C. elegans*. Piwi subfamily members only contain the Piwi domain that includes the canonical catalytic triad, DDH/E and are generally ~40 kDa in size. Members of the Ago subfamily on the other hand, contain two conserved domains called PAZ (Piwi-Argonaute-Zwille) and Piwi and are typically ~100 kDa in size. Among species an organism can contain one or many paralogs of Ago. Fission yeast (*Schizosaccharomyces pombe*) and budding yeast (*Saccharomyces castellii*) both

harbor one gene coding for Argonaute, while fruit fly (*Drosophila melongaster*), humans (*Homo sapiens*), and worm (*Caenorhabditis elegans*), contain five, eight and 27 paralogs, respectively. Generally, multiple paralogs of Argonaute within an organism function within distinct locations (i.e. nucleus or cytosol), vary in nuclease activity or interact with a specific form of non-coding RNA molecule (siRNAs and miRNAs vs. piRNA). In organisms that contain a single Argonaute gene, the protein can be found within the nucleus and cytosol and, participate in transcriptional and post-transcriptional gene regulation mechanisms (Azlan et al., 2016). Not every eukaryote contains an Argonaute gene, for example, baker's yeast (*Saccharomyces cerevisiae*, Jinek and Doudna, 2009).

Among parasitic protozoans, *Trypanosoma brucei* has two genes that code for Argonaute while *Toxoplasma gondii*, *Giardia spp.*, *Entamoeba spp.* and *Trypanosoma. congolense* are predicted to have only one. The presence of Argonaute and small non-coding RNAs amongst these organisms led to interest in developing a genetic tool to target and develop new drugs against parasitic infection (Ullu 2004 and Kolev 2011). *Trypanosoma cruzi*, *Leishmania major*, *Leishmania donovani*, and *Plasmodium falciparum* on the other hand, do not harbor Argonaute coding genes (Ullu 2004 and Kolev 2011).

Our focus is the Apicomplexa parasite *T. gondii* which has infected up to one third of the world's population. Infection in humans and mammals often leads to blindness, encephalitis, severe brain dysfunctions and miscarriages (Kim and Weiss, 2004). The study by Al-Riyahi et al. (2006) showed that *T. gondii* has one Argonaute coding gene with a conserved Piwi domain and a partial PAZ domain (Al-Riyahi et al., 2006). Using RNA Ligase-Mediated Rapid Amplification of cDNA (RLM-RACE),

analysis in type I parasites, TgAgo was found to be encoded by a 2265 nucleotide mRNA which spans five exons that could produce a 58.5 kDa protein. Southern blot analysis determined a single locus for the expression of TgArgonaute. Immunofluorescence assays showed that TgAgo localizes to the cytosol. Despite the smaller size, it was shown that TgAgo is required for double-stranded RNA induced gene silencing even though it lacks a canonical catalytic triad (DDH/E) (Al-Riyahi et al., 2006).

In a later study, Braun et al. (2010) identified a larger TgAgo protein in type II parasites that was predicted to contain additional domains called MID and RGG, and the protein is ~98 kDa (Braun et al., 2010). A full-length HA-Flag-tagged TgAgo overexpression construct composed of the ~7.5 kb genomic sequence was used for the study. They found that HA-Flag-TgAgo had an affinity for small RNAs and polysomes by chromatin immuno-precipitation and chromatographic polyribosome fractionation analysis, respectively. Immunoprecipitation of HA-Flag-TgAgo and mass spectrometry identified several TgAgo interacting proteins involved in RNA metabolism and chromatin-repression complexes, indicating that TgAgo may play a role in translational and transcriptional regulation. HA-Flag-TgAgo was also found to interact with a Tudor-SN (tudor staphylococcal nuclease) and a PRMT1 (arginine methyltransferase), which have been found in the RISC complex of *C. elegans* and *D. melanogaster* (Caudy et al., 2003).

Another study by Musiyenko et al. (2012) investigated the roles of TgAgo in RNA cleavage by antisense guide RNA. A tachyzoite cDNA library from type I *T. gondii* parasites (provided by the AIDS reagent bank), was used as starting material for generating cDNA (2841 nt) used to overexpress an N-terminally Myc-tagged TgAgo.

The construct produced a ~100 kDa Myc-TgAgo. Since TgAgo does not contain the canonical DDH/E catalytic triad, the roles of the interacting proteins TgTSN and TgPRMT1 were investigated in target mRNA cleavage. *In vitro* assays determined that purified Myc-TgAgo alone displayed a weak nuclease activity. When purified Myc-TgAgo and Myc-TgTSN were combined, a greater target mRNA cleavage was achieved (Musiyenko et al., 2012). By *in vitro* methylation assays, it was discovered that TgPRMT1 is required for methylation of the RGG domain of TgAgo, which is used for protein stabilization in addition to recruitment of TgTSN.

The three versions of TgAgo protein are reported in the NCBI protein database (GenBank: ACZ73654.1, ABA70456.1 and ABB05350.1) predicting three possible translational starts sites. Gene ACZ73654.1 identified by Braun et al (2010), contains the first predicted translational start codon in exon 1 (Figure 6.1., ATG1). The predicted coding region spans 7 exons and would produce a ~100 kDa protein which would contain a RGG, PAZ and PIWI domain (protein 1). The second predicted TgAgo, ABA70456.1 identified by Andrews, et al. (2006), has a start codon in exon 3 (ATG 2). Protein 2 would span 5 exons and produce a ~ 82 kDa protein and would contain a PAZ and PIWI domain (Figure 6.1.). The third predicted TgAgo, ABB05350.1 identified by Al-Riyahi et al. (2006) has a start codon in exon 6 (ATG 3). Protein 3 would span 2 exons and produce a ~58 kDa protein and only contain the PIWI domain (Figure 6.1.). The discrepancy could be due to the methodologies used; genomic expression (TgAgo1), a cDNA library (TgAgo2) and RLM-RACE (TgAgo3) or the parasite strains (type I for TgAgo 2 and 3 and type II).

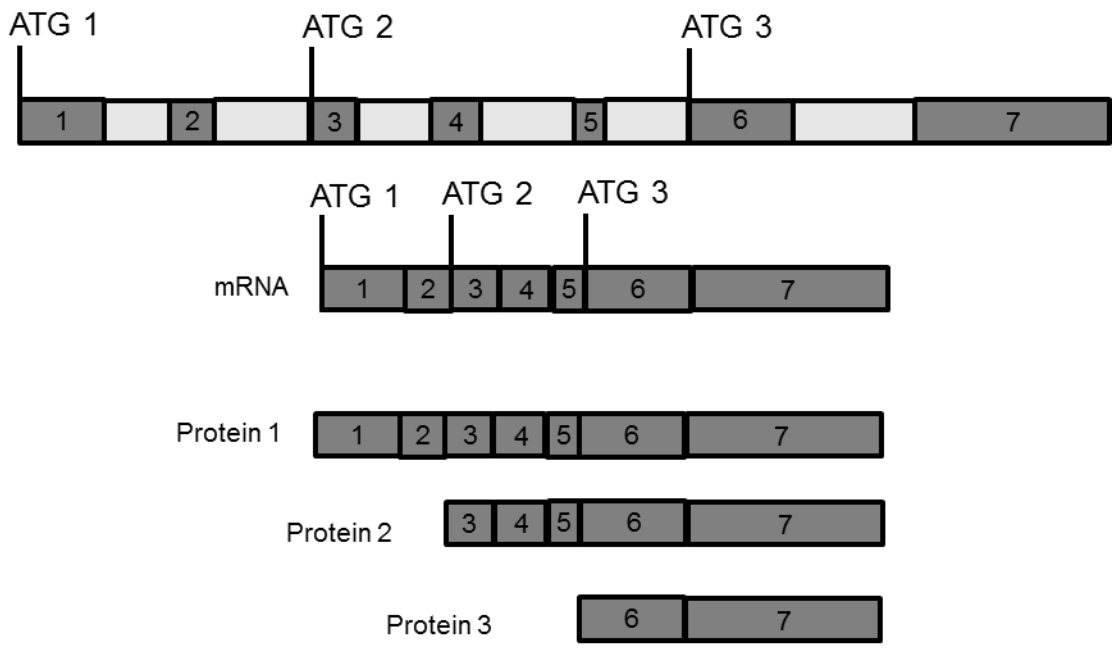


Figure 6.1. TgAgo isoforms could be expressed by multiple start codons to produce three different TgAgo proteins of different sizes.

for TgAgo 1). This study was initiated in an attempt to resolve these discrepancies by determining the expression of the three isoforms of TgAgo and their roles in gene silencing. The role of each TgAgo in gene silencing and the small RNAs they would interact with, could be suggested by the unique domains they contain.

6.2. Materials and Methods

6.2.1. Cell cultures

6.2.1.1. Mammalian cell cultures

Human foreskin fibroblasts (HFF; ATCC, #1041) were grown in Dulbecco's Modified Eagle Media with high D-glucose and L-glutamine (DMEM; Invitrogen, #12100046) supplemented with 10% v/v Cosmic Calf Serum (Hyclone, #H3008704N) and 0.5x Antibiotic–Antimycotic (Invitrogen, #15240-062). Cultures were maintained at 37 °C in 5% CO_2.

6.2.1.2. *Toxoplasma gondii* culture

RHΔHX, T. gondii type I strain, was obtained from NIH AIDS Research and Reference Reagent Program (#2857); Δku80 strain, which is an RHΔHX-derived strain, was obtained from Dr. Vern B. Carruthers (University of Michigan). Unless stated otherwise, both strains were maintained in HFF using Minimum Essential Medium (MEM, Invitrogen, #61100061) supplemented with 1% v/v dialyzed fetal bovine serum (ThermoFisher Scientific, Hyclone, #SH3007903) and diluted (0.5x) Antibiotic-Antimycotic and maintained at 37 °C in 5% CO_2. Selection of transgenic parasites was grown in culture media containing either, mycophenolic acid (25 μg/mL) and xanthine (50 μg/mL) or pyrimethamine (1 μM), for hypoxanthine-xanthine-guanine

phosphoribosyltransferase (HXGPRT) or dihydrofolate reductase (DHFR) resistance, respectively.

6.2.2. Plasmid construction

Transforming plasmids were constructed using pSF-TAP-LIC-HX and pChFP-TAP-LIC-DHFR (obtained from Dr. Vern B. Carruthers, University of Michigan). Each plasmid was linearized by PacI digestion prior to the treatment with T4 DNA polymerase to create compatible ends in the presence of dGTP for a ligation-independent cloning procedure as described by Huynh and Carruthers (2009). TgAgo (1565 bp) amplicon was prepared using gene specific primer pair of TgAgoLICFw/TgAgoLICRv (Table 6.1.), and purified prior to the treatment with T4 DNA polymerase to create compatible ends in the presence of dCTP. Equimolar concentrations of the prepared plasmid and insert were allowed to anneal and subsequently used in bacterial transformations. The resultant plasmids, pAgo-SF-TAP-LIC-Hx and pAgo-ChFP-TAP-LIC-DHFR were analyzed to confirm their sequences via restriction endonuclease digestions. The plasmids were then used in the transformation of *T. gondii*. Transgenic parasites were selected and grown in culturing media containing mycophenolic acid (25 μg/mL) and xanthine (50 μg/mL) or pyrimethamine (1μM).

6.2.3. Dual luciferase assay

For dual luciferase assays, 48 hours post-transfection with 20 μg of Ago expression plasmid, freshly lysed parasites were harvested and lysed with 100 μL of 1x Passive Lysis Buffer (Promega, #E1531) and incubated for 10 minutes at room temperature. Lysates were cleared of debris by centrifugation (10,000 xg for 1 minute),

Table 6.1. Oligonucleotide primers used in the study of TgAgo. Lower case nucleotides were used for LIC cloning.

	Nucleotide sequence
TgAgoLICFw	tacttccaatccaatttaatgcCTCCCCGAAATCGCGTGG
TgAgoLICRv	tcctccacttccaattttagcACAGAACACCATCGGGGTCACC
a	ATCTGGTATTCCG
a’	CGAACACCATCGG
b	ATGGCGTCCAAACCCATTGAA
b’	ACCGGTGTCGACGTCCTC
c	ATGGAGGGATACCAAATGTTCC
c’	CGG GGC TGAATGTCCCTG

and the supernatant was kept on ice and used in the dual luciferase assay, which was carried out in a 2-step fashion as previously described by Crater et al. (2015). Luminescence signals were measured with a 20/20n Luminometer (Turner Biosystems). Rnluc signals were normalized to Ffluc signals for a direct comparison among different independent experiments. The measurement of the dual luciferase expression was tested under at least three independently trials.

6.2.4. Western blot analysis

Cell-free lysates were prepared by the addition of RIPA buffer to collected parasites and sonicated for 30 seconds (x3). Protein concentration was measured using the RC DC Protein Assay according to manufacturer's instructions (Bio-Rad Laboratories, #500-0119). Approximately 20 μg was used for the analysis, of which the lysates were denatured using 6x SDS protein loading buffer and resolved on a 10% SDS-PAGE. The primary antibodies used were α-Flag mouse antisera (1:5,000, Applied Biological Materials Inc., Richmond, B.C, Canada, #G191). The signals were revealed using secondary antibodies: horseradish peroxidase (HRP) conjugated goat anti-mouse antisera (Rockland Immunochemicals Inc., #610-1319) at a 1:10,000 dilution, and a Chemiluminescent HRP Substrate Kit (Millipore, #WBKLS0500).

6.2.5. Immunocytochemistry

HFF monolayers were cultured on coverslips until confluent, infected with ~1 x 10^6 freshly lysed parasites and incubated at 37 °C and 5% CO_2. Infected monolayers on the cover slips were rinsed three times with 1x phosphate buffer saline (PBS), fixed with 3% paraformaldehyde for 5 minutes. Cells were then incubated with Hoechst staining solution (10 μM) at room temperature for 5 minutes, washed twice with 1x PBS and

mounted on glass slides. Image acquisition was carried out using a Leica DMI 6000B inverted fluorescent microscope using HCX PL Apo 100×/1.40–0.70 objective and a Leica DFC 360FX camera with Leica Advanced Fluorescence Application Software V.2.3.0. Post-acquisition image processing was performed with Adobe Photoshop 7.0.

6.2.6. Statistical analysis

Statistical analysis for TgAgoKO gene silencing rescue was completed by the paired student t-test methodology assuming a null hypothesis. All values were calculated from at least three independent tests.

6.3. Results

6.3.1. Multiple proteins from one locus

To begin our investigation whether *T. gondii* expresses multiple TgAgo proteins, we generated transgenic parasites by homologous recombination to have the C-terminus of TgAgo linked to a Strep/Flag (SF) epitope or a cherry fluorescent protein (ChFP). These parasite lines would express TgAgo at the endogenous level and at its natural regulation. To verify TgAgo-SF expression and identify multiple isoforms, SF-tagged parasites were monitored by western blot analysis using anti-Flag antisera (Figure 6.2a). TgAgo-SF expression was monitored over culturing passages to ensure only transgenic parasites were studied. At passages one and two post-transfection with pTgAgo-SF-TAP-LIC-Hx, a band appears at ~58 kDa as reported by Al Riyahi et al. (2006). By passage three another band at ~97 kDa appears, as reported by Braun et al. (2010). A band also appears at ~34 kDa in all passages and in WT parasite lysates (Figure 6.2a). The 58 and 97 kDa proteins are not revealed in WT lysates, which suggests two forms of TgAgo-SF could be produced from the same locus. Localization of TgAgo-ChFP, after at least three passages, was found to be primarily cytosolic, which was in agreement with previous reports (Figure 6.2b. Al Riyahi et al., 2006, Braun et al., 2010, Musiyenko, et al., 2012).

6.3.2. TgAgo is required for gene silencing

To study the effects of TgAgo in gene silencing mechanism, a TgAgo knockout parasite strain from Dr. JC Boothroyd (Stanford University School of Medicine, Stanford, California) was used. The coding region of TgAgo was replaced with the hypoxanthine-xanthine-guanine-phosphoribosyltransferase (HXGPRT) gene, and selected for by

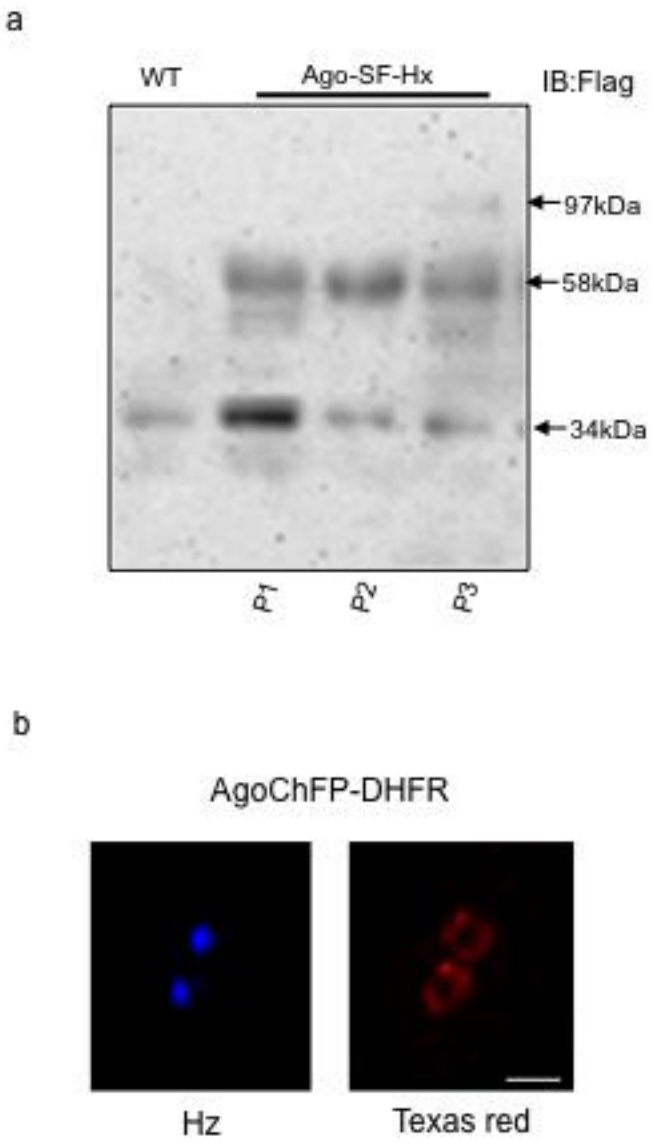


Figure 6.2. Multiple TgAgo proteins. A) Western blot analysis of TgAgo-SF expression at progressive culturing passages (p). Total parasite lysates were analyzed for Flag expression up to 6 days post-transfection. B) TgAgoChFP localization in intracellular parasites. (Hz, Hoechst, Texas red for cherry fluorescent protein. Micrometer bar represents 5 μm.)

mycophenolic acid (25 μg/mL) and xanthine (50 μg/mL). Gene replacement was confirmed by PCR analysis using primers in Table 6.1. Based on their annealing sites, primer sets, a-b', b-a', and b-b' would only produce a PCR product if the HXGPRT gene was properly integrated into the TgAgo locus at 1, 1.4, and 0.5 kb, respectively. Primer set c-c' would anneal within the coding region of TgAgo and was only present in WT parasites (Figure 6.3.).

We used our dual luciferase assay where binding sites for the endogenous miR-4a were located in the 3' UTR of the Rnluc transcript, as described by Crater et al., 2015. Three perfect match sites to miR-4a were used in reactions labeled A, while three mismatch sites were used in reactions labeled B. Rnluc activity decreased to (A) 39 ± 2.0 and (B) 35 ± 4.0 % when normalized to Rnluc activity without binding sites in WT parasites. The same reactions were carried out in TgAgoKO parasites, Rnluc activity was (A) 110 ± 11 and (B) 70 ± 4.0 % (Figure 6.4). Plasmids that could express an Ago protein were introduced via electroporation with the dual luciferase plasmids in an attempt to decrease Rnluc activity within the TgAgoKO parasites. When human Ago 2 (HsAgo2) was introduced, Rnluc activity decreased to (A) 67 ± 4.0 and (B) 36 ± 9.0%. Transfection of the plasmid that would express the 58 kDa, (TgAgo1) decreased Rnluc activity to (A) 87 ± 5.0 and (B) 70 ± 6.0 %. Upon transfection with the construct that contained the genomic sequence of TgAgo (TgAgo2) slightly deceased Rnluc activity to (A) 69 ± 11 and (B) 67 ± 6.0 %. As a control to ensure the decrease in Rnluc activity was due to the introduced plasmid, a GFP expressing plasmid was also tested, Rnluc activity was (A) 110 ± 9.0 and (B) 92 ± 7.0 % (Figure 6.4.).

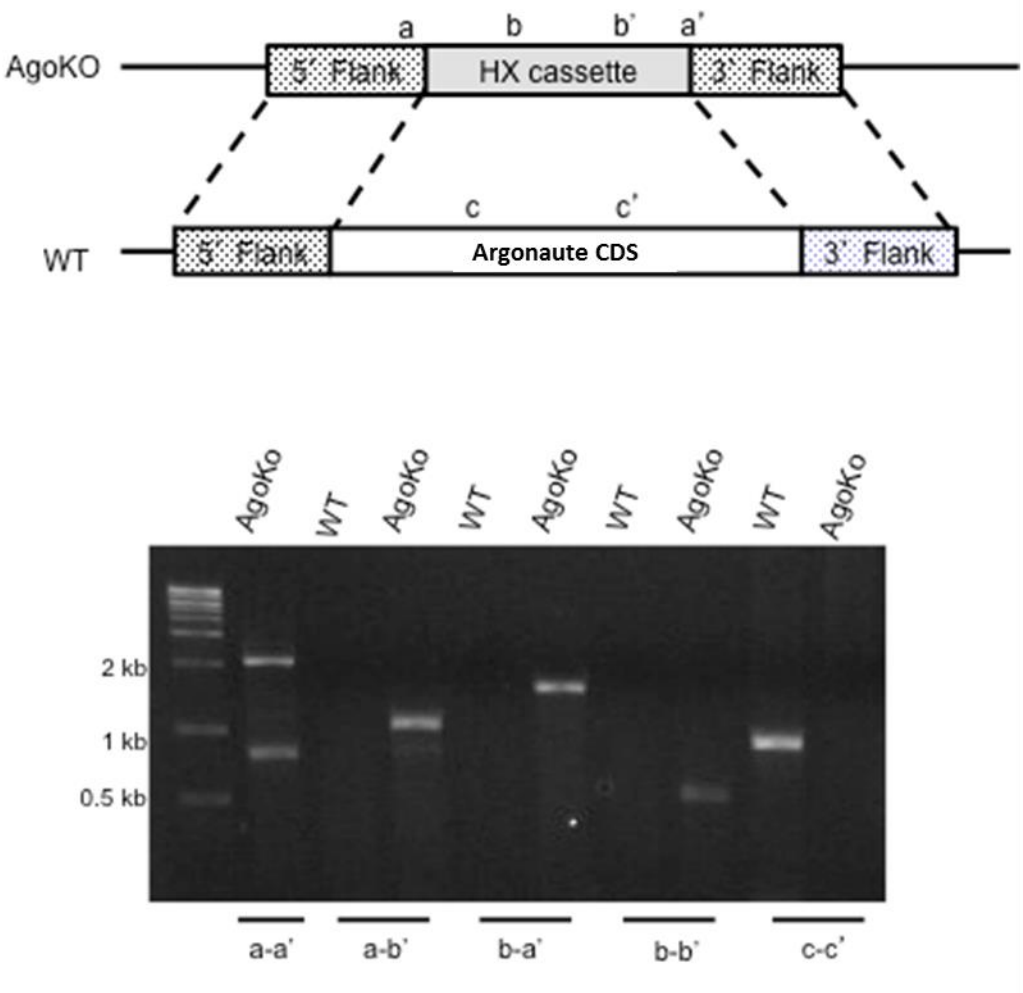


Figure 6.3. Confirmation of TgAgo knockout parasite strain. The coding region of TgAgo was replaced with the HXGPRT gene. Proper gene replacement was confirmed by genomic RNA PCR analysis in TgAgoKO parasites compared to wild-type (WT).

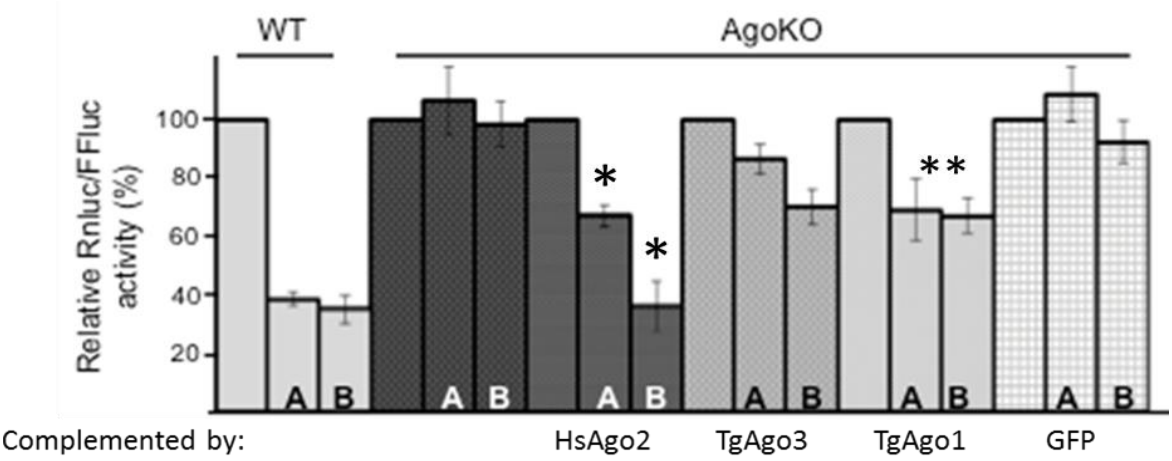


Figure 6.4. Argonaute gene silencing rescue. TgAgo silencing capability was tested with a dual luciferase reporter assay within WT and TgAgoKO parasites. The Rnluc transcripts tested would carry either a perfect match (A) or mismatch (B) interaction for miR-4a (Crater et al., 2015). Rnluc activity in TgAgoKO parasites remained unchanged when compared to Rnluc activity when no miRNA binding sites were tested. A rescue of a silencing effect was observed upon transfection of a plasmid that would express the catalytic active human Ago2 (HsAgo2). When any of the TgAgo expressing plasmids were tested (TgAgo1, TgAgo3), a decrease of Rnluc activity was achieved. A plasmid that would express GFP did not affect Rnluc activity in TgAgoKO parasites. (p-values are the change in TgAgoKO silencing when complemented by a form of Ago for each miRNA binding format from three independent trials, * $p \leq 0.01$ and ** $p \leq 0.02$)

6.4. Discussion

Argonaute is a central protein for which RNA directed gene regulation occurs. It has the ability to confer gene regulation at the transcriptional and post-transcriptional level. Within *T. gondii*, the roles and functions of TgAgo have yet to be fully eluded. Previously, our lab tested the TgAgoKO parasite line for its ability to (i) convert to bradyzoites, (ii) its effects on RNAseIII protein expression and (iii) knockdown a gene of interest (Liwak, 2009). First, TgAgoKO parasites were grown under alkaline conditions for 5 days to induce bradyzoite formation. The cyst walls were detected by immunofluorescence with a *Dolichols*-FITC antibody and at least 100 random vacuoles were counted. Under these conditions, ~70 % of the TgAgoKO parasites formed cysts while only ~14 % of the WT (RH) parasites formed cysts. This suggests TgAgo may play a role in tachyzoite to bradyzoite interconversion. Using an anti-RNAseIII antisera and western blot analysis, it was found that the expression of RNAseIII containing proteins was decreased in the TgAgoKO parasites when compared to WT. Finally, dsRNA for the gene UPRT (uracil phosphoribosyltransferase) was used to test the gene silencing ability in WT and TgAgoKO parasites. UPRT is responsible for the uptake of uracil in the parasites; therefore a radio labeled uracil was used to detect UPRT activity. In WT parasites, the maximal amount of knockdown was ~60 %, however, the TgAgoKO parasites showed no ability to knockdown UPRT. Overexpression of TgAgo3 in these parasites was unable to rescue a knockdown effect (Liwak, 2009). The inability to rescue a gene silencing effect in the TgAgoKO parasites may be related to the type of dsRNA used as an effector. In these experiments the dsRNA was ~ 600 bp. As previously described in Chapter 2, the amount of an induced gene silencing effect could be related to the length or type of RNA. Although a gene silencing effect can be achieved with long

dsRNA, a greater silencing effect was observed when a small dsRNA with a mismatch interaction to the mRNA target was used (Crater at al., 2012).

The presence of two forms of Ago expressed from the same locus in *T. gondii* suggests TgAgo1 and TgAgo3 may be the result of two translational start sites. To determine if ATG1 and ATG3 are translational start codons, we could first use the CRISPR/Cas9 methodology to disrupt each predicted start codon separately in TgAgo-SF parasites. TgAgo-SF expression could be determined by western blot analysis. If ATG1 and ATG3 are bona fide translational start sites, we would expect to see only one isoform expressed (from the site that was not disrupted). An increase in protein expression from the TgAgo isoform could also be possible under these conditions. Next, we could use our dual luciferase reporter system with Tg-miRNA binding sites to monitor gene silencing by (i) Rnluc activity and (ii) qRT-PCR analysis. TgAgo expressed from ATG1 would have the RGG domain required for TSN association. Upon TSN association to TgAgo, we would expect gene silencing to occur by mRNA degradation. TgAgo expressed from ATG3 has only a PIWI domain and we would expect gene silencing to occur by translational repression.

Overall, these data are the first to suggest that *T. gondii* expresses at least two isoforms of Ago that could be required for gene silencing. The mode of gene silencing, mRNA degradation or translational repression, could be determined by the TgAgo isoform. Further investigations into the roles of TgAgo are thereby necessary to understand this unique gene.

6.5. Acknowledgements

This work was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) and the University of Windsor Doctoral Entrance Scholarship to ACP.

6.6. References

Al Riyahi A, Al-Anouti F, Al-Rayes M, Ananvoranich S. Single argonaute protein from *Toxoplasma gondii* is involved in the double-stranded RNA induced gene silencing. Int J Parasitol. 2006 Aug;36(9):1003-14.

Andrews,J., Adams,B., Musiyenko,A. and Barik,S. Two functional Argonaute proteins that differ by a single amino acid mediate RNA interfernce in *Toxoplasma gondii*. 2006. Unpublished.

Azlan A, Dzaki N, Azzam G. Argonaute: The executor of small RNA function. J. of Genetics and Genomics. 2016. 481-494.

Bohmert, K., Camus, I., Bellini, C., Bouchez, D., Caboche, M.and Benning, C. AGO1 defines a novel locus of Ara-bidopsis controlling leaf development. The EMBO Journal, 1998. 17, 170–180

Braun L, Cannella D, Ortet P, Barakat M, Sautel CF, Kieffer S, Garin J, Bastien O, Voinnet O, Hakimi MA. A complex small RNA repertoire is generated by a plant/fungal-like machinery and effected by a metazoan-like Argonaute in the single-cell human parasite *Toxoplasma gondii*. PLoS Pathog. 2010 May 27;6(5):e1000920.

Carthew RW, Sontheimer EJ. Origins and Mechanisms of miRNAs and siRNAs. Cell. 2009 Feb 20;136(4):642-55.

Caudy AA, Ketting RF, Hammond SM, Denli AM, Bathoorn AM, Tops BB, Silva JM, Myers MM, Hannon GJ, Plasterk RH. A micrococcal nuclease homologue in RNAi effector complexes. Nature. 2003 Sep 25;425(6956):411-4.

Crater AK, Cherry A, Holmes M, Kadri D, Ananvoranich. Evaluation of the ability of long and short double-stranded RNAs to induce a gene silencing effect in the parasitic protozoan *Toxoplasma gondii*. 2012. Am. J. Biomed. Sci. 2012 Jan 4(1), 1-13.

Crater AK, Manni E, Ananvoranich S. Utilization of inherent miRNAs in functional analyses of *Toxoplasma gondii* genes. J Microbiol Methods. 2015 Jan;108:92-102.

Faehnle CR, Joshua-Tor L. Argonaute MID domain takes center stage. EMBO Rep. 2010 Aug;11(8):564-5.

Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. Nature. 1998 Feb 19;391(6669):806-11.

Ender C, Meister G. Argonaute proteins at a glance. J Cell Sci. 2010 Jun 1;123(Pt 11):1819-23.

Huynh MH, Carruthers VB. Tagging of endogenous genes in a *Toxoplasma gondii* strain lacking Ku80. Eukaryot Cell. 2009 Apr;8(4):530-9.

Kim K, Weiss LM. *Toxoplasma gondii*: the model apicomplexan. Int J Parasitol. 2004 Mar 9;34(3):423-32.

Jinek M, Doudna JA. A three-dimensional view of the molecular machinery of RNA interference. Nature. 2009 Jan 22;457(7228):405-12.

Liwak, U. *Toxoplasma gondii* differentiation: effects of modifying lactate dehydrogenase and argonaute expression patterns. 2009. Electronic Theses and Dissertations.

Kolev NG, Tschudi C, Ullu E. RNA interference in protozoan parasites: achievements and challenges. Eukaryot Cell. 2011 Sep;10(9):1156-63.

Musiyenko A, Majumdar T, Andrews J, Adams B, Barik S. PRMT1 methylates the single Argonaute of *Toxoplasma gondii* and is important for the recruitment of Tudor nuclease for target RNA cleavage by antisense guide RNA. Cell Microbiol. 2012 Jun;14(6):882-901.

Shabalina SA, Koonin EV. Origins and evolution of eukaryotic RNA interference. Trends Ecol Evol. 2008 Oct;23(10):578-87.

Siomi H, Siomi MC. RISC hitches onto endosome trafficking. Nat Cell Biol. 2009 Sep;11(9):1049-51.

Ullu E, Tschudi C, Chakraborty T. RNA interference in protozoan parasites. Cell Microbiol. 2004 Jun;6(6):509-19.

CHAPTER 7

Discussion and Future goals

7.1. Mechanism of RNA induced gene silencing

7.1.1. Roles and mechanism of miRNAs

These studies have been conducted to understand RNA directed gene regulation mechanism in *Toxoplasma*. The effectiveness of endogenous miRNA is directly related to (i) the abundance of miRNA and (ii) the number of miRNA binding sites within a target transcript. The mode of regulation appears to vary among miRNA families. The miR60 family causes mRNA degradation, while the miR-4 family causes translational repression. Gene expression can also be manipulated by the use of a miRNA inhibitor or mimic. This work shows that miRNA-induced gene silencing could also become a valuable tool in the study of *Toxoplasma*. By using a combination of known genetic methodologies, we took advantage of homologous recombination system offered by the ΔKU80 strain that lack non-homologous end joining (NHEJ) capability to place miR-60a binding sites within the 3' UTR of an essential (LDH1) and nonessential (HoDI) gene to down regulate their expression and monitor the effects. With the use of the miRNA inhibitors and mimics the gene silencing capabilities of miR-60a could be altered and the effects evaluated.

In order to further elucidate the mechanism for RNA induced silencing in *Toxoplasma*, we would need to also investigate the factors involved and underlying mechanism. The presences of putative Dicer and RdRp gene have been predicted in *Toxoplasma* (Braun et al., 2010). The silencing activities of Argonaute, Tudor and PMRT1 have been studied (Musiyenko et al., 2012). However, the interactions among

these proteins, small effector RNAs (siRNA and/or miRNA) and target transcripts collective roles and functions have not yet been studied. Since protein-protein and RNA-protein interactions are vital for a functional gene silencing mechanism, we could employ a BioID and RNA aptamer methodology to examine these factors. BioID is a method to identify proximal and direct protein-protein interactions in a cell system. This is accomplished by the fusion of a biotin ligase enzyme (BirA) to a protein of interest to either the N- or C-terminus. Addition of excess biotin to the cell system leads to biotinyation of proteins either adjacent or directly interacting with the protein of interest (Kim et al., 2016). Biotinyated proteins are then purified on streptavidin-conjugated agarose beads. Interacting and proximal proteins can be identified *via* fluorescence microcopy, western blot analysis or by mass spectrometry (Fairhead and Howarth, 2015).

A methodology to study protein-RNA interactions is by use of the MS2 RNA aptamer to purify RNA-protein complexes (Slobodin and Gerst, 2010). A transcript that contains two MS2 loop repeats would bind to a recombinantly expressed coat protein (CP). The MS2-CP is linked to a streptavidin binding protein that would bind to streptavidin-conjugated beads (Figure 7.1.) similarly to the BioID method, interacting proteins could be identified *via* mass spectroscopy. Since previous results indicate that miR-4a causes translational repression, this methodology could be used with the *Rnluc* transcripts that contain the different miRNA binding formats (perfect vs. mismatch). This would also allow for us to compare the possible different or similar proteins involved in silencing when a perfect match or mismatch interaction is involved. We could expect to find translational machinery proteins as well as RISC components such as TgAgo. Since target mRNA cleavage is not the cause of silencing we would not expect

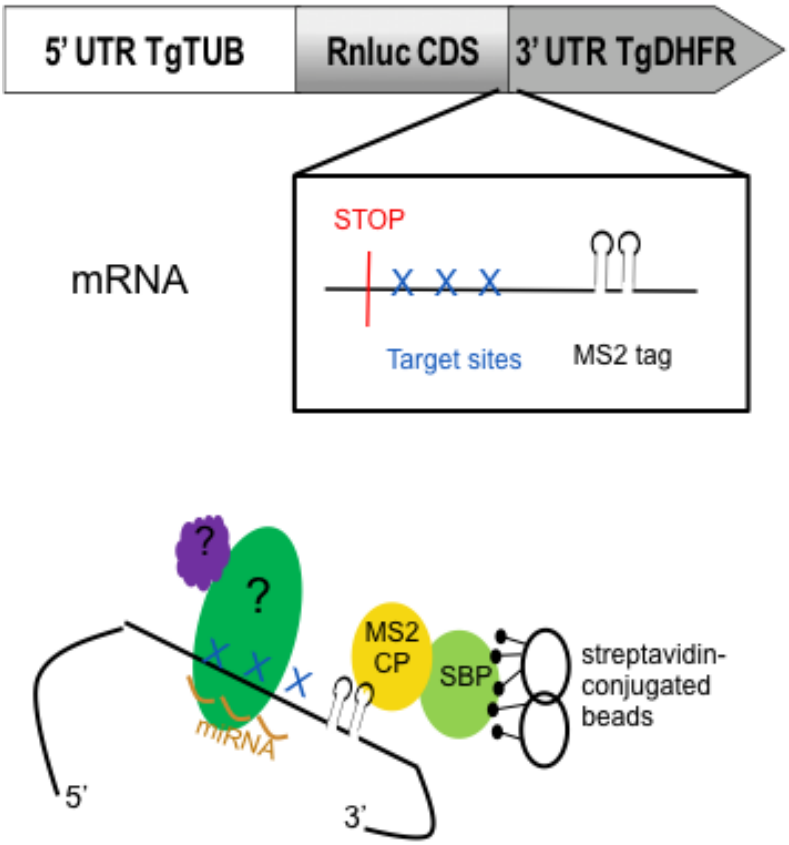


Figure 7.1. A possible approach for determining RNA induced gene-silencing factors by MS2 RNA aptamer. The *Rnluc* transcript would contain the sequence for the MS2 loop following three *Toxoplasma* miRNA binding sites (miR-60a or miR-4a) in the 3' UTR. The secondary structure of the MS2 loops would bind to a MS2 coat protein linked to a streptavidin binding protein that would allow for complex purification by streptavidin-conjugated beads.

to find TgTudor interactions as it appears to be involved in transcript degradation (Musiyenko et al., 2012). Furthermore, the isoform of TgAgo involved could also be identified. As TgTudor is not expected to be involved in translational repression we would expect the smaller isoform of TgAgo to be involved as it does not possess the RGG domain required for TgTudor interaction (Al Riyahi et al., 2006). One possible limitation of the MS2 aptamer would be to elude the proteins involved in mRNA degradation. The transcripts may not be sustained long enough to purify the interacting complex. To test this, we could use the *Rnluc* transcript that contains miR-60a binding sites as previous reports indicate target mRNA degradation occurs for this interaction.

7.1.2. Self-regulating genes and TgUlp1

The most well defined example of a self-regulating gene is the Xist gene involved in cellular differentiation in female mammalians. Tsix is the non-coding RNA gene that is transcribed antisense to the Xist RNA to induce random x inactivation by binding to Xist RNA (Chow et al., 2005). Another example is the *hok/sok* system in *E. coli* during cell division. Expression of the hok gene creates a toxic protein that causes cell death by depolarization of the cell membrane. To prevent accumulation of this deadly protein the *sok* RNA is transcribed from antisense strand to the hok gene. *Sok* regulates hok expression by binding within the *hok* transcript open rereading frame to induce translational repression. The subsequent RNA duplex is recognized by an RNAseIII enzyme and is degraded (Gerdes et al., 1990). These examples of self-regulating genes are both important for cellular differentiation and development. Self-regulating genes are also advantageous when a fast response to external signals is required during stress response (Shimoni et al., 2007). Most self-regulating genes are regulated by a

transcription factor by either a negative or positive feedback loop. For example, human miR-483 has been shown to activate its expression *via* promoter activation through the up-regulation of a transcription factor, USF1 (Emmerling et al., 2016). Self-regulating genes by non-coding RNAs has been observed in other Apicomplexans. For example, the *var* gene family in *Plasmodium*, which is responsible for host cell immune evasion, is regulated by an RNA transcribed from the antisense strand of a conserved intron within the *var* genes (Amit-Avaham et al., 2015). However, at the time of this dissertation, no reports of self-regulating genes by miRNAs have been identified in Apicomplexans. As a self-regulating gene, and as evident by the parasite's inability to complete its lytic cycle asexually, suggests TgUlp1 is likely also important for parasite development and immune response.

7.1.3. Alternative translational starts sites of TgAgo

Initiation of protein translation is initiated by the recruitment of a ribosome at a specific nucleotide sequence called the ribosomal-binding site (RBS). RBS recruitment can be mediated by the 5' cap of mRNAs or in a cap independent mediated mechanism *via* internal ribosome entry sites (IRES). In eukaryotes, the region upstream of the start codon on the mRNA has a consensus sequence (ACCAUGG) called the Kozak sequence. Multiple adenines upstream of the RBS can also enhance ribosomal recruitment (Hellen and Sarnow, 2001, Laursen et al., 2005). Alternative translational start sites within a single transcript can contribute to proteome diversity. Preferential translation is attributed to longer 5' UTRs that are enriched for transcription factors and leaky mRNA scanning. Although the functional significance for alternative start sites is controversial, this mechanism can lead to the production of N-terminal protein variants to promote

diversity in higher eukaryotes (Bazykin and Kochetov, 2011). For example in *A. thaliana*, the mitochondrial and nuclear form of the DNA ligase 1 (AtLIG1) are both translated from a single mRNA from two in-frame start codons (AUG). Similarly, mitochondrial and cytoplasmic isoforms of proteins have been observed in rat with the ornithine decarboxylase-antizyme (Gandre et al., 2003), and the human insulin-degrading enzyme (Leissring et al., 2004) and neuropeptide Y (Kaipio et al., 2005).

The expression of multiple Ago proteins in higher eukaryotes (i. e. human, fruit fly and worm) function within distinct locations (i.e. nucleus or cytosol), vary in nuclease activity or interact with a specific form of non-coding RNA molecule (Azlan et al., 2016). Organisms that contain a single Ago (baker's yeast) can participate in transcriptional and post-transcriptional regulation within the nucleus and cytosol (Jinek and Doudna, 2009). Although three versions of TgAgo can be found within the NCBI protein database, the expression of only two forms has been experimentally identified (GenBank: ACZ73654.1 and ABB05350.1). A ~100 kDa protein that would contain the RGG, PAZ and PIWI domain and ~58 kDa protein that only contains the PIWI domain (Braun et al., 2010 and Al Riyahi et al., 2006).

Toxoplasma contains a Kozak-like sequence (gncAaaATGg, Seeber, 1997). Upon investigation of the TgAgo locus for possible RBS locations, six possible locations were identified with 4 locations within the predicted exons and 2 in the predicated introns (Table 7.1.). Based on sequence similarities and location of the predicted sites, the identified site on exon 6 is the most likely to be a translational start site and corresponds to the ATG3 start site. The protein that would be produced from this site corresponds to

Table 7.1. *Toxoplasma* Kozak-like sequence and possible translational start site on the TgAgo locus. Nucleotide numbers are based on the reported ATG1 start site. Upper case indicates essential nucleotides for RBS recruitment and translation. Perfect match nucleotides to the *Toxoplasma* Kozak-like sequence are indicated by *.

Toxoplasma Kozak-like sequence	gncAaaATGg
Location on TgAgo locus (TGGT1_310160)	**Sequence**
Exon 1 -6-4	tctcttATGa * ***
Intron 4 2418-2428	acgAaaATGg * ******
Exon 6 3490-3500	gccAaaATGt ********
Intron 6 4313-4323	gatAaaATGt ** *****
Exon 7 4593-4603	ggaAatATGg ** ** ****
Exon 7 5058-5068	cgcgacATGg ** * ****

TgAgo3 at ~ 58 kDa. The first predicted start site (ATG1) has the lowest sequence similarity to the Kozak-like sequence and is the least likely possible start site (Table 7.1.). Although western blot analysis indicates the presence of two isoforms of TgAgo, predicted translational start sites indicate one likely start site in exon 6. Taken together, these findings could suggest *Toxoplasma* contains a leaky translation mechanism that produces two forms of TgAgo where one could function within the nucleus and the other in the cytosol.

7.2. MiRNAs as an anti-*Toxoplasma* therapeutic

During host cell invasion by *Toxoplasma*, parasite specific secretory organelles (micronemes and rhoptries) inject effector molecules into the host. These effectors have been shown to interfere with host cell signaling to block apoptosis and progression of the host's cell cycle. Additionally infected host cell expression profiling has shown 15 % of host mRNAs display an alteration (Blader et al, 2001, Nelson et al., 2008). A study used microarray profiling to analyze the change in host cell miRNA expression upon *Toxoplasma* infection. Primary transcripts for miR-17~92 and miR-106b~25 display an increased abundance in infected human foreskin fibroblasts. The mature miRNAs derived from these primary transcripts target mRNAs that encode for pro-apoptotic factors and negative regulators of the cell cycle (Zeiner et al., 2010). Furthermore, a study showed an increase in miRNAs that target immune response factors in *Toxoplasma* infected mouse brains (Cannella et al., 2014). The exact mechanism for this change is not fully understood however, a study by Zeiner et al., (2010) showed the increase the miR-17 ~ 92 primary transcripts is likely due to an increase in promoter activity. The fact that parasite or host miRNAs have not been found locationally different may not infer an exchange of host-parasite miRNAs is not occurring, but the level of change or

location is not currently detectable.

Although systematic siRNA therapies have been used in the clinic treatment of cancer, many different barriers can inhibit their delivery due to unfavorable physicochemical properties and instability of the siRNA (Roberts 2016). To prevent rapid degradation of the siRNA, modified backbones such as those like locked nucleic acids, morpholinos and antagmers are preferred. Non-viral delivery of siRNAs is preferred to prevent potential toxicities associated with viral vectors. Non-viral siRNA vectors involve complexing the siRNA with a positively charge polymer or lipid, conjugation of the siRNA to small molecules such as cholesterol and lipids or encapsulating the siRNA in a nanoparticle (Roberts et al., 2016, Wang et al., 2010). The administration depends on the organs targeted. Generally, locoreginal delivery, such as inhalation is preferred as it has fewer barriers to overcome when compared to systemic delivery, such as intravenous injection (Wang et al., 2011).

Similar methodologies have been utilized to treat intracellular tachyzoites and bradyzoites *in vitro* and *in vivo*. A transductive peptide conjugated to a phosphorodiamidate morpholino oligomer (PMO) was used sequence specifically to inhibit the expression of *Toxoplasma* dihydrofolate reductase (DHFR) and enoyl-ACP reducatase (ENR) enzymes, which are involved in folate and fatty acid synthesis, respectively. Intracellular parasite replication was slowed when either of the mentioned genes was targeted by the peptide-conjugated morpholino (PPMO) in cell culture. The methodology also showed parasite burden was reduced in infected mice that were treated with the DHFR-specific PPMO (Lai et al., 2012). As another possible treatment for anti-*Toxoplasma* therapeutics, it would be interesting to investigate the use of the *Toxoplasma*

specific miRNAs or inhibitors with this methodology. The experiments discussed found a maximal amount of miRNA mimic or inhibitor required for an alerted silencing effect was 20 μg, which translates to 1.23 and 2.35 mM respectively. However, the maximal concentration of the PPMOs discussed earlier was 5 μM indicating treatment by PPMO may have a greater effect at lower concentrations, which would be more advantages to eliminate toxicity and off-targeting effects (Lai et al., 2012). *Toxoplasma* miRNAs have no reported mRNA targets or sequence homology to host miRNAs therefore we would not expect to find an alteration in host gene expression.

7.3. Conclusion and significance

This research has been dedicated to studying the underlying mechanism of RNA induced gene silencing in *Toxoplasma.* Gene regulation is vital for our understanding of the biological processes required for *Toxoplasma* survival and opportunistic life. The first goal was to determine the effectiveness of different effector RNA molecules to induce gene silencing. Next, the roles and effectiveness of endogenous miRNA in *Toxoplasma* were examined. The best gene silencing effect was observed based on the abundance of miRNA present or the number of binding sites on a target transcript. Lastly, TgUlp1 was chosen as a model gene to study the underlying gene silencing mechanism and the biological roles of Tg-miRNAs. In fact, TgUlp1 has proven an excellent model gene to study RNA induced gene silencing. The locus of TgUlp1 contains the highest number of miR-60a binding sites and contains possible precursor miR-60a. It is very likely that–TgUlp1 is a self-regulating gene. Furthermore, the expression of TgUlp1 is critical for SUMOylation and parasite egress. Overall, this dissertation has revealed the importance of proper gene regulation by miRNAs in

Toxoplasma. An alteration in gene expression can be detrimental to the parasite. With these studies, exanimation of genes for quick functional analysis for the development of drugs or use of the Tg-miRNA mimics and inhibitors alone as a drug treatment for *Toxoplasma* infection could be within reach.

7.4. References

Al Riyahi A, Al-Anouti F, Al-Rayes M, Ananvoranich S. Single argonaute protein from *Toxoplasma gondii* is involved in the double-stranded RNA induced gene silencing. Int J Parasitol. 2006 Aug;36(9):1003-14.

Amit-Avraham I, Pozner G, Eshar S, Fastman Y, Kolevzon N, Yavin E, Dzikowski R. Antisense long noncoding RNAs regulate var gene activation in the malaria parasite *Plasmodium falciparum*. Proc Natl Acad Sci U S A. 2015 Mar 3;112(9):E982-91.

Azlan A, Dzaki N, Azzam G. Argonaute: The executor of small RNA function. J Genet Genomics. 2016 Aug 20;43(8):481-94.

Blader IJ, Manger ID, Boothroyd JC. Microarray analysis reveals previously unknown changes in *Toxoplasma gondii*-infected human cells. J Biol Chem. 2001;276:24223–24231.

Braun L, Cannella D, Ortet P, Barakat M, Sautel CF, Kieffer S, Garin J, Bastien O, Voinnet O, Hakimi MA. A complex small RNA repertoire is generated by a plant/fungal-like machinery and effected by a metazoan-like Argonaute in the single-cell human parasite *Toxoplasma gondii*. PLoS Pathog. 2010 May 27;6(5):e1000920.

Cannella D., Brenier-Pinchart M. P., Braun L., VanRooyen J. M., Bougdour A., Bastien O., et al. . (2014). MiR-146a and miR-155 delineate a microRNA fingerprint associated with *Toxoplasma* persistence in the host brain. Cell Rep. 6, 928–937. 10.1016/j.celrep.2014.02.002

Chow JC, Yen Z, Ziesche SM, Brown CJ (2005). Silencing of the mammalian X chromosome. Annual Review of Genomics and Human Genetics. **6**: 69–92.

Emmerling VV, Fischer S, Kleemann M, Handrick R, Kochanek S, Otte K. miR-483 is a self-regulating microRNA and can activate its own expression via USF1 in HeLa cells. Int J Biochem Cell Biol. 2016 Nov;80:81-86.

Fairhead M, Howarth M. Site-specific biotinylation of purified proteins using BirA. Methods Mol Biol. 2015;1266:171-84.

Gandre S, Bercovich Z, Kahana C. Mitochondrial localization of antizyme is determined by context-dependent alternative utilization of two AUG initiation codons.Mitochondrion. 2003;2:245–256.

Gerdes K, Thisted T, Martinussen J (November 1990). Mechanism of post-segregational killing by the hok/sok system of plasmid R1: sok antisense RNA regulates formation of a hok mRNA species correlated with killing of plasmid-free cells. Mol. Microbiol. **4** (11): 1807–18.

Hellen CU, Sarnow P. Internal ribosome entry sites in eukaryotic mRNA molecules. Genes Dev. 2001 Jul 1;15(13):1593-612.

Jan CH, Ruby JG, Friedman R, Bartel DP. Formation, regulation and evolution of *Caenorhabditis elegans* 3'UTRs. (2011) Nature. 469(7328):97-101.

Jinek M, Doudna JA. A three-dimensional view of the molecular machinery of RNA interference. Nature. 2009 Jan 22;457(7228):405-12.

Kaipio K, Kallio J, Pesonen U. Mitochondrial targeting signal in human neuropeptide Y gene. Biochem. Biophys. Res. Commun. 2005;337:633–640.

Kim DI, Jensen SC, Noble KA, Kc B, Roux KH, Motamedchaboki K, Roux KJ. An improved smaller biotin ligase for BioID proximity labeling. Mol Biol Cell. 2016 Apr 15;27(8):1188-96.

Kolev NG, Tschudi C, Ullu E. RNA interference in protozoan parasites: achievements and challenges. Eukaryot Cell. 2011 Sep;10(9):1156-63.

Lai BS, Witola WH, El Bissati K, Zhou Y, Mui E, Fomovska A, McLeod R. Molecular target validation, antimicrobial delivery, and potential treatment of *Toxoplasma gondii* infections. Proc Natl Acad Sci U S A. 2012 Aug 28;109(35):14182-7.

Laursen BS, Sørensen HP, Mortensen KK, Sperling-Petersen HU. Initiation of protein synthesis in bacteria. Microbiol Mol Biol Rev. 2005 Mar;69(1):101-23.

Leissring MA, Farris W, Wu X, Christodoulou DC, Haigis MC, Guarente L, Selkoe DJ. Alternative translation initiation generates a novel isoform of insulin-degrading enzyme targeted to mitochondria. Biochem. J. 2004;383:439–446.

Musiyenko A, Majumdar T, Andrews J, Adams B, Barik S. PRMT1 methylates the single Argonaute of *Toxoplasma gondii* and is important for the recruitment of Tudor nuclease for target RNA cleavage by antisense guide RNA. Cell Microbiol. 2012 Jun;14(6):882-901.

Nelson MM, Jones AR, Carmen JC, Sinai AP, Burchmore R, et al. Modulation of the host cell proteome by the intracellular apicomplexan parasite *Toxoplasma gondii*. Infect Immun. 2008;76:828–844.

Roberts TC, Ezzat K, El Andaloussi S, Weinberg MS. Synthetic SiRNA Delivery: Progress and Prospects. Methods Mol Biol. 2016;1364:291-310.

Ruby JG, Jan, CH, Bartel DP. Intronic microRNA precursors that bypass Drosha processing. (2007) Nature. 448(7149):83-86.

Seeber, F. (1997). Consensus sequence of translational initiation sites from *Toxoplasma gondii* genes. Parasitology Research. **83** (3): 309–311.

Shimoni Y, Friedlander G, Hetzroni G, Niv G, Altuvia S, Biham O, Margalit H. Regulation of gene expression by small non-coding RNAs: a quantitative view. Mol Syst Biol. 2007;3:138.

Slobodin B, Gerst JE. A novel mRNA affinity purification technique for the identification of interacting proteins and transcripts in ribonucleoprotein complexes. RNA. 2010 Nov;16(11):2277-90.

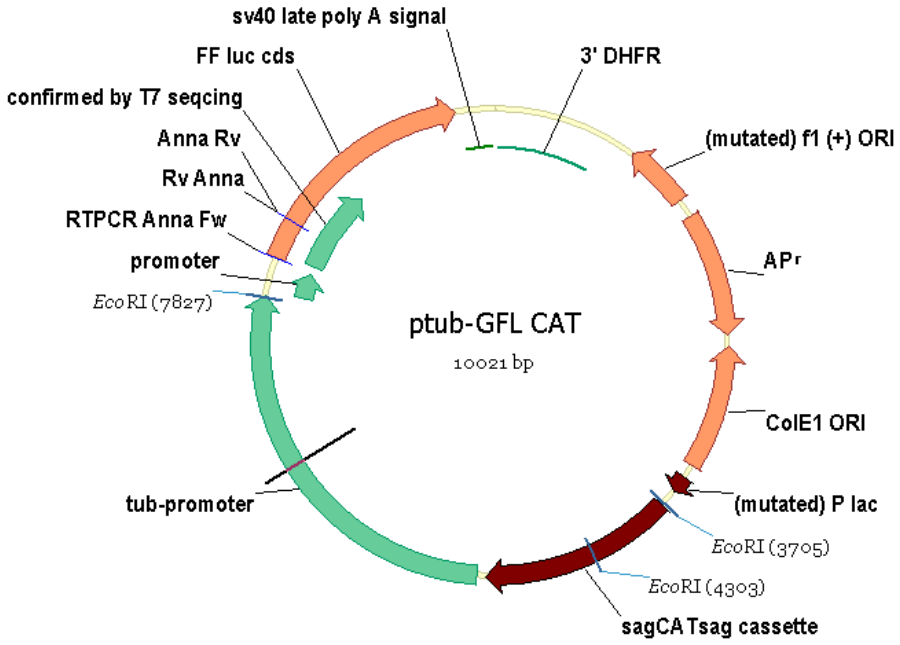
Wang J, Lu Z, Wientjes MG, Au JL. Delivery of siRNA therapeutics: barriers and carriers. AAPS J. 2010 Dec;12(4):492-503.

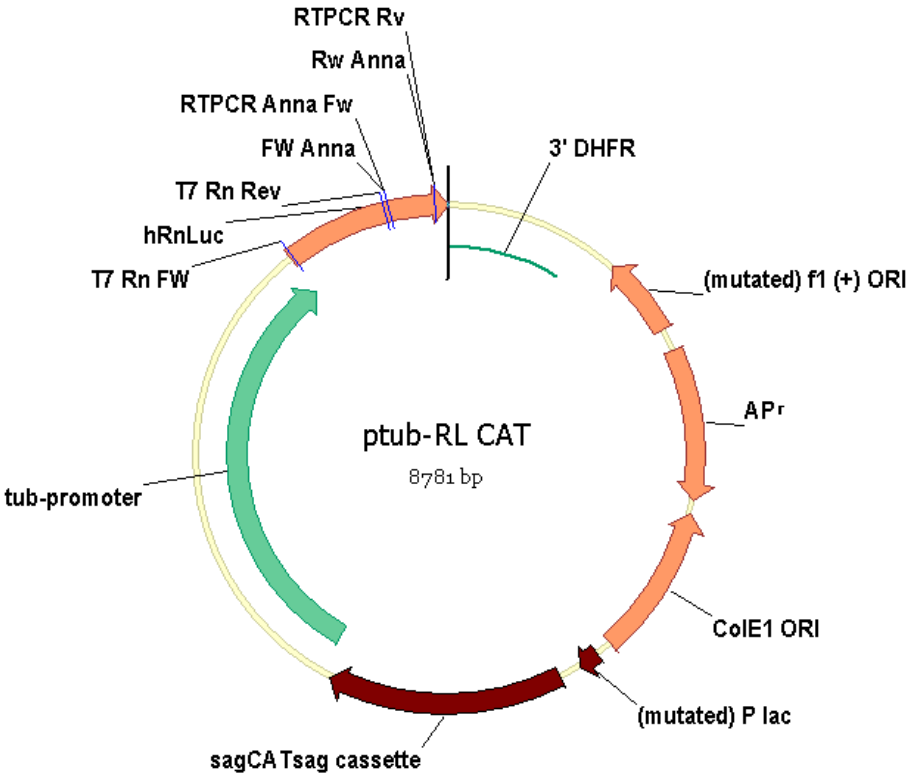
Zeiner GM, Norman KL, Thomson JM, Hammond SM, Boothroyd JC. *Toxoplasma gondii* infection specifically increases the levels of key host microRNAs. PLoS One. 2010 Jan 15;5(1):e8742.

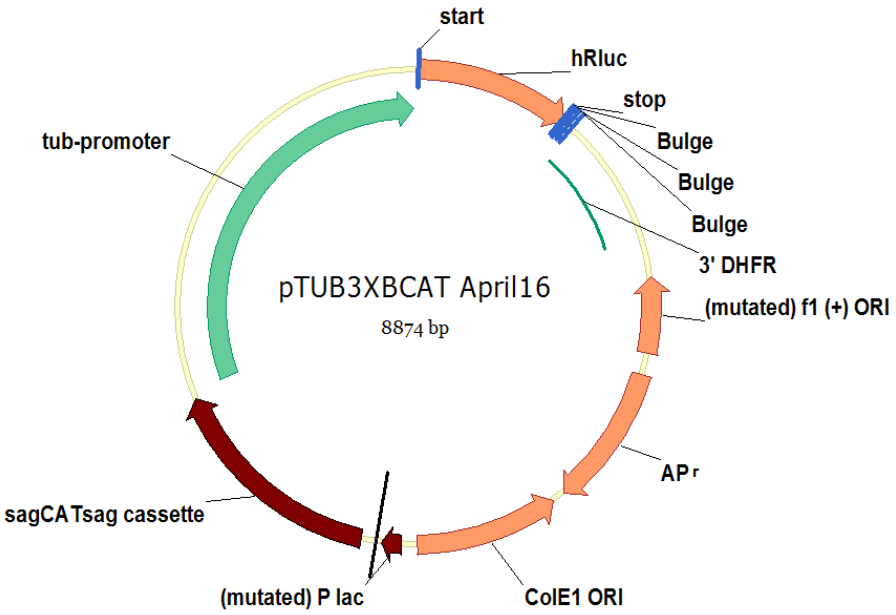
Zhu QH, Springgs A, Matthew L, Fan L, Kennedy G, Gubler F, Helliwell C. A diverse set of microRNAs and microRNA-like small RNAs in developing rice grains. (2008) Genome Research. 18(9):1456-14.

APPENDICES

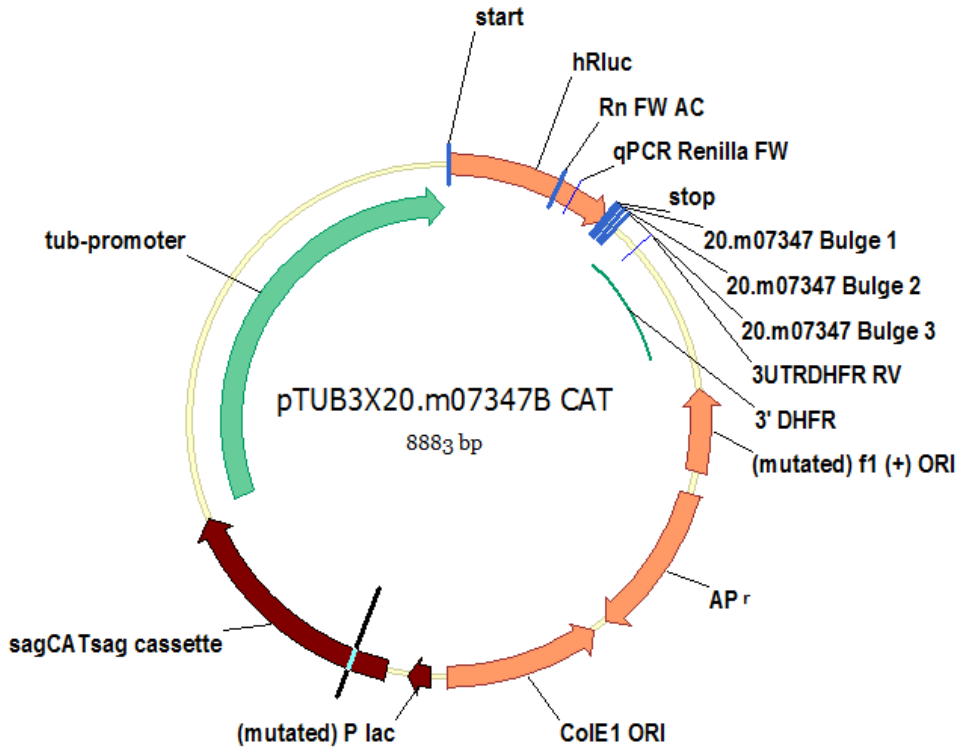
A1. Dual luciferase plasmids for Chapter 2.

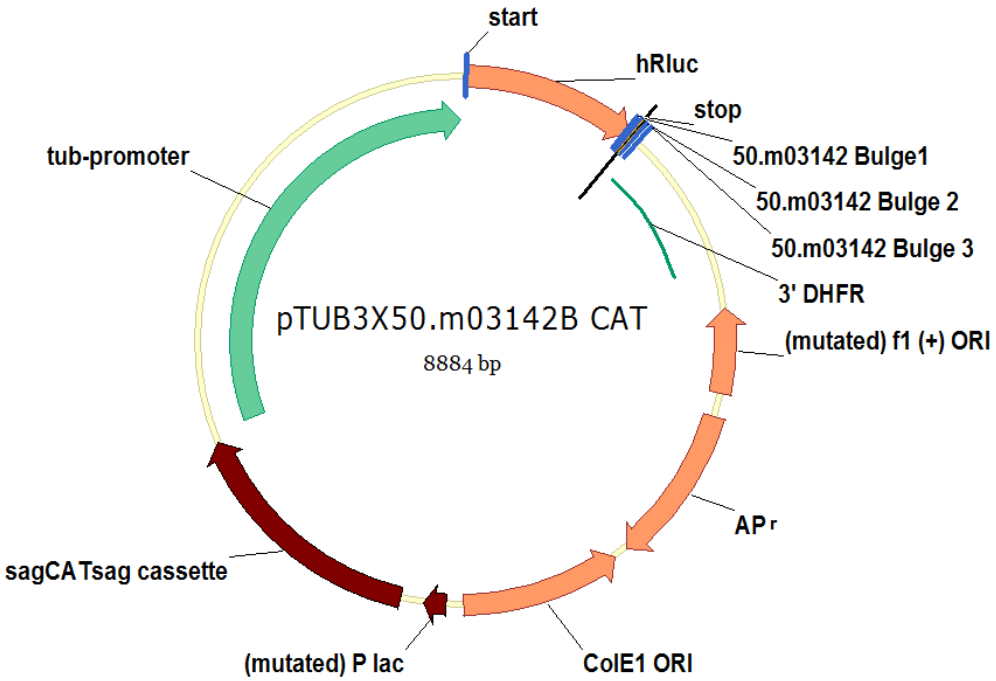


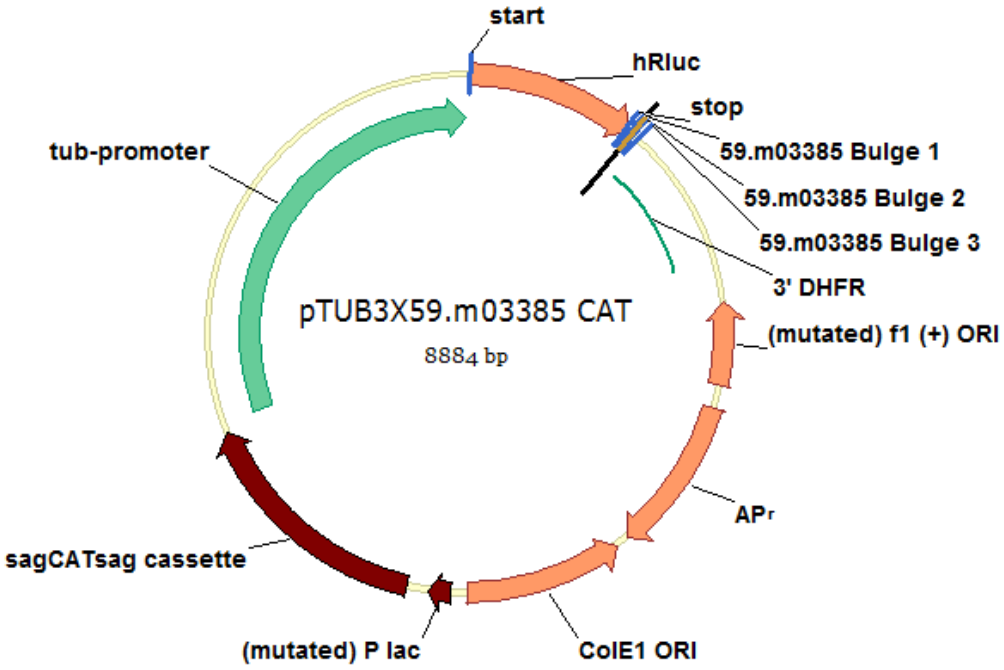



start
hRluc
stop
Bulge
Bulge
Bulge
3' DHFR
(mutated) f1 (+) ORI
AP r
ColE1 ORI
(mutated) P lac
sagCATsag cassette
tub-promoter
pTUB3XBCAT April16
8874 bp

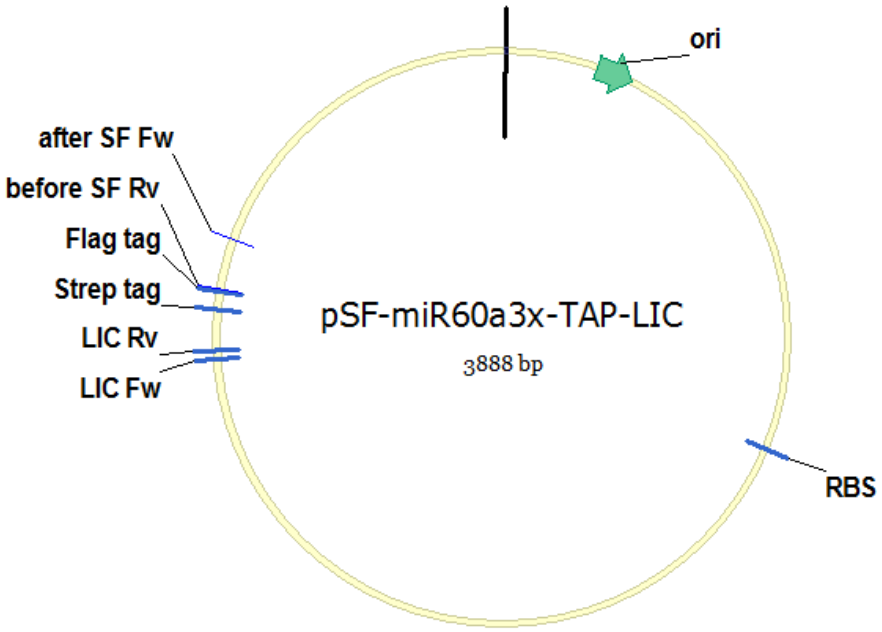
A2. Dual luciferase plasmids with miR-4a binding sites for Chapter 3.



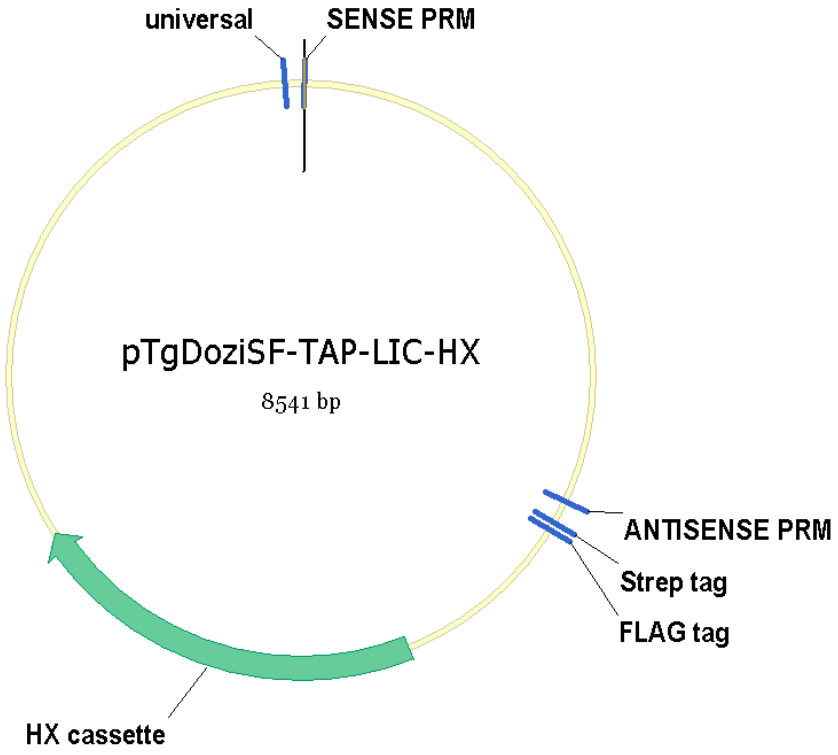


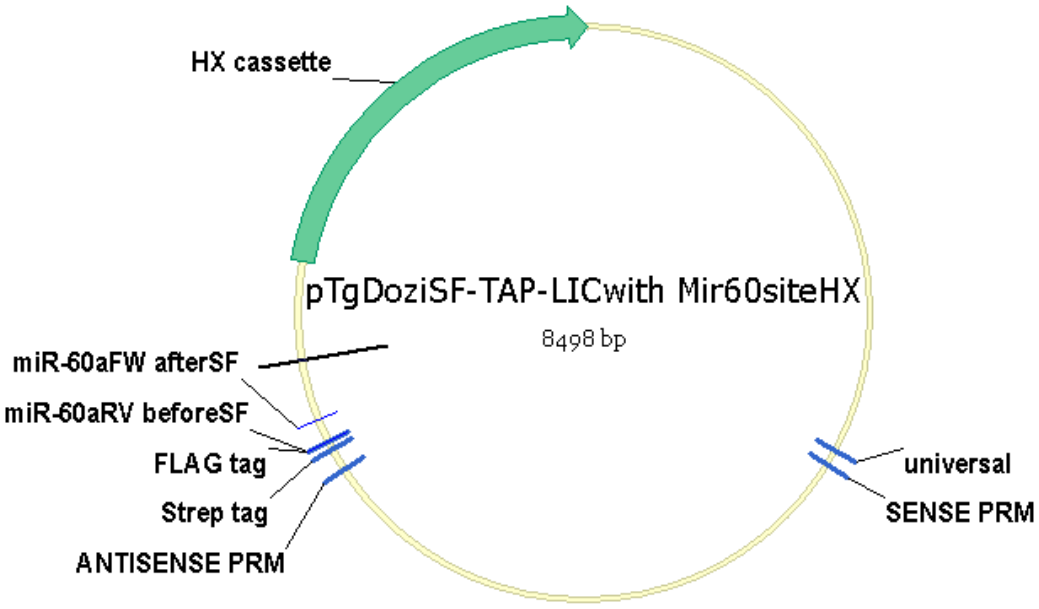

start
hRluc
stop
59.m03385 Bulge 1
59.m03385 Bulge 2
59.m03385 Bulge 3
3' DHFR
(mutated) f1 (+) ORI
APr
ColE1 ORI
(mutated) P lac
sagCATsag cassette
tub-promoter
pTUB3X59.m03385 CAT
8884 bp

A3. LIC plasmid for miR-60a binding sites for Chapter 3

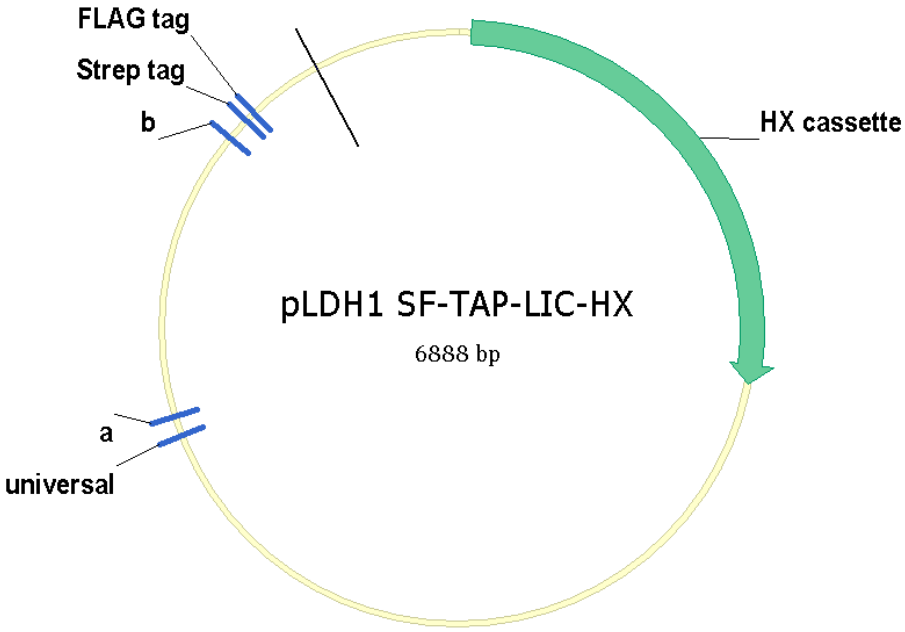


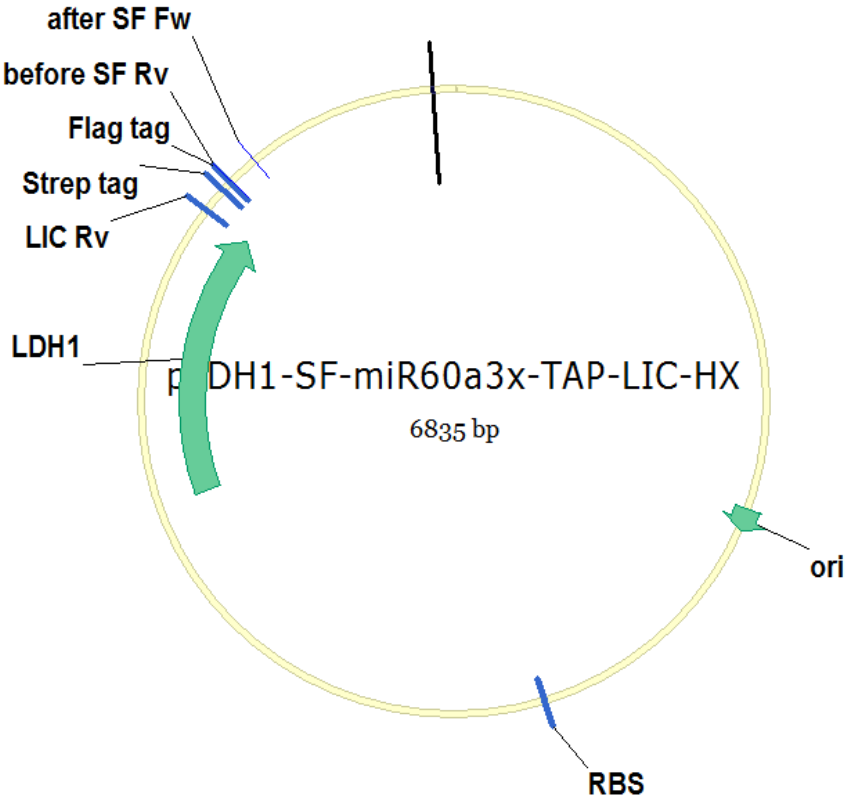
A4. LIC plasmids to create HoDI-SF and HodI-SF-miR60a sites for Chapter 3



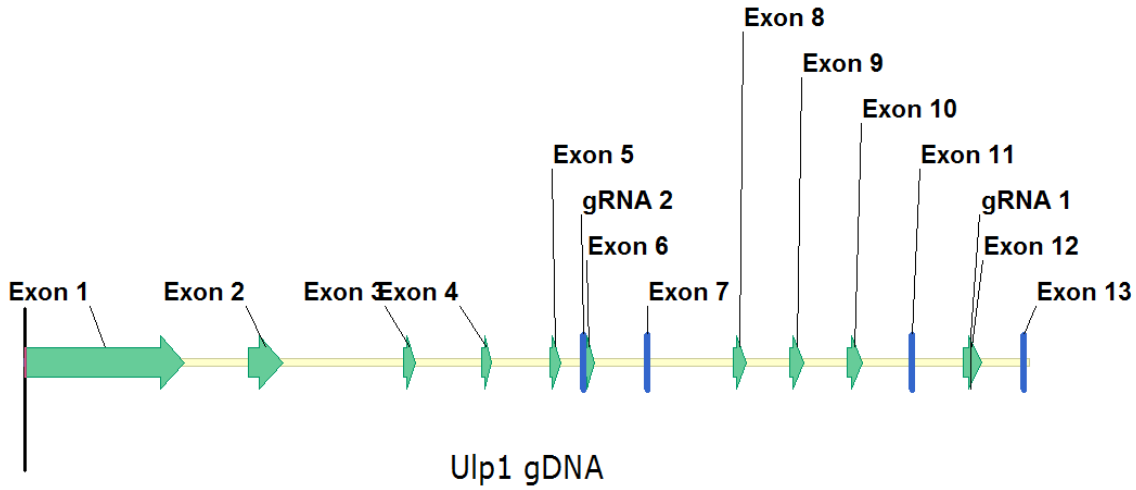


A5. LIC plasmids to generate LDH1-SF and LDH1-SF-miR-60a sites for Chapter 3



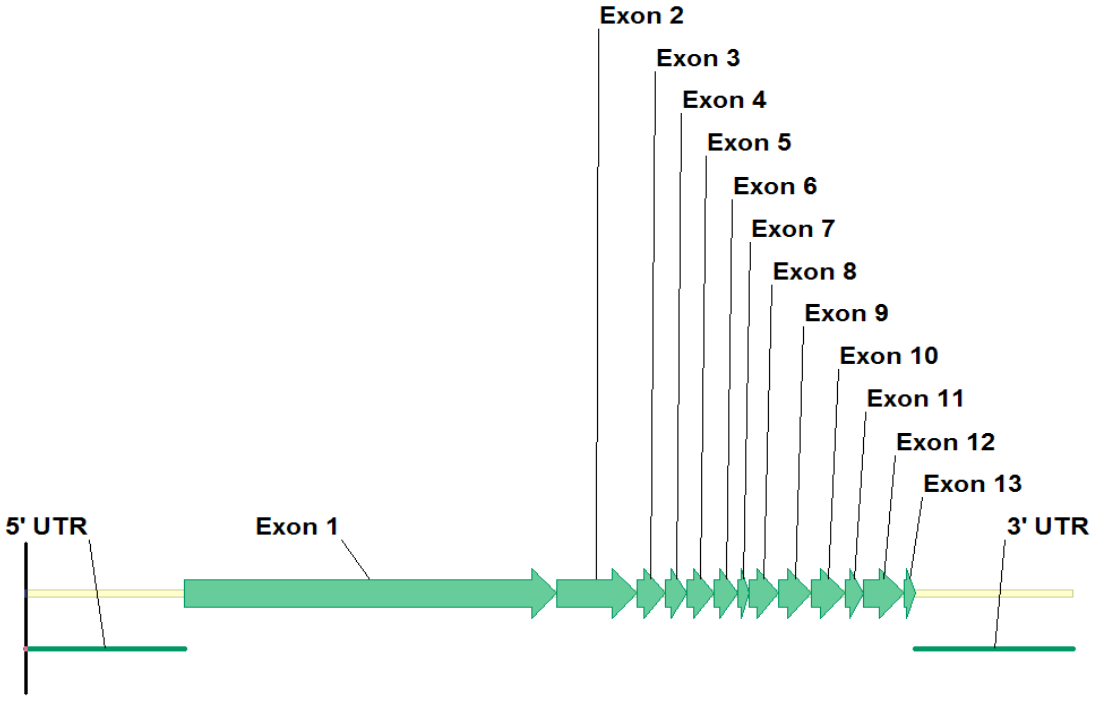


A6. Genomic and cDNA sequences for TgUlp1 (TGGT1_214470) for Chapter 5.



Ulp1 gDNA

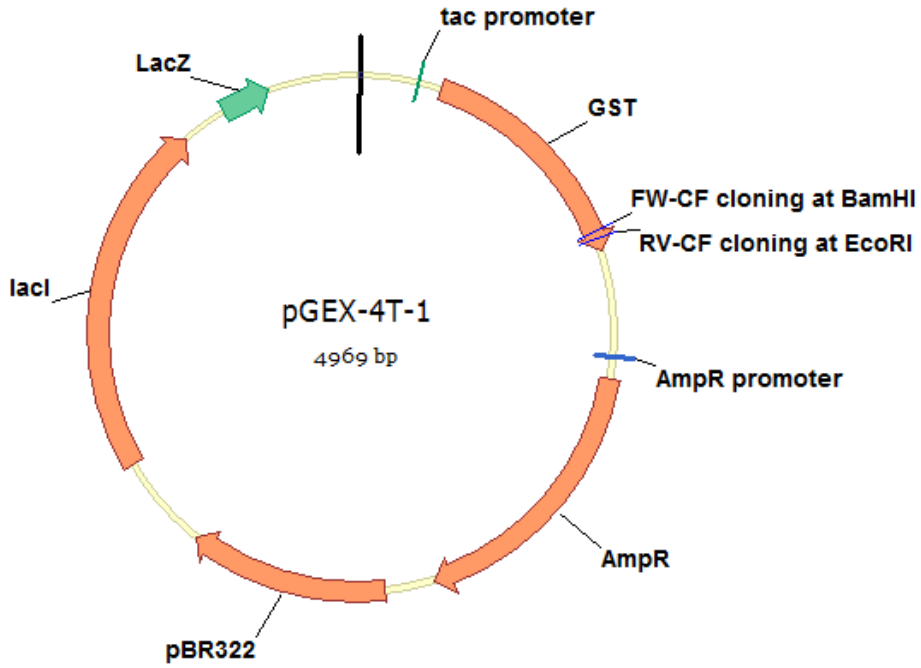
7432 bp

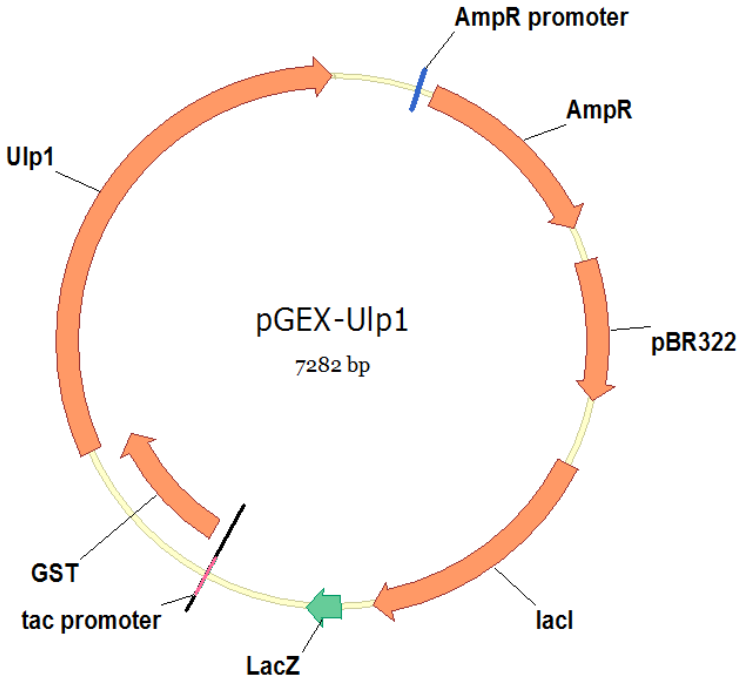


TgUlp1 cDNA

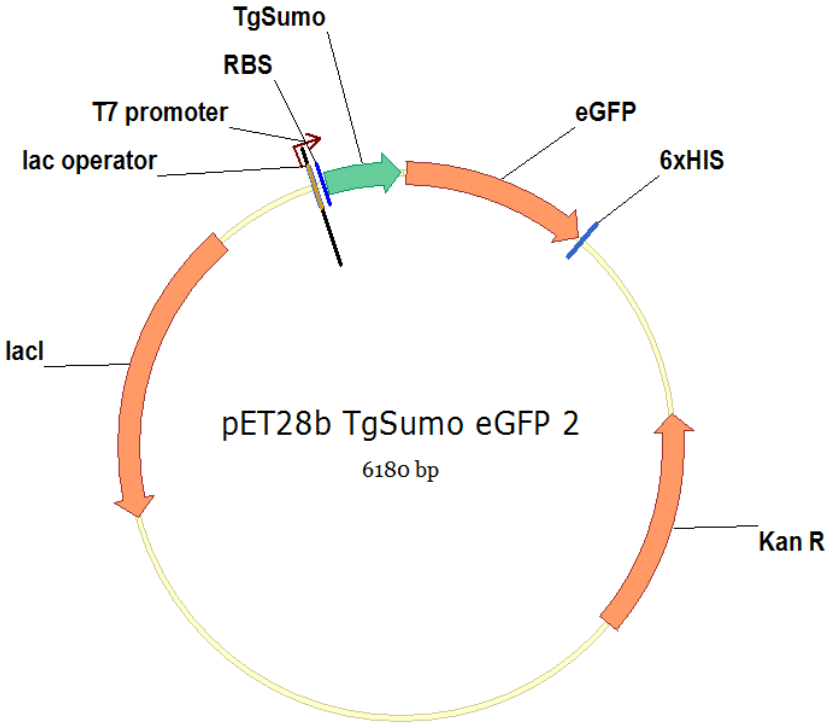
3316 bp

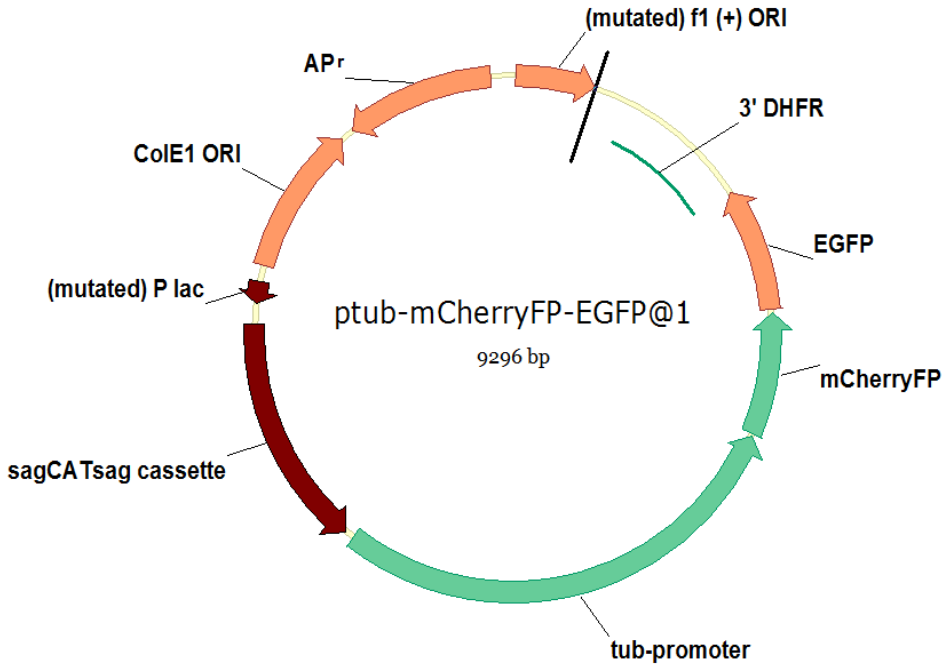
A7. Plasmids for generation of GST-Ulp1 for Chapter 5.



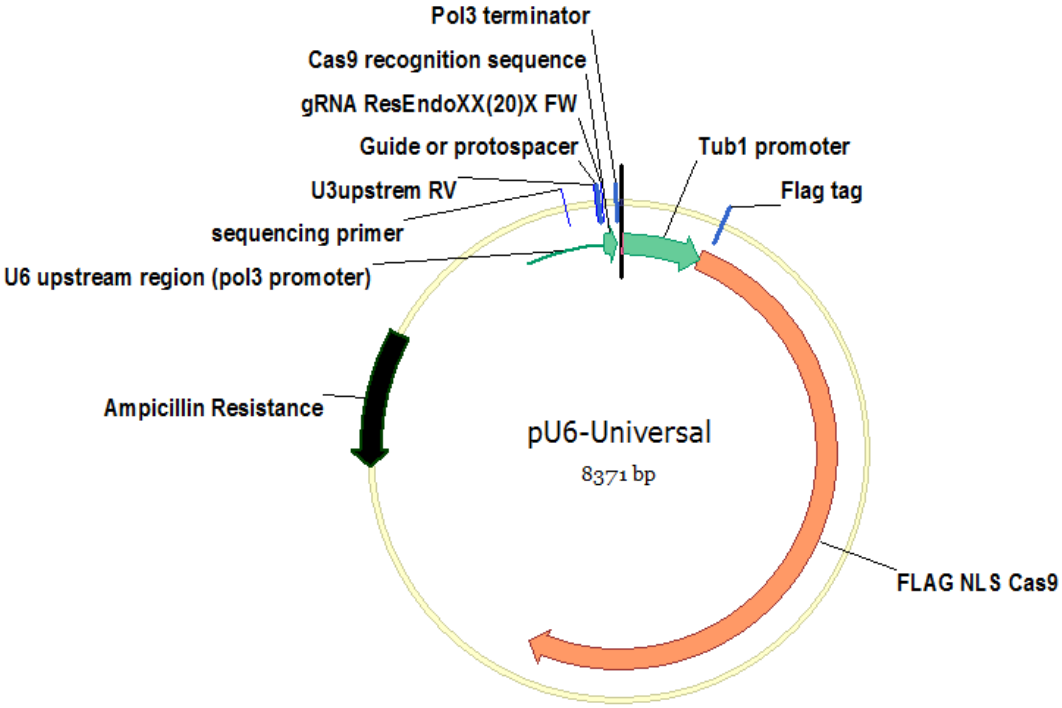


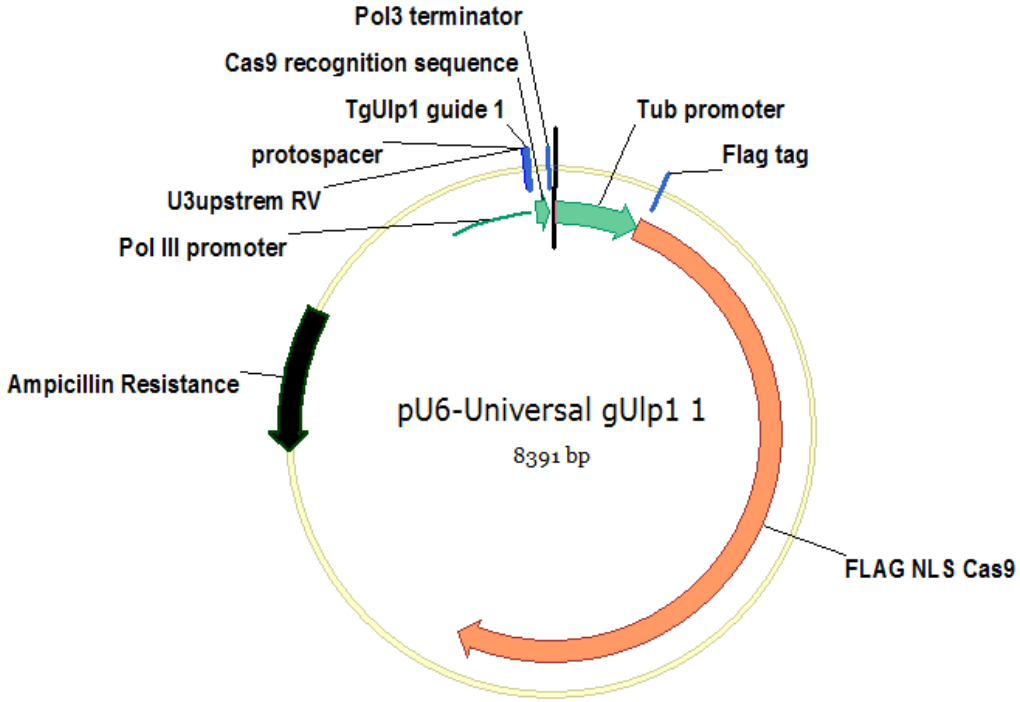
A8. pET28b TgSumoeGFP for Chapter 5 as substrate for TgUlp1 activity assay.

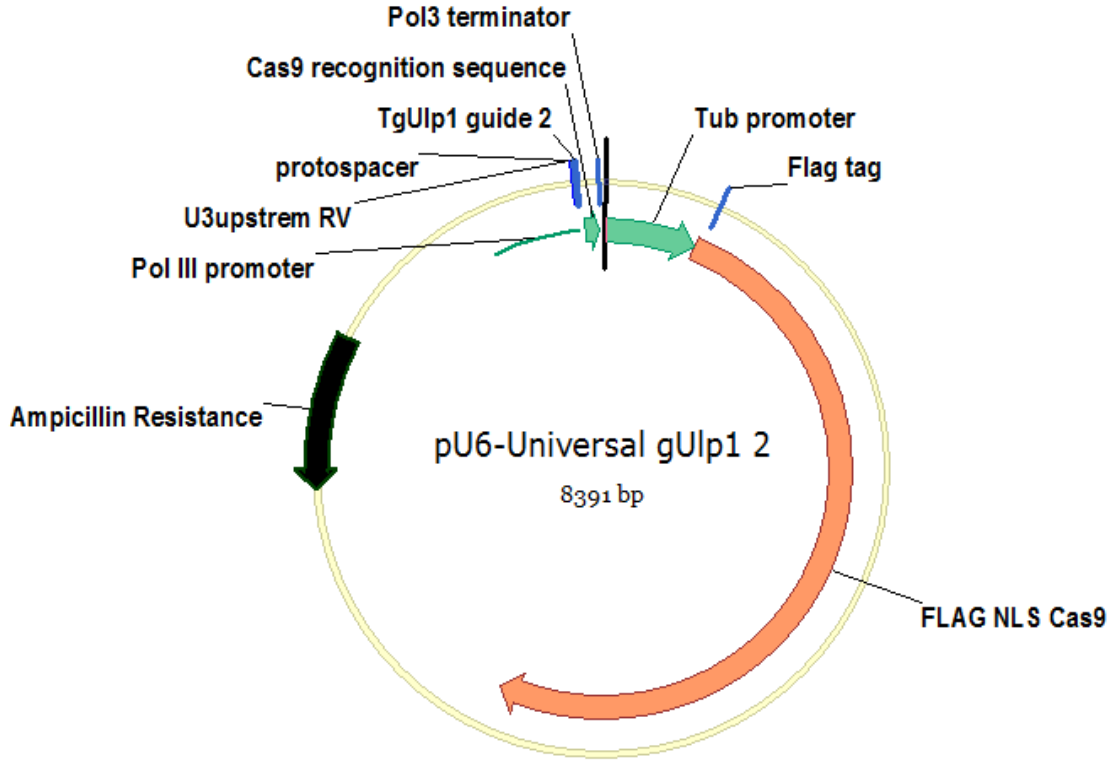




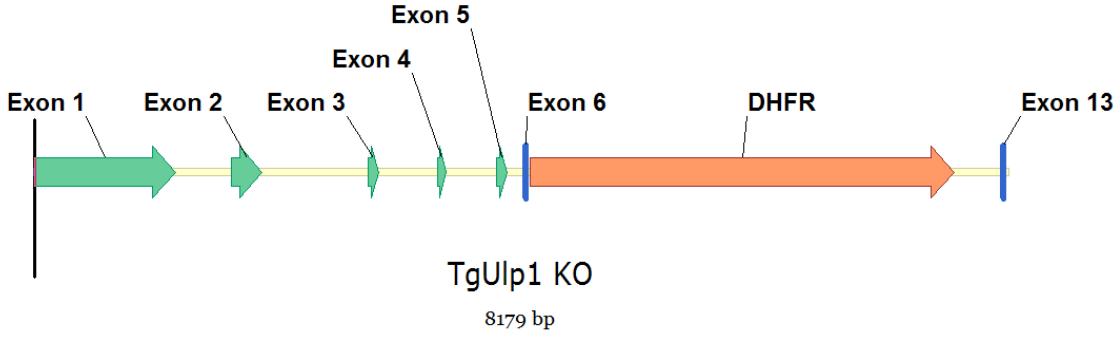
A9. Plasmids used to create TgUlp1 KO for Chapter 5.



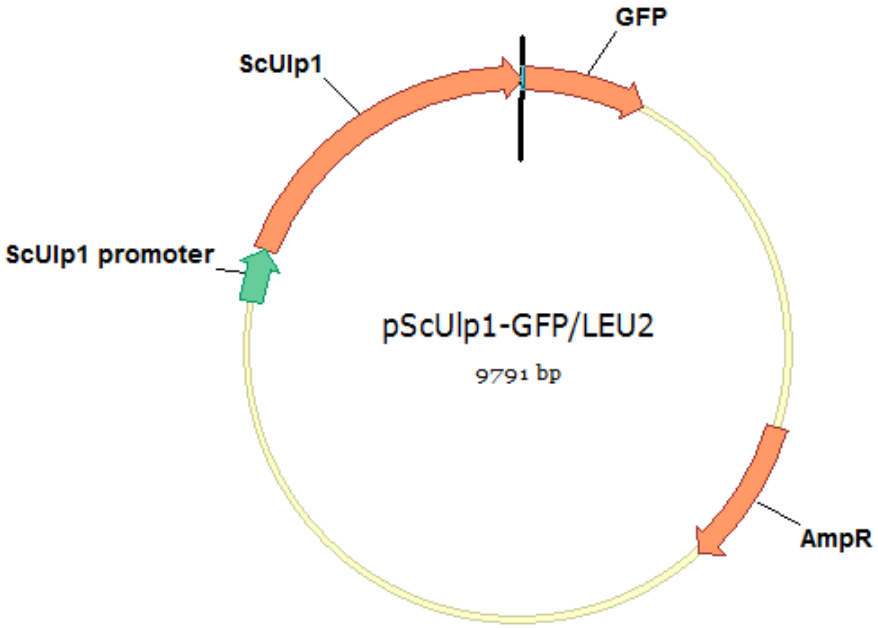


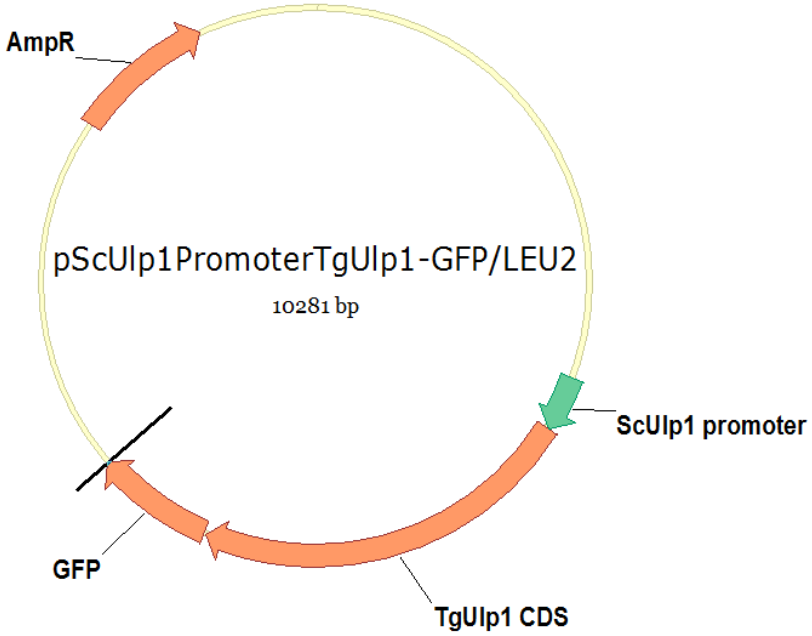


A10. Genomic representation of TgUlp1 KO with DHFR selectable marker insert for Chapter 5.

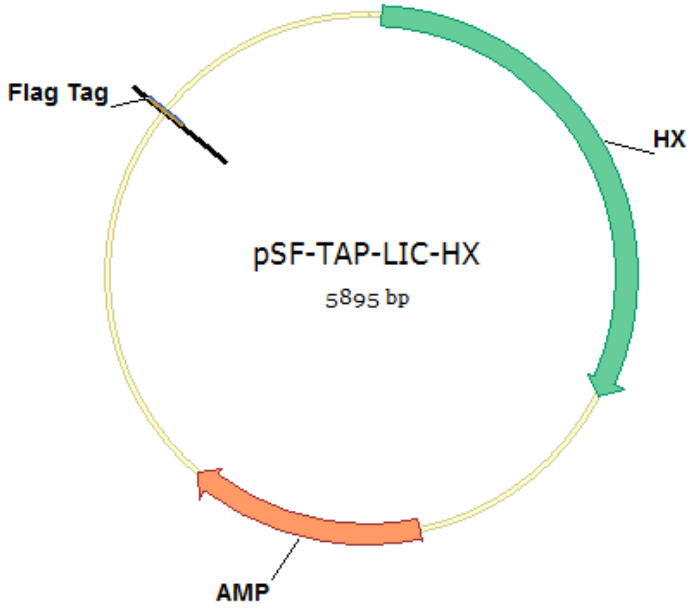


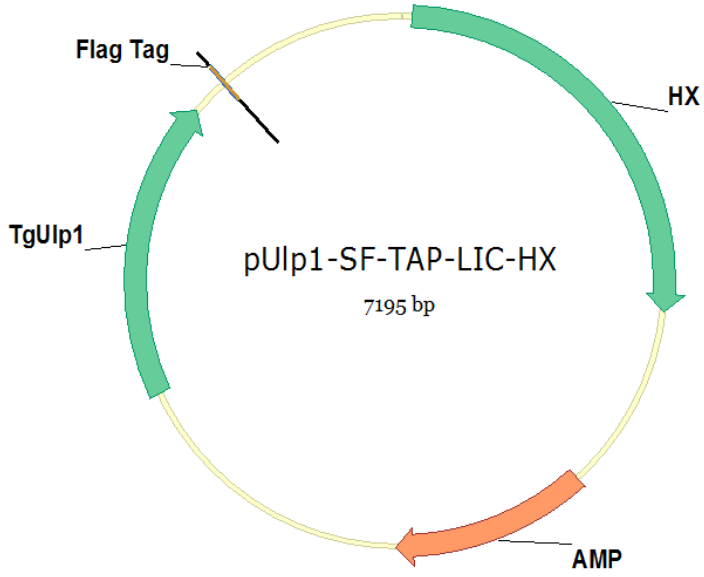
A11. pScUlp1PromoterTgUlp1-GFP/LEU2 used for yeast complementation assays for Chapter 5.



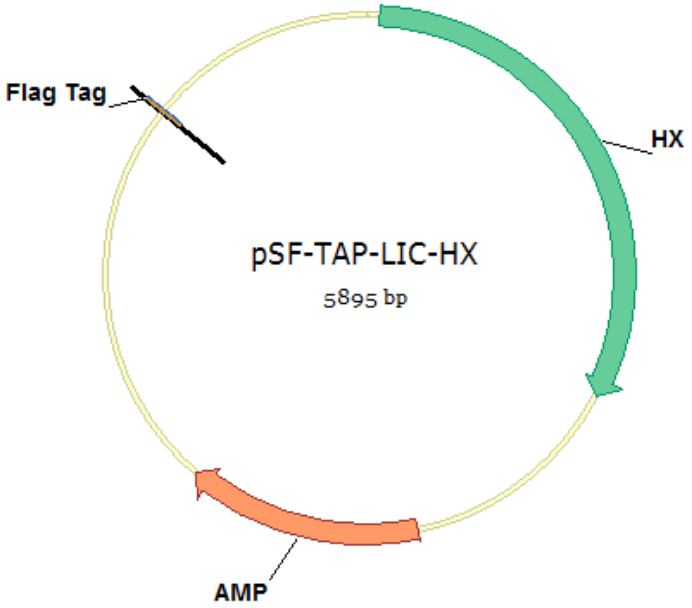


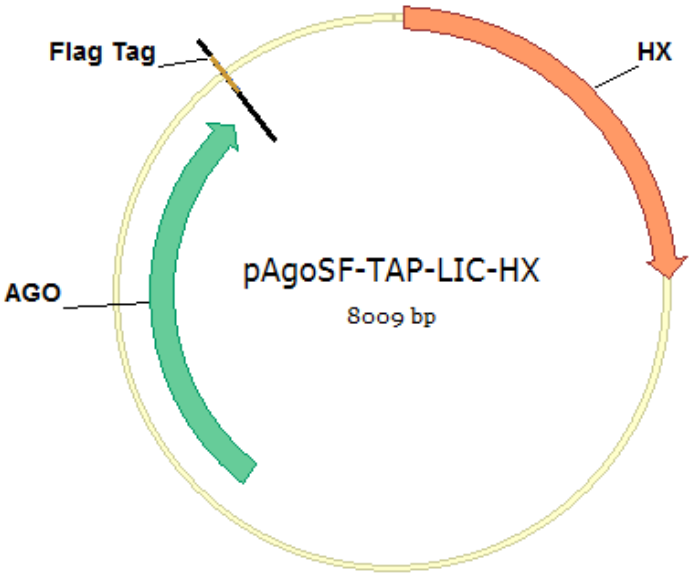
A12. pUlp1-SF-TAP-LIC-Hx plasmid was used to generate TgUlp1-SF transgenic parasites for Chapter 5.



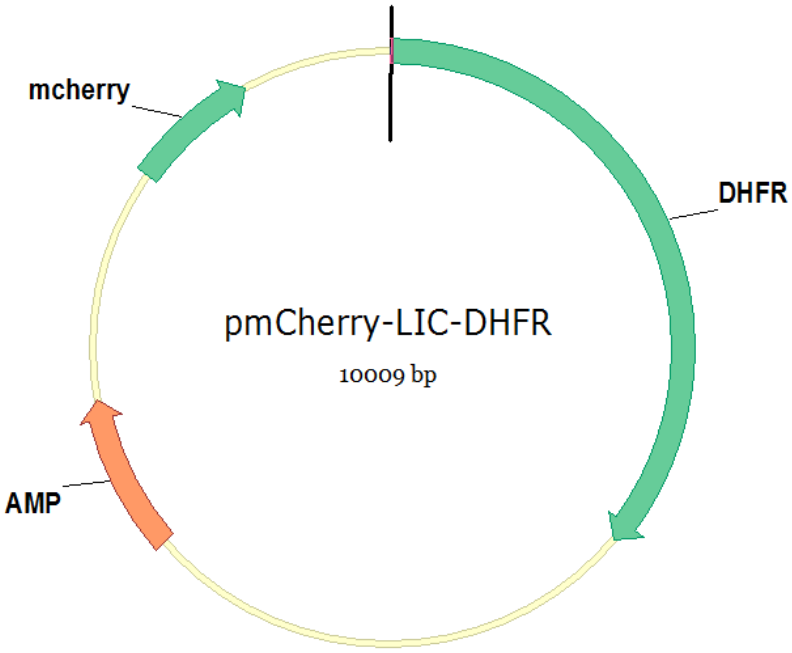


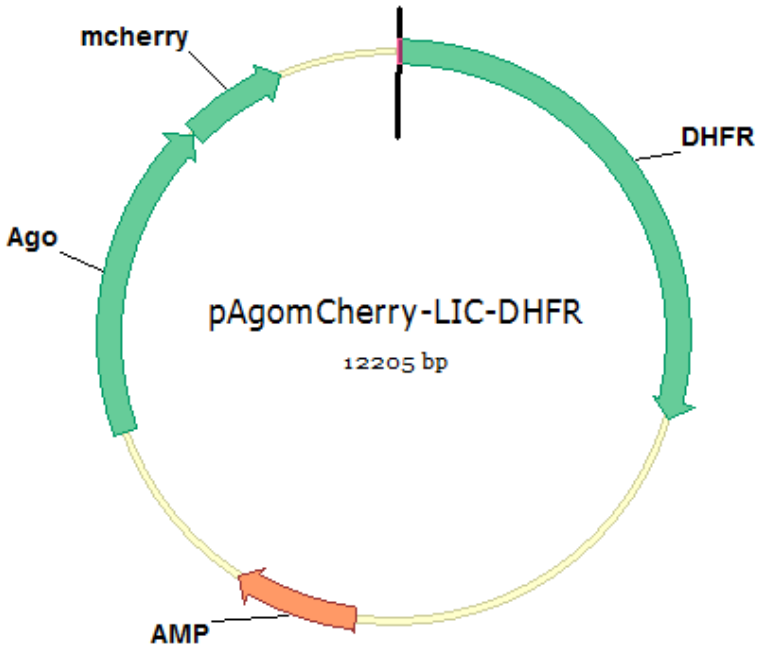
A13. pAgo-SF-TAP-LIC-Hx plasmid was used to generate TgAgo-SF transgenic parasites for Chapter 6.



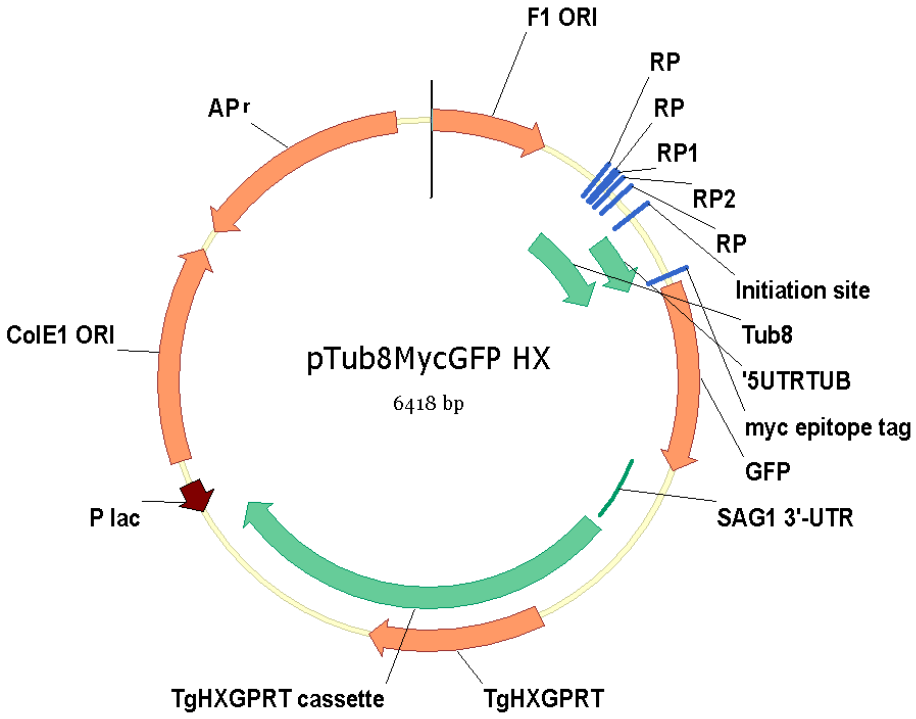


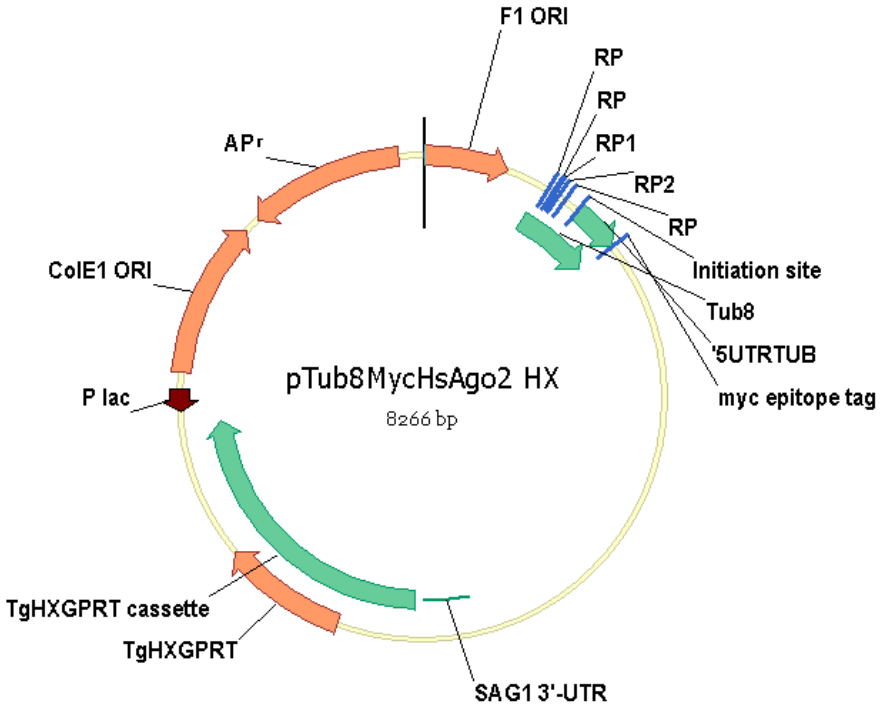
A14. pAgo-mCherry-TAP-LIC-DHFR plasmid was used to generate TgAgo-mCherry transgenic parasites for Chapter 6.





A15. Plasmids to express Argonaute homologs for Chapter 6.





B. Copyright permission

ELSEVIER LICENSE
TERMS AND CONDITIONS

Jan 29, 2017

This Agreement between Anna K Crater ("You") and Elsevier ("Elsevier") consists of your license details and the terms and conditions provided by Elsevier and Copyright Clearance Center.

License Number	4037880302600
License date	Jan 28, 2017
Licensed Content Publisher	Elsevier
Licensed Content Publication	Journal of Microbiological Methods
Licensed Content Title	Utilization of inherent miRNAs in functional analyses of Toxoplasma gondii genes
Licensed Content Author	Anna K. Crater,Emad Manni,Sirinart Ananvoranich
Licensed Content Date	January 2015
Licensed Content Volume Number	108
Licensed Content Issue Number	n/a
Licensed Content Pages	11
Start Page	92
End Page	102
Type of Use	reuse in a thesis/dissertation
Intended publisher of new work	other
Portion	full article
Format	both print and electronic
Are you the author of this Elsevier article?	Yes
Will you be translating?	No
Order reference number	
Title of your thesis/dissertation	RNA Directed gene regulation in Toxoplasma gondii
Expected completion date	Jan 2017
Estimated size (number of pages)	235
Elsevier VAT number	GB 494 6272 12

ELSEVIER ORDER DETAILS
Jan 29, 2017

Order Number
501228958
Order date
Jan 28, 2017
Licensed Content Publisher
Elsevier
Licensed Content Publication
Journal of Microbiological Methods
Licensed Content Title
Antisense technologies in the studying of Toxoplasma gondii
Licensed Content Author
Anna K. Crater,Scott Roscoe,Mikayla Roberts,Sirinart Ananvoranich
Licensed Content Date
Available online 24 December 2015
Licensed Content Volume Number
n/a
Licensed Content Issue Number
n/a
Licensed Content Pages
1
Start Page
End Page
Type of Use
reuse in a thesis/dissertation
Portion
full article
Format
both print and electronic
Are you the author of this Elsevier article?
Yes
Will you be translating?
No
Order reference number
Title of your thesis/dissertation
RNA Directed gene regulation in Toxoplasma gondii
Expected completion date
Jan 2017
Estimated size (number of pages)
235
Elsevier VAT number
GB 494 6272 12

January 28, 2017
American Journal of Biomedical Sciences
Prof. Joseph Bonaventura
Editor-in-Chief

To Whom It May Concern:

Dear Dr. Bonaventure:

I am completing a doctoral dissertation at the University of Windsor entitled "RNA directed gene regulation in *Toxoplasma gondii*." I would like your permission to include in my dissertation the following article:

Crater AK, Cherry A, Holmes M, Kadri D, Ananvoranich A. Evaluation of the Ability of Short and Long Double-stranded RNAs to Induce Homologous Gene Silencing in the Protozoan Parasite, *Toxoplasma gondii*. Am. J. Biomed. Sci. 2012, 4(1), 1-13.

My thesis will be deposited to the University of Windsor's online theses and dissertations repository (http://scholar.uwindsor.ca/etd/) and will be available in full-text on the internet for reference, study and / or copy.

I will also be granting Library and Archives Canada and ProQuest/UMI a non-exclusive license to reproduce, loan, distribute, or sell single copies of my thesis by any means and in any form or format. These rights will in no way restrict republication of the material in any other form by you or by others authorized by you.

Please confirm in writing or by email that these arrangements meet with your approval.
Thank you very much for your attention to this matter.

Sincerely,

Anna K. Crater

VITA AUCTORIS

NAME:	Anna K. Crater-Potter
PLACE OF BIRTH:	Coldwater, Michigan, United States of America
YEAR OF BIRTH:	1985
EDUCATION:	Litchfield High School, Litchfield, MI, 2003
	Spring Arbor University, B.A., Spring Arbor, MI, 2007
	Spring Arbor University, B. Ed., Spring Arbor, MI, 2009