University of Windsor
Scholarship at UWindsor

Electronic Theses and Dissertations | Theses, Dissertations, and Major Papers

4-30-2018

# Investigation of the RNA-interference pathway of Toxoplasma gondii

Farzana Afrin
*University of Windsor*

Follow this and additional works at: https://scholar.uwindsor.ca/etd

## Recommended Citation

Afrin, Farzana, "Investigation of the RNA-interference pathway of Toxoplasma gondii" (2018). *Electronic Theses and Dissertations*. 7445.
https://scholar.uwindsor.ca/etd/7445

This online database contains the full-text of PhD dissertations and Masters' theses of University of Windsor students from 1954 forward. These documents are made available for personal study and research purposes only, in accordance with the Canadian Copyright Act and the Creative Commons license—CC BY-NC-ND (Attribution, Non-Commercial, No Derivative Works). Under this license, works must always be attributed to the copyright holder (original author), cannot be used for any commercial purposes, and may not be altered. Any other use would require the permission of the copyright holder. Students may inquire about withdrawing their dissertation and/or thesis from this database. For additional inquiries, please contact the repository administrator via email (scholarship@uwindsor.ca) or by telephone at 519-253-3000ext. 3208.

# Investigation of the RNA-interference pathway of *Toxoplasma gondii*

By

Farzana Afrin

A Thesis
Submitted to the Faculty of Graduate Studies
through the Department of Chemistry and Biochemistry
in Partial Fulfillment of the Requirements for
the Degree of Master of Science
at the University of Windsor

Windsor, Ontario, Canada

2018

© 2018 Farzana Afrin

**Investigation of the RNA-interference pathway of *Toxoplasma gondii***

by

Farzana Afrin

APPROVED BY:

_____

A. Hubberstey
Department of Biological Sciences

_____

M. Boffa
Department of Chemistry & Biochemistry

_____

S. Ananvoranich, Advisor
Department of Chemistry & Biochemistry

January 10, 2018

## DECLARATION OF ORIGINALITY

I hereby certify that I am the sole author of this thesis and that no part of this thesis has been published or submitted for publication.

I certify that, to the best of my knowledge, my thesis does not infringe upon anyone's copyright nor violate any proprietary rights and that any ideas, techniques, quotations, or any other material from the 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

The RNA-interference pathway is commonly found in eukaryotes, and plays an important role in the inhibition of gene expression. This study focuses on two aspects of the RNA-interference pathway of *Toxoplasma gondii*. The first aspect was to investigate the function of a putative Dicer, an enzyme responsible for the generation of small RNAs (siRNA and miRNA). A parasite strain, whose functional Tg-Dicer expression was abolished, was generated and used in this study in comparison to the parental strain. It was detected that the replication and invasion of the mutated Tg-Dicer strain (TgDicer-mut) is slower than that of the parental strain (TgDicer-wt). Using the dual luciferase assay I detected that TgDicer-mut lacks the ability to silence the expression of *Renilla* (RN) transcript containing Tg-miRNA binding sites. To determine if the inability to silence the RN expression was due to its incapability to process long dsRNA into short dsRNA, TgDicer-mut strain was electroporated with long or short dsRNA complementary to RN transcript. TgDicer-mut can use short dsRNA and cause a decrease in Renilla activity, but not long dsRNA. The study thus confirms that TgDicer is responsible for processing long dsRNA into short dsRNA. The other aspect of this investigation was to verify whether a predicted target of miRNAs named TgHODI (a DEAD-box RNA helicase) could be a target of the two most abundant miRNAs, miR4a and miR60a. A clonal transgenic, flag-tagged TgHODI was used for the study to determine its expression in the presence of miRNA inhibitor. When the inhibitors specific to miR4a and miR60a was used, the expression of TgHODI expression level increased to ~1.8 and ~1.6 times respectively. This indicates that TgHODI may be regulated by miR4a and miR60a.

## DEDICATION

I dedicate my thesis
to Abbu, Ammu and Shoshi.
Ami tomaderke onek bhalobashi.

# ACKNOWLEDGMENTS

In the name of Allah, Most Gracious, Most Merciful. My religion has always been my motivation for learning about the things around me. Allah states in the Holy Quran (50:6-8), "Do they not look at the sky above them? How we have made it, and adorned it and there are no flaws in it? And the Earth - we have spread it out and set thereon mountains standing firm and produced therein every kind of beautiful growth in pairs. To be observed and commemorated by every devotee turning to Allah."

First and foremost, I would like to thank my professor, Dr. Sirinart Ananvoranich. Thank you for being a great mentor and for always being there when I needed help. I have learned so much from you. Thank you for being patient with me and for pushing me to do better. I appreciate all the time and effort you have spent with me.

I would like to thank the graduate students in Dr. Ananvoranich's lab who have taught me so much more than just science. Thank you, Anna, for your guidance and for being a great female mentor. Thank you, Emad, for teaching me the basics in the lab. Thank you, Mike and Scott, for being so encouraging. Thank you, Amran, for being such a great friend.

I would like to thank the undergraduate students in my lab. Thank you, Mikayla, for being my little sister for the past three years and for being a source of comfort during my rough days. Thank you, Nadya, for encouraging me to grow taller. Thank you, Karlee, for all your funny stories. Thank you, Ambreen, for pretending my stories were funny. Thank you, Rachel, for all those PCRs. Thank you, Iulian, for all of those interesting discussions. Thank you, Andrei, for encouraging me in so many ways.

Thank you, Hyder, Terence, Bei, and Artur, for teaching me about research when I was an undergraduate student. Thank you, Krithika, Justin, and Mariam, for always encouraging me and teaching me to be optimistic. I would like to thank all the other graduate students in the department for your advice and encouragement. I would also like to thank all of the professors in the Biochemistry department for allowing me to use equipments in their labs. Thank you, Drs. Mutus, Pandey and Lee, for teaching me interesting graduate courses. I learned so much from you.

I would like to thank the support staff in the department for all their help. We all know the department could not function without you. Thank you, Marlene, for your support throughout my master's program and especially during my thesis writing. Thank you, Cathy and Elizabeth, for everything that you do. I would also like to thank the people in my committee, Drs. Hubberstey, Boffa, and Ananvoranich, for their advice and suggestions.

Lastly I would like to thank my family who have always supported me. Thank you, dad, for sitting with me in the lab on those late nights. Thank you, mom, for believing in me when I did not. Thank you, Fahmida, the best sister in the world, for all your motivation and support.

# TABLE OF CONTENTS


DECLARATION OF ORIGINALITY ..... III

ABSTRACT ..... IV

DEDICATION ..... V

ACKNOWLEDGMENTS ..... VI

LIST OF FIGURES ..... X

LIST OF TABLES ..... XI

LIST OF ABBREVIATIONS ..... XII

Chapter I ..... 1

Literature Review ..... 1

1.1 *Toxoplasma gondii* ..... 1

1.2 Life Cycle of *T. gondii* ..... 1

1.3 Parasite Multiplication ..... 5

1.4 Toxoplasmosis ..... 6

1.5 *T. gondii* and Other Apicomplexa ..... 6

1.6 The Genetic Manipulation of *T. gondii* ..... 8

1.6.1 A Selection Marker, Hypoxanthine Xanthine Guanine Phosphoribosyl Transferase ..... 8

1.6.2 Dual luciferase Assay ..... 9

1.7 CRISPR-Cas9 Technology ..... 10

1.8 RNA Induced Post-transcriptional Gene Regulation in Eukaryotes ..... 11

1.8.1 Pre-mRNA Processing ..... 11

1.8.2 RNA-Induced Gene Silencing ..... 11

1.8.3 MicroRNA ..... 12

1.8.4 MicroRNA Biogenesis ..... 13

1.9 Ribonucleases ..... 16

1.9.1 RNase III Family ..... 17

1.9.2 Dicer and its Conserved Domains ..... 20

1.9.3 PAZ domain ..... 21

1.9.4 Helicase Domain ..... 21

1.9.5 Double-Stranded RNA Binding Domain ..... 22

1.10 Potential MiRNA Target ..... 24

1.12 Research Objective ..... 25

Chapter II ..... 26
Materials and Methods ..... 26
2.1 Human Fibroblast and Parasite Cultures ..... 26
2.2 Plasmid Construction ..... 26
2.3 Generation of Truncated TgDicer Parasite Using CRISPR-Cas9 Technology ..... 27
2.4 Transfection by Electroporation ..... 28
2.5 Dual Luciferase Assay ..... 28
2.6 Small RNA Isolation and Imaging ..... 29
2.7 *In vitro* Transcription of RNA ..... 30
2.8 Western Analysis ..... 31
Chapter III ..... 32
Results and Discussion ..... 32
3.1 Bioinformatic Analysis of Putative Dicer in *T. gondii* ..... 32
3.2 Generation and Analysis of a Transgenic *T. gondii* Strain, Whose Tg-Dicer Expression is Abrogated. ..... 39
3.3 Loss-of-Function Effect of Tg-Dicer in its Doubling Time ..... 44
3.4 The Effect of Transgenic Dicer on the Biogenesis of miRNA ..... 48
3.5 The Ability of Truncated Dicer In dsRNA Processing ..... 52
3.6 Identification of TgHODI as a MicroRNA Target ..... 56
Chapter IV ..... 63
Future Directions and Conclusion ..... 63
REFERENCES ..... 67
APPENDICES ..... 74
VITA AUCTORIS ..... 77

# LIST OF FIGURES

Figure 1.1: *T. gondii* Morphology ……..…3

Figure 1.2: Life Cycle of *T. gondii*……..……4

Figure 1.3: MiRNA biogenesis……...……..…..15

Figure 1.4: The predicted mechanism for RNA cleavage of Aa-RNase III…….…...19

Figure 1.5: The domains of *Giardia intestinalis* Dicer……...23

Figure 3.1: Comparing the size and the domains of Tg-Dicer with other Dicers…….....…..33

Figure 3.2: Clustal Analysis of TgDicer Helicase ATP Binding Domain……..35

Figure 3.3: Clustal Analysis of TgDicer Helicase C-terimnal Domain…….…....….36

Figure 3.4: Clustal Analysis of TgDicer RNase IIIa Catalytic Domain…….……37

Figure 3.5: Clustal Analysis of TgDicer RNase IIIb Catalytic Domain.…….……...…...…38

Figure 3.6: Representation of Tg-Dicer locus in the parental strain and in the transgenic parasite following the CRISPR-CAS9 genome editing …….…….……40

Figure 3.7: Screening for Parasite Clones Producing YFP…….…..…41

Figure 3.8: Analysis of transgenic TgDicer-mut strain…….……43

Figure 3.9: Loss-of-function analysis of TgDicer on the parasite's doubling time...……..….47

Figure 3.10: Dual Luciferase of TgDicer-mut and TgDicer-wt…….....……..….....49

Figure 3.11: Polyacrylamide gel of TgDicer-mut and TgDicer-wt…….…….……..…51

Figure 3.12: Dual Luciferase Assay of Exogenous Long DsRNA and Other Small RNA on TgDicer-wt and TgDicer-mut…….....…….....……...….54

Figure 3.13: Dual Luciferase to determine whether anti-4a and anti-60a is able to prevent gene silencing.…….....…….....…….....……..…….……..60

Figure 3.14: Western Analysis of TgHODI-SF with miR4a and miR60a inhibitors. …….…….62

Figure A1**:** An Illustration of pU6-Universal plasmid and pU6-gRNADicer…….…...75

Figure A2**:** An Illustration of pDicer YFP-LIC-HX ……..…76

# LIST OF TABLES

Table 1.1: RNase III consensus sequence in different species……..……..…..….18

Table 3.1: Predicted miR4a binding sites on TgHODI…….....……..….....……57

Table 3.2: Predicted miR60a binding sites on TgHODI……...…….....…57

Table A1: List of all oligonucleotides used in this study.…….……..…...…….74

# LIST OF ABBREVIATIONS

| | |
|---|---|
| ATP | adenosine triphosphate |
| bp | base pair |
| BSA | bovine serum albumin |
| cDNA | complementary DNA |
| DMEM | Dulbecco’s Modified Eagle Medium |
| DNA | deoxyribonucleic acid |
| dNTP | deoxyribonucleoside triphosphate |
| DPBS | Dulbecco’s phosphate-buffered saline |
| dsRNA | double-stranded RNA |
| dFBS | dialyzed fetal bovine serum |
| FF | Firefly luciferase |
| HFF | human foreskin fibroblasts |
| HODI | Homologue of DOZI |
| HRP | horseradish peroxidase |
| HXGPRT | hypoxanthine-xanthine guanine phosphoribosyltransferase |
| miRNA | microRNA |
| mM | millimolar |
| MPA | mycophenolic acid |
| mRNA | messenger ribonucleic acid |
| μM | micromolar |
| PAGE | polyacrylamide gel electrophoresis |
| PBS | phosphate buffered saline |

| | |
|---|---|
| PCR | polymerase chain reaction |
| PV | parasitophorous vacuoles |
| PMSF | phenylmethylsulfonylfluoride |
| RISC | RNA induced silencing complex |
| RN | *Renilla* luciferase |
| RNA | ribonucleic acid |
| RNAi | RNA interference |
| RNase | Ribonuclease |
| rRNA | ribosomal ribonucleic acid |
| siRNA | small interfering RNA |
| TAE | tris-acetate-EDTA |
| Taq | *Thermus aquaticus* |
| TBE | tris-borate-EDTA |
| TEMED | N,N,N',N'-tetramethylethylenediamine |
| tRNA | transfer ribonucleic acid |
| UTR | untranslated region |
| PTGR | post-transcriptional gene regulation |

# Chapter I

# Literature Review

## 1.1 *Toxoplasma gondii*

*Toxoplasma gondii* is a unicellular parasite that was first discovered in 1908 in the tissue of *Ctenodactylus gundi*, a north African rodent. The genus *Toxoplasma* was derived from the Greek word *toxo* depicting its crescent shape (Figure 1.1) and the word *plasma* meaning creature. *T. gondii* is the only species in this genus, and was named after its host (Nicolle and Manceaux, 1908; Splendore, 1908). This parasite's ability to infect is not limited to just rodents. This obligate intracellular protozoan can infect any nucleated warm blooded animal cell including humans. 30 % of the human population worldwide is estimated to be infected by this parasite (Montoya and Liesenfeld, 2004). Infected individuals with a healthy immune system usually do not show any major symptoms. However, immunocompromised individuals infected by *T. gondii*, may show clinical severity such as brain inflammation, eye infection or lung infection (Araujo et al., 1987).

## 1.2 Life Cycle of *T. gondii*

The life cycle of *T. gondii* consists of both sexual and asexual reproduction as shown in Figure 1.2. Sexual reproduction can only occur in the intestines of felines (definitive host), whereas asexual reproduction can occur in birds and mammals (intermediate host). Asexual reproduction of *T. gondii* has two parts; the tachyzoite and the bradyzoite phase. When the parasite is in its rapidly growing tachyzoite phase, it is able to invade any nucleated cell through attachment and penetration of the cell

membrane of the host using gliding motility. Gliding motility is the ability to move without the use of cilia or flagella (Russell and Sinden, 1981). The actin filaments found beneath the inner membrane in the parasite are the main components that are needed to perform the gliding motions (Wetzal et al., 2003; Dobrowolski and Sibley, 1996). This type of invasion is in contrast to other pathogens which enter a host cell through phagocytosis (Falkow et al., 1992). Rhoptries and micronemes are secretory organelles in the apical complex and which secrete contents, allowing for the formation of moving junctions which propels the parasite into the host. and for the building of the parasitophorous vacuole (Huynh and Carruthers, 2006). These organelles are shown in Figure 1.1. After invasion, parasite is enclosed in a parasitophorous vacuole (PV) which is derived from the host cell membrane but it also consists of lipids and proteins from the parasite. The PV acts as a transport interface between the parasite and the host organism (Suss-Toby et al., 1996). The parasite then multiplies and the parasites egress out of the host cell, after which they invade new host cells.

In sexual reproduction, (Figure 1.2) *T. gondii* enters the feline when the feline ingests infected intermediate hosts such as rats and birds containing *T. gondii* cysts. The parasites are released and infect the cat intestinal epithelial cells. The parasites then undergo sexual reproduction which produces oocysts. The oocysts leave the cat through the feces after which they sporulate (Dubey, 1998). If other animals and birds come across mature oocysts, they become infected and the parasite undergoes asexual reproduction.

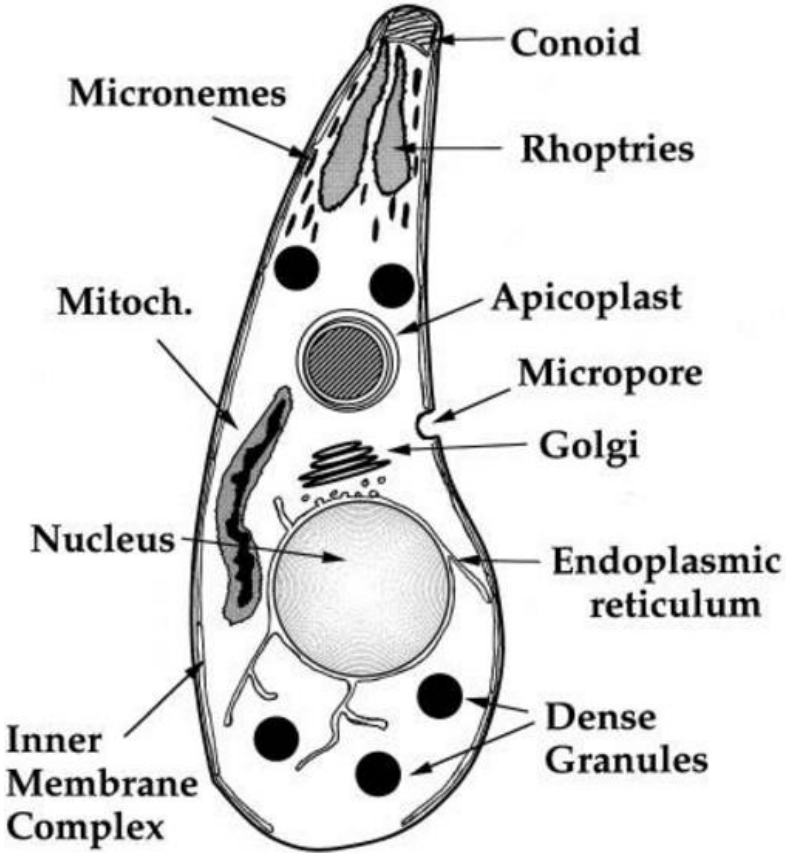



**Figure 1.1: *T. gondii* Morphology**

Major organelles in *T. gondii* are the nucleus, mitochondria, Golgi apparatus, endoplasmic reticulum and the apicoplast. *T. gondii* contains other organelles such as micronemes, rhoptries, and the conoids which aid in parasitic invasion of host cells (Black and Boothryord, 2000).

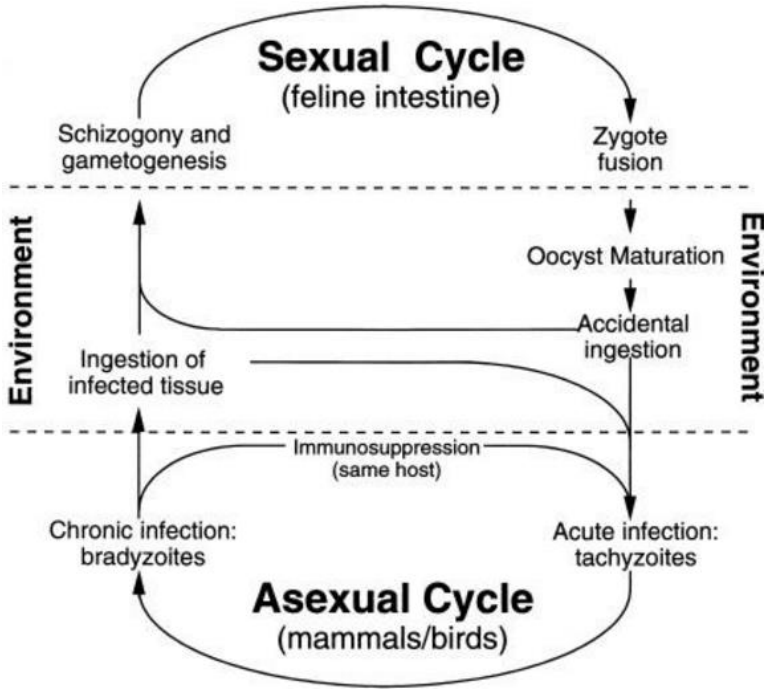



**Figure 1.2 Life Cycle of *T. gondii***

*T. gondii* has a complex life cycle consisting of a sexual reproductive cycle in the intestines of felines and an asexual reproductive cycle in any other warm-blooded nucleated cell (Dubey, 1998).

## 1.3 Parasite Multiplication

Tachyzoites and bradyzoites multiply inside the PV through the process of endodyogeny, producing two identical daughter cells. Endodyogeny begins with the production of an inner membrane complex in the middle of the cell which later produces the two compartments. The nucleus, mitochondria and other organelles divide and separate into the two compartments. A cleavage furrow forms and continues down the cell until the two daughter parasites are separated (Endo et al., 1982).

Tachyzoites multiply rapidly and, after several rounds of replication, cause the host cell to lyse and release the parasites through a process called egress. The free parasites are able to attach to another host cell and continue this cycle. The host immune system can quickly clear the tachyzoites from the host system. But this parasite can also convert to the bradyzoite phase causing chronic infection (Endo et al., 1982).

Bradyzoites are slow replicating forms of *T. gondii* which usually reside in tissue cysts of the muscle or central nervous system of the host organism. The bradyzoites are encased by a glycosylated cyst wall, which consists of parasite and host derived materials and protects the parasite from the host immune system (Ferguson et al., 1987). If the host organism becomes immunocompromised, the bradyzoites can differentiate back into the tachyzoite form and can cause damage in the host.

## 1.4 Toxoplasmosis

Humans acquire *T. gondii* through oral ingestion of contaminated food. Most infected individuals do not show any symptoms; although some may develop some flu like symptoms, headaches and swollen lymph nodes. However, toxoplasmosis of infants born from newly infected mothers may be affected with blindness, physical disability, or mental disability (Montoya et al., 2004). Toxoplasmosis of immunocompromised adults, can cause encephalitis, eye infection, lung infection or even death (Araujo et al., 1987).

## 1.5 *T. gondii* and Other Apicomplexa

The phylum Apicomplexa consists of a large number of protists, that are characterized by their elongated shape (Figure 1.1) with an apical end consisting of specific organelles which is termed the apical complex (Chobotar et al., 1982). A few of these organelles, such as micronemes, rhoptries, dense granules and connoids, are mainly responsible for attachment and invasion (Aikawa, 1988). Another structure which separates apicomplexans from the other phyla is the organelle called the apicoplast which was evolutionarily derived from the chloroplast but lost its photosynthetic abilities (Kohler et al., 1997). Although the function of this organelle is not clear, this organelle is crucial for parasite survival (Fichera et al., 1997).

Other members of this phylum are the *Plasmodium spp., Eimeria spp., Theileria spp., and Cryptosporidium spp.* which are all pathogenic agents. For example, *Plasmodium falciparum* causes malaria in humans and is transmitted through carrier mosquitoes (Martens and Hall, 2000). *Theileria spp.* and *Eimeria spp*. cause diseases in farm animals (Altay et al., 2008; Soulsby et al., 1982). *Cryptosporidium spp.* is another

apicomplexan parasite; it causes diarrhea, and is the leading cause of waterborne disease in the U.S. (Morrissette, 2002).

The apicomplexan parasite *Hammondia hammondi* is the closest relative to *T. gondii* with more than 95 % genetic synteny (Walzer et al., 2013). This high percent of genetic synteny means that the order of the genes in the chromosomes of these two organisms is about 95% similar, indicating recent divergence. They are both protozoa which can reproduce sexually only in feline intestines. Felines that are infected with either *T. gondii* or *H. hammondi* shed their respective oocysts which are indistinguishable from each other morphologically. These parasites are so similar that up until 1975, *H. hammondi* was thought to be another strain of *T. gondii* (Heydorn et al., 2001; Dubey et al., 2003). A noteworthy difference between the two parasites is in their ability to infect humans and other intermediate hosts. *T. gondii* is known to infect at least 30% of humans, while *H. hammondi* is not known to infect humans at all. Also, *H. hammondi* cannot infect birds, and is avirulent in mice. *T. gondii* has a wider variety of intermediate hosts and can infect a greater population percentage of intermediate hosts sharing with *H. hammondi*.

These apicomplexan parasites are harmful to the health of human and livestock. Study of parasite invasion and propagation in host organisms is important for discovering a cure. *T. gondii* can serve as the model organism because it is easily cultured in the laboratory, and its haploid genome allows for transient and stable transfection and selection of genetically altered clones (Black and Boothroyd, 1998).

## 1.6 The Genetic Manipulation of *T. gondii*

Transfections by DNA and RNA are performed with introducing vectors by electroporation. Stable transformations are fairly easy due to the parasite's haploid genome allowing for the selection of stable transgenic clones by homologous and non-homologous integrations (Donald and Roos, 1994). To select for transgenic parasites, there are a number of selection markers, such as uracil phosphoribosyltransferase and hypoxanthine xanthine guanine phosphoribosyl transferase, which are commonly used.

### 1.6.1 A Selection Marker, Hypoxanthine Xanthine Guanine Phosphoribosyl Transferase

*T. gondii* does not have the ability to produce purines through the *de novo* synthesis pathway. It thus relies on the purine salvage pathway for survival (Krug et al., 1989). Wildtype *T. gondii* has two enzymes that can produce XMP (xanthine monophosphate) which is a precursor to the purine GMP (guanine monophosphate). The first enzyme is hypoxanthine-xanthine guanine phosphoribosyl transferase (HXGPRT) which can convert xanthine to xanthine monophosphate (XMP). XMP later gets converted into GMP. The second enzyme is inosine monophosphatase which can convert inosine monophosphate (IMP) into XMP which later becomes GMP. If one of the two enzymes becomes inactive then the parasite depends on the other functional enzyme for survival. A common transgenic strain of *T. gondii* which is used for genetic manipulation is RHΔHX strain. In this strain of parasite, the HXGPRT gene is knocked out forcing the parasite to rely on inosine monophosphatase. To make genetic alterations in RHΔHX strain, HXGPRT can be integrated into the RHΔHX genome as a selectable marker in the process of knocking out genes or introducing other genes in the genome. To select for

parasites with HXGPRT, mycophenolic acid (MPA) and xanthine is used. MPA is used to inhibit inosine monophosphatase while xanthine is used as a substrate for the HXGPRT to produce XMP. Parasites which were transformed with the HXGPRT gene, survived in exposure to MPA and xanthine. The parasites which were not transformed in its genome would not survive because the MPA inhibited inosine monophosphatase which was the only enzyme available for the conversion of xanthine to XMP.

### 1.6.2 Dual luciferase Assay

The dual-luciferase assay is generally performed using two reporter plasmids that express *Renilla* luciferase and firefly luciferase. Firefly luciferase is a 61 kDa enzyme which oxidizes the substrate luciferin causing the emission of light (Wood et al., 1984). This reaction also requires ATP, $Mg^{2+}$ and $O_2$. The 36 kDa *Renilla* luciferase found in *Renilla reniformis* uses colenterazine as a substrate and oxidizes it to produce light (Matthew et al., 1977). Both of these enzymes do not have any post-translational modifications (Sherf et al., 1996). The emission of light is directly proportional to the amount of active luciferase. The dual luciferase system was used in the study of gene silencing caused by short and long double stranded RNA in *T. gondii* (Crater et al., 2012). Other applications of this system includes the study of promoters, intracellular signaling and mRNA processing (Solberg et al., 2013; Cheng et al., 2010)

## 1.7 CRISPR-Cas9 Technology

The Clustered-Regularly-Interspaced-Short-Palindromic-Repeats (CRISPR) technology is a method of genetic manipulation in many organisms (Mali et al., 2013). This technology is derived from the adaptive immunity of bacteria and archaea which uses this phenomenon to protect itself from foreign viral nucleic acids (Barrangou et al., 2007). The main components of the CRISPR gene editing technology consists of a Cas9 endonuclease, the 20-30 nts guide RNA and the RNA scaffold (Cong et al., 2013). The guide RNA and the Cas9 enzyme are brought together by the RNA scaffold, where the guide RNA base pairs with the genomic DNA, leading to a double stranded cleavage by the Cas9 endonuclease activity (Jinek et al., 2012). The ability of the CRISPR-Cas9 system to target a specific point on the genome allows for efficient and precise genome editing through the deletion, addition, or replacement of genes (Wang et al., 2014). Genes encoding proteins can be altered to produce a dysfunctional or an endogenously tagged protein (Lackner et al., 2015). The CRISPR system can be used to alter many target genes in parallel allowing for the identification of genes that play an important role in the phenotype of interest (Shalem et al 2014). An alternate form of the Cas9 enzyme which does not have endonuclease activity, can be used as a transcriptional activator or a repressor to a locus of interest (Konermann et al., 2013). This CRISPR-Cas9 system has been successfully used by researchers in many model organisms, including *T. gondii* (Shen et al., 2014; Sidik et al., 2014).

## 1.8 RNA Induced Post-transcriptional Gene Regulation in Eukaryotes

### 1.8.1 Pre-mRNA Processing

After the RNA polymerase II transcribes a pre-mRNA, there are many steps of mRNA processing before a mature mRNA is ready for translation. Messenger RNA processing allows for proper stability and translation of mRNA. There are three main steps. First step is the addition of a 5' cap to the 5' end of the pre-mRNA. A GTP is added to the 5' region of the pre-mRNA in the reverse orientation, after which methyl groups are added. This cap is important for the proper alignment of mRNA during translation (Cooper, 2000). Second step is the addition of the poly A tail. During this polyadenylation process the pre-mRNA is cleaved at the 3' end and a poly-A polymerase adds a 200 nucleotide poly-A tail to the transcript. The poly A tail is important in mRNA stability and in the translation process (Cooper, 2000). Another very important process in mRNA processing is the removal of introns by the process of splicing. Introns are noncoding genes are spaced throughout the coding regions called exons. The introns are spliced off by spliceosomes which loop the intron into a circle and cut it out after which the adjacent exons are joined together (Cooper, 2000).

### 1.8.2 RNA-Induced Gene Silencing

RNA-induced gene silencing is a cellular mechanism that regulates gene expression. RNA silencing usually involves a non-coding RNA molecule, which binds to other effector molecules such as proteins and cofactors (Catalanotto et al., 2000). After this, the RNA can base pair with an mRNA causing mRNA degradation or translational

repression through the help of these effector molecules (Arasu et al., 1991). The RNA guide involved in this silencing process can be classified in two types of RNA; long non-coding RNA and short non-coding RNA. Long non-coding RNA is usually more than 200 nucleotides to a few thousand nucleotides and short non-coding RNA are usually between 20 - 200 nts (Miller, 2014). There are two main types of short RNA which are responsible for gene regulation at the mRNA level in somatic cells; microRNA (miRNA) and short-interfering RNA (siRNA). Generally in animal cells, miRNAs are small dsRNA (~22 nts) which arise from hairpin loop (~ 70 nts) structures and partially pair with their target (Okamura et al., 2004). On the other hand siRNAs, which are also small dsRNA (~22 nts) arise from long dsRNA and they have almost perfect pairing with their target. We are going to be focusing on miRNAs since that is the topic of this study.

### 1.8.3 MicroRNA

MicroRNAs are double-stranded non-coding RNAs which are about 22-nt long (Figure 1.3). They are produced by the cleavage of a hairpin structure by RNase III enzymes (Ambrose et al., 2003). MiRNAs are important in post-transcriptional gene regulation. MiRNAs are loaded onto the effector complex referred to as the RISC factor, where they guide the effector complex to their mRNA targets. MiRNA can mainly bind to the mRNA through its first 8 nucleotides. This complementary binding of the miRNA to the mRNA can cause translational repression or mRNA degradation (Ambrose et al., 2003).

### 1.8.4 MicroRNA Biogenesis

The synthesis of miRNA begins in the nucleus where the primary transcript, called primary miRNA (pri-miRNA), is transcribed by RNA polymerase II or III as shown in Figure 1.3 (Lee et al., 2004). There are many sources of the pri-miRNAs. Pri-miRNA may arise from introns of coding and non-coding transcripts, in which case the pri-miRNA would share the promoter for the host gene. The pri-miRNA may also have its own promoter. The locus for transcription could contain only one pri-miRNA, or multiple pri-miRNA units. As a result, multiple miRNAs can be produced at the same time (Faller et al., 2008). Nuclear processing of the stem loop structure of pri-miRNA occurs where it is cleaved by the microprocessor complex, producing the precursor miRNA (pre-miRNA). The microprocessor complex consists of Drosha and its cofactor. Drosha is an RNase III enzyme, which contains two tandem RNase III domains and the double stranded RNA binding domain. Drosha cleaves the double-stranded RNA hairpin structure producing a two nucleotide over-hang. Cleavage of the pri-miRNA from two helical turns from the stem loop produces the pre-miRNA, which is about 65 bp long (Lee et al., 2003). The cofactor in the microprocessor consists of DiGeorge syndrome critical region gene 8 (DGCR8) in humans. DGCR8 is thought to be important in substrate recognition (Shiohama et al., 2003). There are homologous cofactors in other organisms (Denli et al., 2004).

The pre-miRNA is then exported out of the nucleus through nuclear pores by the nuclear transport receptor, exportin-5 (Lund et al., 2004). Once the pre-miRNA is in the cytoplasm, it comes into contact with another RNase III, called Dicer, for proper binding and cleavage of the pre-miRNA. The product of Dicer is a miRNA duplex around 22nt

with a 3' overhang at both ends (Lee et al., 2003). The passenger strand of the miRNA duplex is degraded, and the single stranded miRNA is loaded onto the Argonaute which is in the RNA Induced Silencing Complex (RISC) (Catalanotto et al., 2000). The RISC complex is guided to specific mRNA which can base pair with the miRNA seed region.

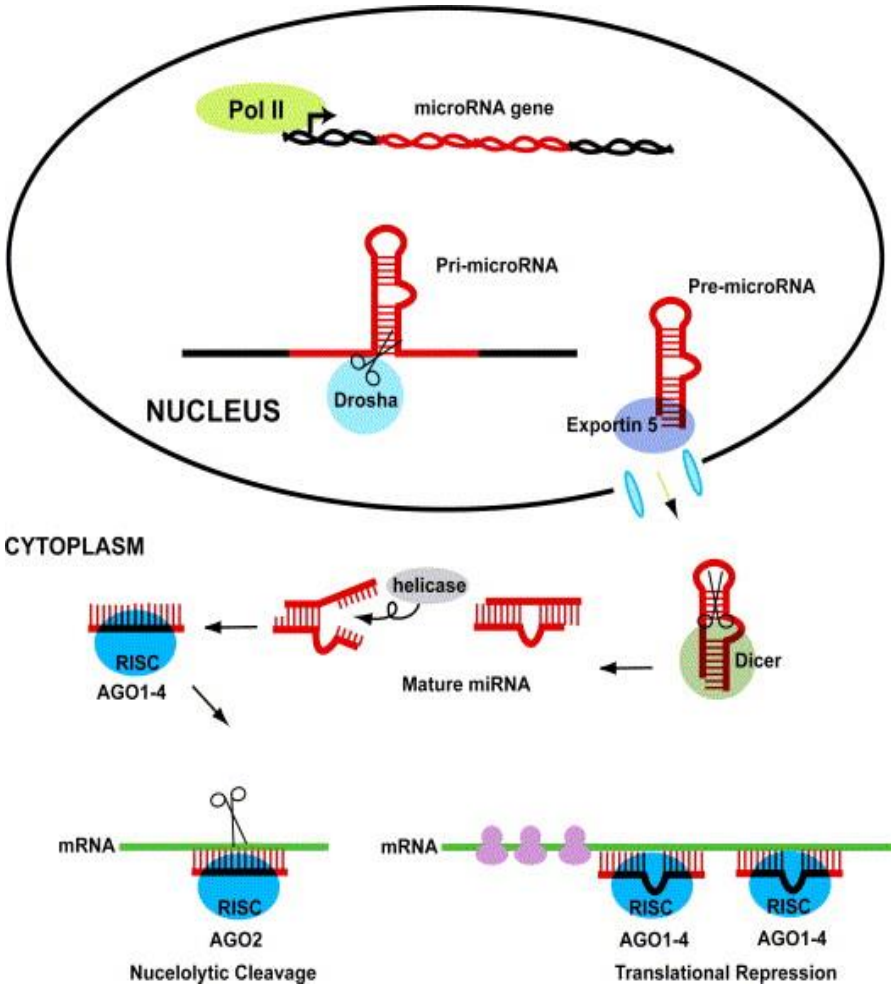



**Figure 1.3: MiRNA Biogenesis**

In the nucleus, the RNA polymerase II transcribes the pri-miRNA which is cut by Drosha producing a pre-miRNA. This product is exported through Exportin 5 into the cytoplasm, where Dicer cleavage action produces a mature miRNA duplex. One of the strands of the miRNA duplex associates with the RISC complex and guides the RISC complex to a mRNA target. The mRNA is either cleaved or is translationally repressed. Image is obtained from *Hammond, S. M*. 2005.

## 1.9 Ribonucleases

Ribonucleases are responsible for processing RNA precursors to a mature form, for producing alternate forms of RNA, or for the degradation of the RNA (Cudney et al., 1988; Canistraro et al., 1991). RNases can be divided into two separate categories; exoribonucleases and endoribonucleases. Exoribonucleases are mainly involved in RNA degradation and most of them remove nucleotides from the extremities (3' end or 5' end) of the RNA molecule.

Endoribonucleases cleave RNA within the strand. Some of these enzymes can cleave single-stranded RNA, while others cleave double-stranded RNA. Endoribonucleases usually cleave RNA using divalent cations as a cofactor. These enzymes usually produce a RNA product with a 3' hydroxyl and 5' phosphate. There are many different types of endoribonucleases, such as RNase H I, RNase H II, RNase E, RNase G, RNase P, RNase I and RNase III. RNase H I is a type of endoribonuclease that can cleave RNA in a RNA-DNA hybrid. It plays a role in DNA replication by removing the RNA primers in Okazaki fragments (Court et al., 1993). RNase H II is responsible for the removal of misincorporated ribonucleotides (Nomura et la., 1988). RNase E is part of the degradosome which is involved in the degradation of RNA (Carpousis et al., 2002). RNase G can also play a role in the degradation of mRNAs and is responsible for producing a mature rRNA (Li & Deutscher, 1996). RNase P plays a role in producing mature tRNAs (Altman et al., 1992). RNase I is different from other endoribonucleases as it does not require a divalent cation. It produces a 3' phophoryl product where all the other endonucleases produce a 3' hydroxyl (Sakamoto et al., 1983).

### 1.9.1 RNase III Family

There are three classes of RNase III. Class I consists of the simplest RNase III enzyme found in bacteria and yeast. These enzymes contain only one RNase III domain and one dsRNA binding domain (dsRBD) which is responsible for binding to dsRNA (Meador et al., 1990). Class II enzymes consist of Drosha and other Drosha homologs. These enzymes have two RNase III domains. Drosha may also contain a dsRBD, a proline rich region and a serine arginine rich domain (Wu et al., 2000). Class III enzyme is Dicer which is characterized by its two RNase III domains, along with a PAZ domain, and helicase domain (Carmel et al., 2004).

All RNases III contain at least one RNase III domain (Meader et al., 1990). The consensus sequence of RNase III domains is outlined in Table 1.1. The letters in red indicate the catalytic residues; glutamic and aspartic acids. RNA cleavage is predicted to occur through a single step, SN2 reaction, as shown in Figure 1.4. This reaction is dependent on two divalent metals such as $Mg^{2+}$, which are stabilized by the glutamic and aspartic acid residues. One of the metals activates a water molecule which carries out a nucleophilic attack on the phosphate of the RNA backbone. This nucleophilic attack produces two RNA molecules; one with a 3' hydroxyl group and another with a 5' phosphate.

**Table 1.1 :** RNase III consensus sequence in different species.

| | | RNase III Consensus Sequence | |
|---|---|---|---|
| Class I | Aa-RNase III<br>RNase Domain | 37 ETLEFLGD | 107 DVFE |
| | Hh-RNase III<br>RNase Domain | 3402 QRLEFLGD | 3564 DVVE |
| Class II | Hs-Drosha<br>RNase Domain 1<br>RNase Domain 2 | 966 ERLEFLGD<br>1144 QRMEFLGD | 1073 DLRE<br>1217 DHHE |
| | Mm-Drosha<br>RNase Domain 1<br>Rnase Domain 2 | 965 ERLEFLGD<br>1143 QRMEFLGD | 1072 DLRE<br>1143 DHHE |
| Class III | Gi-Dicer<br>RNase Domain 1<br>RNase Domain 2 | 333 QRLELLGD<br>642 QKLELLGD | 404 DMYE<br>716 DTFE |
| | Hs-Dicer<br>RNase Domain 1<br>RNase Domain 2 | 1303 ERLEMLGD<br>1693 QRLEFLGD | 1551 DCVE<br>1800 DIFE |
| | At-DCL1<br>RNase Domain 1<br>RNase Domain 2 | 1375 ERAELLGD<br>1594 QRLEFVGD | 1504 DVVE<br>1693 DIVE |

This table outlines the RNase III consensus sequence of seven different organisms. Class I RNase III consists of Aa-RNase III (Erdmann et al., 2012) and Hh-RNase III which are from *Aquifex aeolicus* and *Hammondia hammondi*. Class II consists of Drosha homologs such as Hs-Drosha (Kwon et al., 2016) and Mm-Drosha (Kwon et al., 2016) which are from *Homo sapiens* and *Mus musculus*. Class III contains Dicer homologs such as Gi-Dicer (Du et al., 2008), Hs-Dicer (Fortin et al., 2002) and At-DCL1 which are from *Giardia intestinalis*, *Homo sapiens*, and *Arabidopsis thaliana* respectively. The consensus sequences for Hh-RNase III and At-DCL1were determined using bioinformatics.

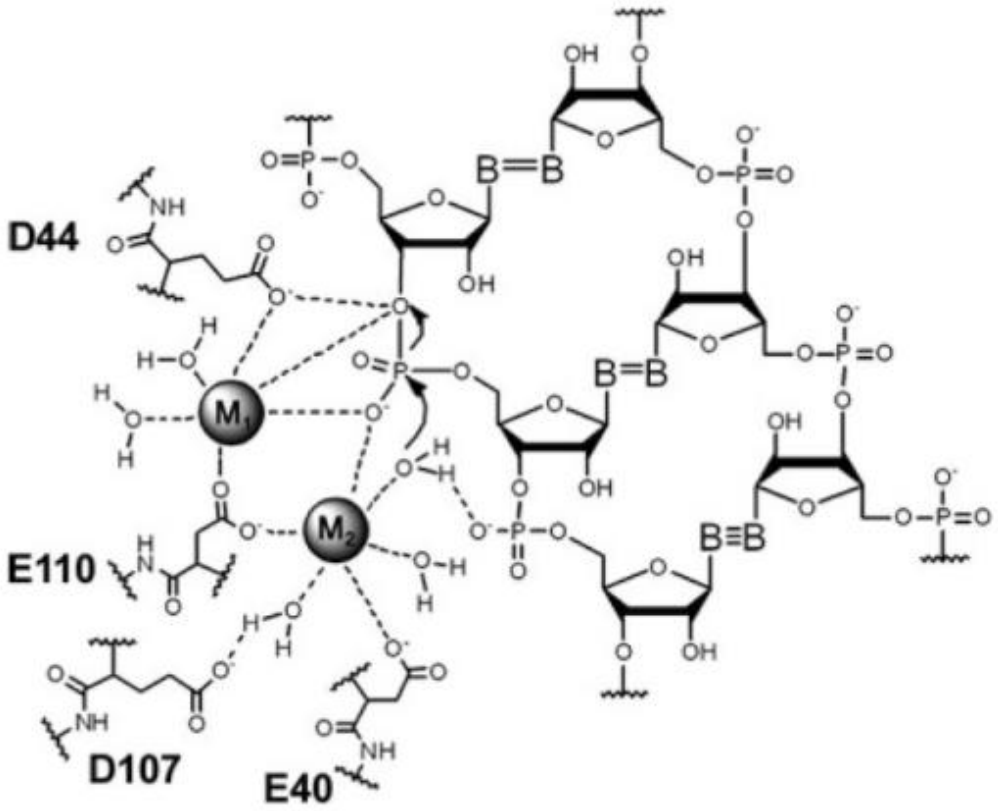



**Figure 1.4: The predicted mechanism for RNA cleavage of Aa-RNase III**

The metals are depicted by $M_1$ and $M_2$ and these metals are coordinated around water molecules, glutamic acid, and aspartic acid residues (E40, D44, D107, E110). $M_2$ activates a water molecule which generates a nucleophilic attack on a phosphate of the RNA backbone. This creates a 5' phosphate and 3' hydroxyl molecule producing a cleavage on one of the strands of the RNA duplex. This image was obtained from Erdmann et al., 2012.

### 1.9.2 Dicer and its Conserved Domains

Dicer is present in most eukaryotes. Plants are known to contain multiple Dicer homologs. For example *Arabidopsis thaliana* contains four dicer-like-proteins and these proteins have been thoroughly studied (Garcia-Ruiz et al., 2010). The different Dicer-like-proteins in *A. thaliana* have divergent roles in which certain Dicer-like-proteins cleave long dsRNA while others cleave miRNA hairpins. On the other hand mammals, such as humans and mice, have one Dicer gene. This single Dicer can cleave long dsRNA and hairpin stem-loop RNA structures. Deletion of the Dicer gene can cause major cognitive damage to mice and is sometimes lethal (Shin et al., 2009).

The canonical *H. sapiens* Dicer (Hs-Dicer), contains multiple domains; two RNase III domains, PAZ domain, dsRNA binding domain and the helicase domain (Zhang et al., 2004). All of these domains are highly conserved and crucial for endonuclease activity. A full length crystal structure of the Hs-Dicer has not been produced to date due to its large size (200 kDa), so the crystal structure of *G. intestinalis* Dicer (Gi-Dicer) (in Figure 1.5) will be used to illustrate some of the following Dicer domains.

As shown in Figure 1.5, Gi-Dicer contains tandem RNase III domains (RNase IIIa and RNase IIIb) which form an intramolecular dimer. Each RNase III domain is responsible for the cleavage of one strand of the dsRNA (Zhang et al., 2004). RNase IIIa domain processes the RNA strand which has the protruding 3'-OH, while RNase IIIb processes the opposite strand containing the 5'- phosphate (Zhang et al., 2004). Other than being the catalytic center, the RNase III domains are also reported to interact with the PIWI domains of Argonaute (Tahbaz et al., 2004).

### 1.9.3 PAZ domain

The PAZ domain of Dicer is important for anchoring one end of the dsRNA helix (Figure 1.5). The PAZ domain contains a conserved pocket which associates with 7 base pairs at the free end of the dsRNA containing the 3' overhang. It acts like a ruler between it and the RNase III catalytic sites. So the PAZ domain allows for the Dicer to be able to produce accurate and constant sized product of small RNA (Zhou et al., 2006).

### 1.9.4 Helicase Domain

Many enzymes that associate with DNA or RNA contain a helicase domain. Helicase domains are known for giving an enzyme the ability to unwind or remodel dsDNA or dsRNA (Soifer et al., 2008). The helicase domain of Dicer has been known to contribute to the binding of Dicer substrates; pre-miRNA and pre-siRNA. Of the two possible substrates, Dicer catalyzes the production of miRNA from pre-miRNA hairpins faster than the production of siRNA from pre-siRNA. The helicase domain may be responsible for this phenomenon because the deletion of the helicase domain enhances the cleavage of pre-siRNA (Tsutsumi et al., 2011). The Hs-Dicers and other Dicer-like proteins are known to localize in the cytoplasm. But there is also growing evidence that the Hs-Dicer can shuttle between the nucleus and the cytoplasm. The helicase domain is thought to play a role in the transportation of Hs-Dicer from the nucleus to the cytoplasm (Doyle et al., 2013). Another surprising role of the helicase domain of Dicer may be to transport the miRNA from Dicer to the RISC (Soifer et al., 2008).

### 1.9.5 Double-Stranded RNA Binding Domain

The double stranded RNA binding domain (dsRBD) is an auxiliary domain for substrate recognition of the dsRNA but it is not necessary for cleavage. But in the absence of the PAZ domain, the dsRBD becomes necessary for cleavage (Chakravarthy et al., 2010). Other than binding to dsRNA, there is evidence showing that the dsRBD of Hs-Dicer has a nuclear localization signal which implies that it may play a role in nucleocytoplasmic shuttling (Doyle et al., 2013).

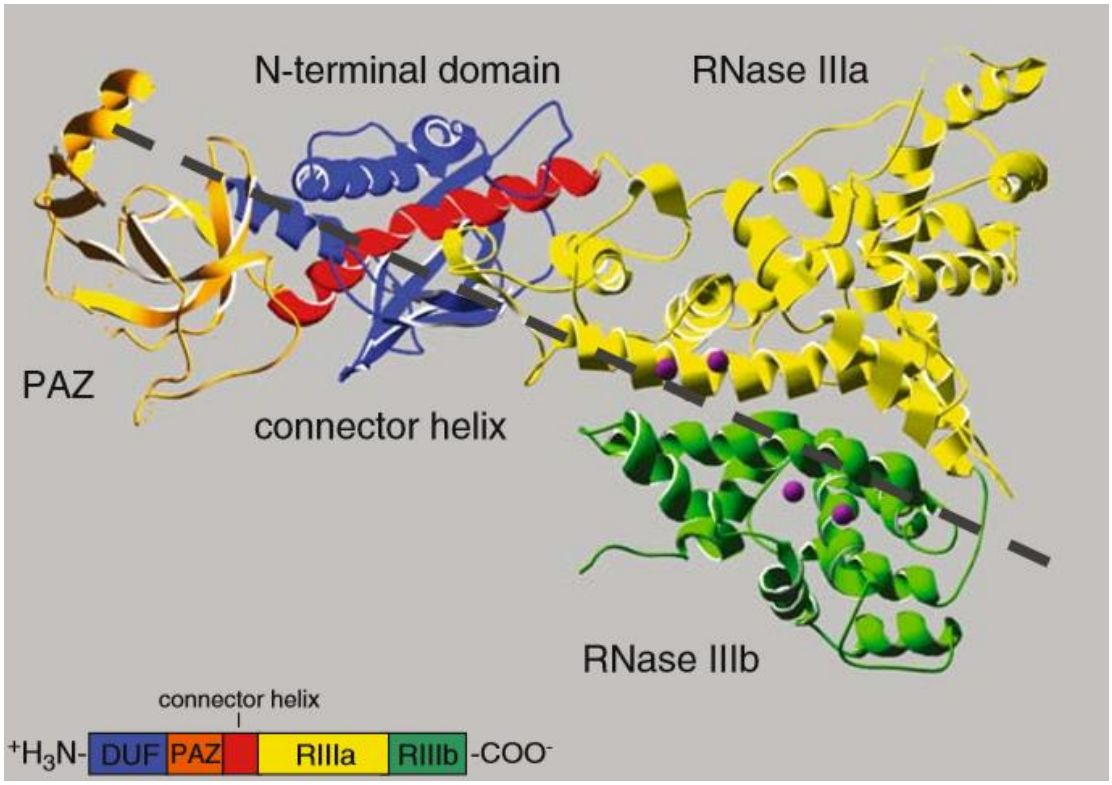



**Figure 1.5: The domains of *Giardia intestinalis* Dicer**

The crystal structure of *Giardia intestinalis* Dicer is depicted with its domains (Macrae et al., 2006). The predicted dsRNA is shown by a dashed line. The PAZ domain in orange is used to anchor the 3' overhang of the double stranded RNA. The RNase IIIa (yellow) and RNase IIIb (green) domains are responsible for cleaving the dsRNA and producing a 3' overhang on the other end of the RNA. The N-terminal DUF domain and the connector helix connecting the PAZ domain and RNase III a domain is coloured in blue and red. This image was obtained and modified from Paddison, 2008.

## 1.10 Potential MiRNA Target

After Dicer cleaves the pre-miRNA, a miRNA duplex is produced where one of the strands of the duplex, allows the RISC complex to bind to a mRNA and cause post-transcriptional silencing. Via base pair interaction, miRNAs bind their mRNA target and either cause the degradation of the mRNA or cause the mRNA to be translationally repressed. The canonical seed pairing of around 6 nt at the 5' end of the miRNA is generally thought to contribute significantly to the miRNA target binding. The seed paring may be a 6-mer where the base pairing may range from nucleotides 1-6, 2-7, or 3-8. A 7-mer ranges from nucleotides 1-7 or 2-8 and an 8-mer ranges from nucleotides 1-8. All of these types of base parings in the seed region may contain a mismatch. For example if a 7-mer which ranges from 2-8 contains a mismatch at nt 5 it would be called a 7-mer with one mismatch (Seok et al., 2016). It is difficult to predict the targets of miRNAs because, although it base pairs with its target, it base pairs partially.

## 1.12 Research Objective

The goal of this study was to gain a better understanding of RNA-based post-transcriptional gene regulation in *T. gondii*. The first objective was to study the function of Dicer by generating a transgenic Dicer knockout strain of *T. gondii* (TgDicer-mut). This TgDicer-mut was compared to the wild-type strain (TgDicer-wt) to determine the importance of Dicer in the duplication rate of the parasite. The dual luciferase system was used to test whether endogenous miRNA-directed gene silencing can occur in TgDicer-mut. This reporter system was also used to test exogenous short and long dsRNA-directed gene silencing in the both the wild-type and transgenic strain.

The second objective of this study was to identify and verify a miRNA target in *T. gondii*. This was done by using base pair prediction of the *TgHODI* to determine possible binding sites. MiRNA inhibitors were introduced to determine if the level of a flag tagged TgHODI is affected, through the use of western analysis.

# Chapter II

# Materials and Methods

## 2.1 Human Fibroblast and Parasite Cultures

Dulbecco's Modified Eagle Medium (DMEM) supplemented with D-glucose (25 mM), L-glutamine, 10% cosmic calf and 0.5x antibiotic-antimycotic was used to grow and maintain Human foreskin fibroblast (HFF) cells. These cells were kept in this media in 5% $CO_2$ at 37$^{o}$C for optimal growth.

RHΔHX type I parasites (TgDicer-wt) were cultured in the HFF cells and maintained in ED1 media which consisted of Minimal Essential Media Eagle (MEM) with 1% dialiyzed fetal bovine serum (dFBS) and 0.5x antibiotic-antimycotic. The transgenic strains (TgDicer-mut and TgHODI-SF-HX) were maintained in ED1 media with the addition of mycophenolic acid (25 mg/mL) and xanthine (25 mg/mL).

## 2.2 Plasmid Construction

The parental pU6-Universal plasmid was used to construct the CRISPR-Cas9 plasmid to target the *Tg-Dicer* gene. The pU6-universal plasmid was obtained from S. Lourido from the Whitehead Institute for Biomedical Research via Addgene (Cat #52694). An illustration of pU6-Universal is shown in the Appendices (Figure A1). A PCR reaction was performed using two primers, gRNADicerEcoN1 and U3upstrem, to introduce specific binding sites of gRNA to the *Tg-Dicer* locus. A list of all the primers used in this study is in the Appendicies (Table A1). The 8.9 kb amplified product was purified by gel extraction (QiaexII Gel Extraction kit (Qiagen #20021)) following the

manufacturer's protocol. The linear fragments were ligated by a T4 DNA ligase (Fermentas #02076201) and a restriction digest was performed for analytical purposes. This new plasmid targeting the *Tg-Dicer* gene was named gRNADicer (Figure A1, Appendices).

### 2.3 Generation of Truncated TgDicer Parasite Using CRISPR-Cas9 Technology

The CRISPR-Cas9 system was used for the insertion of a YFP HX gene at the *Tg-Dicer* locus. The gRNADicer plasmid contained the DNA sequence for the two main components of the CRISPR-Cas9 system; the guide RNA and Cas9 endonuclease. The 20 nt guide RNA sequence targeting the *Tg-Dicer* gene in exon 5 (Figure 3.6) is driven by the U6 promoter. The Cas9 gene in this plasmid is driven by the TgTub1 promoter. The resultant truncated parasite is named TgDicer-mut.

The YFP HX insertion contained a gene for the yellow fluorescent protein along with an HXGPRT cassette. The YFP gene was inserted at the same reading frame as the *Tg-Dicer* gene for easy detection of a transgenic parasite clone. The HXGPRT cassette is a transcription unit for the expression of hypoxanthine xanthine guanine phosphoribosyl transferase and was used as a selection marker for transgenic parasites containing this gene. The YFP HX insertion was made from a previously constructed plasmid called pDicerYFP-HX (Figure A2, Appendices). To make this YFP HX amplicon, the following oligonucleotides were used; P3_Fw and Dicer_Homology_HX_Rev. For proper integration into the *Tg-Dicer* locus, this PCR product contained homologous regions at both ends; a 100 bp homology at the 5' end of the amplicon and a 20 bp homology at the 3' end. The amplification of this plasmid would produce a 3.6 kb PCR product.

Around 1 x $10^6$ freshly lysed RHΔHX parasites were electroporated with the 100 µg of CRISPR-Cas9 plasmid (gRNADicer) and 40 µg of the YFP HX amplicon. Microscopy was used to check for YFP signal which showed that some parasites were fluorescent. To isolate a single fluorescent clone, a serial dilution on a 96 well plate was performed. Eight single clones were isolated and out of which, only three were fluorescent. Only one of those three clones showed verification of the transgenic TgDicer-mut through a PCR analysis. The sequence of the oligonucleotide for the PCR analysis are shown in Appendix A.

## 2.4 Transfection by Electroporation

To transfect *T. gondii* around $10^6$ freshly lysed parasites were electroporated using a BTX ECM 630 at 1500 volts (25 Ω, and 25 µF). Each electroporation contained the parasites in 400 µL of electroporation mixture (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) (Roos et al., 1994). The parasites were electroporated in 4 mm-gap cuvettes after which they were subcultured into 40 mm plates with confluent HFF cells in ED1 media.

## 2.5 Dual Luciferase Assay

The dual luciferase assay was performed using the parasite strains; RHΔHX parasites, TgDicer-mut, and TgHODI-SF. Approximately $10^6$ freshly lysed parasites were electroporated with the 5µg each of the FF and RN plasmids. Both of the FF and RN plasmids express transcripts containing a 5' UTR derived from tubulin and a 3' UTR derived from DHFR. The FF transcripts had no miRNA binding sites, and served as the internal control. Four different RN constructs were used; RnoB, Rn60a, Rn4a and RnLet7

where the transcript for RnoB did not contain any miRNA binding sites. Rn60a, Rn4a and RnLet7 produced transcripts which contained three miR60a, miR4a and let7 miRNA binding sites respectively, in the 3'-UTR of the RN transcripts. Parasites were electroporated with the reporter constructs. They were also electroporated with or without RNA effectors and harvested two days later.

The parasites lysates were prepared by using 100 µL of Passive Lysis Buffer (Promega, #E1531) for 15 minutes at room temperature then centrifuging at 10,000 g for 2 minutes to separate the supernatant from the cell lysates. 20 µL of the supernatant was added separately to FF and RN reaction mixtures. The FF mixture consists of 20 mM Tricine, 10 mM $MgSO_4$, 5 mM DTT, 250 µM ATP and 250 µM Coenzyme A and 200 µM D-luciferin (Sigma Aldrich, #L9504)). The RN mixture consists of 100 mM $K_2HPO_4/KH_2PO_4$ (pH 7.6) 500 mM NaCl, 1 mM EDTA, 0.02% BSA and 0.1 µM Coelenterazine (Santa Cruz Biotechnology Inc, #sc-205908). A 20/20n Luminometer (Turner Biosystem) was used to measure the luminescence signals. Two trials were performed for each dual luciferase assay.

## 2.6 Small RNA Isolation and Imaging

Small RNA from RHΔHX and TgDicer-mut was isolated from $2 \times 10^6$ freshly lysed parasites using the RNAzol RT (Molecular Research Center, Inc., #RN 190) according to the manufacturer's protocol for miRNA isolation. The RNA was separated according to size on a 15% Tris-borate urea polyacrylamide gel and Sybr Gold Nucleic Acid Stain (Molecular Probes, S11494) to reveal the RNA bands.

## 2.7 *In vitro* Transcription of RNA

In the *in vitro* transcription of the miRNA inhibitors (anti-4a and anti-60a), 2 μM of the miRNA inhibitor template (miR4a_upper and mir60a upper) and 2 μM of T7promoterGG were mixed and heated to 75°C. It was then slowly cooled and left on ice for the oligonucleotides to anneal. One fifth of the annealed template was added to a reaction mixture containing 5 units of T7 RNA polymerase, 80 mM Hepes-KOH at pH 7.5, 24 mM $NaCl_2$, 2 mM spermidine, 40 mM DTT, 10 mM rNTPs, and 1 unit of pyrophosphatase. Then this 50 μL reaction was incubated at 37 °C for 16 hours. Reactions were then extracted with one volume of phenol-chloroform mixture (1:1), and resultant RNAs were precipitated and quantified using the Thermo Scientific NanoDraop2000. The RNA bands were separated by gel electrophoresis to check for integrity.

The miRNA is a small RNA duplex containing some mismatches. It was made by separately *in vitro* synthesizing each strand of the miRNA. The primers T7promoterGG and Sense_Tpl was used to synthesize one of the strands of the miRNA. T7promoterGG along with Antisense_mismatch_tpl was used to make the other strand of the miRNA. Then the two strands of RNA were heated to 75°C and slowly cooled and left on ice for the RNA to anneal. The siRNA was created through the same method except the Antisense_mismatch_tpl was replaced with Antisense_match_tpl. The long dsRNA was made by PCR amplifying pTubRnluc using the following primers: FW_RnLuc_1 and RV_RnLuc_1. These two primers contain T7promoterGG binding sequence and a similar protocol as described above was used to synthesize the long dsRNA.

## 2.8 Western Analysis

Parasites were collected, and 50 µL of RIPA buffer was used for lysing. Then the samples were sonicated for 10 seconds and then incubated on ice. The RC DC Protein Assay (Bio-Rad Laboratories, Cat #500-0119) was used to measure the protein concentration. A 6x SDS protein loading dye was used to denature the lysates, and the samples were separated by size on a 10% SDS-PAGE for 1 hour. Resolved proteins were transferred on to a nitrocellulose membrane. The nitrocellulose blots were blocked with 5% (w/v) skim milk in TBST for 30 minutes. The blots were incubated with primary antibodies in 2% (w/v) skim milk in TBST for 1 hour. The primary antibody which was used for the Flag tagged HODI is mouse anti-FLAG (1:5,000) (ABM Inc. #G191) and rabbit anti-LDH1 (1:2,000). The western blot was then washed 3 times for 5 minutes with TBST and incubated with their proper secondary antibodies for 1 hour at room temperature. The secondary antibodies used were: horseradish peroxidase (HRP) anti-mouse (1:10,000) (Rockland, #19072) and HRP-anti-rabbit (1:10,000) (Rockland, #15949). The blots were then washed again and Chemiluminescent HRP Substrate Kit (Millipore) was used to visualize the bands. The FluorChem Q Imager (Alpha Innotech) with AlphaView-FluorChem Q software was used in the visualization process.

# Chapter III

# Results and Discussion

## 3.1 Bioinformatic Analysis of Putative Dicer in *T. gondii*

*T. gondii* is predicted to contain only one Dicer-like protein (Gene ID: TGGT1_267030). This putative Dicer (Tg-Dicer) contains two RNase III domains and a RNA helicase domain and lacks the double-stranded RNA binding (DSRB) domain and the PIWI-ARGONAUTE-ZWILLE (PAZ) domains (Braun et al., 2010). This pattern of domains is very similar to the Dicer-like protein 1 (DCL1) of *Chlamydomonas reinhardtii* as shown in Figure 3.1. Tg-Dicer and the DCL1 enzyme in *C. reinhardtii* have a 48 % identity. Another similarity between these two proteins is that Tg-Dicer and *C. reinhardtii* DCL1 are also significantly larger than Dicer homologs from *H. sapien, A. thaliana, S. pombe* and *G. intestinalis*. For example, Tg-Dicer is predicted to be 498.5 kDa while the Hs-Dicer is only 200 kDa. *Tg-Dicer* gene is located in chromosome IX and is 21,503 bp long. It contains 15 exons and the transcript length is 13,491 bp.

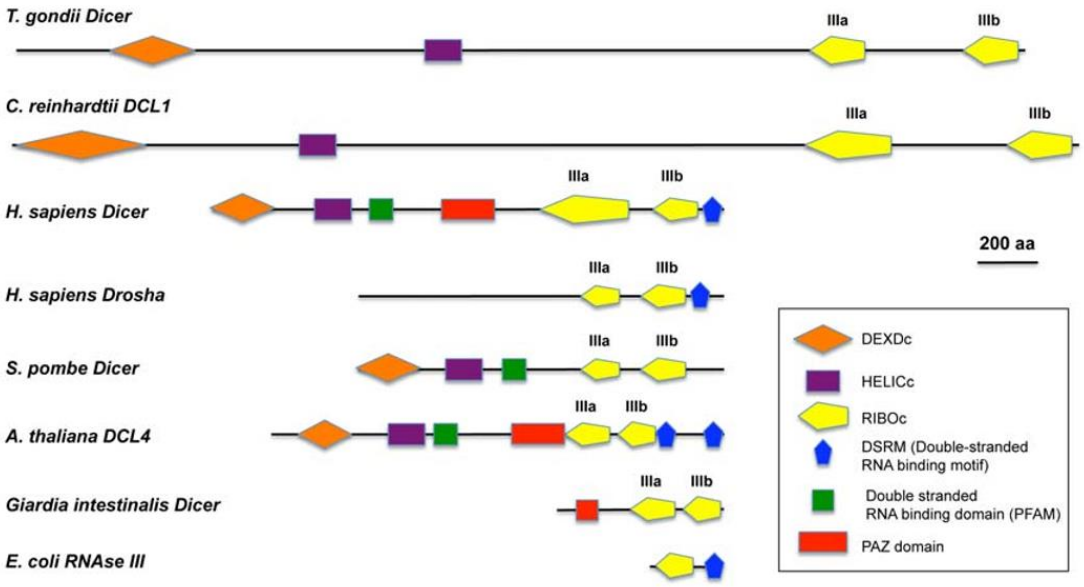



**Figure 3.1: Comparing the size and the domains of Tg-Dicer with other Dicers**

Tg-Dicer is much larger compared to the other well known Dicer homologs. Tg-Dicer contains two RNase III domains (red) and the Helicase domain (orange and purple). It lacks a PAZ (red) and a dsRNA binding domain (blue). *C. reinhardtii* DCL1 appears to have a similar size and it also consists of similar domains as Tg-Dicer. The image was obtained from Braun et al., 2010.

A clustal analysis of each of Tg-Dicer domains was performed using the amino acid sequences of three well studied Dicer-like proteins; Hs-Dicer, At-DCL1 and Gi-Dicer. The percent identity of Tg-Dicer with Hs-Dicer, At-DCL1 and Gi-Dicer are as follows; 48%, 38 % and 43% respectively. Figures 3.2-3.5 show the clustal analysis of Tg-Dicer where the motifs are highlighted in yellow. An asterisk indicates a fully conserved residue, two dots indicate strongly similar properties and one dot indicates a weakly conserved residue. The clustal analysis of the helicase domain is separated into the Helicase ATP binding domain shown in Figure 3.2 and the C-terminal helicase domain in Figure 3.3. Since Gi-Dicer does not contain a helicase domain, Figure 3.2 and 3.3 only has alignment of Tg-Dicer with Hs-Dicer and At-DCL1. In Figure 3.1 the conserved DEAH motif of the helicase domain is highlighted in yellow. The RNase IIIa domain of Tg-Dicer is aligned with the RNase IIIa domains of all three of the other homologous Dicers in Figure 3.4 and RNase IIIb of Tg-Dicer is aligned with the RNase IIIb domains of the other Dicers in Figure 3.5. The conserved signature motifs of both RNase III domains are highlighted in yellow.

## Helicase ATP Binding Domain

```
Tg-Dicer      RDTRGQLQAWPYQYRVYRRATLENTVVALPTGTGKTLVAAMVVHTWLANVF----VADKL 569
Hs-Dicer      AIHDNIYTPRKYQVELLEAALDHNTIVCLNTGSGKTFIAVLLTKELSYQIRGDFSRNGKR
At-DCL1       KEKVVEEQARRYQLDVLEQAKAKNTIAFLETGAGKTLIAILLIKSVHKDLMSQ--NRKML
                        **  : . *  .**:. * **:***::* :: :    ::

Tg-Dicer      VIFLAPTTQLARQQHRAISTVLSLLRDPAGAEASRGRLAVRLPGVLEKERRTRGRVASRD 629
Hs-Dicer      TVFLVNSANQVAQQVSAVR-------------------------------------------
At-DCL1       SVFLVPKVPLVYQQAEVIR-------------------------------------------
               :**. ..  . **  .:

Tg-Dicer      DDDGERAYFRSVLRVDWSHPKYWAAFYSALSEYKAALLDLQRLPAAPPAAFDLRRFHTGR 689
Hs-Dicer      ------------------------------------------------------------
At-DCL1       ------------------------------------------------------------


Tg-Dicer      LLAPYRFVDSQRAALLRESFAEPSSEEANVQAASLEDSEGERSGNREGEREEDENEQDVS 749
Hs-Dicer      -----------------------------------THSDLKVGEYSNLEVNA---SWTK
At-DCL1       -----------------------------------NQTCFQVGHYCGEMGQD---FWDS
                                                  : *.  .   :       .

Tg-Dicer      EVEHQAVCWPRVFVMTPDKFRNLACRGLLRIDRIGLLIFDEAHTVGRPLRSLGAYATILR 809
Hs-Dicer      ERWNQEFTKHQVLIMTCYVALNVLKNGYLSLSDINLLVFDECHLAIL----DHPYREIMK
At-DCL1       RRWQREFESKQVLVMTAQILLNILRHSIIRMETIDLLILDECHHAVK----KHPYSLVMS
              .  :: .   :*::**     *:  .. : :. *.**::**.* .         *  ::

Tg-Dicer      VFLHLCQEEAKPRILGLTASPVPDGKLGFFDLESELKRAMRSLETIFQARIVKADT---- 865
Hs-Dicer      LCENC---PSCPRILGLTASILNGKCDPE-----ELEEKIQKLEKILKSNAETATDLVVL
At-DCL1       EFYHTTPKDKRPAIFGMTASPVNLKGVSS---QVDCAIKIRNLETKLDSTVCTIKDRKEL
                 :       * *:*:*** :            :    ::.**. :.:   .

Tg-Dicer      ------------ESAPGIRDWSQGQPD---------------------------------- 880
Hs-Dicer      DRYTSQPCEIVVDCGPFTDRSGLYER---LLMELEEA---------LNFINDCNISVH--
At-DCL1       EKHVPMPSEIVVEYDKAATMWSLHETIKQMIAAVEEAAQASSRKSKWQFMGARDAGAKDE
                           :       .  :
```

**Figure 3.2 :Clustal Analysis of Tg-Dicer Helicase ATP Binding Domain**

A clustal analysis of Tg-Dicer Helicase ATP binding domain with Hs-Dicer and At-DCL1 shows homology. The DEAH box of the Helicase is highlighted in yellow.

**Helicase C-terminal Domain**

```
Tg-Dicer      QGASWLF---------GIGKASQGRSS------------------CLDRLRAMHTVWGKL 1729
At-DCL1       AAEKVAA---------EVGKPENGNAHDEMEEGELPDDPVVSGGEHVDEVIGAAVADGKV
Hs-Dicer      ASLDLKFVTPKVIKLLEILRKYKPYERQQFESVEWY----------NNRNQDNYVSWSDS
               . .            : :  :                       :.     .  ..

Tg-Dicer      EKKT-QKLNAQKHRP-STTPQSHSPSLASS-LSSSLSSSLSSSASSSASSSASSSLSSSL 1786
At-DCL1       TPKV-QSLIKLLLK------YQHTADFRAIVFVERVVAALVLPKV----FAELPSLSFIR
Hs-Dicer      EDDDEDEEIEEKEKPETNFPSPFTNILCGIIFVERRYTAVVLNRLIKEAGKQDPELAYIS
                .  :.      :        .:  : .  : .   :::             .*:

Tg-Dicer      SSSLSSSASSSASW----GDGGKGVWRLMRRLGWGAPLQLLVSTSVLEEGIDVPACNLVV 1842
At-DCL1       CASMIGHNNSQ-------EMKSSQMQDTISK-FRDGHVTLLVATSVAEEGLDIRQCNVVM
Hs-Dicer      SNFITGHGIGKNQPRNKQMEAEFRKQEEVLRKFRAHETNLLIATSIVEEGVDIPKCNLVV
              .  : .   ..                : :       **::**: ***:*:  **:*:

Tg-Dicer      QMDMPSSLVRFVQAKGRARKKPAEFVVLCPDSGPCGASW-PDASLPFSSSFSS------- 1894
At-DCL1       RFDLAKTVLAYIQSRGRARKPGSDYILMVERGNVSHAAF-LRNARNSEETLRKEAI----
Hs-Dicer      RFDLPTEYRSYVQSKGRARAPISNYIMLADTDKIKSFEEDLKTYKAIEKILRNKCSKSVD
              ::*: .    ::*::****   ::::::   .              .. : .

Tg-Dicer      -----------SSSFFPFSSSLQPQT------PDCTVSSPSLL----------------- 1920
At-DCL1       -ERTDLSHLKDTSRLISID--AVPGTVYKVEATGAMVSLNSAVGLVHFYCSQLPGDRYAI
Hs-Dicer      TGETDIDPVM-------DDDDVFPPYVLRPDDGGPRVTINTAIGHINRYCARLPSDPFTH
                              .    *         .  *:  : :
```

**Figure 3.3: Clustal Analysis of Tg-Dicer Helicase C-terminal Domain**

A clustal analysis of Tg-Dicer C-terminal domain with Hs-Dicer and At-DCL1 shows homology. The sequence of the Tg-Dicer Helicase C-terminal domain ranges from residue 1697 to 1920.

**RNase IIIa Catalytic Domain**

**Figure 3.4: Clustal Analysis of Tg-Dicer RNase IIIa Catalytic Domain**

A clustal analysis of Tg-Dicer RNase IIIa Catalytic Domain with Gi-Dicer, Hs-Dicer and At-DCL1 shows homology. The Catalytic Residues are highlighted in yellow.

**RNase IIIb Catalytic Domain**

**Figure 3.5: Clustal Analysis of Tg-Dicer RNase IIIb Catalytic Domain**

A clustal analysis of Tg-Dicer RNase IIIb Catalytic Domain with Gi-Dicer, Hs-Dicer and At-DCL1 shows homology. The Catalytic Residues are highlighted in yellow.

## 3.2 Generation and Analysis of a Transgenic *T. gondii* Strain, Whose Tg-Dicer Expression is Abrogated.

To investigate the significance of Tg-Dicer and its role in post-transcriptional gene regulation, we created a transgenic parasite strain whose expression of functional Dicer is abolished by an insertion mutagenesis. The CRISPR-CAS9 methodology was used to generate this parasite strain by disrupting the coding sequence of Tg-Dicer locus by inserting a DNA fragment in exon 5 of *Tg-Dicer*. A schematic diagram of the *Tg-Dicer* gene in the parental parasite strain and in the transgenic TgDicer-mut parasite strain is shown in Figure 3.6. In the wildtype *Tg-Dicer*, the exons are depicted with grey boxes. The helicase ATP binding domain, the helicase C terminal domain, and two RNAse III domains are highlighted in orange, purple and yellow respectively. The site of CRISPR-CAS9 cleavage is shown with an arrow on exon 5. The guide RNA binding site is also highlighted in red in this DNA sequence. The DNA fragment, which is to be inserted, is shown below the parental strain. This DNA fragment consists of an open reading frame of YFP and a transcription unit encoding HXGPRT, a selectable marker. The 5' end of the insertion contains a 100-bp homologous region to the end of exon 5 while the 3' end of the insertion contains a 20-bp homologous region to intron 5 which is not labeled but it is indicated by a grey line next to exon 5. Since the insertion is in the middle of the gene and upstream of the catalytic RNase III domains, this insertion would disrupt the amino acid sequence of Tg-Dicer protein, and give a non-functional Tg-Dicer protein.

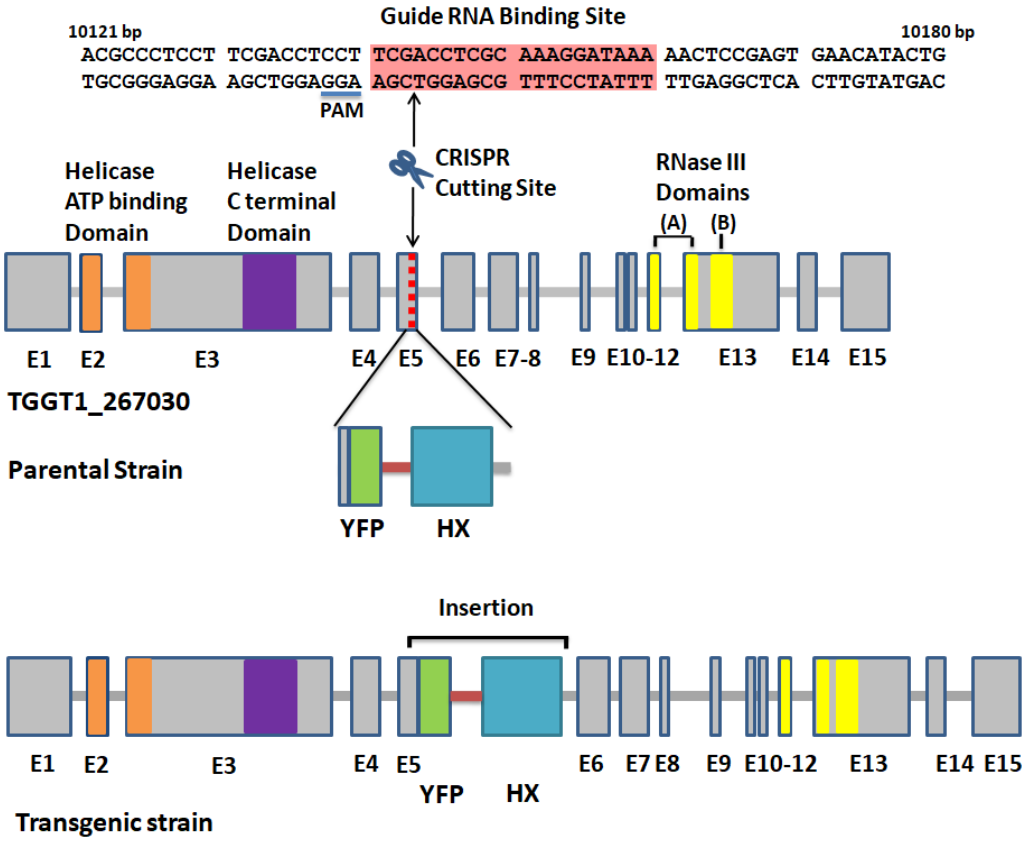



**Figure 3.6: Representation of Tg-Dicer locus in the parental strain and in the transgenic parasite following the CRISPR-CAS9 genome editing.**

The Tg-Dicer gene ID is TGGT1_267030 and this gene contains the helicase ATP binding domain, the helicase C terminal domain and two RNAse III domains shown in orange, purple and yellow respectively. E1 through E15 represents the 15 exons in *Tg-Dicer*. The CRISPR cutting site is shown in a red dotted line on exon 5 (E5). The 20 bp sequence for the CRISPR binding site is highlighted in red above the *Tg-Dicer* gene with an arrow pointing at the nucleotide bases where the cleavage is predicted to occur. The PAM (Protospacer Adjacent Motif) is depicted with a blue underline. The insertion contained a yellow fluorescent protein (YFP) and a HXGPRT cassette labeled as HX and they are shown in green and blue respectively. The 5' and 3' end of this amplicon contains homologous regions to the *Dicer* gene. The TgDicer-mut gene locus with the insertion of the YFP HX gene is shown at the bottom of the figure and is labeled as the transgenic strain.

Following transfection, transformed parasites were subjected to a limiting dilution to obtain a clonal strain expressing YFP. All clonal strains which were fluorescent were selected (Figure 3.7). Three clones were initially selected for further PCR analysis.

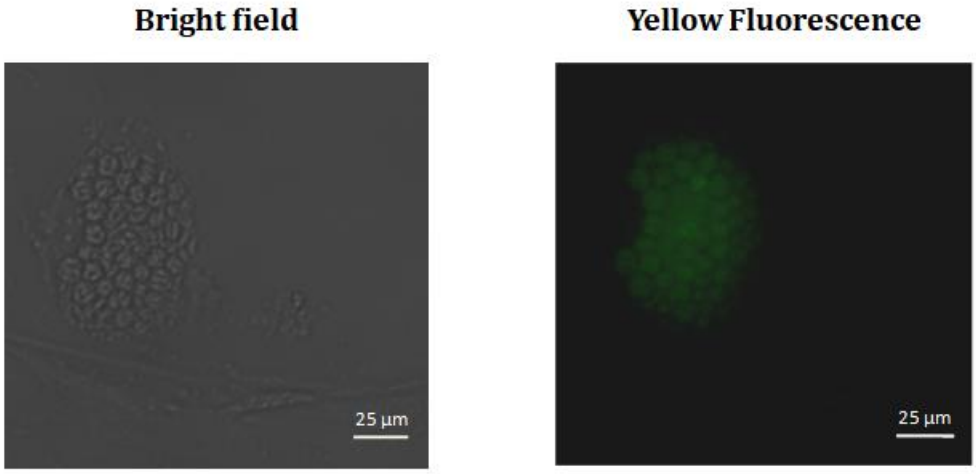



**Figure 3.7: Screening for Parasite Clones Producing YFP**

A representative image of the screening for parasites expressing YFP. A PV containing many transgenic parasites are shown with 25 µm bar at the bottom of the images. A Leica DMI 6000B inverted fluorescent microscope was used with a HCX PL Apo 100×/1.40–0.70 objective using the Leica Advanced Fluorescence Application Software V.2.3.0.

To confirm that DNA insertion is at the locus of *Tg-Dicer*, PCR analysis was performed using oligonucleotides specific to the locations indicated on Figure 3.8. The analysis was conducted using genomic DNA isolated from the transformed parasites with YFP signals, and its parental strain, which was referred to as TgDicer-wt. Out of three positive YFP clones, only one had positive confirmation through PCR analysis. Reactions A, B, and C aimed to show the integration of the YFP HX fragment in the genomic locus of the TgDicer, when the clonal line was correctly targeted. A representative image of the PCR analysis was shown in Figure 3.8. This clonal line was referred to as TgDicer-mut.

Reaction A showed the 5' integration of the YFP HX gene in the genome, while Reaction B showed the 3' end integration. As shown in Figure 3.8, Reaction A produced a 1.5-kb band for TgDicer-mut while TgDicer-wt did not have any bands. Reaction B produced a 2.3-kb band for TgDicer-mut while TgDicer-wt did not have any bands. Reaction C contains primers which are both outside the insertion, creating a large band for the TgDicer-mut (5-kb) while the TgDicer-wt has a much smaller band (1.5-kb). The last PCR with GAPDH primers served as a positive control in this experiment to show that appropriate genomic DNA, quality, and quantity were used.

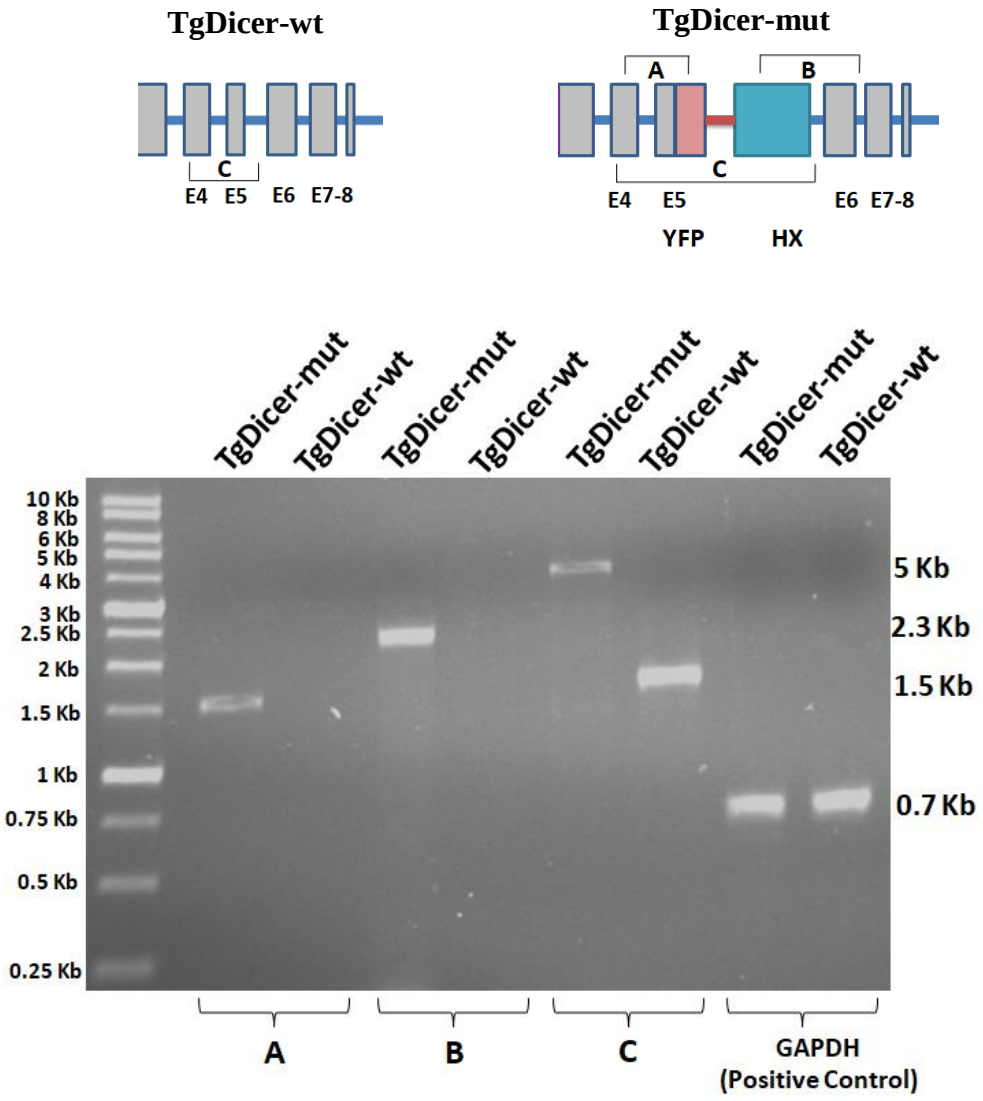



**Figure 3.8 Analysis of transgenic TgDicer-mut strain**

A schematic diagram of the genome for TgDicer-wt and TgDicer-mut is shown at the top. PCR analysis of transgenic parasite and parental strain. PCR A and B showed the 5' and 3' integration of the YFP HX in the genome of TgDicer-mut compared to the parental line. PCR C contains primers outside of the insertion so the parental strain contains a much smaller band than the TgDicer-mut. The GAPDH PCR served as a positive control.

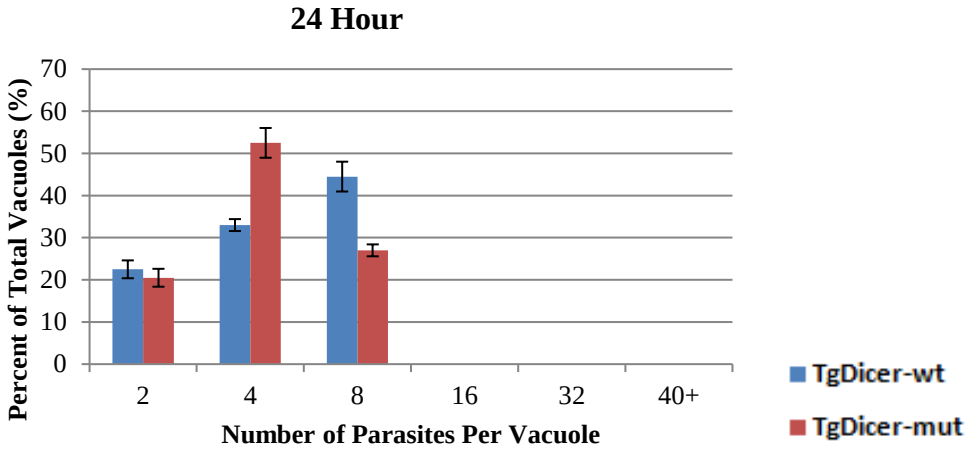
### 3.3 Loss-of-Function Effect of Tg-Dicer in its Doubling Time

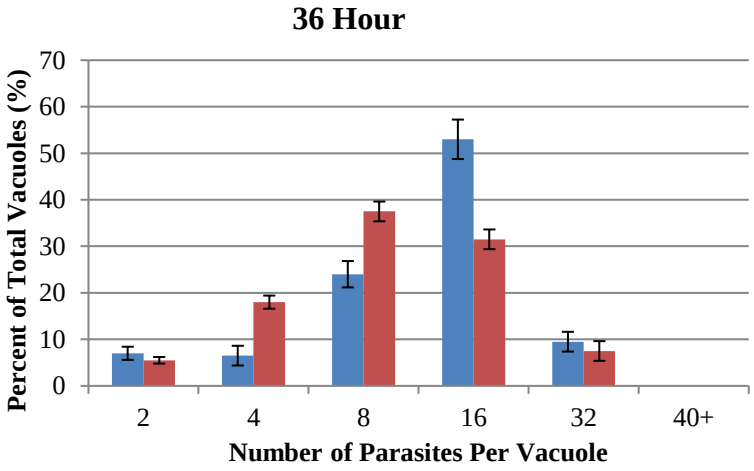
To determine if the abolishment of functional Dicer protein affects the growth of the parasite, an assay for the parasite's doubling time was carried out in the transgenic TgDicer-mut, and compared to its parental strain. As *T. gondii* multiplies via endodyogeny within the parasitophorous vacuole, we therefore counted the number of parasites per vacuole at three different time points; 24, 36 and 48 hours (Figure 3.9). At 24 hour post infection of the parental strain, the majority of vacuoles had 8 parasites ($2^3$) in each vacuole. In TgDicer-mut, majority of vacuoles contained 4 parasites ($2^2$) per vacuole. This indicates that the duplication rate of the parental strain is approximately 8 hours while the duplication rate of TgDicer-mut is 12 hours.

At 36 hours the majority of the parental strain had 16 parasites ($2^4$) per vacuole and the majority of the TgDicer-mut contained 8 parasites ($2^3$) per vacuole. This means that the duplication rate of TgDicer-mut is ~12 hours while the duplication rate of the parental strain is ~9 hours. So, similar to the 24 hour time point, at 36 hours, the duplication rate of TgDicer-mut is slower than the parental strain. At the 48 hour time point, around 25 % of the wildtype strain had more than 40 parasites per vacuole but there were not any TgDicer-mut which had more than 40 parasites per vacuole. This pattern indicated that the parental strain has a faster doubling time than the TgDicer-mut. The duplication rate of TgDicer-mut is 12 hours while the duplication rate of the parental strain is 9 hours.

Notably, when comparing *T. gondii* to its closes relative, *H. hammondi, H. hammondi* genome does not contain a Dicer gene. A major difference between the two parasites is the extent of the virulence in intermediate hosts. As described in the literature

review, *T. gondii* is more successful at having a wider range of hosts and affects a higher population among the common hosts sharing with *H. hammondi*. Therefore it can be speculated that the lack of a Dicer gene in *H. hammondi* may play a role in its promiscuity.


24 Hour
Percent of Total Vacuoles (%)
70
60
50
40
30
20
10
0
2
4
8
16
32
40+
Number of Parasites Per Vacuole
TgDicer-wt
TgDicer-mut


36 Hour
Percent of Total Vacuoles (%)
70
60
50
40
30
20
10
0
2
4
8
16
32
40+
Number of Parasites Per Vacuole

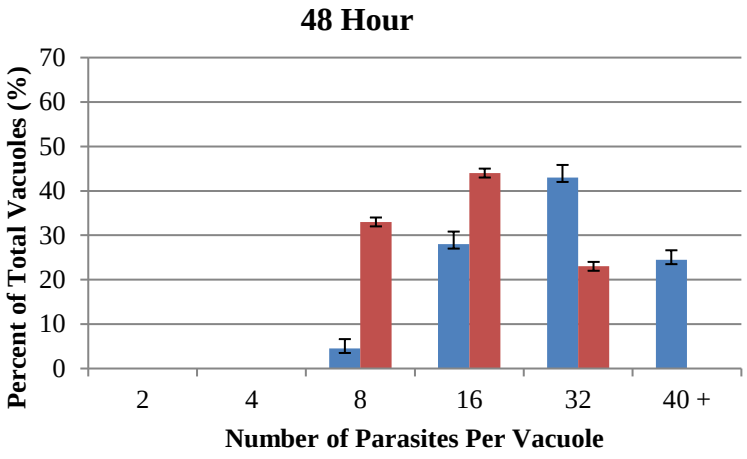
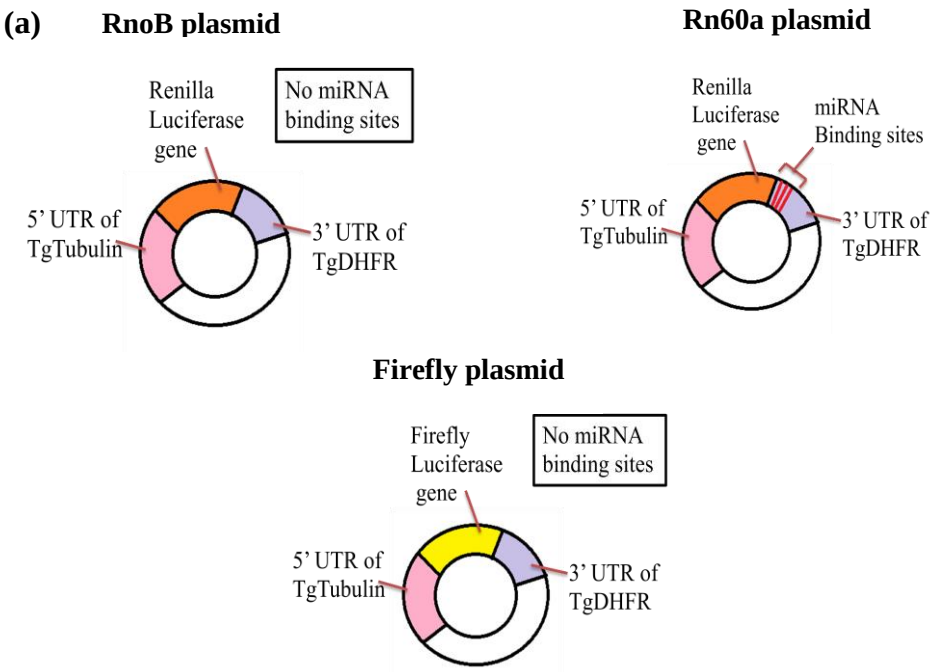

48 Hour
Percent of Total Vacuoles (%)
70
60
50
40
30
20
10
0
2
4
8
16
32
40 +
Number of Parasites Per Vacuole

**Figure 3.9: Loss-of-function analysis of Tg-Dicer on the parasite's doubling time**
TgDicer-wt and TgDicer-mut were allowed to infect and divide in host cells for 24, 36, and 48 hours after which the number of parasites in 100 vacuoles were counted. Each time point had two trials and the results show that TgDicer-mut duplication is slower than the TgDicer-wt.

## 3.4 The Effect of Transgenic Dicer on the Biogenesis of miRNA

The removal of a functional Tg-Dicer could affect the parasite's ability to produce mature miRNAs. The dual luciferase system was used to test whether miRNAs are being produced in the TgDicer-mut parasites. In this system, the expression of both *Renilla* luciferase and Firefly luciferase transcripts were under the control of tubulin promoter, allowing for the same level of transcription of these two genes. To ensure similar mRNA stability, *T. gondii* dihydrofolate reductase 3' UTR was used for both Firefly luciferase and *Renilla* luciferase transcripts. The transcript of the RN containing miRNA binding sites for miR60a in the 3'UTR (Rn60a) is used to compare with a RN construct which does not contain any binding sites (RnoB). Firefly luciferase was used as the internal control for transfection efficiency. The plasmids with the three RN and FF luciferase constructs are shown in Figure 3.10a.

When the dual luciferase system was introduced to TgDicer-wt and TgDicer-mut parasites, the activity of Renilla without any miRNA binding site was comparable in both strains which indicated that both strains can express the reporter genes. When RN with miR60a binding sites was introduced to TgDicer-wt, we see a 60 % decrease in RN expression when compared to RN with no binding sites. When this system was introduced into TgDicer-mut, RN with miRNA binding sites showed maximum Renilla activity. This experiment showed that in wildtype parasites, the RN expression is affected if the RN transcript contains miRNA binding sites. But in TgDicer-mut, the presence of miRNA binding sites in the RN transcript did not affect RN activity, indicating a lack of gene silencing ability for this strain.



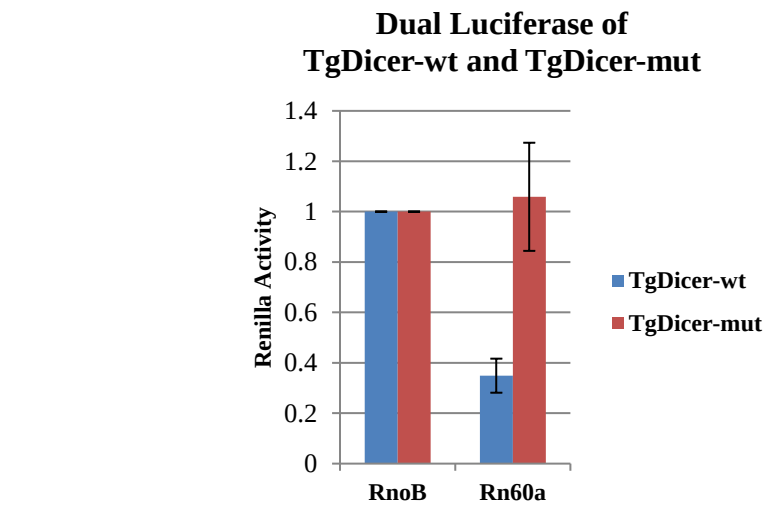




**Figure 3.10: Dual Luciferase of TgDicer-mut and TgDicer-wt**

(a) The plasmids which encode for the dual luciferase transcripts contain a 5' UTR of TgTubulin and a 3' UTR of TgDHFR. RnoB and Firefly plasmids produce transcripts which do not contain any miRNA binding sites. Rn60a transcript contains three miR60a binding sites in its 3'UTR. (b) A dual luciferase assay of TgDicer-wt and TgDicer-mut with RnoB transcript showed that both of the strains, were able to express RN. Rn60a transcript caused a significant decrease in RN expression in the TgDicer-wt when compared to RnoB. But in TgDicer-mut, Rn60a had the same level of RN activity as RnoB.

To further investigate the differences in the profile of miRNAs in the TgDicer-mut and TgDicer-wt, small RNAs were isolated from both TgDicer-wt and TgDicer-mut. The RNA samples were fractionated on a 15% Tris-borate urea polyacrylamide gel. The resolved RNAs were revealed using Sybr Gold staining. No difference was detected in the RNA from the two strains as shown in Figure 3.11. This may seem like contrasting data when it is compared to the dual-luciferase assay results in Figure 3.10 (b) but it must be noted that 75% of the total small RNA consists of mostly rRNA and tRNA in *T. gondii* and other organisms. (Braun et al., 2010; Aravin et al. 2005). This polyacrylamide gel analysis was not sensitive enough to detect miRNA levels. To fully determine if mature miRNAs are being produced in the TgDicer-mut, a northern blot should be performed which will be discussed in a later chapter for Future Directions.

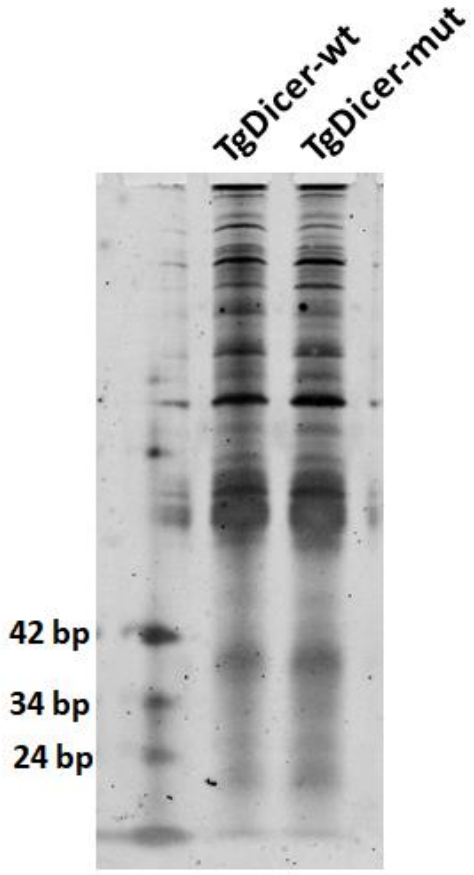



**Figure 3.11: Polyacrylamide gel of TgDicer-mut and TgDicer-wt**

A 15% Tris-borate urea polyacrylamide gel was used to resolve the RNAs from TgDicer-mut and TgDicer-wt and the RNA bands were revealed using Sybr Gold staining. No difference was detected in the RNA from the two strains.

## 3.5 The Ability of Truncated Dicer In dsRNA Processing

To investigate the role of Dicer in processing long dsRNA and RNA silencing activity, a dual luciferase assay was performed using TgDicer-wt and TgDicer-mut. A FF construct, which does not contain miRNA binding sites, was used as the internal control. The transcript of a RN construct contained three miRNA binding sites for the *H. sapien* let-7 miRNA in the 3' UTR, and referred to as RnLet7. The let-7 miRNA has not been found in *T. gondii* (Crater et al., 2012). This RnLet7 construct was used as the target transcript for testing the silencing ability of TgDicer-wt and TgDicer-mut. In addition to the plasmid constructs, four types of RNA effectors were introduced into the two parasite strains: siRNA, miRNA, long dsRNA, and RNase III treated dsRNA.

Figure 3.12 shows the dual luciferase assay performed with and without these RNA effectors using TgDicer-mut and TgDicer-wt. When no RNA was added to either TgDicer-wt and TgDicer-mut, RN expression can be observed indicating that both the strains are able to express RN luciferase. The introduction of siRNA caused the RN expression to decrease to about 60 % and 30 % of maximum RN expression in TgDicer-wt and TgDicer-mut respectively. MiRNA also caused a decrease in RN expression to about 30 % and 60 % in TgDicer-wt and TgDicer-mut respectively. In the introduction of long dsRNA, we see that TgDicer-wt strain had a lowered RN activity to about 40 %. But in TgDicer-mut, the long dsRNA did not have any effect and we see maximum RN activity. The RNase III treated dsRNA caused both of the strains to have about 30 % RN activity and it served as a control to show that if the long dsRNA is processed into smaller RNA it can cause a decrease in RN expression in TgDicer-wt and TgDicer-mut.

This experiment showed that Tg-Dicer is required for the processing of long dsRNA into smaller RNA in order for exhibiting the ability to suppress RN expression. We further postulate that Tg-Dicer is responsible for producing mature siRNA for gene silencing and that it may participate in RISC function. The finding that let7-siRNA and let7-miRNA exhibited differences in inducing the silencing effect on the target transcripts in TgDicer-mut and TgDicer-wt parasite strains, indicate that Tg-Dicer might play a different role in miRNA- and siRNA-RISC complex. The lack of function Tg-Dicer could impair miRNA-RISC activity. We could also extract from this experiment that a catalytically functional Dicer is not required for the RISC activity because gene silencing occurred when mature, small RNAs are provided for the transgenic parasites. The Tg-Dicer without RNase III domains lacks the ability to process long dsRNA into short dsRNA. The truncated Tg-Dicer with the helicase domain may be sufficient to allow active RISC activity.

A

| Type of RNA | Location of target | Binding region (nucleotides) |
|---|---|---|
| let7-siRNA | 3' UTR | 943-965, 972- 994, 1000 -1022 |
| let7-miRNA | 3' UTR | 943-965, 972- 994, 1000 -1022 |
| Long dsRNA | Coding sequence | 12 - 607 |

| | | |
|---|---|---|
| let7-siRNA | 5′ aaugagguaguucaauaggcugugccuu 3′<br>3′ uuccacuccaucaaguuauccgacacaa 5′ | Sense Strand<br>Antisense Strand |
| let7-miRNA | 5′ aaugagguaguucaauaggcugugccuu 3′<br>3′ uuccacuccauc auccgacacaa 5′<br>ggucc | Sense Strand<br>Antisense Strand |

B

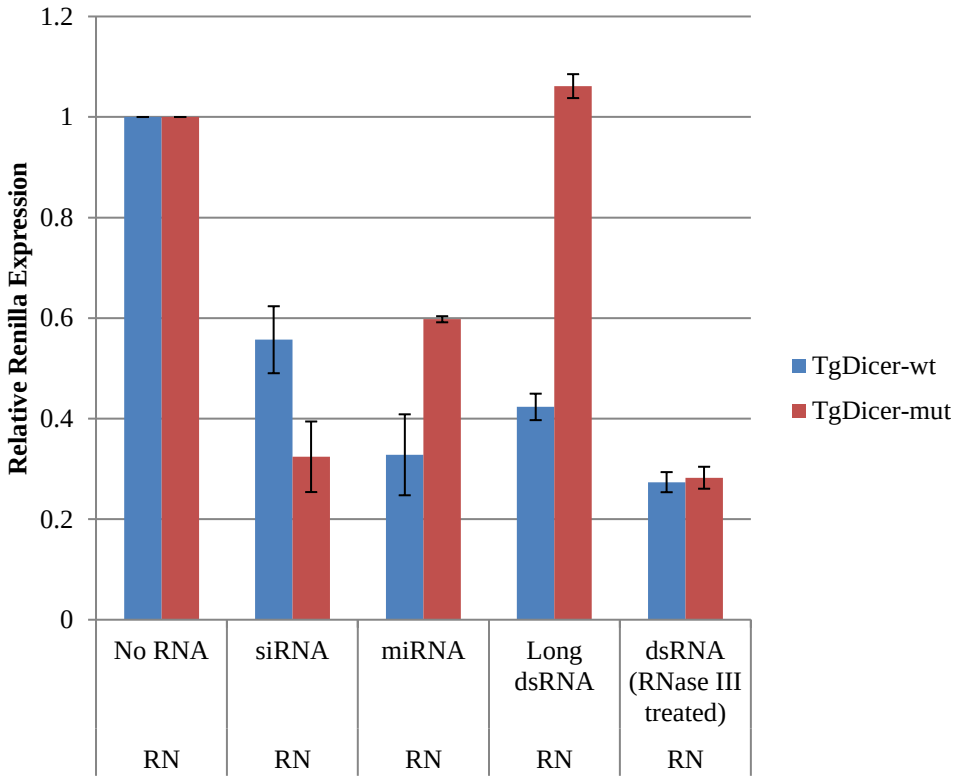


**Figure 3.12: Dual luciferase assay of exogenous long dsRNA and other small RNA on TgDicer-wt and TgDicer-mut**

(A) A schematic diagram of a *Renilla* luciferase (RN) transcript is shown (Crater et al., 2012). The RN transcript contains a 5' UTR of TgTubulin, RN coding sequence, and 3' UTR of TgDHFR. The RN coding sequence is between nucleotides 1 and 933. This RN construct contained three let7-miRNA binding sites at the 3' UTR of this transcript and is called RnLet7. This RN transcript was the target of four RNA effectors; let7-miRNA, let7-siRNA, long ds-RNA, and dsRNA treated with RNase III. The binding region for the miRNA and siRNA are in the 3' UTR of the transcript. The binding region for long dsRNA and dsRNA treated with RNase III, is the coding sequence for RN transcript. The sequences for let7-miRNA and let7-siRNA show that the siRNA is perfectly complementary to its target while the miRNA contains mismatches (B) TgDicer-wt and TgDicer-mut were electroporated with 1 µg FF plasmid encoding for FF luciferase and 1 µg RnLet7 plasmid encoding RN luciferase. 48 hours after electroporation, a dual luciferase assay was performed on TgDicer-wt and TgDicer-mut, with and without 4 µg of RNA effectors. The control did not contain any effector RNA, which resulted in maximum RN expression in both TgDicer-wt and TgDicer-mut. SiRNA caused the RN expression to decrease to about 60 % and 30 % when compared to maximum RN expression in TgDicer-wt and TgDicer-mut respectively. MiRNA caused about 30 % and 60 % RN expression in TgDicer-wt and TgDicer-mut parasites respectively. In the introduction of long dsRNA, TgDicer-wt strain had a lowered RN activity to about 40 %. But in TgDicer-mut, the long dsRNA caused maximum RN activity. The RNase III treated dsRNA which served as a positive control, caused both of the strains to have 30 % RN activity. Two trials were performed in this experiment.

## 3.6 Identification of TgHODI as a MicroRNA Target

The other goal of my study was to validate a predicted target of miRNA in *T. gondii*. Previous studies indicated over 80 genes as putative target of miR60a and almost 20 genes as putative target of miR4a (Braun et al., 2010) Among these putative genes is a DEAD-box RNA helicase (gene ID is TGGT1_313010), named TgHODI. This gene has been a study subject of our group. TgHODI is able to bind to the backbone of RNA spanning any five consecutive nucleotides and causes the unwinding of RNA (Russell et al., 2013). Previous studies showed that TgHODI is significant in the regulation of mRNA during stress response and more specifically it is a component of stress granules (Cherry and Ananvoranich, 2014). Through this study, we investigate whether the mRNA of TgHODI is regulated by miRNAs. MiR4a and miR60a binding sites in *TgHODI* transcript were analyzed using Vector NTI. The following tables (Table 3.1 and 3.2) show the possible binding sites for miR4a and miR60a respectively in the *TgHODI* transcripts.

**Table 3.1**: **Predicted miR4a Binding Sites on TgHODI**. The predicted base pairing is highlighted in yellow and the G:U wobble pairs are highlighted in green. The number of complementary base pairing is indicated.

| Pred. # | Seed base pairing | miR4a 3' UUACUGAUGUCGAAGGUUCGUUUGUA5' | Location on TgHODI mRNA | # of Comp. bp |
|---|---|---|---|---|
| **1** | 1-6<br>No mismatch | mRNA 5' CCAAUUGCUCUCGCCGGGAAAACAU 3'<br>miR4a 3' UUACUGAUGUCGAAGGUUCGUUUGUA 5' | Exon 2<br>400-425 | 8 |
| **2** | 2-8<br>GU Wobble | mRNA 5' GGAAGAAUUCUGGACCUGGCGAACAA 3'<br>miR4a 3' UUACUGAUGUCGAAGGUUCGUUUGUA 5' | Exon 3<br>685 - 710 | 12 |
| **3** | 2-8<br>GU Wobble | mRNA 5' AUGGUGGUCAUGGACGAAGCAGACAA 3'<br>miR4a 3' UUACUGAUGUCGAAGGUUCGUUUGUA 5' | Exon 3<br>739 - 764 | 14 |
| **4** | 2-7<br>No mismatch | mRNA 5' CGGUCACGGUCAAAGACUUCAAACAC 3'<br>miR4a 3' UUACUGAUGUCGAAGGUUCGUUUGUA 5' | Exon 3<br>857 - 882 | 7 |
| **5** | 1-8<br>1 mismatch | mRNA 5' AAGCUUCAAAUCAACCAAGCAAUCAU 3'<br>miR4a 3' UUACUGAUGUCGAAGGUUCGUUUGUA 5' | Exon 4<br>1006 - 1031 | 17 |
| **6** | 2-8<br>1 mismatch | mRNA 5' AUGCGUGACUUGCGAGAAGCAAAGAA 3'<br>miR4a 3' UUACUGAUGUCGAAGGUUCGUUUGUA 5' | 3' UTR<br>1430 - 1455 | 12 |
| **7** | 1-8<br>1 mismatch | mRNA 5' CCUCGCGACAGCUAACUGGCUAACAU 3'<br>miR4a 3' UUACUGAUGUCGAAGGUUCGUUUGUA 5' | 3' UTR<br>1531 - 1556 | 15 |
| **8** | 1-7<br>1 mismatch | mRNA 5' AUUCGUUCAAACCCAACUCCCAACAU 3'<br>miR4a 3' UUACUGAUGUCGAAGGUUCGUUUGUA 5' | 3' UTR<br>1684 - 1709 | 11 |
| **9** | 2-8<br>1 mismatch | mRNA 5' UGUUUCGUUUUCGGUCCGGCAAAGAG 3'<br>miR4a 3' UUACUGAUGUCGAAGGUUCGUUUGUA 5' | 3' UTR<br>1786 - 1811 | 12 |
| **10** | 2-8<br>1 mismatch | mRNA 5' CACAUGUAAGCGGGUCCGGCAAAGAG 3'<br>miR4a 3' UUACUGAUGUCGAAGGUUCGUUUGUA 5' | 3' UTR<br>1858 - 1883 | 12 |

**Table 3.2: Predicted miR60a binding sites on TgHODI**. The predicted base pairing is highlighted in yellow and the G:U wobble pairs are highlighted in green. The number of complementary base pairing is indicated.

| Pred. # | Seed base pairing | miR60a 3' UCAUACCUAAAGCAUGGCUGACACA 5' | Location on TgHODI mRNA | # of Comp. bp |
|---|---|---|---|---|
| **1** | 2-8<br>1 mismatch | mRNA 5' UCGACCUUGUUCUUGGCGUCUGUGA 3'<br>miR60a 3' UCAUACCUAAAGCAUGGCUGACACA 5' | 5' UTR<br>-293 to -269 | 13 |
| **2** | 1-6 | mRNA 5' CCCUCUUCUGGGUUGUCUUCUGUGU 3'<br>miR60a 3' UCAUACCUAAAGCAUGGCUGACACA 5' | 5' UTR<br>-133 to -109 | 11 |
| **3** | 2-8<br>1 mismatch | mRNA 5' CGAAGACGAAAGGCACGGACUUUGA 3'<br>miR60a 3' UCAUACCUAAAGCAUGGCUGACACA 5' | Exon 2<br>296 to316 | 12 |
| **4** | 1-7<br>1 mismatch | mRNA 5' GUACAACCCGGUGCAUGUACUCUGU 3'<br>miR60a 3' UCAUACCUAAAGCAUGGCUGACACA 5' | Exon 3<br>651 to675 | 11 |
| **5** | 1-7<br>1 mismatch | mRNA 5' CACGACCGAGAGGCAGAAACUGUUU 3'<br>miR60a 3' UCAUACCUAAAGCAUGGCUGACACA 5' | 3' UTR<br>1766 to 1790 | 9 |

The bioinformatic approach showed that the *TgHODI* transcript contains binding sites for both of the two most abundant miRNAs in *T. gondii*; miR60a and miR4a. Table 3.1 shows ten possible miR4a binding sites in *TgHODI* transcript; of which five possible binding sites are in the exons and five possible binding sites the 3'-UTR. Table 3.2 shows a total of five possible miR60a binding sites in *TgHODI* transcript of which two binding sites are in the 5'-UTR, two in the exons, and one in the 3'-UTR. To test if miR60a or miR4a regulates the expression of *TgHODI*, miRNA inhibitors of each miRNA were used. The inhibitors used here are single-stranded small RNAs (20-30 nt) containing complementary sequence to the miRNA guide strand.

A dual luciferase assay was conducted to determine the effects of miRNA inhibitors in the *T. gondii* system. As shown in Figure 3.13, the transcript for RnoB, which does not contain for any miRNA binding sites, showed maximum RN activity. The transcript for Rn4a which contained three miR4a binding sites had a decreased RN activity to about 55% of the maximum, indicating that the expression of RN luciferase was affected due to the presence of the miR4a binding sites in the 3' UTR of the RN transcript. The results showed that the expression was reduced in a half when compared to the transcript without miR4a binding sites. The expression of RN with miR4a binding sites was altered when 20 µg of anti-4a was introduced to the Rn4a system. There was an increase in RN activity to about 80 %. 40 µg of anti-4a resulted in maximum RN activity. This indicated that the RN expression can be recovered when anti-4a is present. The let-7 miRNA was used as a negative control to show sequence specific silencing of anti-4a and anti-60a. Since there are no let-7 binding sites on Rn4a, when 40 µg of anti-let-7 (let-7 miRNA inhibitors) was introduced into the Rn4a system, the RN activity was the same as

RN activity of Rn4a with no RNA effectors. The Rn60a is RN transcript with three miR60a binding sites in its 3' UTR and this construct had about 30 % of maximum RN activity. After 20 µg of anti-60a was introduced to the Rn60a system, there was an increase in *Renilla* activity to about 70 % and 40 µg of anti-60a resulted in maximum *Renilla* activity. The anti-let7 miRNA did not change the RN activity of Rn60a. This data indicates that anti-60a can inhibit the gene silencing of RN with miR60a binding sites.

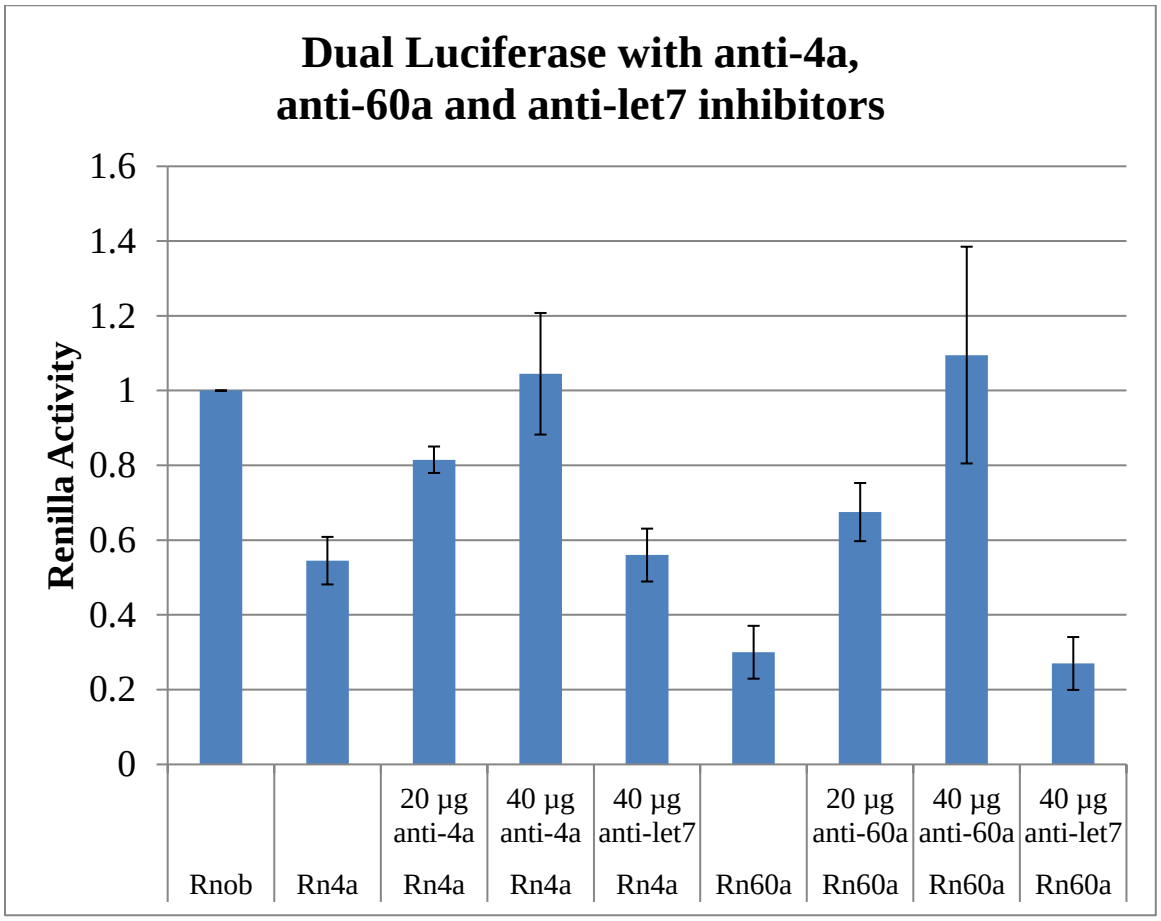



**Figure 3.13: Dual luciferase assay to determine whether miR4a inhibitor and miR60a inhibitor can prevent gene silencing.**

Two trials were conducted for a dual luciferase reporter assay performed in TgHODI-SF strain of *T. gondii*. Parasites which were electroporated with RnoB (which contains no miRNA binding sites) showed maximum *Renilla* activity. Rn4a and Rn60a are RN constructs which contain three miR4a and three miR60a binding sites respectively. Parasites with transcripts containing Rn4a and Rn60a showed a decrease in RN activity to about 55% and 30 % respectively. 20 µg of anti-4a and anti-60a caused an increase in RN activity of the Rn4a and Rn60a transcripts to about 80 % and 70 % respectively. Adding 40 µg of anti-4a and anti-60a caused maximum RN expression in Rn4a and Rn60a. 40 µg of anti-let7 was used as the negative control and did not change the *Renilla* activity of Rn4a and Rn60a.

The dual luciferase assay showed that miRNA inhibitors caused an increase in the expression of RN which contained miRNA binding sites. To determine if the miR4a and miR60a inhibitors could affect the TgHODI protein level a western analysis was performed. A strain of *T. gondii* containing a flag-epitope tagged HODI (TgHODI-SF) (Cherry and Ananvoranich, 2014) was used, to allow for the quantification of TgHODI expression by western analysis. This strain of parasites was individually electroporated with 40 µg of miR60a inhibitor (anti-60a), miR4a inhibitor (anti-4a) or let-7 inhibitor (anti-let7). Two trials were performed and in each experiment. The western blot analysis results are shown in Figure 3.14. Same amounts of total parasite lysate was analyzed where the TgHODI-SF protein was revealed by anti-Flag antibody at 62 kDa and TgLDH1(lactate dehydrogenase 1) with a band at 35 kDa was used as the internal control. Parasites without miRNA inhibitors had TgHODI protein expression level normalized to 1. The introduction of anti-60a and anti-4a caused an increase of the TgHODI protein level by 1.6 and 1.8 respectively. While anti-let7 did not cause the protein level of TgHODI to change. This indicates that TgHODI may be controlled by miR4a and miR60a.

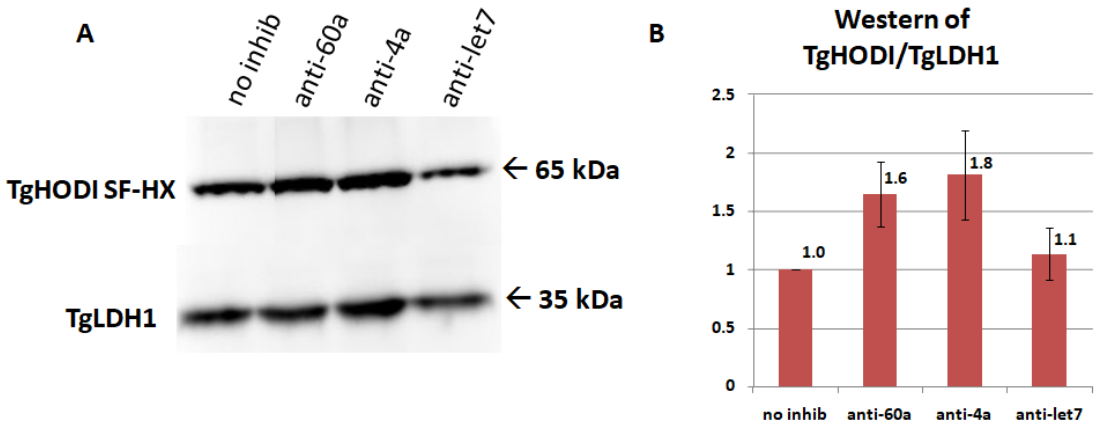



**Figure 3.14 Western analysis of TgHODI-SF with miR4a, miR60a and let-7 inhibitors.**

The expression of TgHODI-SF (62 kDa) was analyzed using anti-Flag antibody at 62 kDa and the internal control TgLDH1 at 35 kDa was analyzed using anti-LDH1. Quantification of the bands showed that 40 µg anti-60a and anti-4a caused an increase in the protein level of TgHODI-SF to about 1.6 and 1.8 while anti-let7 did not cause an increase in TgHODI-SF level.

# Chapter IV

## Future Directions and Conclusion

Dicer is a conserved enzyme in many eukaryotic cells including mammals, plants, and protozoa (Shin et al., 2009; Garcia-Ruiz et al., 2010; Macrae et al., 2006). The conserved function of Dicer homologs is to produce mature miRNAs and siRNAs, which makes this enzyme a fundamental component of post-transcriptional gene regulation. There are many types of post-transcriptional gene regulation in *T. gondii*. There is evidence that certain genes in *T. gondii* contain cis-acting elements which form hairpin structures allowing for mRNA stability (Holmes and Ananvoranich, 2014). Other work from this group shows that *T. gondii* contains a poly-A binding protein which is also known for its role in mRNA stability. It is established that long and short dsRNA can induce gene silencing in this parasite (Crater et al., 2012; Crater et al., 2015)

Prior to my study, no putative Dicer protein in *T. gondii* has been characterized. My study has given an insight on the significance of Tg-Dicer in the replication of the parasite and its role in gene silencing. My findings showed that the parasite strain, TgDicer-mut, whose Tg-Dicer expression was abolished, replicated slower than the parental strain. Furthermore, TgDicer-mut strain exhibited an impaired gene silencing activity. Here I was unable to detect mature miRNAs being made in the TgDicer-mut strain. A Northern blot analysis should be performed, probing for miR60a and miR4a; the two most abundant miRNAs in *T. gondii.* If miRNAs are not being made, there will be an accumulation of pre-miR60a and pre-miR4a which is a larger RNA species (> 70 bp) and

a lack of mature miR60a and miR4a (around 22 bp) in the TgDicer-mut when compared to the wildtype strain.

To study the activity of Dicer in more detail, the Tg-Dicer can be endogenously tagged to visualize Tg-Dicer protein localization through microscopy. Studying the localization of Dicer is important, as the canonical Dicer was originally thought to localize in the cytoplasm, but recent studies have shown that Dicer can localize in the nucleus where it promotes the formation of heterochromatin (White et al., 2014). Tagging the endogenous Dicer can also allow us to determine any associating complexes through immunoprecipitation. Knowledge of the associating proteins would be important because Dicers in other organisms associate with many other proteins to carry out their function (Margis et al., 2006). This is especially important for Tg-Dicer because Dicers in other organisms which lack the PAZ domain are known to require other proteins for proper RNA catalysis. For example R2D2 in *D. melanogaster* is needed for the activity of a Dicer homolog named Dicer-2 (Liu et al., 2003). The identification of interactive proteins can help determine the enzymatic mechanism of Tg-Dicer cleavage.

Moreover, biochemical analysis of Tg-Dicer can be performed using exogenous expression of the Tg-Dicer protein in *E. coli*. The protein can be used in the determination of the ability of Tg-Dicer to bind to miRNAs. A gel shift assay can be utilized with labeled miRNAs. I would expect Tg-Dicer to be able to bind to miRNA and the sample with Tg-Dicer and miRNAs should move slower down the gel when compared to a sample with just miRNA.

Exogenous expression of the Tg-Dicer gene in *E. coli* can also be used to study Tg-Dicer cleavage ability. *In vitro* synthesized pre-miRNA and long dsRNA can be introduced to the recombinant Tg-Dicer. Then a northern blot can be used to probe for the precursors and the Tg-Dicer product. One sample containing just the two precursor RNAs and another sample with the precursor RNA and Tg-Dicer should be run on a gel. If the Tg-Dicer is able to catalyze the precursor RNAs, this would result in smaller RNA strands which would be detected lower on the gel than RNA precursors.

The TgDicer-mut strain used in this study still has the Tg-Dicer helicase domain. The helicase domain in other homologous dicers may be involved in the transfer of the miRNA duplex from Dicer to Argonaute proteins (Soifer et al., 2008). Results from section 3.5 shows that TgDicer-mut is still able to carry out gene silencing when small RNA are present. This indicates that the small RNA can be loaded into the RISC complex in this TgDicer-mut strain. To determine if the helicase domain is responsible in RISC assembly, a full knock out can be used where the whole *Tg-Dicer* gene is deleted. Then the small RNA can me introduced into the new knockout strain. If this new strain can still show gene silencing then it would mean there are other factors which contribute to RISC loading. But if it is no longer able to show gene silencing then it would mean that the helicase region was important in RISC loading.

The second objective of this study was to investigate a potential miRNA target; TgHODI. TgHODI is an important DEAD-Box RNA helicase enzyme involved in the post-transcriptional gene regulation of *T. gondii*. TgHODI could be important in mRNA translational repression during stressful environments (Cherry and Ananvoranich, 2014). I speculate that TgHODI protein levels are affected by miR4a and miR60a. Future work

should determine the mRNA levels of TgHODI, with and without miR4a and miR60a inhibitors to determine if translational repression or mRNA decay is occurring. It may also be interesting to determine if miR4a and miR60a is up-regulated in *T. gondii* when it is under stress or during the bradyzoite stage. To determine the miRNA target location, base substitution mutations can be done at the predicted miRNA binding sites. Then miR4a inhibitors can be introduced into the new strain of parasites. If the mutation of a certain binding site causes the inhibitors to not have an effect on the TgHODI-SF levels, this would indicate that, this binding site is responsible for miR4a binding. The dual luciferase assay can also be used to determine the miR4a and miR60a target location, where the predicted binding sites can be incorporated into the 3' UTR of the reporter system. The reporter system can be tested to see if the reporter activity changes when there are predicted miR4a and miR60a binding sites.

To conclude, this thesis aimed to study at an upstream factor of miRNA production, Tg-Dicer, and a downstream target of miRNA; TgHODI. It was determined that a knockout of Tg-Dicer caused a phenotypic change in the parasite along with indication of loss of gene suppression ability. It was also determined that TgHODI may in fact be regulated by miRNAs. Further studies linking these phenomena to *T. gondii* will provide a better understanding of *T. gondii* post-transcriptional gene regulation.

# REFERENCES


Aikawa, M. (1988). Fine structure of malaria parasites in the various stages of development. *Malaria: Principles and Practice of Malariology. Churchill Livingstone, Edinburgh, United Kingdom*, 97-129.

Altay, K., Aydin, M. F., Dumanli, N., & Aktas, M. (2008). Molecular detection of Theileria and Babesia infections in cattle. *Veterinary Parasitology*, *158*(4), 295-301.

Altman, S., & Kirsebom, L. (1999). Ribonuclease P. *COLD SPRING HARBOR MONOGRAPH SERIES*, *37*, 351-380.

Ambros, V., Lee, R. C., Lavanway, A., Williams, P. T., & Jewell, D. (2003). MicroRNAs and other tiny endogenous RNAs in C. elegans. *Current Biology*, *13*(10), 807-818.

Arasu, P., Wightman, B., & Ruvkun, G. (1991). Temporal regulation of lin-14 by the antagonistic action of two other heterochronic genes, lin-4 and lin-28. *Genes & Development*, *5*(10), 1825-1833.

Araujo, F. G., & Remington, J. S. (1987). Toxoplasmosis in immunocompromised patients. *European Journal of Clinical Microbiology & Infectious Diseases*, *6*(1), 1-2.

Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D., & Horvath, P. (2007). CRISPR provides acquired resistance against viruses in prokaryotes. *Science*, *315*(5819), 1709-1712.

Black, M. W., Arrizabalaga, G., & Boothroyd, J. C. (2000). Ionophore-resistant mutants of *Toxoplasma gondii* reveal host cell permeabilization as an early event in egress. *Molecular and Cellular Biology*, *20*(24), 9399-9408.

Black, M. W., & Boothroyd, J. C. (1998). Development of a stable episomal shuttle vector for *Toxoplasma gondii*. *Journal of Biological Chemistry*, *273*(7), 3972-3979.

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

Blaszczyk, J., Tropea, J. E., Bubunenko, M., Routzahn, K. M., Waugh, D. S., Court, D. L., & Ji, X. (2001). Crystallographic and modeling studies of RNase III suggest a mechanism for double-stranded RNA cleavage. *Structure*, *9*(12), 1225-1236.

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 pathogens*, *6*(5), e1000920.

Brawerman, G. (1993). mRNA degradation in eukaryotic cells: an overview. In *Control of Messenger RNA stability*, 149-159.

Cannistraro, V. J., & Kennell, D. A. V. I. D. (1991). RNase I*, a form of RNase I, and mRNA degradation in *Escherichia coli*. *Journal of Bacteriology*, *173*(15), 4653-4659.

Carmell, M. A., & Hannon, G. J. (2004). RNase III enzymes and the initiation of gene silencing. *Nature Structural &Molecular Biology*, *11*(3).

Carpousis, A. J. (2002). The Escherichia coli RNA degradosome: structure, function and relationship to other ribonucleolytic multienyzme complexes. *Biochemical Society Transaction, 30(2), 150-155*

Catalanotto, C., Azzalin, G., Macino, G., &Cogoni, C. (2000). Transcription: Gene silencing in worms and fungi. *Nature*, *404*(6775), 245-245.

Chakravarthy, S., Sternberg, S. H., Kellenberger, C. A., & Doudna, J. A. (2010). Substrate-specific kinetics of Dicer-catalyzed RNA processing. *Journal of Molecular Biology*, *404*(3), 392-402.

Cheng, Z., Garvin, D., Paguio, A., Stecha, P., Wood, K., & Fan, F. (2010). Luciferase reporter assay system for deciphering GPCR pathways. *Current Chemical Genomics*, *4*, 84.

Cherry, A. A., & Ananvoranich, S. (2014). Characterization of a homolog of DEAD-box RNA helicases in *Toxoplasma gondii* as a marker of cytoplasmic mRNP stress granules. *Gene*, *543*(1), 34-44.

Chobotar, B., & Scholtyseck, E. (1982). Ultrastructure. *The biology of the Coccidia*, 101-165.

Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A., & Zhang, F. (2013). Multiplex genome engineering using CRISPR/Cas systems. *Science*, *339*(6121), 819-823.

Cooper, T. A. (2002). mRNA splicing: regulated and differential. *eLS*.

Court D. (1993). RNA processing and degradation by RNase III, p 71–116. In Elasco JG and Brawerman G (ed), Control of Messenger RNA Stability. Academic Press, New York, N.Y. 71-117

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), 1-13.

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.

Cudny, H., Zaniewski, R., & Deutscher, M. P. (1981). Escherichia coli RNase D. Catalytic properties and substrate specificity. *Journal of Biological Chemistry*, *256*(11), 5633-5637.

Denli, A. M., Tops, B. B., Plasterk, R. H., Ketting, R. F., & Hannon, G. J. (2004). Processing of primary microRNAs by the Microprocessor complex. *Nature*, *432*(7014), 231-235.

Dobrowolski, J. M., & Sibley, L. D. (1996). *Toxoplasma* invasion of mammalian cells is powered by the actin cytoskeleton of the parasite. *Cell*, *84*(6), 933-939.

Donald, R. G., & Roos, D. S. (1994). Homologous recombination and gene replacement at the dihydrofolate reductase-thymidylate synthase locus in *Toxoplasma gondii*. *Molecular and Biochemical Parasitology*, *63*(2), 243-253.

Doyle, M., Badertscher, L., Jaskiewicz, L., Güttinger, S., Jurado, S., Hugenschmidt, T., Kutay, U., & Filipowicz, W. (2013). The double-stranded RNA binding domain of human Dicer functions as a nuclear localization signal. *RNA*, *19*(9), 1238-1252.

Du, Z., Lee, J. K., Tjhen, R., Stroud, R. M., & James, T. L. (2008). Structural and biochemical insights into the dicing mechanism of mouse Dicer: a conserved lysine is critical for dsRNA cleavage. *Proceedings of the National Academy of Sciences*, *105*(7), 2391-2396.

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

Dubey, J. P., & Sreekumar, C. (2003). Redescription of *Hammondia hammondi* and its differentiation from *Toxoplasma gondii*. *International Journal for Parasitology*, *33*(13), 1437-1453.

Dubremetz, J. F. (2007). Rhoptries are major players in Toxoplasma gondii invasion and host cell interaction. *Cellular microbiology*, *9*(4), 841-848.

Endo, T., Sethi, K. K., & Piekarski, G. (1982). *Toxoplasma gondii*: Calcium lonophore A23187-mediated exit of trophozoites from infected murine macrophages. *Experimental Parasitology*, *53*(2), 179-188.

Erdmann, V. A., & Barciszewski, J. (Eds.). (2012). *From nucleic acids sequences to molecular medicine*. Springer Science & Business Media. 19-45

Falkow, S., Isberg, R. R., & Portnoy, D. A. (1992).The interaction of bacteria with mammalian cells. *Annual Review of Cell Biology*, *8*(1), 333-363.

Ferguson, D. J. P., & Hutchison, W. M. (1987). An ultrastructural study of the early development and tissue cyst formation of Toxoplasma gondii in the brains of mice. *Parasitology Research*, *73*(6), 483-491.

Fichera, M. E., & Roos, D. S. (1997). A plastid organelle as a drug target in apicomplexan parasites. *Nature*, *390*(6658), 407-409.

Fortin, K. R., Nicholson, R. H., & Nicholson, A. W. (2002). Mouse ribonuclease III. cDNA structure, expression analysis, and chromosomal location. *BMC genomics*, *3*(1), 26.

Frenkel, J. K., & Dubey, J. P. (1975). *Hammondia hammondi*: a new coccidium of cats producing cysts in muscle of other mammals. *Science*, *189*(4198), 222-224.

Frenkel, J. K., Dubey, J. P., & Miller, N. L. (1970). *Toxoplasma gondii* in cats: fecal stages identified as coccidian oocysts. *Science*, *167*(3919), 893-896.

Garcia-Ruiz, H., Takeda, A., Chapman, E. J., Sullivan, C. M., Fahlgren, N., Brempelis, K. J., & Carrington, J. C. (2010). *Arabidopsis* RNA-dependent RNA polymerases and dicer-like proteins

in antiviral defense and small interfering RNA biogenesis during Turnip Mosaic Virus infection. *The Plant Cell*, *22*(2), 481-496.

Hamilton, A. J., & Baulcombe, D. C. (1999). A species of small antisense RNA in posttranscriptional gene silencing in plants. *Science*, *286*(5441), 950-952.

Hammond, S. M. (2005). Dicing and slicing. *FEBS Letters*, *579*(26), 5822-5829.

Heydorn, A., &Mehlhorn, H. (2001).Further remarks on *Hammondia hammondi* and the taxonomic importance of obligate heteroxeny. *Parasitology Research*, *87*(7), 573-577.

Holmes, M., Itaas, V., & Ananvoranich, S. (2014). Sustained translational repression of lactate dehydrogenase 1 in Toxoplasma gondii bradyzoites is conferred by a small regulatory RNA hairpin. *The FEBS journal*, *281*(22), 5077-5091.

Huynh, M. H., & Carruthers, V. B. (2006). Toxoplasma MIC2 is a major determinant of invasion and virulence. *PLoS pathogens*, *2*(8), e84.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., & Charpentier, E. (2012). A programmable dual-RNA–guided DNA endonuclease in adaptive bacterial immunity. *Science*, *337*(6096), 816-821.

Köhler, S., Delwiche, C. F., Denny, P. W., Tilney, L. G., Webster, P., Wilson, R. J. M., Palmer, J.D., & Roos, D. S. (1997). A plastid of probable green algal origin in Apicomplexan parasites. *Science*, *275*(5305), 1485-1489.

Konermann, S., Brigham, M. D., Trevino, A. E., Hsu, P. D., Heidenreich, M., Cong, L., Platt, R.J., Scott, D.A., Church, G.M. & Zhang, F. (2013). Optical control of mammalian endogenous transcription and epigenetic states. *Nature*, *500*(7463), 472-476.

Krug, E. C., Marr, J. J., & Berens, R. L. (1989). Purine metabolism in *Toxoplasma gondii*. *Journal of Biological Chemistry*, *264*(18), 10601-10607.

Kwon, S. C., Nguyen, T. A., Choi, Y. G., Jo, M. H., Hohng, S., Kim, V. N., & Woo, J. S. (2016). Structure of human DROSHA. *Cell*, *164*(1-2), 81-90.

Lackner, D. H., Carré, A., Guzzardo, P. M., Banning, C., Mangena, R., Henley, T., ... & Bürckstümmer, T. (2015). A generic strategy for CRISPR-Cas9-mediated gene tagging. *Nature communications*, *6*, 10237.

Lee, Y., Ahn, C., Han, J., Choi, H., Kim, J., Yim, J., Lee, J., Provost, P., Rådmark, O., & Kim, V. N. (2003). The nuclear RNase III Drosha initiates microRNA processing. *Nature*, *425*(6956), 415-419.

Lee, Y., Kim, M., Han, J., Yeom, K. H., Lee, S., Baek, S. H., & Kim, V. N. (2004). MicroRNA genes are transcribed by RNA polymerase II. *The EMBO Journal*, *23*(20), 4051-4060.

Li, Z., & Deutscher, M. P. (1996). Maturation pathways for E. coli tRNA precursors: a random multienzyme process in vivo. *Cell*, *86*(3), 503-512.

Liu, Q., Rand, T. A., Kalidas, S., Du, F., Kim, H. E., Smith, D. P., & Wang, X. (2003). R2D2, a bridge between the initiation and effector steps of the Drosophila RNAi pathway. *Science*, *301*(5641), 1921-1925.

Lund, E., Güttinger, S., Calado, A., Dahlberg, J. E., & Kutay, U. (2004).Nuclear export of microRNA precursors. *Science*, *303*(5654), 95-98.

MacRae, I. J., Zhou, K., Li, F., Repic, A., Brooks, A. N., Cande, W. Z., Adams, P.D., & Doudna, J. A. (2006). Structural basis for double-stranded RNA processing by Dicer. *Science*, *311*(5758), 195-198.

Mali, P., Esvelt, K. M., & Church, G. M. (2013).Cas9 as a versatile tool for engineering biology. *Nature Methods*, *10*(10), 957-963.

Margis, R., Fusaro, A. F., Smith, N. A., Curtin, S. J., Watson, J. M., Finnegan, E. J., & Waterhouse, P. M. (2006). The evolution and diversification of Dicers in plants. *FEBS Letters*, *580*(10), 2442-2450.

Martens, P., & Hall, L. (2000). Malaria on the move: human population movement and malaria transmission. *Emerging infectious diseases*, *6*(2), 103.

Matthews, J. C., Hori, K., & Cormier, M. J. (1977). Purification and properties of Renilla reniformis luciferase. *Biochemistry*, *16*(1), 85-91.

Meister, G., Landthaler, M., Patkaniowska, A., Dorsett, Y., Teng, G., & Tuschl, T. (2004). Human Argonaute2 mediates RNA cleavage targeted by miRNAs and siRNAs. *Molecular Cell*, *15*(2), 185-197.

Meador, J., & Kennell, D. (1990). Cloning and sequencing the gene encoding *Escherichia coli* ribonuclease I: exact physical mapping using the genome library. *Gene*, *95*(1), 1-7.

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

Morrissette, N. S., & Sibley, L. D. (2002). Cytoskeleton of apicomplexan parasites. *Microbiology and Molecular Biology Reviews*, *66*(1), 21-38.

Nardelli, S. C., Che, F. Y., de Monerri, N. C. S., Xiao, H., Nieves, E., Madrid-Aliste, C., Angel, S.O., Sullivan, W.J., Angeletti, R.H., Kim, K., & Weiss, L. M. (2013). The histone code of *Toxoplasma gondii* comprises conserved and unique posttranslational modifications. *MBio*, *4*(6), e00922-13.

Nicolle, C., Manceaux, L. (1908). Sur une infection a corps de Leishman (on organismes voisons) du gondi. *CR Acad Sci*, *147*, 736.

Nomura, T., & Ishihama, A. (1988). A novel function of RNase P from Escherichia coli: processing of a suppressor tRNA precursor. *The EMBO Journal*, *7*(11), 3539.

Okamura, K., Ishizuka, A., Siomi, H., & Siomi, M. C. (2004). Distinct roles for Argonaute proteins in small RNA-directed RNA cleavage pathways. *Genes & development*, *18*(14), 1655-1666.

Paddison, P. J. (2008). RNA interference in mammalian cell systems. In *RNA Interference* (pp. 1-19).Springer Berlin Heidelberg.

Remington, J. S., Thulliez, P., & Montoya, J. G. (2004).Recent developments for diagnosis of toxoplasmosis. *Journal of Clinical Microbiology, 42*(3), 941-945.

Russell, D. G., & Sinden, R. E. (1981). The role of the cytoskeleton in the motility of coccidian sporozoites. *Journal of Cell Science*, *50*(1), 345-359.

Russell, R., Jarmoskaite, I., & Lambowitz, A. M. (2013). Toward a molecular understanding of RNA remodeling by DEAD-box proteins. *RNA Biology*, *10*(1), 44-55.

Sakamoto, H., Kimura, N., & Shimura, Y. (1983). Processing of transcription products of the gene encoding the RNA component of RNase P. *Proceedings of the National Academy of Sciences*, *80*(20), 6187-6191.

Seok, H., Ham, J., Jang, E. S., & Chi, S. W. (2016). MicroRNA target recognition: insights from transcriptome-wide non-canonical interactions. *Molecules and Cells*, *39*(5), 375.

Shiohama, A., Sasaki, T., Noda, S., Minoshima, S., & Shimizu, N. (2003). Molecular cloning and expression analysis of a novel gene DGCR8 located in the DiGeorge syndrome chromosomal region. *Biochemical and Biophysical Research Communications*, *304*(1), 184-190.

Shalem, O., Sanjana, N. E., Hartenian, E., Shi, X., Scott, D. A., Mikkelsen, T. S.,Heckl, D., Ebert, B.L., Root, D.E., Doench, J.G.,& Zhang, F. (2014).Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science*, *343*(6166), 84-87.

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), e01114-14.

Sherf, B. A., Navarro, S. L., Hannah, R. R., & Wood, K. V. (1996). Dual-Luciferase® reporter assay: An advanced co-reporter technology integrating firefly and Renilla luciferase assays. *Promega Notes*, *57*(2).

Shin, D., Shin, J. Y., McManus, M. T., Ptáček, L. J., & Fu, Y. H. (2009). Dicer ablation in oligodendrocytes provokes neuronal impairment in mice. *Annals of Neurology*, *66*(6), 843-857.

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

Soete, M., Camus, D., & Dubrametz, J. F. (1994).Experimental induction of bradyzoite-specific antigen expression and cyst formation by the RH strain of *Toxoplasma gondii* in vitro. *Experimental Parasitology*, *78*(4), 361-370.

Soifer, H. S., Sano, M., Sakurai, K., Chomchan, P., Sætrom, P., Sherman, M. A., Collingwood, M.A., Behlke, M.A. & Rossi, J. J. (2008). A role for the Dicer helicase domain in the processing of thermodynamically unstable hairpin RNAs. *Nucleic Acids Research*, *36*(20), 6511-6522.

Solberg, N., & Krauss, S. (2013). Luciferase assay to study the activity of a cloned promoter DNA fragment. *Gene Regulation: Methods and Protocols*, 65-78.

Soulsby, E. J. L. (1982). *Helminths, arthropods and protozoa of domesticated animals* (No.Ed. 7). Bailliere Tindall.
Splendore, A. (1908). Un nuovo protozoa parassita de' conigli incontrato nelle lesion anatomiche d'una malattia che ricorda in molti punti il Kala-azar dell'uomo. *Revista da Sociedade Scientifica de São Paulo.* 3: 109-12.

Suss-Toby, E., Zimmerberg, J., & Ward, G. E. (1996).*Toxoplasma* invasion: the parasitophorous vacuole is formed from host cell plasma membrane and pinches off via a fission pore. *Proceedings of the National Academy of Sciences*, *93*(16), 8413-8418.

Tahbaz, N., Kolb, F. A., Zhang, H., Jaronczyk, K., Filipowicz, W., &Hobman, T. C. (2004).Characterization of the interactions between mammalian PAZ PIWI domain proteins and Dicer. *EMBO Reports*, *5*(2), 189-194.

Tsutsumi, A., Kawamata, T., Izumi, N., Seitz, H., & Tomari, Y. (2011).Recognition of the pre-miRNA structure by Drosophila Dicer-1. *Nature Structural &Molecular Biology*, *18*(10), 1153-1158.

Walzer, K. A., Adomako-Ankomah, Y., Dam, R. A., Herrmann, D. C., Schares, G., Dubey, J. P., & Boyle, J. P. (2013). Hammondia hammondi, an avirulent relative of *Toxoplasma gondii*, has functional orthologs of known *T. gondii* virulence genes. *Proceedings of the National Academy of Sciences*, *110*(18), 7446-7451.

Wang, T., Wei, J. J., Sabatini, D. M., & Lander, E. S. (2014). Genetic screens in human cells using the CRISPR-Cas9 system. *Science*, *343*(6166), 80-84.

Wetzel, D. M., Håkansson, S., Hu, K., Roos, D., & Sibley, L. D. (2003). Actin filament polymerization regulates gliding motility by apicomplexan parasites. *Molecular biology of the cell*, *14*(2), 396-406.

White, E., Schlackow, M., Kamieniarz-Gdula, K., Proudfoot, N. J., & Gullerova, M. (2014). Human nuclear Dicer restricts the deleterious accumulation of endogenous double-stranded RNA. *Nature Structural and Molecular Biology*, *21*(6), 552.

Wood, K. V., De Wet, J. R., Dewji, N., & DeLuca, M. (1984). Synthesis of active firefly luciferase by in vitro translation of RNA obtained from adult lanterns. *Biochemical and Biophysical Research Communications*, *124*(2), 592-596.

Wu, H., Xu, H., Miraglia, L. J., & Crooke, S. T. (2000). Human RNase III is a 160-kDa protein involved in preribosomal RNA processing. *Journal of Biological Chemistry*, *275*(47), 36957-36965.

Zhang, H., Kolb, F. A., Jaskiewicz, L., Westhof, E., & Filipowicz, W. (2004). Single processing center models for human Dicer and bacterial RNase III. *Cell*, *118*(1), 57-68.

Zhou, G. Q., Zhang, Y., Ferguson, D. J., Chen, S., Rasmuson- Lestander, Å., Campbell, F. C., & Watt, S. M. (2006). The Drosophila ortholog of the endolysosomal membrane protein, endolyn, regulates cell proliferation. *Journal of Cellular Biochemistry*, *99*(5), 1380-1396.

# APPENDICES

**Table A1:** List of all oligonucleotides used in this study.

| Oligonucleotide Name | Oligonucleotide sequence |
|---|---|
| gRNADicerEcoNI | TTTATCCTTTGCGAGGTCGAGTTTTAGAGCTAGAAATAGCAAG |
| newP3 without HX | TTTCCATTTTTCGTCCTGGCG |
| Dicer Homology HX RV | ACCTGCGGCCTCACTGACTTATTAAGTTGGGTAACGCCAG |
| Primer A FW (Dicer_Validate_FW) | CGGTGGCGCCAGGAAAAG |
| Primer A RV (pEGFP-N1) | GCCGTCCAGCTCGACCAG |
| Primer B FW (Dicer Confirm RV) | CATATATGTATTACATGCAC |
| Primer B RV (HX Confirm FW) | CTTCGTCGGCTTCAGCATTG |
| Primer C FW (Dicer_Validate_FW) | CGGTGGCGCCAGGAAAAG |
| Primer C RV (Dicer_Validate_RV) | CGTTTACTCCGAACGCGC |
| GAPDH_Fw | GGTGTTCCGTGCTGCGAT |
| GAPDH_Rv | GCCTTTCCGCCGACAAT |
| T7promoterGG | TAATACGACTCACTATAGG |
| Sense_Tpl (Let7perfectmarchUpper) | AATGAGGTAGTTCAATAGGCTGTGCCTATATGAGTCGTATTA |
| Antisense_match_Tpl (Let7perfectmarchLower) | AACACAGCCTATTGAACTACCTCACCTATATGAGTCGTATTA |
| Antisense_mismatch_Tpl (Let7mismatchLower_5loop) | AACACAGCCTACCTGGCTACCTCACCTATATGAGTCGTATTA |
| FW_RnLuc_1 | TAATACGACTCACTATAGGTGTACGACCCCGAGCAAC |
| RV_RnLuc_1 | TAATACGACTCACTATAGGTAGGCAGCGAACTCCTCA |
| antiMIR60a_tpl | ACACAGTCGGTACGAAATCCATACTccTATATGAGTCGTATTA |
| antiMIR4a_tpl | ATGTTTGCTTGGAAGCTGTAGTCATTccTATATGAGTCGTATTA |

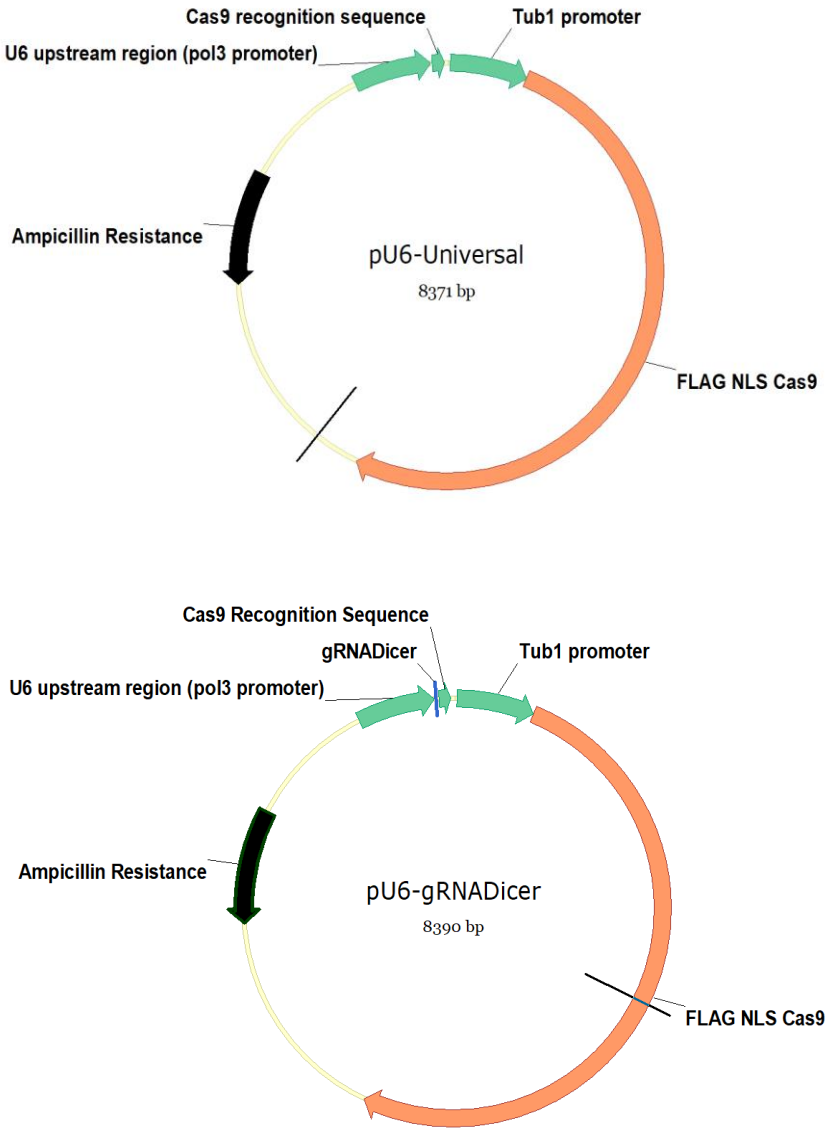



**Figure A1:** An Illustration of pU6-Universal plasmid and pU6-gRNADicer

This U6-Universal plasmid was obtained from S. Lourido from the Whitehead Institute for Biomedical Research via Addgene (Cat #52694). It contains the Tub1 promoter for the Cas9 endonuclease. It also contains a U6 upstream region for the binding of RNA polymerase III. The pU6-gRNADicer contains the same components except the gRNA for the Tg-Dicer locus was integrated after the U6 upstream promoter region.

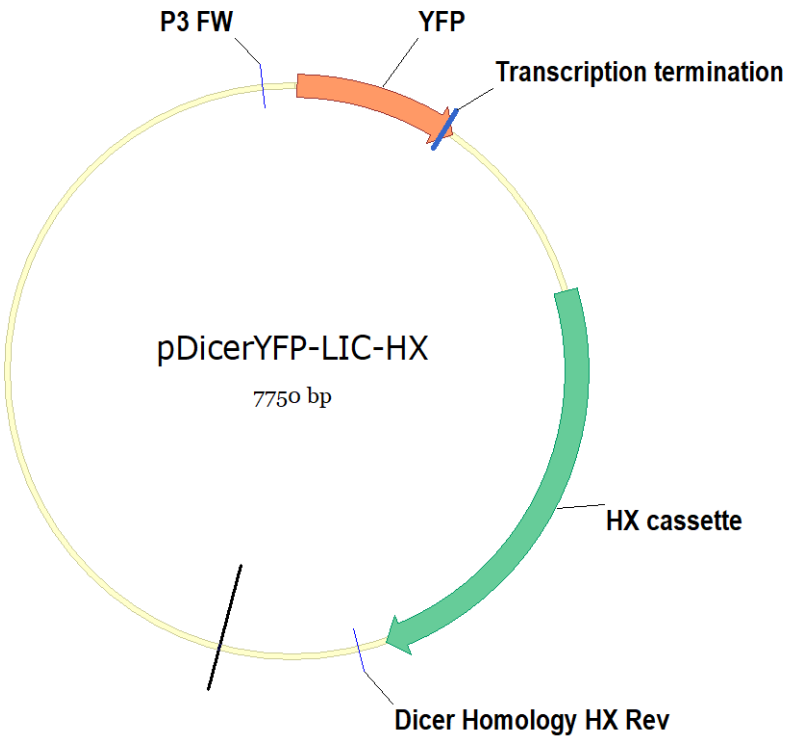



**Figure A2:** An Illustration of pDicer YFP-LIC-HX

The YFP HX insertion was amplified from the plasmid named pDicer_YFP_LIC_HX using two oligonucleotides; P3_FW and Dicer_Homology_HX_Rev. This insertion contains the HX (Hypoxanthine-xanthine -guanine phosphoribosyl transferase) cassette and the gene encoding for the Yellow Fluorescent Protein (YFP).

# VITA AUCTORIS

| | |
|---|---|
| Name: | Farzana Afrin |
| Place of Birth: | Dhaka, Bangladesh |
| Year of Birth: | 1992 |
| Education: | Vincent Massey Secondary School<br>Windsor, ON<br>2006-2010 |
| | University of Windsor, Windsor, ON<br>2010-2014<br>B. Sc. Honours Biochemistry |
| | University of Windsor, Windsor, ON<br>2014- Present<br>M. Sc. Biochemistry |