2017

Investigation of the argonaute protein variants in Toxoplasma gondii and the contribution of argonaute to RNA silencing

Md Amran Hossain
University of Windsor

Investigation of the argonaute protein variants in *Toxoplasma gondii* and the contribution of argonaute to RNA silencing

By

Md Amran Hossain

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

2017

Investigation of the argonaute protein variants in *Toxoplasma gondii* and the contribution of argonaute to RNA silencing

by

Md Amran Hossain

APPROVED BY:

_____

W. Crosby
Department of Biological Sciences

_____

M. Boffa
Department of Chemistry & Biochemistry

_____

S. Ananvoranich, Advisor
Department of Chemistry & Biochemistry

January 25, 2017

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

Argonaute protein is a vital component in the RNA interference pathway and post-transcriptional gene regulation in eukaryotes and the protein was recently identified in *Toxoplasma gondii*. However, the molecular mechanism by which the argonaute protein participates in regulating gene regulation pathways is unclear. This study was aimed to investigate the presence of argonaute protein variants in *T. gondii* (TgAgo), and its contribution in post-transcriptional gene silencing pathway. Furthermore, the study was designed to elucidate the influence of TgAgo on parasite growth and bradyzoite development. Using PCR analysis, I did not detect any transcript variants of TgAgo. A single transcript of TgAgo was identified to have 2,232 nucleotides encompassing 5 exons. Protein immunoblot assay showed a single protein of ~58.5 kDa. Gene silencing assays demonstrated that knockout of argonaute expression (AGOKO) reduced the gene silencing ability of an RNAi-like mechanisms in *T. gondii*, suggesting the importance of TgAgo in post transcriptional gene regulation pathways. *T. gondii* growth pattern was unaffected by the removal of Ago expression. However, bradyzoite formation was increased in AGOKO strains. The study suggests that TgAgo is a vital component of bradyzoite formation in the life cycle of *T. gondii*.

DEDICATION

I dedicate this thesis to my loving parents

ACKNOWLEDGMENTS

All praises belong to Almighty Allah, the most merciful, the most beneficent and the kindest for giving me the opportunity, courage and energy to carry out and complete the entire thesis work.

I am very grateful and deeply indebted to my supervisor Dr. Sirinart Ananvoranich, for her inspiration, extensive support, constant supervision and scholastic guidance to complete the research work. I would like to convey my appreciation to my committee members Dr. Michael Boffa and Dr W. Crosby for their constructive comments and inputs.

My deepest appreciation goes to my lab mates Anna Crater Potter, Farzana Afrin, Scott Roscoe and also my previous lab mates Michael Holmes and Emad Manni for their helpful hands that supported me during my research work. I would also like to appreciate Mikayla Roberts and Nadya Morrow for their support and obviously the beautiful moments we have shared together during my work. Special thanks to our graduate secretary Marlene Bezaire for her help.

I would like to express gratitude to all other teachers, lab mates and other staff of the department of Chemistry and Biochemistry, University of Windsor for their constant help and sincere cooperation during the entire study period.

I express my gratitude to my siblings, friends and well-wishers who inspired me for higher studies.

Finally my greatest gratitude to my beloved parents Md Ali Hossain Dewan and Taslima Begum; they are my source of inspiration and encouragement.

TABLE OF CONTENTS

Chapter 1: Literature Review

Chapter 2: Materials and Methods

Chapter 3: Results

LIST OF FIGURES

LIST OF APPENDICES

Appendix A - Table

Appendix B - Supplementary Figures

LIST OF ABBREVIATIONS

AGO	Argonaute
AGOKO	Argonaute knockout strain
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
eIF4E	Initiation factor 4E
FITC	Fluorescein isothiocyanate
HFF	Human foreskin fibroblasts
HIV/AIDS	Acquired immunodeficiency syndrome
HXGPRT	Hypoxanthine-Xanthine guanine phosphoribosyltransferase
IFA	Immunofluorescence assay
Luc	Luciferase
miRNA	MicroRNA
mM	Millimolar
MPA	Mycophenolic acid
mRNA	Messenger ribonucleic acid

NFBS	Newborn fetal bovine serum
Oligo	Oligonucleotide
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
PCR	Polynerase chain reaction
PV	Parasitophorous vacuoles
PMSF	Phenylmethylsulfonylfluoride
RISC	RNA induced silencing complex
RNA	Ribonucleic acid
RNAi	RNA interference
RNase	Ribonuclease
RPMI	Roswell Park Memorial Institute
siRNA	small interfering RNA
TAE	Tris-Acetate-EDTA
Taq	Thermus Aquaticus
TBE	Tris Borate EDTA
TEMED	N,N,N',N'-Tetram ethylethylenediam ine
UTR	Untranslated region
PTGR	Post transcriptional gene regulation

Chapter 1
Literature Review

1.1 Discovery of the *Toxoplasma gondii*

Toxoplasma gondii is a ubiquitous, obligate intracellular protozoan parasite belonging to the phylum Apicomplexa. Other pathogenic members in this phylum include *Plasmodium*, *Eimeria*, *Neospora*, *Babesia*, *Theileria* and *Cryptosporidium* (Kim and Weiss, 2004; Echeverria *et al*., 2010). *T. gondii* has the ability to infect a variety of warm-blooded vertebrates including humans, and causes frequent infections in wild, domesticated and companion animals (Hunter and Sibley, 2012). One third of the human population worldwide is anticipated to have *T. gondii* infection which can lead to the development of toxoplasmosis (Hoffmann *et al*., 2012).

T. gondii was first discovered by Nicolle and Manceaux in 1908 in tissues of a species of a North African hamster-like rodent (*Ctenodactylus gundi*; Tenter *et al*., 2000). In the late 1960s, an infectious stage of *T. gondii* was found in cat faeces, which was recognized as infectious to other hosts (intermediate hosts). Following that, in 1970, sexual stages were first defined in the small intestine of cats and the two phase (Sexual and asexual) life cycle of *T. gondii* was defined (Frenkel *et al*., 1970; Tenter *et al*., 2000; Dubey, 2009).

1.2 Life cycle

T. gondii has the ability to infect and replicate within virtually any nucleated mammalian or avian cell. The life cycle of *T. gondii* is facultatively heteroxenous, whereas it can complete its life cycle without intermediate host. The life cycle comprises two phases; the sexual phase within the feline (definitive host) and the asexual phase that takes place within non-feline hosts (intermediate host) (Tenter *et al*., 2000; Black and Boothroyd, 2000). The life cycle is summarized schematically in Figure 1.1. *T. gondii* exhibits three infectious forms namely tachyzoites, bradyzoites (in tissue cysts), and sporozoites (in oocysts) throughout its life cycle in different hosts (Dubey, 2006).

1.3 Mode of transmission

T. gondii can be transmitted through vertically as well as horizontally. Tachyzoite forms can transmit vertically from mother to foetus through placenta during pregnancy. However, horizontal transmission occurs via ingesting tissue cysts from undercooked meat, consuming oocysts from contaminated food or drink, or accidental ingestion of oocysts from the environment. Furthermore, tachyzoites may transmit during blood transfusion from an infected donor to recipient and also during tissue transplantation. Breastfeeding is another source of *T. gondii* transmission from the infected mother to the offspring (Tenter *et al*., 2000; Dubey, 2009).

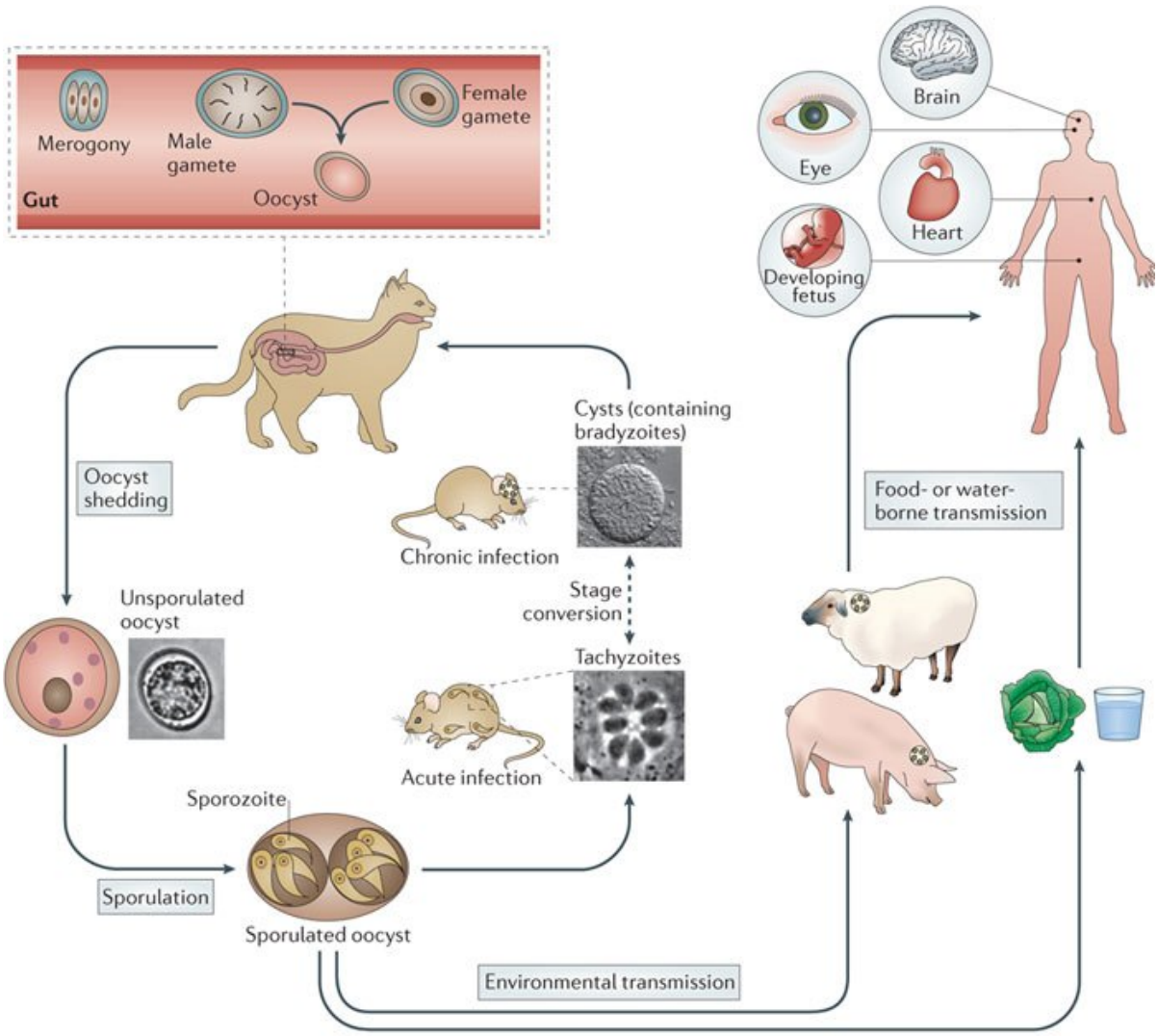


Figure 1.1: Diagram representing the life cycle of *T. gondii.* A sexual cycle takes place within intestinal cells of felines where male and female gametes are formed. Male gamete fertilizes the female gamete and giving rise to the oocysts, which are shed through faeces and contaminate the external environment (food, water) and provides a route of infection for intermediate hosts. An asexual cycle occurs within the intermediate hosts (mammals, birds, and almost in any warm-blooded organisms) where they can produce acute infection by rapidly growing tachyzoites or can differentiate into slow growing bradyzoites within the tissue cysts, leading to chronic infection. Ingestion of tissue cysts is also the primary source of *T. gondii* infection for both definitive and intermediate hosts (Hunter and Sibley, 2012).

1.4 Sexual life cycle

The sexual cycle occurs in the definitive host (felines). Felines can be infected with any of the three infectious stages of *T. gondii*, i.e. tachyzoites, bradyzoites (in tissue cysts), and sporozoites (in oocysts; Dubey, 2006), although infection with bradyzoites is more pathogenic than infection with either oocysts or tachyzoites (Dubey, 1998; Dubey, 2006). After entering into the host, bradyzoites reproduce in the enterocytes by endodyogeny, followed by repeated endopolygeny. Afterwards, it develops to gametes (gamogony) and initiates the sexual cycle. The male gamete fertilizes the female gamete, and two walls are formed around the fertilized zygote to form the oocyst (Dubey, 2016). Subsequently, unsporulated oocysts are released into the environment through faeces. Sporogony formation occurs outside the host and leads to the development of highly infectious and extremely stable oocysts containing two sporocysts where each sporocysts contains four sporozoites (Tenter *et al*., 2000; Boothroyd and Grigg, 2002).

1.5 Asexual life cycle

The intermediate host (humans, animals) can be infected via any of the three infectious stages of *T. gondii*. In the intermediate host, *T. gondii* undergoes the asexual replication. Upon ingestion of oocyst, the sporozoites form parasitophorous vacuoles (PV) in the small intestinal simple columnar epithelial cells (enterocytes) without infection and a few sporozoites enter in the lamina propria cells. The sporozoites first enter in the enterocytes but develop in the lamina propria cells. In the lamina propria the sporozoites convert to tachyzoites and then infect the enterocytes, as well as the capillary endothelial cells, macrophages, plasma cells, lymphocytes, neutrophils, eosinophils,

smooth muscle cells and fibrobroblasts (Dubey *et al*., 1998). On the other hand, if bradyzoites are ingested, a proportion of bradyzoites are destroyed in the lumen of the gut and others convert to tachyzoites in the small intestinal lamina propria and migrate to extraintestinal organs. In the intermediate hosts, *T. gondii* undergoes two phases of asexual development. In the first phase, tachyzoites (or endozoites) multiply swiftly by repeated endodyogeny in the host cells and eventually rupture the host cells and migrate to attach, invade and colonize a new host cell (Dubey *et al*., 1998) and this repeated cycle is known as the lytic cycle of *T. gondii* (Black and Boothroyd, 2000; Figure 1.2). Tachyzoites spread throughout all tissues in the body but differentiate into the dormant bradyzoite stage upon exposure to a potent host immune response (Skariah *et al*., 2010) and begin the second phase of development which results in the formation of tissue cysts. Within the tissue cyst, bradyzoites (or cystozoites) multiply slowly by endodyogeny (Tenter *et al*., 2000). Cyst formation is predominantly found in the central nervous system (CNS), eye and also in the skeletal and cardiac muscles. Cysts may also be found in visceral organs, such as lungs, liver, and kidneys. Tissue cysts are the terminal life-cycle stage in the intermediate host where they may persist for life or may break down the cysts and bradyzoites can transform to tachyzoites which may re-infect the host cells and again transform to bradyzoites within new tissue cysts (Tenter *et al*., 2000; Cenci-Goga *et al*., 2011).

1.6 Tachyzoite to bradyzoite interconversion

Tachyzoites and bradyzoites are two discrete stages in the asexual cycle of *T. gondii* (Tenter *et al*., 2000; Dubey *et al*., 1998). Tachyzoites are the rapidly growing form

of the parasite and exhibits acute infection in an immunocompromised host. Tachyzoites are often crescent shaped of approximately 5 µm long and 2 µm wide (Black and Boothroyd, 2000; Dubey *et al*., 1998) and enter into host cells through the initial recognition of the cell surface glycoproteins of the host cells by the help of tachyzoite specific surface antigen SAG1 (Black and Boothroyd, 2000).

Bradyzoites are the dormant form of *T. gondii*, grows slowly and generate tissue cysts within the host and causes the chronic infection. Bradyzoites remain intracellular and divide by the endodyogeny process. Tissue cysts can vary in sizes based on the number of bradyzoites contains in the cyst (Dubey *et al*., 1998). The tissue cyst develops by modifying the parasitophorous vacuole into an elastic, glycoprotein-rich protective cyst wall which is the key structural feature for bradyzoites conversion (Tomita *et al*., 2013). This cyst wall is easily detectable in culture by fluorescently labeled *Dolichos biflorus* lectin which binds the glycoprotein CST1 (Zhang *et al*., 2001; Tomita *et al*., 2013). Moreover, bradyzoites are influenced to up-regulate the stage-specific transcription of genes such as the heat shock protein bradyzoite antigen 1 (BAG1) (Bohne *et al*., 1997), the metabolic enzymes lactate dehydrogenase isoenzyme 2 (LDH2) and enolase isoenzyme 1 (ENO1) to adapt to adverse conditions. On the other hand, down regulate the tachyzoite-specific lactate dehydrogenase isoenzyme 1 (LDH1), enolase isoenzyme 2 (ENO2), and surface antigen 1 (SAG1) proteins (Yang and Parmley, 1997; Dzierszinski *et al*., 2001; Ferguson, 2004).

In vitro studies for understanding the mechanism behind the interconversion between tachyzoite and bradyzoite are convenient than *in vivo* studies (Skariah *et al*., 2010). Getting large amounts of encysted bradyzoites *in vivo* for study is challenging, but

in vitro studies allow for large scale simultaneous conversion to bradyzoites, providing an understanding of the process in detail and study of the mechanisms involved. Furthermore, the *in vitro* cysts have been shown to be similar to *in vivo* cysts.

There are a variety of methods that can be utilized to initiate *in vitro* bradyzoite conversion from tachyzoite. Modifying the culture conditions or chemical treatments are the effective methods assaying for differentiation. The methods includes, changing the pH of the culturing media (pH 8.0-8.2), raising the incubation temperature from 37°C to 43°C or sodium arsenite treatment (Soete *et al*., 1994). Moreover, the treatment with interferon (IFN)-γ (Bohne *et al*., 1993), mitochondrial inhibitors (antimycin A, myxothiazol, oligomycin; Bohne *et al*., 1994) and increasing the levels of cyclic nucleotides (Kirkman *et al*., 2001) can also facilitates into bradyzoites differentiation.

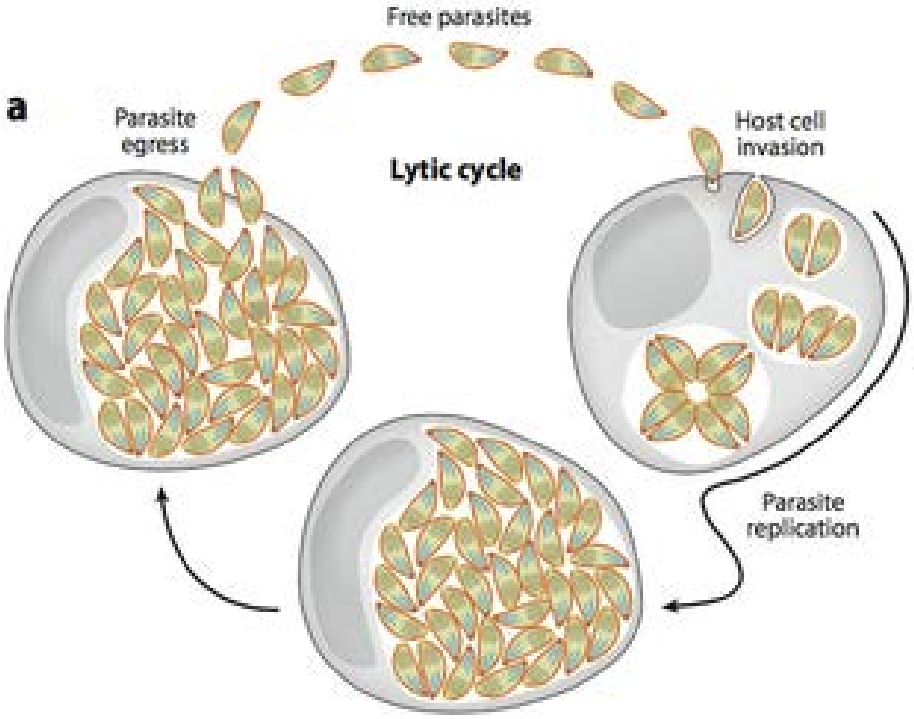


Figure 1.2: Lytic cycle of *T. gondii*. Extracellular parasites first attach and invade the host cell. After entering the host cell, the parasites become surrounded by parasitophorous vacuole (PV). In the PV the parasites undergo several rounds of replication until lysed from the host cells, motile parasites are released which migrate towards and invade a new host cell (Blader *et al*., 2015).

1.7 *T. gondii* as a model apicomplexan organism

There are several pathogenic members in the phylum Apicomplexa which cause different diseases. *Plasmodium* species infect human red blood cells causing malaria (Cowman and Crabb, 2002). *Eimeria* and *Cryptosporidium* cause enteric disease, while *Neospora* and *Theileria* are systemic veterinary pathogens. Experimenting with these parasites is highly challenging due to the difficulty of culturing and maintaining them in the laboratory (Kim and Weiss, 2004). However, *T. gondii* is experimentally tractable to

maintain in the lab, as well as readily amenable to genetic manipulation. This allows for high efficiency transient and stable transfection, as well as provides the opportunity to do research with diverse cell markers (Roos *et al*., 1995; Kim and Weiss, 2004). Moreover, *T. gondii* is relatively convenient for advanced microscopic analyses, which also makes *T. gondii* a good model apicomplexan organism (Kim and Weiss, 2004).

1.8 Genotypes and Clonal Lineages of *T. gondii*

T. gondii has three clonal lineages, known as types I (RH and GT-1), type II (ME49 and its derivatives PDS, PLK, PTG) and type III (CTG and VEG) respectively. Genome sequence analyses indicated the sequences similarities are high (99% or more between strains) for the 3 major lineages (Howe and Sibley, 1995; Su *et al*., 2003; Saeij *et al*., 2005). The strains are found in a variety of hosts from various geographic regions positioned primarily in North America and Europe (Howe and Sibley, 1995). Among them, type I strains are most pathogenic and can cause 100% mortality in mouse model with a small number of parasites and significantly higher levels of parasitema. However, type I strains have less potential to form cysts in culture and in animals. Additionally, type I strains are significantly more common in human congenital and ocular toxoplasmosis. On the contrary, type II strains are less virulent, but are the most prevalent clinical isolates, responsible for almost two-thirds of human cases of toxoplasmosis and considerably more common in AIDS patients with toxoplasmosis. Besides this, type II strains produce a higher number of cysts and are associated with reactivation of chronic infection. Type III strains are also less virulent and are more common in animal than

human hosts (Derouin and Garin, 1991; Howe and Sibley, 1995; Boothroyd and Grigg, 2002; Kim & Weiss, 2004).

1.9 Regulation of gene expression

Gene expression is the collective process that includes transcriptional, post-transcriptional and post-translational regulatory events to synthesize the functional gene product from the genetic information. The gene products can be regulatory molecules, enzymes or structural protein molecules to serve the major structural elements of living systems (Petsko and Ringe, 2004). The regulation of gene expression is imperative to ensure the proper level of gene expression and maintain accurate biological functions, cellular differentiation, morphogenesis and the versatility and adaptability necessary for the cellular survival (Halbeisen *et al*., 2008).

Transcriptional gene regulation is achieved through the direct interaction among the controlling factors (transcription factors), where these controlling factors are proteins in nature (Starr and Starr, 2014). Regulatory DNA binding sites or elements are known as enhancers, insulators and silencers. Furthermore, the transcription factors can be modified by intracellular signals through protein post-translational modification processes including phosphorylation, acetylation, or glycosylation. These modulated control factors influence the ability of transcription factors to interact with the transcription machinery to facilitate transcriptional regulation. In addition to this, epigenetics has a great influence on gene expression, which relies on non-sequence based changes in DNA structure such as DNA methylation and histone modification that alter the accessibility of DNA to transcription factors and modulate transcription levels (Starr

and Starr, 2014; Phillips, 2008). In the DNA methylation process, methyl groups are added to the fifth carbon atom of a cytosine ring (5-methylcytosine) of eukaryotic DNA by DNA methyltransferases. These altered cytosine residues usually positioned immediately next to a guanine base (CpG methylation) and two methylated cytosines placed diagonally to each other on opposite strands of DNA and can lead to inappropriate regulation of gene expression (Phillips, 2008). Histone modifications include lysine acetylation, lysine and arginine methylation, serine and threonine phosphorylation, and lysine ubiquitination and sumoylation, which can modify chromatin structure into active euchromatin (DNA is accessible for transcription) or inactive heterochromatin (DNA is more compact and less accessible for transcription) and regulate transcription (Vaquero *et al*., 2003; Starr and Starr, 2014). Transcriptional gene regulation occurs in different tissues, developmental stages, and under different environmental conditions, however the process is relatively slow and insufficient to rapidly and precisely modulate gene expression compared to post-transcriptional gene regulation (PTGR; Hall *et al*., 2005; Moore, 2005).

Post-transcriptional regulation is the regulation of gene expression after synthesis of the mRNA, but before translation. mRNA biogenesis is regulated by the formation of mRNPs (mRNA binding proteins; Moore, 2005). The components of the mRNP guide the transcript to specific biological processes by recognizing the *cis*-acting elements of the transcript (typically at the 5’ and 3’ UTR of the transcript), and its subsequent ability to recruit certain RNA-binding proteins (RBPs) through their RNA recognition motif (RRM; Moore *et al*., 2005). Therefore, all PTGR (post-transcriptional gene regulation)

mechanisms depend on the interaction between *cis*-acting elements encoded within the transcript and the *trans*-acting factors that bind and recognize them (Moore, 2005).

The pre-mRNA is synthesized in the nucleus and undergoes several post-transcriptional modifications to become a mature mRNA before being exported to the cytoplasm. Major modifications include (i) acquisition of a 5' cap structure; (ii) splicing out of introns; and (iii) addition of a polyadenylated (poly A) tail at the 3′ end, followed by nuclear export to cytoplasm.

The addition of the 5' cap structure with pre-mRNAs is a vital co-transcriptional maturation process (Furuichi and Shatkin, 1989). The 5’ end m7G (7-methylguanosine) capping is the most universally recognizable *cis*-acting element (Topisirovic *et al*., 2011), which protects the mRNA from degradation by 5’ to 3’ exonucleases (Moore, 2005; Carmody and Wente, 2009; Furuichi and Shatkin, 1989). 5’ capping is carried out by a series of enzymatic reactions, which start when the primary transcript is 20-30 nucleotides long. The enzymes involved in this process include nucleotide triphosphatase, guanylyltransferase, and N7-methyltransferase (Topisirovic *et al*., 2011). The m7G cap then recruits the cap binding complex for transporting the transcript to the cytoplasm (Proudfoot *et al*., 2002).

Cleavage of 3’ end and subsequent polyadenylation at the 3’ end of the pre-mRNA is also an important process for maturation and mRNA stability (Darnell, 2013). The presence of the polyadenylation signal sequence near the 3' end of the pre mRNA allows the cleavage and polyadenylation specificity factors (CPSF) to bind to pre-mRNA (Gruber *et al*., 2014). Following that the pre mRNA is cleaved at a site present between a highly conserved AAUAAA sequence and a uracil (U) or GU-rich motif, located further

downstream on the pre-mRNA (Proudfoot *et al*., 2002). The transcript then recruits nuclear poly (A) polymerase (PAP) within the cleavage reaction complex which catalyzes the addition of the poly (A) tail (Gruber *et al*., 2014; Proudfoot *et al*., 2002). This template-independent polymerase generates a tail of between 200 to 250 adenosines which protects the 3' end from ribonuclease digestion. The Poly (A) tail then binds with nuclear poly (A)-binding protein to export mRNA from the nucleus to the cytoplasm (Gruber *et al*., 2014) where it is capable of either undergoing active translation or becoming translationally repressed. Many genes have more than one polyadenylation site. Aalternative polyadenylation controls the inclusion or exclusion of *cis*-acting elements, including potential microRNA-binding sites, within its 3’ untranslated region (UTR) which can restrain the translatability and stability of the transcript (Gruber *et al*., 2014; Sandberg *et al*., 2008; Liu *et al*., 2007). On the other hand, alternative polyadenylation may produce multiple transcripts (mRNA) from a single gene, which can code for a different protein though this is the less common process (Shen *et al*., 2008).

Splicing out of introns from pre mRNA is the final step to produce mature mRNA. Predominantly splicing occurs after complete synthesis of transcript with polyadenylation. However, in some cases splicing can occur co-transcriptionally (Darnell, 2013). The splicing process involves the participation of a donor site (5' end of the intron) with GU consensus sequence, a branch site (near the 3' end of the intron) with a highly conserved adenosine (A) nucleotide involved in lariat formation, and an acceptor site (3' end of the intron) with AG consensus sequence (Proudfoot *et al*., 2002). Most introns follow the GU-AG rule while there are some exceptions with the ends GC-AG or AU-AC (Lewin *et al*., 2004; Proudfoot *et al*., 2002). The splicing reaction is catalyzed by

spliceosome, a large RNA-protein complex. Spliceosome are composed of protein (U2 small nuclear RNA auxiliary factor 1 (U2AF35), U2AF2 (U2AF65) and SF1) and five small nuclear ribonucleoproteins (snRNPs: U1, U2, U4, U5, and U6) that recognize splice sites in the pre-mRNA sequence (Lewin *et al*., 2004; Proudfoot *et al*., 2002; Black, 2003). The spliceosome reactions involves remove the intron as a form of lariat structure and eventually connects the adjacent exons with the linkage of 5' and 3' exons (Lewin *et al*., 2004; Darnell, 2013).

Alternative splice sites may generate a series of different transcripts and can produce multiple protein isoforms from a single genomic locus (Matlin *et al*., 2005; Hassan *et al*., 2012). It is expected that 95% of the transcripts originating from genes with multiple exons undergo alternative splicing, and this occurs in a tissue-specific manner and/or under specific cellular conditions (Pan *et al*., 2008). However, alternative splicing may down regulate gene expression post-transcriptionally by improper splicing of the transcript which can target mRNAs for nonsense-mediated decay (Matlin *et al*., 2005; Lewis *et al*., 2003). Studies have revealed that *T. gondii* utilizes functional splicing machinery and contain the canonical 5' GU-AG 3' splice junction. Moreover, recent studies have discovered that *T. gondii* also uses alternative splicing to regulate gene expression post-transcriptionally (Suvorova and White, 2014).

The mature transcript is exported to the cytoplasm through the nuclear pores, and undergoes protein synthesis (Köhler and Hurt, 2007). In most circumstances, mRNA assembles into ribonucleoprotein (RNP) complexes which recruit their exporters by class-specific adaptor proteins to transport the transcript (Köhler and Hurt, 2007). Mature mRNA consists of three parts: 5' untranslated region (5' UTR), protein-coding region or

open reading frame (ORF), and 3' untranslated region (3' UTR). The coding region encodes the information for protein synthesis. The mRNA codes for a single protein sequence, it is known as monocistronic and is common in eukaryotes. Where a given mRNA codes for multiple proteins, it is known as polycistronic and is common in prokaryotes (Lewin *et al*., 2004). The regulation of gene expression is also controlled through the processes of (i) translation, (ii) mRNA degradation (decay) and (iii) mRNA storage (Erickson and Lykke-Andersen, 2011).

During translation, the transcript is recruited into the translational machinery. Recognition of m7G 5'-cap structure by the eukaryotic initiation factors (eIFs) which form the eIF4F complex is the most common approach of translational initiation to produce the protein in eukaryotes. The eIF4F complex is comprised of eIF4E (cap binding protein), eIF4A (DEAD-box helicase) and eIF4G (scaffolding protein). eIF4E interacts with the 5’ cap of the transcript and also with eIF4G (López-Lastra *et al*., 2005; Sonenberg and Hinnebusch, 2009). eIF4G interacts with PABP (Poly A binding protein) and potentially circularizes the mRNA in the closed loop model (Malys and McCarthy, 2011; Wells *et al*., 1998). Afterwards, the closed loop structure recruitments the pre-initiation complex which contains the small ribosomal 40S subunit and the initiation factors including Met-tRNAi and eIF2-GTP (Sonenberg and Hinnebusch, 2009). Upon recruitment of the pre-initiation complex, the scanning process moves along the mRNA chain toward its 3'-end to reach the start codon (typically AUG). Following the recognition of the start codon eIF2-GTP is hydrolyzed to eIF2-GDP and recruits the 60S ribosomal subunit. The translational process continues until the stop codon, where the termination signal dissociates the ribosome (Sonenberg and Hinnebusch, 2009).

Cap-independent translation initiation is an alternate translational process for specific mRNAs in eukaryotes which occurs during cellular stress when overall translation is reduced. In cap-independent translation, recognition of the 5’ cap is not necessary to initiate the scanning from the 5' end of the mRNA to the start codon for translational start. Instead an internal ribosome entry site (IRES), aids in the assembly of the translational machinery at a position close to the initiation codon (López-Lastra *et al*., 2005).

Gene expression can be regulated either through the global translational repression of most of the transcripts or by transcript-specific translational repression by selectively removing the transcripts from the translational pool. Cytoplasmic mRNA turnover is an important characteristic feature for cellular physiology and is a key factor in gene expression. The *cis*-acting elements reside within the 5’ or 3’ UTR of the mRNA have potential roles in cytoplasmic mRNA turnover by the recruitment and placement of *trans*-acting factors (Moore, 2005). RNA interference (RNAi) is a potential RNA-mediated gene silencing pathway that occurs post transcriptionally to regulate gene expression either through mRNA degradation or mRNA suppression from translation (Saurabh *et al*., 2014).

1.10 Discovery and general regulation of RNA interference (RNAi) pathways

RNA interference (RNAi) is an evolutionary conserved mechanism wherein small silencing RNAs play a vital role on their regulatory targets (mRNA) to reduce the expression of target genes (Ghildiyal and Zamore, 2009; Figure 1.3). Andrew Fire and Craig C. Mello were awarded the Nobel Prize in Physiology or Medicine in 2006 for

work on RNA interference in the nematode worm *Caenorhabditis elegans*, which was first published in 1998. In their study, it was discovered that double stranded RNA (dsRNA) has the ability to trigger RNAi (Fire *et al*., 1998). Among protozoan parasites, RNAi mechanisms were first discovered in *Trypanosoma brucei* (Ngo *et al*., 1998). From its discovery, RNAi and its regulatory roles was appreciated for its potential in suppression of gene expression and accepted as a precise, efficient, stable and improved technology for gene suppression analysis (Saurabh *et al*., 2014). MicroRNAs (miRNAs) and short interfering RNAs (siRNAs) are components of an RNA-based mechanism of gene regulation found in eukaryotes (Carthew, 2006; Ghildiyal and Zamore, 2009).

MiRNAs are endogenous small noncoding RNAs approximately 22 nucleotide long that derive from distinctive hairpin precursors in animals and plants and can play an important role in gene expression (Bartel, 2004; Wienholds *et al*., 2005). *lin-4,* is the first identified miRNA in *C. elegans* in 1993 which regulates post embryonic development (Lee *et al*., 1993; Wightman *et al*., 1993), and *let-7*, is the second discovered miRNA was also from the same worm (Reinhart *et al*., 2000). miRNAs are important for their vital role in multiple biological processes. miRNAs regulate gene expression by base-pairing to target gene mRNAs (Zhang *et al*., 2007). It has been suggested that around 30% to 60% of mRNAs are regulated post-transcriptionally via miRNA (Friedman *et al*., 2009). Computational approaches and next generation sequencing provide a way to predict the miRNA targets which may provide the proper understanding of miRNA regulatory activities in different biological systems (Xu *et al*., 2014; Ekimler and Sahin., 2014).

During biogenesis, primary miRNAs (pri-miRNAs) are transcribed by RNA polymerase II (Pol II). The length of pri-miRNA varies from hundreds to thousands of

nucleotides (Cai *et al*., 2004; Lee *et al*., 2004). miRNAs can be transcribed from the introns of both protein coding and non-coding host genes (Lagos-Quintana *et al*., 2003). After synthesis, pri-miRNAs are bind with DGCR8/Pasha (a protein that contains two double-stranded RNA binding domains) through distinct stem-loop structures in the pri-miRNAs. Drosha (a RNase III endonuclease) then cleaves the flanks of pri-miRNA to liberate a 60 to 70 nucleotide stem-loop to generate the precursor miRNA (pre-miRNA). The pre-miRNA contains a 5′phosphate and a 2 nucleotide 3′overhang (Gregory *et al*., 2004; Denli *et al*., 2004; Han *et al*., 2004; Landthaler *et al*., 2004). miRNAs can also be produced from mirtrons (miRNAs found in introns) which are identical to pre-miRNAs. Mitrons are processed without Drosha (Ender and Meister, 2010), though they're processed by the splicing machinery (Czech and Hannon, 2011). Following their biogenesis, pre-miRNAs are binds with the Exportin-5 in the presence of Ran-GTP cofactor and exports from the nucleus into the cytoplasm. (Yi *et al*.2003).

In the cytoplasm, the 2-nucleotide 3' overhang of pre-miRNAs is recognized by Dicer (a second RNase III endonuclease) for further processing. Dicer cleaves the hairpin structure and releases a 20-22 nucleotide length of double stranded miRNA, called the miRNA (guide strand):miRNA*(passenger strand) duplex (Grishok *et al*., 2001; Hutvagner *et al*., 2001; Ketting *et al*., 2001; Zhang *et al*., 2004). Finally, the duplex is unwound by ATP-dependent helicase into single stranded mature miRNA (guide strand) and a miRNA* (passenger strand). Passenger strand is degraded while guide strand is incorporated into a ribonucleoprotein effector RNA-induced silencing complex (RISC) which induces post-transcriptional gene silencing (Schwarz *et al*., 2003; Hammond, 2005; Khvorova *et al*., 2003). Mass spectrometric analyses revealed RISC components

in *T. gondii* which are similar to components characterized in human and Drosophila (Braun *et al*., 2010). In *T. gondii*, miRNAs are reported as highly dynamic translational regulators which regulate mRNA with perfect or near perfect complementary bindings (Braun *et al*., 2010)

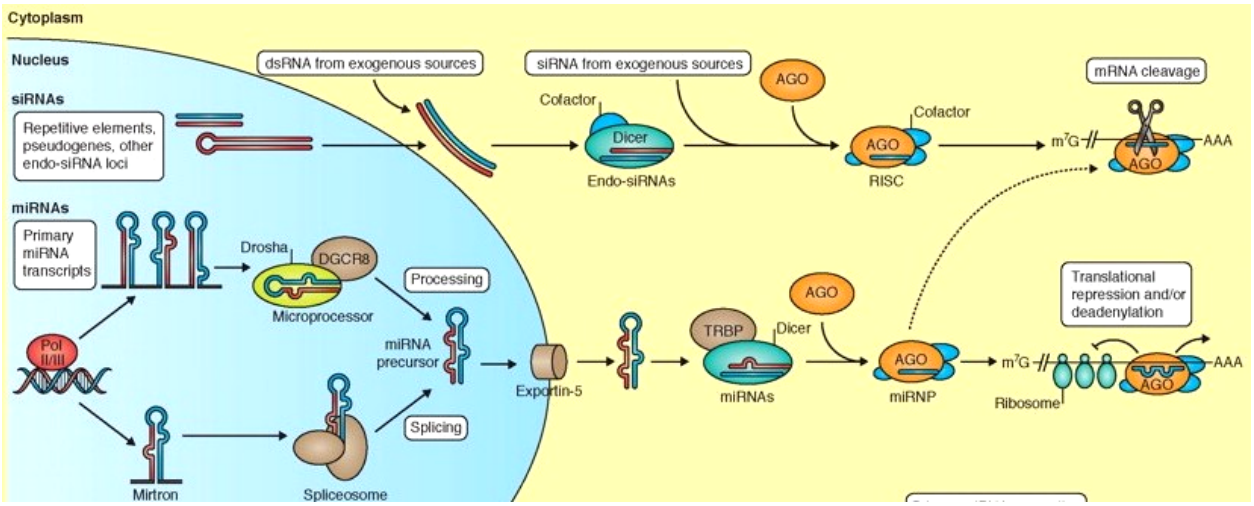


Figure 1.3: The RNA interference pathway of gene silencing. MicroRNAs (miRNAs) are generally transcribed by RNA polymerase II and can also be produced from mirtrons. siRNAs are synthesized from long dsRNAs either from exogenous sources or endogenous sources. These miRNAs and siRNAs are incorporated into RNA-induced silencing complex (RISC) and induce post-transcriptional gene silencing (Ender and Meister, 2010).

SiRNAs are synthesized from long dsRNAs either from exogenous sources (exosiRNAs) or endogenous sources (endo-siRNAs; Ender and Meister, 2010; Ghildiyal and Zamore, 2009). Exogenous sources can be viral RNAs or artificially injected perfect base-paired dsRNA into the cytoplasm. The endogenous siRNAs can be derived from transposable elements, natural antisense transcripts, long inter-molecularly paired hairpins and pseudogenes. Endo-siRNAs have been found in plants and in animals such

as *C. elegans*, flies and mice. siRNAs are processed in a similar way to miRNAs, however siRNAs are processed by Dicer instead of Drosha (Ender and Meister, 2010). Afterwards, like miRNA processing, the guide strand of the siRNA duplex is incorporated into RNA-induced silencing complex (RISC) and induces post-transcriptional gene silencing (Ghildiyal and Zamore, 2009; Iwasaki and Toomari, 2009; Ghildiyal and Zamore, 2009; Du and Zamore, 2005).

miRNAs recognize their targets through base pairing between the 5′ end of the miRNA (i.e. nucleotides 2–8, the seed region) and complementary sequences in the 3′untranslated regions (3′ UTRs) of the target mRNAs and destabilized mRNAs. This is the major component of post transcriptional repression (Baek *et al*.2008; Timothy, 2007). In animal cells, an imperfect complementary sequence alignment between miRNA and its targeted mRNA can facilitate translational repression. The number of miRNA binding sites in a targeted mRNA often associates with the level of repression (Cuellar and McManus, 2005). Moreover, miRNA activity may diversify in a cell type–specific manner by RNA editing of a miRNA primary transcript. The middle of the 5' proximal half "seed" region is the most important site for miRNA hybridization to targets (Kawahara *et al*., 2007). The mode of miRNA mediated gene regulation depends on the characteristics of the target recognition by miRNAs. The regulatory mechanism include, 1) degradation of mRNA by sequence specific cleavage 2) inhibition of translation without sequence-specific mRNA degradation which includes inhibition of translational initiation, inhibition of translation elongation or premature termination of translation (ribosome drop-off), 3) mRNA de-adenylation and degradation, and/or 4) mRNA sequestration (Eulalio *et al*., 2008; Timothy, 2007; Engels and Hutvagner, 2006).

1.11 Argonaute protein

Argonaute protein is the central component of the RNA-induced silencing complex (RISC; Rand *et al*., 2004; Hutvagner and Simard, 2008). Argonaute proteins bind microRNAs (miRNAs) and small interfering RNAs (siRNAs), which directs argonaute protein to their specific targets through sequence complementarity and regulate specific gene expression (Rand *et al*., 2004; Song *et al*., 2004; Meister, 2013; Garcia-Ruiz *et al*., 2015).

1.12 Structure of the argonaute protein

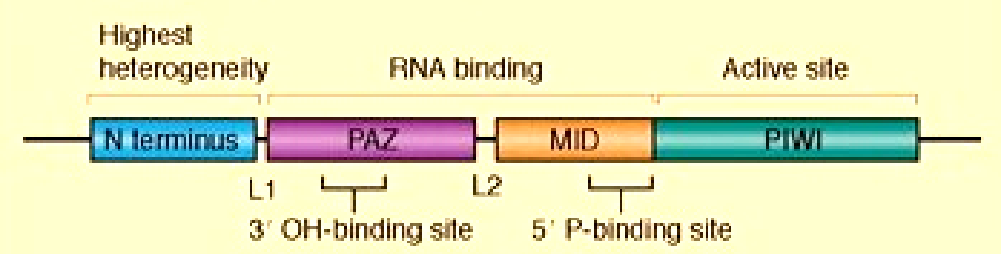


Figure 1.4: Illustration of argonaute. Major domains and their core functions are shown (Ender and Meister, 2010)

The argonaute (AGO) protein has four characteristic domains: N- terminal, PAZ, Mid and a C-terminal PIWI domain (Hutvagner and Simard, 2008; Figure 1.4). The PAZ domain recognizes the characteristic 3’ overhang of both miRNAs and siRNAs by a low affinity sequence-independent manner (Lingel *et al*., 2004; Ma *et al*., 2004). The Mid domain anchors the 5′-phosphate of the siRNA and miRNA (Ma *et al*., 2005). Argonaute

protein from human, mammals, chordates (such as zebra-fish), Drosophila (AGO1), *C. elegans* (ALG-1 and ALG-2) contains the MC domain in the Mid domain, which encompasses significant similarity to the m^7G cap binding motif of eIF4E, the translation initiation factor (Kiriakidou *et al*., 2007).

The PIWI domain exhibits the RNase-H-like fold (Asp-Asp-Glu/Asp) with the same cleavage activity (Song *et al*., 2004). Additionally, cleavage component of argonaute have a more relaxed catalytic centre (Asp-Asp-Asp/Glu/His/Lys), where two aspartates are very crucial to convey cleavage activity (Hutvagner and Simard, 2008). However, the PIWI domain of TgAgo lacked the canonical catalytic centre but R-R-E (Arg-Arg-Glu) residues may be responsible for weaker cleavage activity in TgAgo (Musiyenko *et al*., 2012). The most characteristic feature of the cleavage products from PIWI domain is that it contains a 3'OH group and 5' phosphate which are also similar to the products of RNase-H-like processing. Furthermore, structural studies revealed that the small RNA-mediated cleavage of the target RNA always occurs at a fixed place (Yuan *et al*., 2005). Cleavage occurs between the tenth and eleventh nucleotides of the guide strand (counting from the anchored 5′-end) which precedes the scissile phosphate on the target RNA (Haley and Zamore, 2004). Some organisms may lack the PAZ domain and have only the PIWI domain for argonaute protein, for instance, *Trypanosoma cruzi* and *Leishmania major*, where these argonaut proteins are unnecessary for RNAi pathway (Hutvagner and Simard, 2008).

1.13 Argonaute protein family

Argonaute proteins are categorized into three paralogous groups: 1) argonaute-like proteins similar to *Arabidopsis thaliana* AGO1; 2) PIWI-like proteins, closely related to *D. melanogaster* PIWI; and the newly recognized, *C. elegans*-specific group 3 argonautes (Yigit *et al*., 2006; Hutvagner and Simard, 2008). Argonaute-like and PIWI-like proteins are found in eukaryotes, but the number of genes for argonaute varies in different species (Hutvagner and Simard, 2008). The human genome encodes for 8 argonaute genes where 4 are argonaute-like and 4 are PIWI-like. Similarly, *D. melanogaster* encodes 5 argonaute genes (2 argonaute-like and 3 PIWI-like). On the other hand, *A. thaliana* has 10 and *Schizosaccharomyces pombe* has 1 argonaute encoding genes but all are argonaute like. *C. elegans* has about 26 argonaute genes where 18 of them are group 3 argonautes but 5 are argonaute like and 3 are PIWI-like argonautes. However, organisms with both argonaute-like and PIWI-like paralogs, regulate two types of function. For example, PIWI-like proteins are related to transcriptional silencing such as transposon silencing in the nucleus, whereas argonaute-like proteins are responsible for post transcriptional gene silencing mechanisms in the cytoplasm (Hutvagner and Simard, 2008).

1.14 Argonaute function

Argonaute proteins (either argonaute-like or PIWI-like) are essential for RISC in small non-coding RNA-mediated silencing pathways in fission yeast, fungi, plants, flies and mammals (Irvine *et al*., 2006; Hutvagner and Simard, 2008). Argonaute-like protein (Ago2) in Drosophila and mammals contribute to siRNA maturation by removing the

non-active passenger strand and initiating sequence specific cleavage of the target RNAs (Matranga *et al*., 2005). PIWI-like proteins, by contrast, may participate in the maturation of repeat-associated small interfering (rasi) RNAs and Piwi-interacting RNAs (piRNAs) in flies and mammals (Gunawardane *et al*., 2007; Brennecke *et al*., 2007). However, some argonaute proteins, for instance human argonaute‑3 (AGO3), are not able to facilitate miRNA related cleavage of target RNA, though they have the intact PIWI domain with canonical active site. This suggests that not only the active PIWI domain but also other factors yet to be discovered may be necessary for cleavage activity of argonaute proteins (Hutvagner and Simard, 2008). Besides this, argonaute proteins are associated with developmental process in diverse organisms (Carmell *et al*., 2002). The reduction of argonaute expression in Drosophila and *C. elegans* produces defective germlines (Cox *et al*., 2000; Tijsterman *et al*., 2002).

1.15 Argonaute protein in *T. gondii*

The argonaute protein of *T. gondii* can be categorized into Ago-like subfamilies because its similarity to Arabidopsis Ago1 (Al Riyahi *et al*., 2006). The *T. gondii* argonaute sequence shows greater similarity with Arabidopsis Ago1 than Drosophila PIWI. *T. gondii* argonaute protein called TgAgo is approximately 58.5 kDa, encoding 524 amino acids (GenBank accession no. DQ196314). This is smaller than argonaute proteins reported from other eukaryotic organisms but similar in size to argonaute from archaea. Amino acid sequence analysis revealed that TgAgo contains a conserved PIWI domain but lacks PAZ domain. Despite its small size, TgAgo is recruited into the RISC or effector complexes for gene silencing. However, other research groups have reported

different lengths of the argonaute protein including 743 aa (GenBank accession no. DQ177874), and 946 aa (GenBank accession no. GU046561).

1.16 *T. gondii* argonaute (TgAgo) sequence analysis

According to the ToxoDB.org (ToxoDB 31 9 Mar 17), gene TGGT1_310160 (GT1), TGME49_310160 (ME49), and TGVEG_310160 (VEG) encode argonaute protein in *T. gondii* in three archetype strains, type 1 (GT1), type 2 (ME49), and type 3 (VEG), respectively. The gene is located on chromosome XI, and the primary transcript (pre-mRNA) is predicted to have seven exons and six introns with 5346 nucleotides total. The mature transcript is predicted to contain 2841 nucleotides (exon 1: 372 + exon 2: 204 + exon 3: 267 + exon 4: 258 + exon 5: 150 + exon 6: 625 + exon 7: 965 nucleotides) and encodes 946 amino acid residues resulting in a predicted molecular weight of the argonaute protein of ~ 105 kDa.

However, the National Center for Biotechnology Information (NCBI) database (www.ncbi.nlm.nih.gov) has documented three different lengths of the *T. gondii* argonaute protein reported from three different research groups. The Hakimi research group (marked as number 1 in Figure 1.5) reported that argonaute consists of 946 amino acids (GenBank accession no. GU046561) and contains all 7 exons with a predicted polypeptide molecular weight of ~105 kDa as reported in ToxoDB.org (Braun *et al*., 2010). The Barik research group (marked as number 2 in Figure 1.5) reported the size of argonaute consists of 743 amino acids (GenBank accession no. DQ177874), where the primary transcript (exons + introns) contains 3940 nucleotides and the mature transcript is 2232 nucleotides long and comprises 5 exons (Figure 1.5), with a calculated molecular

weight of 83.5 kDa. This latter author did not mention its molecular weight in any literature. The Ananvoranich research group (marked as number 3 in Figure 1.5) reported the expression of a short version of argonaute consisting of 524 amino acids (GenBank accession no. DQ196314). Where the pre-mRNA sequence is 2564 nucleotides and protein coding sequence is 1575 nucleotides with only 2 exons (Figure 1.5). In addition, this report indicated the 5'UTR (368 bp) and 3' UTR (621 bp) and found the predicted molecular weight of argonaute to be 58.5 kDa (Al Riyahi *et al*., 2006). Notably, all research groups reported the same sequences at the 3' end of the transcript but their observations vary at the 5' end (Figure 1.5). Moreover, the predicted FASTA amino acid sequences also revealed that they vary from N- terminal end but C- terminal end have similar sequences for the functional domains of the active protein (Figure 1.6). It is possible that the argonaute protein encoding gene has more than one gene product at the protein level, which may be a result of an alternative splicing mechanism of mRNA from argonaute encoding gene.

Reported lengths of argonaute protein from *T. gondii*

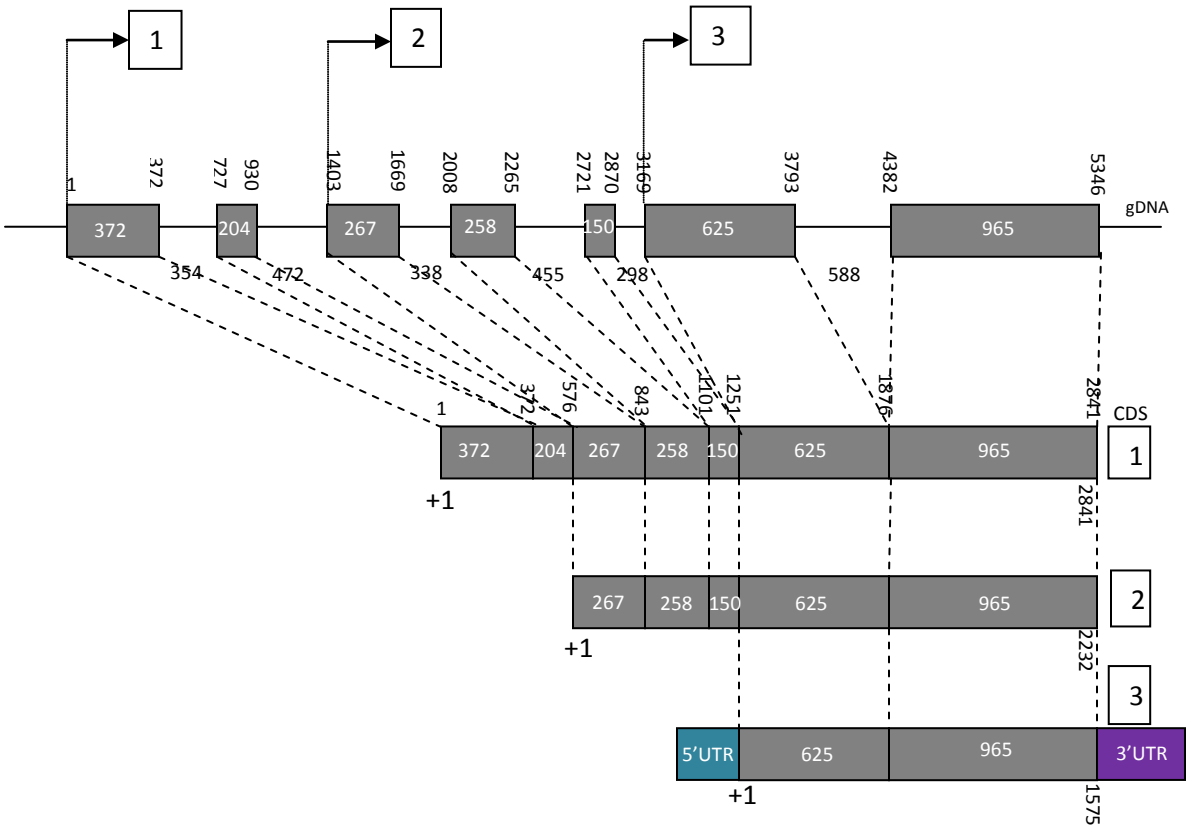


Figure 1.5: Different lengths of argonaute protein from *T. gondii* was documented in National Center for Biotechnology Information (NCBI). +1, indicate the first nucleotide of ORF for argonaute mentioned from different research groups. The largest reported length consists of 7 exons (CDS marked as 1) but another described size does not include exon 1 and 2 (CDS marked as 2). The smallest size reported has only exon 6 and 7, though the UTR's were indicated (CDS marked as 3). CDS: Coding sequences.

FASTA format of amino acid sequences for argonaute protein from *T. gondii*

MNGGGRGRDGSRGRGGDSRAGSNPYGRGFGSGDHRGGNRGGGGPFGGGGRGGMG
GGGFGSVLPVLRGPLKCLSNHCELVVSNRRQAVAVWKHFVDFKARAEPGKKRMLSYE
ARREILDEFLDGLVALGHLARDERALILYDGDHLLVCMKKLEYIEMLRGVSLSIHPQQS
VIMANLGEQLSMQNSATEEVAQILQLIMNHAAVMEGYQMFHGVNFFKDMPENFVPL
RGPNKAFQIWDGFSQAVAPYQTAGTMSWNAVFNLRACTSTKAIPLVKYIEIQASAIAKR
MVDLSTEKGCRALMADAEVMRRLNRRLRGTKLESFHILNRDTQQPEKRVYKLKELM
TFSASSEESRFDLQDGRRVTVLQHFKETYPHASGIPPFQPLINTRSKDRPAYLPVSIVTL
KHQPAREGVTEEDRAQVAEHMIMPPRNRVAKTEQLLSLVFGPQGHSAPKVLDAFGVIL
QLEAKTAQGRVLESPLIKYREVPGGGPGRSSKTVRPAQGDWNLRDAAFCRGTVCKVW
ALYSFVETSQNVIENLARVLKTQGAKYGVNLTTKPGLGAYTRNDRRPMLDQFAAFVAL
AKTKGCELLFVILNERVSLDIYQIVKSCTDVNFPSQCLNGRHKCIDAIFRGADNPNPQYF
ANVMSKVNMKLQGVNQTLEADIIKQEIGTDKSTLVLAVETSFFANPTKTSPPPTAPIVC
ACTGNMDDDLGAFGHAVCVESRKHPIVTDIGSMFKTILSYRKTTKNWPARIIYLRSAT
TEAHFPLVLAGEIRAIEELYVRENRSKPRILAVAVQRRQQTRLFPTKEMQAQGNNLPPG
FLLANSLQHPGHFRNFLLISHKALQGTARPTRYYILRDDANRDMEKVAQLMYSLCHVY
GRCQRAVSIPAPLYYAELLAARAQSYMKVGMRRERNIDIDDLSHLSGEAGEKMLTETR
HFADEYLRSTAAKVTPMVFC

Figure 1.6: FASTA format of amino acid sequences for argonaute protein from *T. gondii*. Orange square indicating start codon (M= Methionine) for different lab. Light green highlighted sequences indicates PAZ domain, reported in 2 labs (indicated 1 and 2 in Figure 1.5). Violet color sequences PIWI + MID domain for argonaute reported from all labs.

1.17 Research goals and hypothesis

The current study was aimed to elucidate if variants of the argonaute protein are expressed by *T. gondii*. Argonaute cDNA was analyzed to determine the argonaute transcript variants. Protein immunoblot analysis was performed to determine the variants of the argonaute protein. C-terminal FLAG tagged stable transgenic parasites (pAGO-SF-TAP-LIC-HX) were generated via a genetic approach called homologous recombination, where TgAGO was expressed with a FLAG under the control of its own promoter and thus expressed at the similar level as its parental strain.

The second objective of the study was to determine the contribution of argonaute protein to post transcriptional gene silencing in *T. gondii*, as well as to determine the value of argonaute proteins on the parasite growth and on interconversion ability. This aspect was performed by comparing the properties of wild type *T. gondii* stains with an argonaute knock-out (AGOKO) parasite strain.

Chapter 2
Materials and Methods

2.1 Cell and Parasite Culture

2.1.1 Mammalian cell culture

Human foreskin fibroblasts (HFF, ATCC-1041) were used to maintain intracellular *T. gondii*. The HFF host cells were grown and maintained in Dulbecco's Modified Eagle Medium (DMEM) with high D-glucose and L-glutamine (Invitrogen, #12100046) supplemented with 10% cosmic calf serum (ThermoFisher Scientific, Hylcone, # H3008704N), and 5μg/ml antibiotic-antimycotic (Invitrogen, # 15240-062) at 37°C in 5% CO_2.

2.1.2 *T. gondii* culture

Confluent monolayers of HFF host cells were infected with *T. gondii* and sustained using Minimum Essential Medium (MEM, Invitrogen, #61100061) supplemented with 1% (v/v) dialyzed fetal bovine serum (dFBS) ((ThermoFisher Scientific, Hyclone, #SH3007903) and 5μg/ml antibiotic-antimycotic (Invitrogen, # 15240-062) incubated at 37°C in 5% CO_2.

RHΔHX is a *T. gondii* type 1 strain commonly used in laboratories which lacks the gene HXGPRT (Donald *et al.*, 1996). RHΔHX was acquired from NIH AIDS Research and Reference Reagent Program (#2857). ΔKu80 strain (lacking the Ku80 gene) is a RHΔHX-derived strain and was obtained from Dr. Vern B. Carruthers

(University of Michigan, USA). A null-TgAgo mutant (AGOKO) strain, also a RHΔHX-derived strain, was obtained from Dr. Boothroyd, Stanford University School of Medicine. These strains were maintained in the culture media without selection. However, TgAgo-SF-TAP-LIC-HX transgenic parasites were generated by electroporation of ΔKu80ΔHX strain with plasmid (pAGO-SF-TAP-LIC-HX) containing HXGPRT and maintained in the culture media containing 25 μg/ml mycophenolic acid and 50 μg/ml xanthine.

2.2 Plasmid generation

2.2.1 Plasmid construction

pAGO-SF-TAP-LIC-HX plasmids were previously constructed using pSF-TAP-LIC-HX plasmid (obtained from Dr. Vern B. Carruthers, University of Michigan). The coding sequences of TgAgo (Gene ID TGGT1_310160) were PCR amplified from the ΔKu80 parasite's genome using the oligonucleotide primers. The construct was analyzed by restriction endonucleases.

2.2.2 Transformation

Plasmids were transformed into DH5α competent bacterial cells and plasmid was isolated and restriction digested with appropriate restriction enzymes to confirm the desired plasmid construct. Briefly, all plasmids were transformed into DH5α competent *E. coli* cells by the heat shock method. 25 μL of competent cells were incubated with 2 μL of plasmid DNA on ice for 30 mins, followed by heat shock at 42°C for 45 seconds, and incubating on ice for another 2 min. Afterward, 500 μL of pre-warmed (Room temperature) SOC (Super Optimal Broth with catabolite-repression) media (2% (w/v)

bacto-tryptone, 0.5% (w/v) bacto-yeast extract, 0.05% (w/v) NaCl, 10 mM $MgCl_2$, 2.5 mM KCl and 20 mM glucose) was added to obtain maximal transformation efficiency of *E. coli*. The reaction was incubated at 37°C for 1 hour with 225 rpm agitation. The cells were then plated on an LB (Luria-Bertani) agar plates containing ampicillin (1% (w/v) bacto-tryptone, 0.5% (w/v) bacto-yeast, 1% (w/v) NaCl, 1.5% (w/v) bacto-agar and 50 mg/L antibiotic) and grown at 37°C overnight.

2.2.3 Plasmid isolation by miniprep

A single colony was picked from LB agar plate using a sterile toothpick and transferred into 3 mL of LB broth supplemented with ampicillin antibiotic and grown at 37°C overnight with 225 rpm agitation. An aliquot of the culture (1.5 ml) was centrifuged at 12,000 rpm for 1 min and the pellet was re-suspended in 100 µL of re-suspension solution (50 mM glucose, 25mM Tris-HCl (pH 8.0), 10 mM EDTA). Afterwards, the bacterial cells were lysed using 200 µL of lysis buffer solution (0.2N NaOH, and 1% SDS). Next, 150 µL of precipitation buffer solution (3 M NaOAc, pH 5.2) was then added, the sample was mixed by inversion and centrifuged at 12,000 rpm for 15 min at 4º C. The clear supernatant was transferred into new tube and combined with 400 µL of TE-buffered phenol:chloroform, vortexed for 30 seconds, and centrifuged at 12,000 rpm for 1 min at room temperature. The top aqueous phase was transferred into a fresh tube and the plasmid DNA was precipitated with 800 µL of 95% ethanol and centrifuged at 12,000 rpm for 15 min at 4°C. The pellet was washed with 75% ethanol, air dried, re-suspended in 30 µL of TE buffer containing 20 µg/ml RNase A and incubated at 37°C for 15 min. The plasmid was stored at -20°C.

2.2.4 Restriction digestion

Plasmids were confirmed by restriction digestion (according to manufacturer's protocol) with the appropriate restriction enzymes for a minimum of 2 hours at 37°C. The restriction fragments were resolved on 1% (w/v) agarose gels (0.4 g agarose in 40 mL of 1x TAE buffer, and 0.02% (w/v) ethidium bromide). The samples were mixed with DNA gel loading Dye (6x: 0.25% (w/v) bromophenol blue, 0.25% (w/v) xylene cyanol FF, and 40% (w/v) sucrose in water) .The gel was electrophoresed for 25 mins at 140 V in the miniature horizontal gel system with 1x TAE buffer. Gels were visualized and imaged using an Alpha imager Imaging System.

2.3 Transfection of plasmid in *T. gondii* by electroporation

Freshly lysed ΔKu80 parasites (~ 2 x 10^6 parasites) were collected and centrifuged at 5000 rpm for 5 minutes and the pellet was resuspended with 400 μL electroporation mixture (cytomix) buffer (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), 2 mM ATP and 5 mM glutathione were added to the mixture prior to electroporation (Roos *et al*., 1995). Ago-SF-TAP-LIC-HX plasmids were digested with AflII restriction enzyme to linearize the plasmid and 50 μg of linearized plasmids were added to the mixture and transferred to 4 mm-gap cuvette. The instrument used for electroporation was BTX ECM 630 (Voltage: 1500, Resistance: 25 Ω, Capacitance: 25 μF). Immediately after electroporation, the parasites were cultured in confluent HFF monolayers at 37°C in 5% CO_2 using MEM with 1% dialyzed fetal bovine serum (dFBS) and antibiotic-antimycotic. 24 hours post infection, the media was replaced with MEM

supplemented with 1% dFBS, 0.5x antibiotic antimycotic, 25 μg/ml mycophenolic acid and 50 μg/ml xanthine to select the transgenic parasites.

Parasites which integrate the plasmid in their genome (stable transgenic parasites) were isolated according to the protocol from Rommereim *et al*., 2013. Freshly lysed parasites were counted using a haemocytometer and was infected the HFF confluent cells with 1 parasite per 200 μL of culturing media in each well of 96-well plates by serial dilution. The wells were then incubated at 37°C in 5% CO_2 for 5 to 6 days and were observed under microscope for 5 to 6 days for the well contains a single parasitophorous vacuole. When the parasites lysed out, they were sub-cultured to 6 well plates. The transgenic parasites were then confirmed by western blotting analysis.

2.4 Western blotting

Western blot analyses were performed using the lysates prepared from freshly lysed parasites. Freshly lysed parasites were pelleted by centrifugation at 500 rpm for 5 minutes and resuspended in 40 μL 1x TE buffer containing 1 mM phenylmethylsulfonyl fluoride (PMSF) and aprotinin (protease inhibitor; 1 μg/mL). The mixture was then sonicated for 1 min, kept on ice and centrifuged at 13,000 rpm for 1 min at room temperature. The supernatant was separated from the cellular debris and protein concentration measured using the Bradford reagent according to manufacturer's instructions. Absorbance was measured at 595 nm on the Genesys 10 UV/Visible spectrophotometer.

Cell lysates were denatured in 6x SDS protein loading dye (12% (w/v) SDS, 0.06% (w/v) bromophenol blue, 47% (v/v) glycerol, 6% (v/v) 1M Tris-pH 6.8 and 12.5% (v/v) β mercaptoethanol). This mixture was then heated for 15 minutes and resolved on a

10% SDS-PAGE at 140 V for 1 h 30 mins in a 1x SDS-PAGE running buffer (25 mM Tris-base (pH 8.3), 250 mM glycine, and 0.1% SDS) using a vertical gel electrophoresis system, and subsequently transferred to a nitrocellulose membrane at 100 V for 1 h in 1x transfer buffer (20% (v/v) methanol, 0.29% (w/v) glycine, 0.59% (w/v) Tris-base, 0.037% SDS). After transfer, the membrane was blocked in 5% (w/v) non-fat milk in 1x Tris-buffered saline containing Tween-20 (TBST; 0.068 M NaCl, 8.3 mM Tris-base (pH 7.6), 0.1% Tween) for 1 h at 4°C temperature. Blocked membranes were then probed with the rabbit anti-FLAG (1:5,000) (Frogga) primary antibodies in 2% (w/v) non-fat milk in TBST overnight at 4°C. The membrane was washed 3 times for 5 min with 1xTBST at room temperature. Afterwards, the membrane was incubated in horseradish peroxidase-conjugated goat anti-rabbit (1:20,000) secondary antibody for 1 hr at room temperature, and washed again 3 times for 5 min each. Chemiluminescent detection was performed using the chemiluminescent HRP substrate kit (Millipore, #WBKLS0500) and visualized using the FluorChem Q Imager (Alpha Innotech) with AlphaView-FluorChem Q software.

2.5 Genomic DNA isolation

Genomic DNA was isolated using DNAzol Reagent (Molecular Research Center, Inc) per manufacturer's instructions. After collection of egressed parasites from culture plates, parasites were pelleted at 5000 rpm for 5 mins and re-suspended in 100 µL of distilled H_2O mixed with 100 µL of DNAzol Reagent and lysed by gentle pipetting. 200 µL of phenol chloroform was subsequently added to the tube and centrifuged for 1 min at 12,000 rpm at room temperature. The supernatant was then transferred to a fresh tube.

DNA was precipitated by the addition of 20 µL of 3M NaOAc (pH 5.2) and 800 µL of cold 95% ethanol and mixed by inversion. The mixture then incubated at -80°C for 15 mins, followed by centrifugation at 12,000 rpm for 15 mins at 4°C to pellet the DNA. The DNA precipitate was washed with 500 µL of 75% ethanol, air dried for 15 mins, and resuspended in 30 µL of TE buffer containing RNase A (20 µg/ml) and the tube incubated for 15 minutes at 37°C. Finally, the genomic DNA was stored at -20°C.

2.6 RNA isolation

Total RNA was isolated from freshly egressed RHΔHX parasites using the TRI reagent (Molecular Research Center, Inc., # TR 118). Parasites were centrifuged at 5000 rpm for 5 mins and the parasite pellets were re-suspended in 500 µL of TRIzol reagent for homogenization. Homogenized sample then incubated for 5 minutes at room temperature to allow complete dissociation of nucleoprotein complex. Afterwards, 100 µL of phenol chloroform was added and mixed by inversion. The mixture was incubated for 2-3 minutes at room temperature, followed by centrifugation at 12,000 rpm for 15 min at 4°C and the aqueous supernatant transferred to a fresh tube. After that, 500 µL of 100% isopropanol was added to the aqueous phase and mixed by inversion for homogenization. The tube was incubated at room temperature for 10 minutes then centrifuged for 20 minutes at 12,000 rpm at 4°C. The supernatant was removed and the precipitated RNA was washed with 500 µL of 75% ethanol, air dried for 15 mins, and resuspended in 30 µL of DEPC-treated water. RNA was treated with 1 unit of RQ1 RNase free DNase (Promega, #M6101) for 30 minutes at 37°C. Treated RNA samples were quantified by Nanodrop and then electrophoresed on 1 % agarose gel to verify for

the presence of distinct rRNAs and mRNA smear bands, and the absence of DNA templates. Samples were stored for further uses at −20°C. The RNA was used for reverse transcriptase reaction.

2.7 Reverse transcription reaction (to produce cDNA)

The reverse transcriptase reaction was performed as per manufacturer's instructions (M-MuLV Reverse transcriptase; BioBasic Inc). About 2 µg of total RNA were annealed with 250 ng of gene specific primer (TgArg3'end: 5'GCTTCGCCACTGAGGTGAG3'). A negative control reaction excluding reverse transcriptase enzyme was also performed (NoRT). The extension temperature was 42°C for 1 hour.

2.8 Polymerase chain reaction (PCR)

PCR was performed per the manufacturer's protocol (*Taq* DNA polymerase). Reaction volume of 25 µL containing 2.5 µL (x10 concentration) of *Taq* DNA polymerase buffer, 1.25 µL of 50 mM $MgCl_2$, 2 µL of 2.5 mM dNTPs, 1 µL of 10 µM primers (forward and reverse), 0.2 µL of *Taq* DNA polymerase enzyme, 1 µL template (1 µg) and nuclease free dH_2O volume to 25 µL mixed in a PCR tube. The PCR reaction tubes were placed in the 96-well BioRad T100TM Thermal Cycler for 30 cycles (Denaturation, Annealing, Extension/Elongation temperature and duration were fixed according to protocol and length of PCR product).

2.9 RNA interference assay (Dual Luciferase assay)

To do the gene silencing assay, a dual luciferase plasmid system was employed. The system consists of two plasmids referred to as Firefly luciferase (*Ffluc*) and Renilla luciferase (*Rnluc*) (Crater et al., 2012). To ensure the equal level of transcript production and stability of the transcript, the same promoter (TgTubulin) and same 3'UTR was used for construction of both of the plasmids (Crater *et al*., 2012). The Renilla plasmid contained the microRNA (miRNA) binding site whereas Firefly plasmid did not. Three different *Rnluc* plasmids were used where one does not exhibit binding site, another two *Rnluc* transcript carrying three binding sites (complementary nucleotide sequence to the guide strand) for one of two abundant endogenous *Tg*-miRNAs species, namely *Tg*-miR-60a (pTubRnMIR-60a3x) and *Tg*-miR-4a (pTubRnMIR-4a3x). Following transfection by electroporation with equal concentrations (1µg) of luciferase plasmid *Rnluc* (pTubRn) and *Ffluc* (pTubFf)), the parasites were cultured in confluent HFF monolayers grown in 60 mm tissue culture plates using using MEM media with 1% dFBS and antibiotic-antimycotic at 37°C in 5% CO_2. 48 hours post infection, the freshly egressed parasites were harvested, and lysed with 100 µL of 1x Passive Lysis Buffer (Promega, #E1531) and incubated at room temperature for 10 minutes. The clear supernatant was collected after brief centrifugation at 12,000 rpm for 2 mins and the supernatant was kept on ice and used for the dual luciferase assay. For a *Ffluc* assay, 100 µL of freshly prepared reaction mixture containing 200 µM D-luciferin (Sigma Aldrich, #L9504), 20 mM Tricine, 10 mM $MgSO_4$, 5 mM DTT, 250 µM ATP and 250 µM Coenzyme A was mixed with 20 µL of lysate and was incubated for 10 seconds at room temperature. Similarly, for *Rnluc* assays, 100 µL of freshly prepared reaction mixture containing 0.1 µM

Coelenterazine (Santa Cruz Biotechnology Inc, #sc-205908), 100 mM K_2HPO_4/KH_2PO_4 (pH 7.6) 500 mM NaCl, 1 mM EDTA, and 0.02% BSA was mixed with 20 μL of lysate and was incubated for 10 seconds at room temperature. Immediately after a10 second incubation, the luminescence signals from both reactions were measured with a 20/20n Luminometer (Turner Biosystems; Crater *et al*., 2012). Three independent experiments were performed.

2.10 Growth rate assay

HFF cells were grown on coverslips until confluent. Coverslips with confluent host cell monolayers were infected with 20 μL of freshly lysed parasites and incubated at 37°C in 5% CO_2 with MEM media supplemented with L-glutamine and 1% fetal bovine serum (FBS) (Hyclone) and 0.5% antibiotic antimitotic (Invitrogen, # 15240-062). The parasites were allowed to infect the HFF on the cover-slips and the infected cells were grown for either 24 or 36 h for the determination of growth. The cover-slips were fixed with 3% paraformaldehyde and then Hoechst stained (concentration 3.25 μM).

2.11 Bradyzoite differentiation assay (Alkaline treatment)

For bradyzoite differentiation assays, confluent HFF cells were grown on coverslips and infected with about 20 μL of freshly lysed parasite (tachyzoite form) and incubated at 37°C in 5% CO_2 for 8 hours in MEM media supplemented with L-glutamine and 1% fetal bovine serum (FBS; Hyclone) and 0.5% antibiotic-antimycotic (Invitrogen, # 15240-062). MEM media was subsequently replaced with RPMI-1640 media

(Invitrogen) at pH 8.3, and incubated at 37°C without CO_2 to induce bradyzoite formation. RPMI media was changed every 48 hours and for 5 days (Soete *et al*., 1994).

2.12 Fluorescence microscopy

Coverslips containing *T. gondii* infected HFF cells were rinsed three times with 1x Dulbecco's phosphate buffer saline (DPBS), fixed with 3% (w/v) paraformaldehyde for 5 minutes and permeabilized with 0.25% Triton X 100 in PBS for 10 minutes. Coverslips were then washed three times with 1x PBS and subsequently incubated in a blocking solution of 5% (w/v) bovine serum albumin (BSA) in PBS with gentle agitation for 1 hour at room temperature. Coverslips were then incubated in a humidified chamber with FITC conjugated *Dolichos biflorus* agglutinin (1:300) and diluted in 2% (w/v) BSA for 1 hour for the visualization of cyst structures. Cells were stained with 3.25 μM Hoechst stain (Sigma-Aldrich) for 10 mins at room temperature with gentle agitation. Afterwards, the coverslips were washed again 3 times with 1x PBS and mounted on glass slides with aqueous mounting medium (Fluoromount). Imaging was performed using a Leica DMI 6000 fluorescence microscope, with a Leica DFC 360FX camera and a Leica STP6000 control board. Fields of fluorescent cells were positioned at the 40X objective and images were obtained under the PL Apo 100x/1.40 oil objective and a 1.52 refraction index. The blue filter was used in the detection of Hoechst 33342, the green filter was used to detect green florescent cyst wall with *Dolichos biflorus* lectin stain utilizing Leica Application Software (LAS).

2.13 PCR cloning

PCR products (Result section; Fig 3.3 B7) were purified using a gel extraction technique (QIAEX, Cat. No. 20021) verified by gel electrophoresis and quantified by Nanodrop. pGEM-T Easy Vector (Promega) used for PCR cloning, according to manufacturer's protocol. 50 ng of linearized vector and 6 ng of PCR products were added to the tube containing ligation buffer with T4 DNA Ligase. After mixing, the reactions were incubated overnight at 4°C. After ligation, the recombinant DNA molecules were then transformed into competent DH5α *E. coli* cells and cultured on LB-ampicillin-X-Gal plates for blue/white color screening. The white colonies (contain inserts) were picked and cultured in LB broth with ampicillin and plasmids isolated to verify by restriction digestion that the plasmid contained the insert. The verified plasmids were then subjected to DNA sequencing.

Chapter 3
Results

The first section of results describes an investigation of the expression of argonaute protein in *T. gondii*. A stable transgenic parasite strain of C- terminal FLAG-tagged argonaute was generated for the investigation. The second section highlights the functional contribution of the argonaute proteins to post-transcriptional gene silencing in *T. gondii*, as well as the influence of argonaute protein on parasite growth and interconversion. The latter study was performed using wild type *T. gondii* and argonaute knock-out (AGOKO) parasite strains.

3.1 Investigation of the argonaute variants in *T. gondii*

3.1.1 Analysis of argonaute transcripts from *T. gondii*.

cDNA coding for argonaute protein was obtained by reverse transcription of total RNA by M-MuLV reverse transcriptase, using argonaute gene specific primers (TgArg3'end and AGO Exon 5_Rv) as shown in Figure 3.1 A. TgArg3'end primer binds to the transcript in exon 7 and amplified full length of cDNA to exon 1. And AGO Exon 5_Rv primer used to further verify the 5' end of the transcript which binds to the transcript in exon 5 and amplified to exon 1. To verify the quality of cDNA and to ensure the cDNA was not contaminated with genomic DNA, an end-point PCR was performed with a specific primer set (indicated as "a" in Figure 3.1 A) using cDNA sample and no-RT control. A PCR product was observed from the cDNA sample (L1; Figure 3.1 B1) but not from negative control (L9; Figure 3.1 B1). Eight different overlapping primer sets (indicated as a, b, c, d, e, f, g, h) were then used to detect different regions of the

transcript (Figure 3.1 A). The PCR product from primer set “a” showed a single band at the predicted length of 860 bp (L1; Figure 3.1 B1), and no alternative band or form was observed from this region. Similarly, for primer set “b” the predicted product was detected (816 bp) but a very short product was also seen that not correlate with any length of alternative form of transcript (L2; Figure 3.1 B1). Primer set “c” amplified a single product which corresponding to the predicted length of 298 bp with no alternative band or form for this primer set (L3; Figure 3.1 B1). Primer set “d” and primer set “e” produced predicted PCR products indicated with one low intensity band suggesting the products were either derived from alternative form of transcript or from the different length of cDNA product amplified from same gene specific primer (TgArg3'end) or from the degraded cDNA product or primer non-specific attachment to the template (L4, L5; Figure 3.1 B1). To confirm these results an independent PCR analysis was conducted with the same primer sets (“d” and “e”) from different cDNA sample and even using different PCR reagent (L1 and L2; Figure 3.1 B2, L1; Figure 3.1 B3 and L1; Figure 3.1 B6). PCR product from primer set “d” were also compared between the cDNA sample from TgArg3'end and AGO Exon 5_Rv gene specific primers (L1 and L2; Figure 3.1 B2). PCR product from primer set “d” shown a prominent band at the predicted position (570 bp) (L1 and L2; Figure 3.1 B2 and L1; Figure 3.1 B6). On the other hand, primer set “e” which overlaps the region for primer set “c and d” amplified a single predicted length (738 bp) of PCR product (L1; Figure 3.1 B3) which strongly suggested that there was no alternative form of transcript for this region. Therefore , the primer sets “a”, “b”, “c”, “d” and “e” confirmed the length of the sequences from exon 3 to exon 7 with no alternative form of the transcript.

Primer sets “f”, “g” and “h” were used to analyze the transcript from exon 1 to the 5’ junction of exon 3 (L6, L7, L8; Figure 3.1 B1). Primer set “f” showed a shorter band (~500 bp) than the expected band of 869 bp (L6; Figure 3.1 B1) whereas primer set “g” showed a larger band (~500 bp) than the expected band of 393 bp (L7; Figure 3.1 B1). Even though, primer set “f” overlaps the region of primer set “g” (Figure 3.1 A), both primer sets produced similar size product. Noticeably, primer set “f” and “g” were used to amplify the product from genomic DNA to ensure the primer set is amplifying the predicted product from genomic DNA (L2, L1; Figure 3.1 B4). The lengths of the amplified PCR product were expected (L2, L1; Figure 3.1 B4). Afterwards, Several PCR reactions were conducted for these primer sets using cDNA from TgArg3'end and AGO Exon 5_Rv to verify the data but only the unexpected length of PCR products were amplified (L3, L4, L5, L6; Figure 3.1 B2, L2 and L1; Figure 3.1 B5, L2; Figure 3.1 B6) which did not even correlates with any alternative form of the transcript. Finally, primer set “h” which overlaps the primer sets “f” and “g” region showed a shorter band (~600 bp) than the expected size of 1179 bp (L8; Figure 3.1 B1, L7 and L8; Figure 3.1 B2). A new PCR was conducted using newly amplified cDNA sample with different brand end-point PCR reagent to crosscheck if there were multiple bands which could indicate an alternative form of the transcript (L1; Figure 3.1 B7). Multiple bands were detected for the primer set “h” suggesting the variation in this region may correlate with alternative forms of argonaute mRNA. PCR products that appeared at around 1200, 900, 700 bp (L1; Figure 3.1 B7) were isolated for DNA cloning and sent for nucleotide sequencing. The sequences did not have *T. gondii* argonaute in them, but sequences showed similarity with sequences from *T. gondii* Zinc finger domain containing protein mRNA,

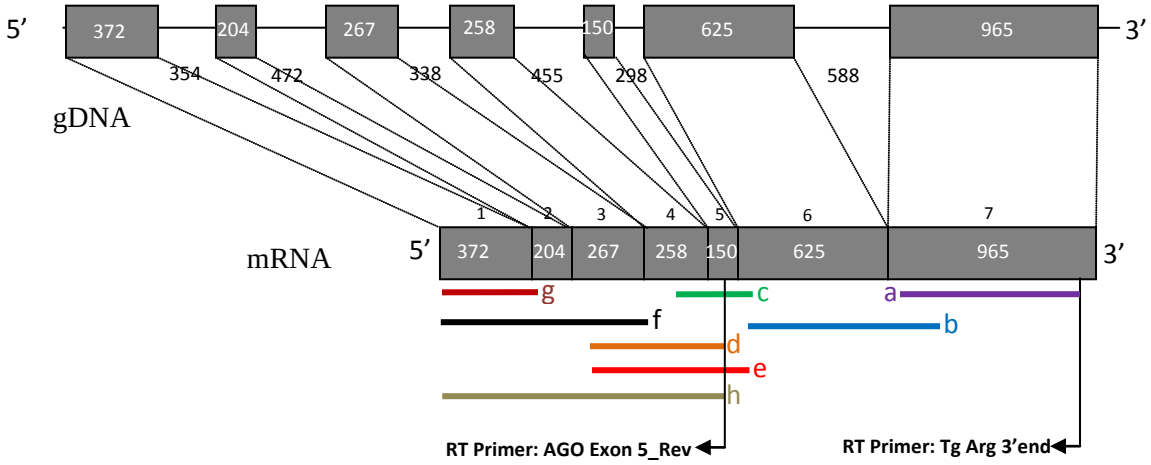
Wuchereria bancrofti genome, the *T. gondii* genome (VEG, chromosome III, V, VI, IX, XI), a *T. gondii* putative glycoprotein mRNA, the *T. gondii* Tubulin beta chain gene CDS, a *Neospora caninum* Liverpool hypothetical protein mRNA, a *Neospora caninum* Liverpool genome sequence (chromosome IX), a *T gondii* serine/ threonine protein kinase putative mRNA and a *T. gondii* hypothetical protein mRNA. However, a portion of the sequences from the transcript which was common to all the related CDS reported in different lab was amplified by PCR and also sent for sequencing, which showed the same sequence with *T. gondii* argonaute.

3.1.2 Generation of transgenic *T. gondii* strains expressing recombinant tagged TgAgo protein at physiological levels

To determine if *T. gondii* expresses different variants of the argonaute protein, a streptavidin-FLAG (SF; 21 a.a., ~2.5 kDa) were tagged C-terminally to create TgAgo-SF strains via a homologous recombination (knock-in) in the ΔKu80ΔHX parasite strain. The ΔKu80 strain allows for only homologous recombination (Huynh and Carruthers, 2009). PCR analysis was performed to isolate and identify parasite clones stably expressing SF, using primer set “a” and “b” (a+b; Figure: 3.2A1). To monitor the expression of recombinant protein, positive clones and parental strains were analyzed (western blot analysis) using α- FLAG antibody. Several western blot analyses were performed using protein isolated from different passage of the parasite (Figure: 3.2 B1-B4). When parasite lysates were analyzed with α-FLAG antibody a clear prominent band was observed at ~62 kDa for FLAG tagged argonaute from all of the western blot membrane. No band was observed for parental strain (ΔKu80ΔHX), which was used as a negative control

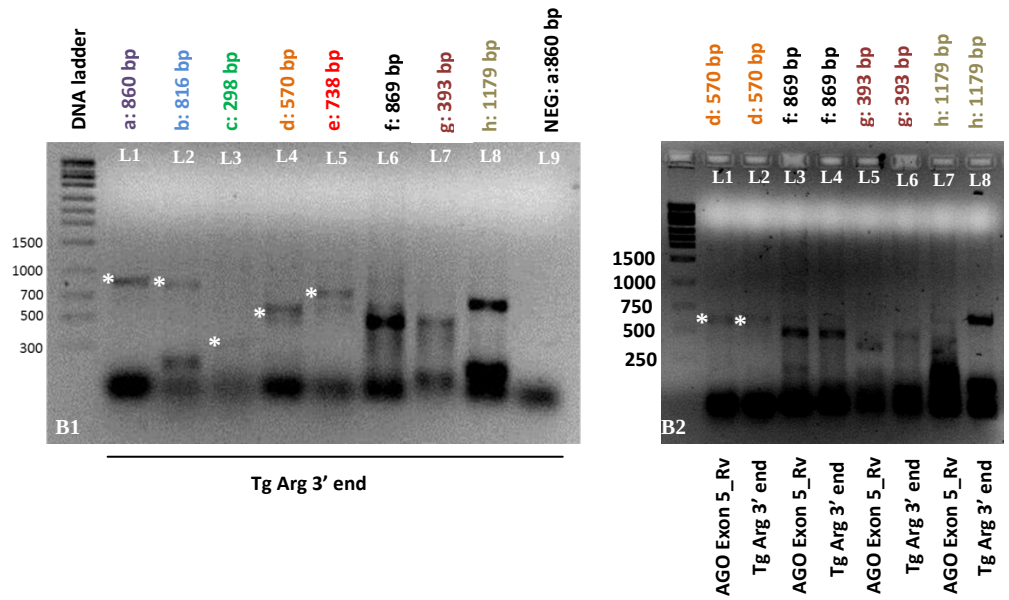
where argonaute was not tagged (Figure: 3.2 B1-B4).The molecular weight of the streptavidin FLAG (SF) TAP is ~2.5 kDa and the molecular weight corresponding to the argonaute protein itself is ~58.5 kDa. A band at ~23 kDa (Fig 3.2 B3) might be the degraded product.

A



B

.



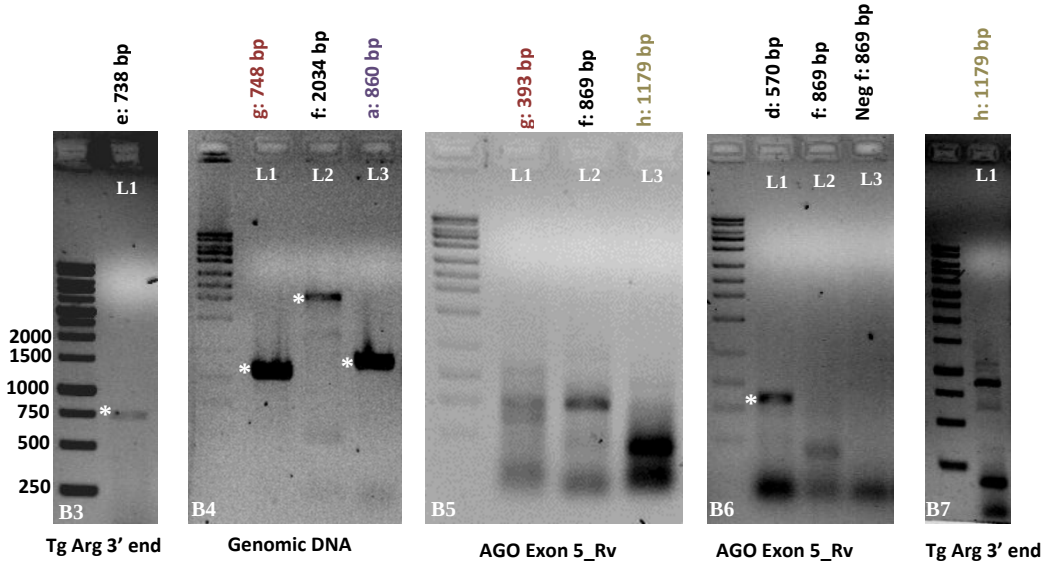


Figure 3.1: Detection of argonaute transcripts from *T. gondii*. A) The full length of argonaute transcript (Exon 1 -7 and their length indicated) from Toxodb.org. The gene specific primers used for reverse transcriptase reaction as well as 8 overlapping primer sets indicated as a, b, c, d, e, f, g, h. B) PCR product ran on 1% agarose gel. In "B1" gel image shows all PCR products for 8 overlapping primer sets including the sample from negative reverse transcriptase reaction. In B2 gel image shows the PCR product using cDNA from TgArg3'end and AGO Exon 5_Rv gene specific primers, B4 shows the PCR product for primer set "a, f, g" from genomic DNA. B3, B5, B6 and B7 show the PCR product for primer set "d, e, f, g, h" after running a new PCR using different gene specific primers and also using differ PCR enzyme. Two different types of 1Kb ladder were used for the gel images where in B1 the lower band starts from 300 bp but for B2-B7 the lower band starts from 250 bp.

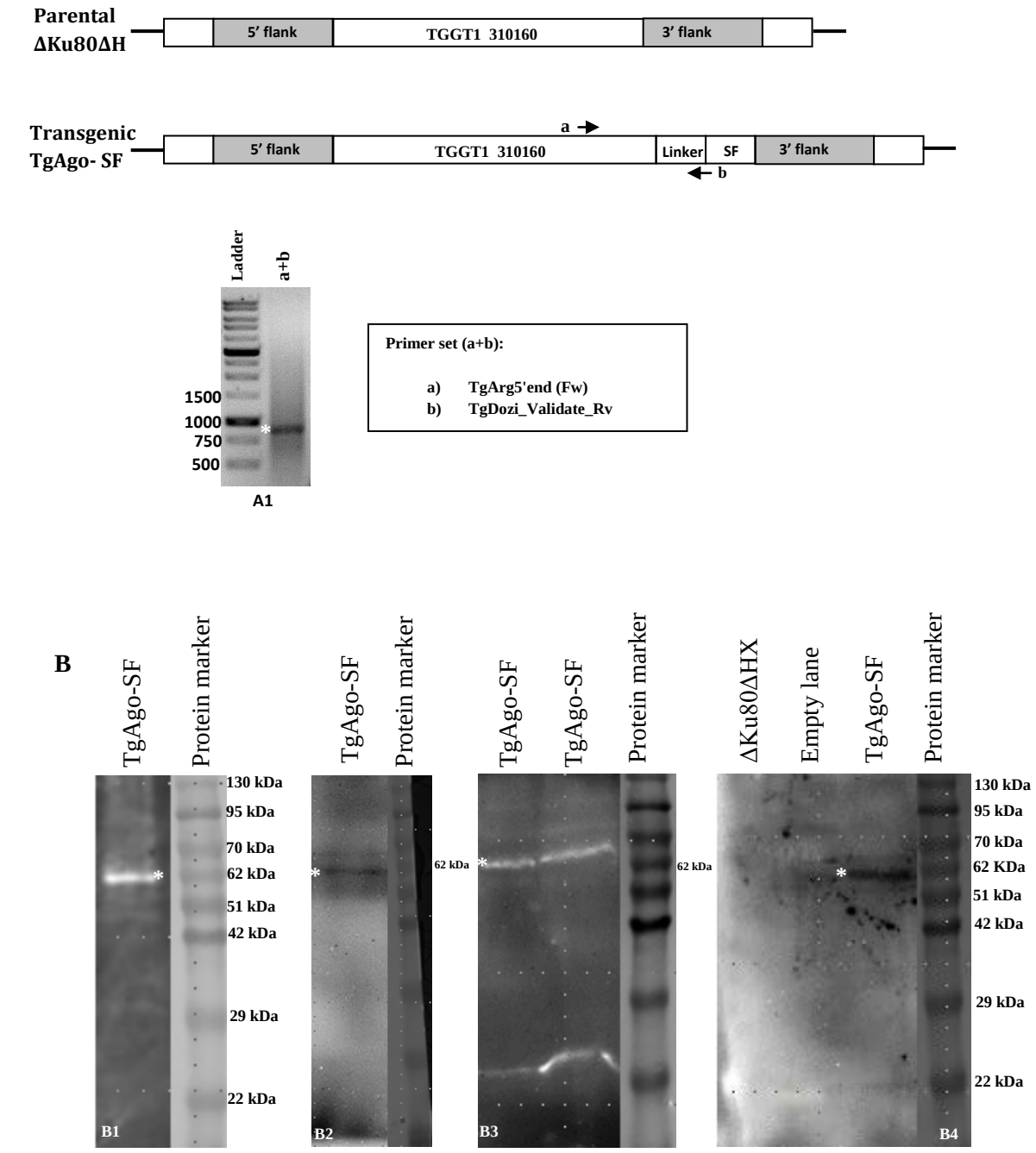


Figure 3.2: Investigation of the size of the argonaute protein from *T. gondii*. ΔKu80 parasites were genetically modified to contain transgenic sequences allowing for the analysis of endogenous TgAgo protein. A) Illustration of parental and transgenic genomic arrangement B) Expression of TgAgo-SF in transgenic parasites. Western blot image shows the band for FLAG tagged argonaute protein from *T. gondii* and no band for wild type *T. gondii* (ΔKu80ΔHX) as negative control.

3.2 The contribution of argonaute protein to post-transcriptional gene silencing pathway in *T. gondii*, and effect on parasites growth and interconversion ability.

3.2.1 Verification of argonaute gene knock out from T. *gondii* genome.

The AGOKO clonal strain was obtained from Dr. Boothroyd (Stanford University School of Medicine). The clonal strain was analysed to confirm that the argonaute encoding sequence was replaced by the targeted plasmid insertion. Three primer sets (Figure 3.3) were used in end-point PCR analysis to evaluate for the replacement of the argonaute gene by the Hypoxanthine-Xanthine guanine phosphoribosyltransferase (HXGPRT) cassette. The primer set c+c' corresponded to the deleted region, where a PCR product was not observed from AGOKO strain but was observed from the parental strain, thus identifying argonaute knock out clones (Figure 3.3). Furthermore, PCR using primer sets a+a' and b+b', which corresponded to the inserted region, were positive from argonaute disrupted clones (AGOKO), but were negative in RHΔHX strain. These results suggested that the HXGPRT cassette was inserted correctly into the argonaute locus (Figure 3.3).

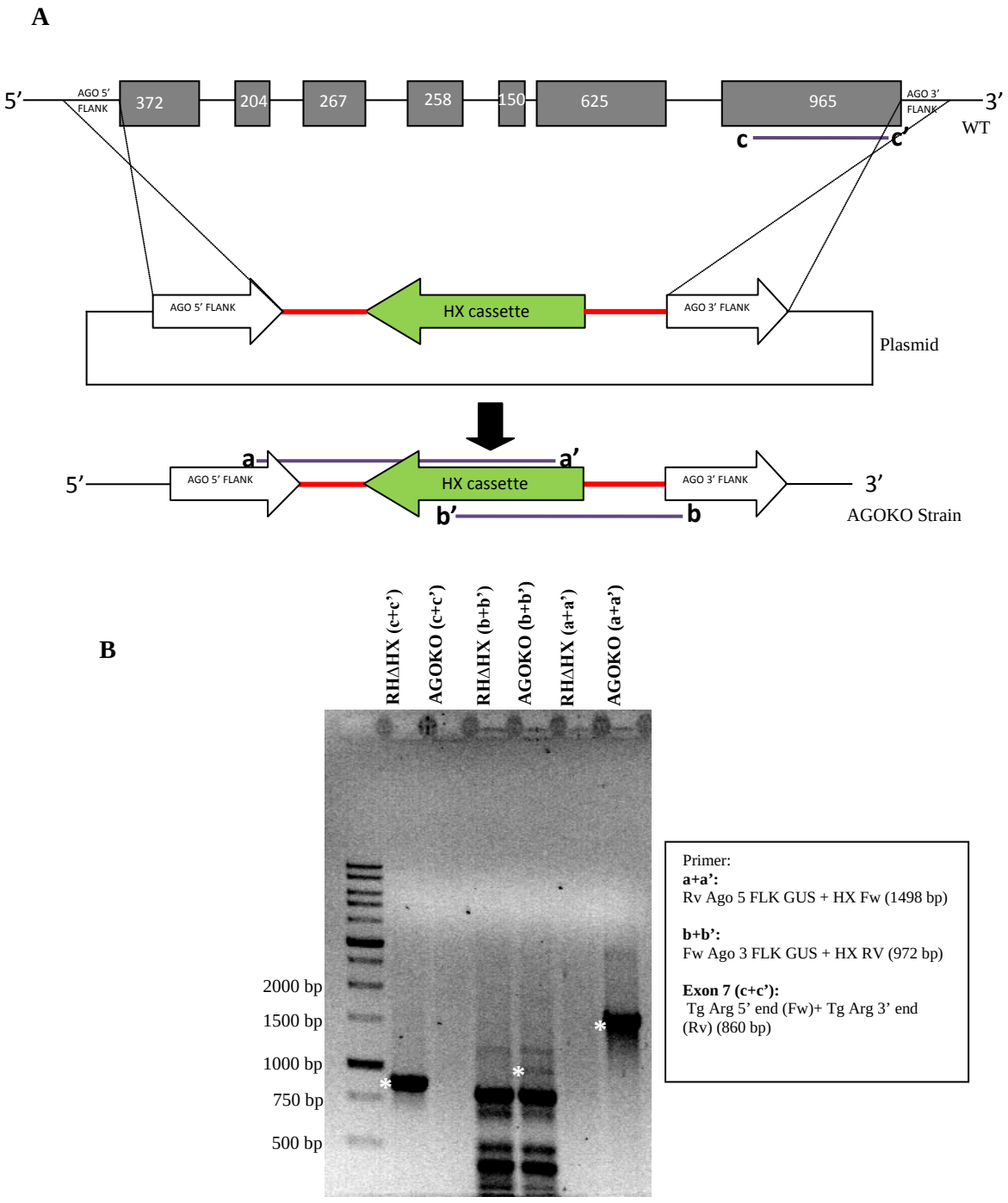


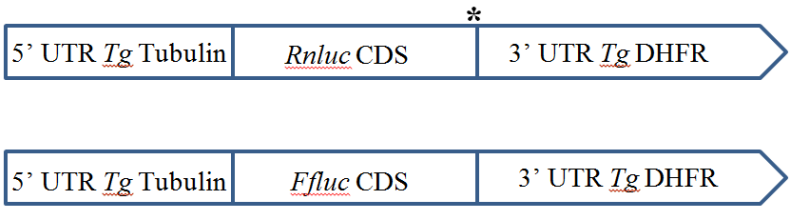
Figure 3.3: Confirmation of argonaute gene knocked out from *T. gondii* genome. A) Plasmid constructs for AGOKO (see appendix to see actual plasmid contruct) from *T. gondii* genome. Plasmid targets a 5.3 kb deletion of the argonaute gene via integration of the HXGPRT cassette (backward orientation) into the RHΔHX strain. B) Primer sets were used to evaluate the parentral strain (RHΔHX) and AGOKO for replacement of the argonaute gene with HX cassette. (Designated PCR product were indicated with asteric* mark)

3.2.2 Gene silencing assay analysis using reporter silencing assay (Dual Luciferase assay).

Since argonaute proteins are the core component in the RNA interference (RNAi) gene silencing pathway (Rand *et al*., 2004; Song *et al*., 2004), the removal of the lone argonaute encoding gene should alter post transcriptional gene regulation in *T. gondii*. This hypothesis was tested by comparing the gene silencing ability of the argonaute knockout strain (AGOKO) with the wild type parasite (RHΔHX). A dual luciferase reporter system was used to evaluate this hypothesis. In the dual luciferase assay, two reporter transcripts were used, namely Firefly luciferase *(Ffluc*) and Renilla luciferase (*Rnluc*) (Figure 3.4 A). *Ffluc* transcript lacking a microRNA binding site and was used for internal control for transfection efficiency among various experiments. On a parental construct of an *Rnluc* transcript carrying the *Tg*-miRNA binding sites (complementary nucleotide sequence to the guide strand) for one of two abundant endogenous *Tg*-miRNAs species, namely *Tg*-miR-60a and *Tg*-miR-4a. The *Rnluc* transcript with the *Tg*-miR-60a binding site was indicated as Rn60a and Rn4a with *Tg*-miR-4a binding site. The guide strands (25-26 nucleotides) and seed sequences (underlined 6 nucleotides) of *Tg*-miR-60a and *Tg*-miR-4a are shown in Figure 3.4 B. The relative expression of the *Rnluc* transcripts with *Tg*-miRNAs binding site was measured and normalized against the *Rnluc* reporter transcript without any binding sites for *Tg*-miRNAs (Rnob). RHΔHX revealed that transcripts with the *Tg*-miR-60a binding site showed a relative expression of 28 ± 23%, reflecting a silencing effect of 72% whereas the transcript with a *Tg*-miR-4a binding site showed a relative expression to 55 ± 16%, reflecting a 45% silencing effect, which is in agreement with previous studies (Crater *et al*., 2015). On the other hand, in

AGOKO parasites, transcripts with *Tg*-miR-60a binding site showed a relative expression of around 98 ± 4%, reflecting only a 2% silencing effect. The transcript with *Tg*-miR-4a binding site showed a relative expression of around 86 ± 15%, reflecting a silencing effect of 14%. In conclusion, it was found that the silencing ability of the parasites was reduced for both *Tg*-miR60a and for *Tg*-miR-4a targets in the absence of argonaute expression (Figure 3.4 C).

A



B

Tg-miRNA (name and sequence)	Predicted gene target and interaction (Gene ID and sequence)
Tg-miR-60a acacagucgguacgaaauccauacu	5' AGUAUGGAAUUCGUACCGACUGUGU 3' 3' ucauaccuaaagcauggcugacaca 5'
Tg-miR-4a augcuuugcuuggaagcuguagucauu	TGGT1_203470 (A) 5' AAUGACUACAGCUUCCAAGCAAACAU 3' 3' uuacugaugucgaaggcucguuugua 5'

C

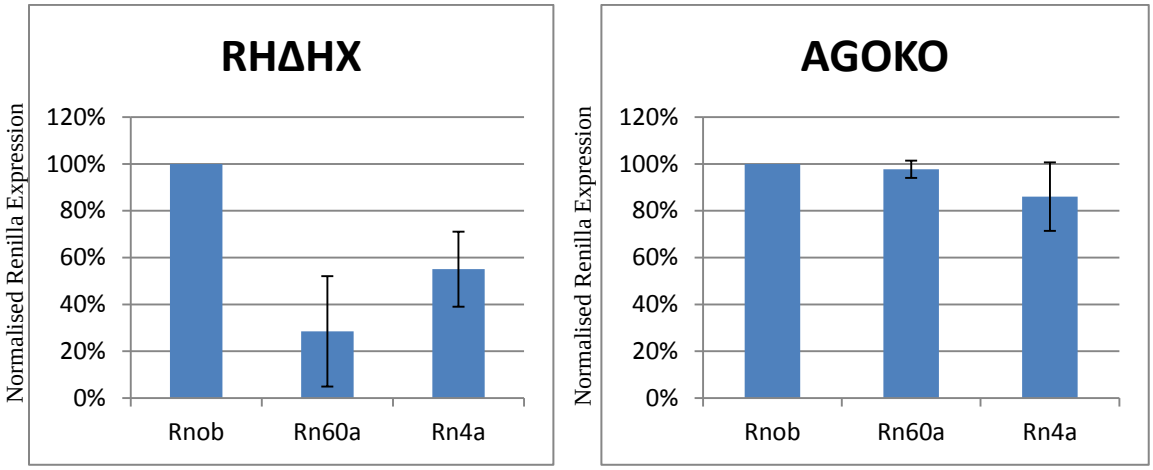


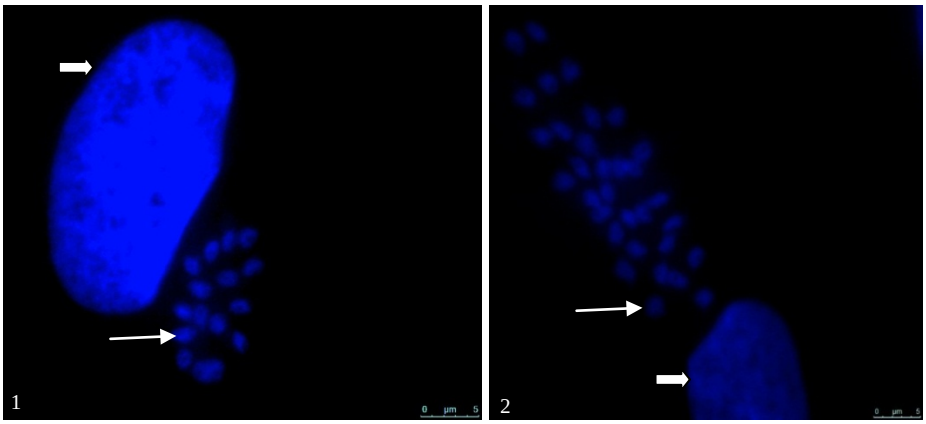
Figure 3.4: Gene silencing assay. A) Graphical representation of transcripts generated from the dual luciferase reporter system used in this study. Asterisk (*) mark on Renilla luciferase *(Rnluc)* transcript indicates the microRNA binding site. Firefly transcript does not contain microRNA binding site. B) *Tg*-mir60a and *Tg*-mir4a binding sequences. The guide strands (25-26 nucleotides) and seed sequences (underline marked) of *Tg*-miR-60a and *Tg*-miR-4a are shown. Three different Renilla constructs were used, where one has no binding site (Rnob), another one has binding site for *Tg*-miR-60a (Rn60a) and the third one has for *Tg*-miR-4a (Rn4a). C) Dual Luciferase assay for RHΔHX and AGOKO. Relative Renilla expression was normalized against Renilla with no binding site (Rnob) expression. Concentration used for electroporation: *Ffluc*: *Rnluc* = 1:1; 1 µg of each construct. Analyses were performed in triplicate with standard deviation shown (error bars).

3.2.3 The role of argonaute on growth rate and bradyzoite differentiation in *T. gondii*.

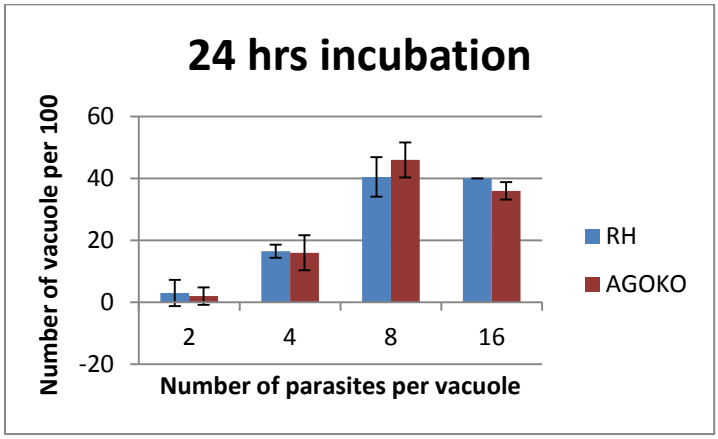
3.2.3.1 Growth rate assay:

The study was design to determine if the ablation of argonaute expression has any effect on parasite growth. To assess growth, parasite numbers per vacuole were counted after 24 hours and 36 hours following infection (Figure 3.5 A). After 24 hours of infection, 40 ± 6% of vacuoles showed 8 parasites per vacuole for the RHΔHX strains, but AGOKO parasites exhibited 46 ± 5%. Furthermore, the number of vacuoles with 16 parasites/vacuole from RHΔHX and AGOKO was 40 ± 0% and 36 ± 3% respectively (Figure 3.5 B). This gives an average parasite doubling time of approximately 6-8 hours for both strains. Around 16% of the vacuoles contained 4 parasites and a relatively few number of vacuoles contained 2 parasites per vacuole for both RHΔHX and AGOKO strains. After 36 hours of infection, RHΔHX showed 29 ± 1% of vacuoles contained 16 parasites per vacuole and 26 ± 3% in AGOKO. In addition, 64% of vacuoles in RHΔHX and 68% of vacuoles in AGOKO were found to contain 32 parasites per vacuole (Figure 3.5 C). These figures were correlated with a doubling time of 6-8 hour for both RHΔHX and AGOKO strians. In summary, this study revealed that altered argonaute expression did not significantly affect parasitic growth.

A



B



C

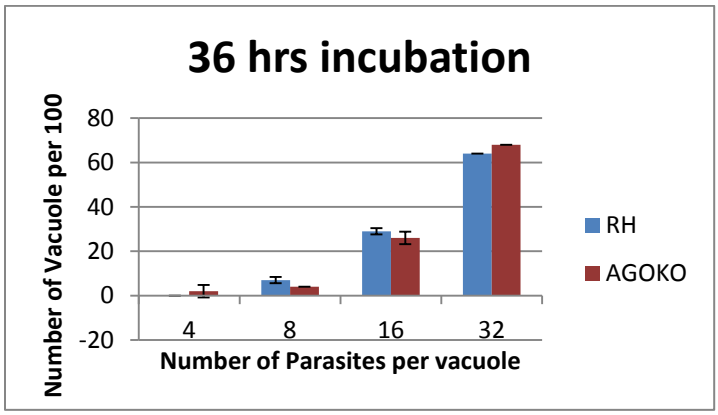
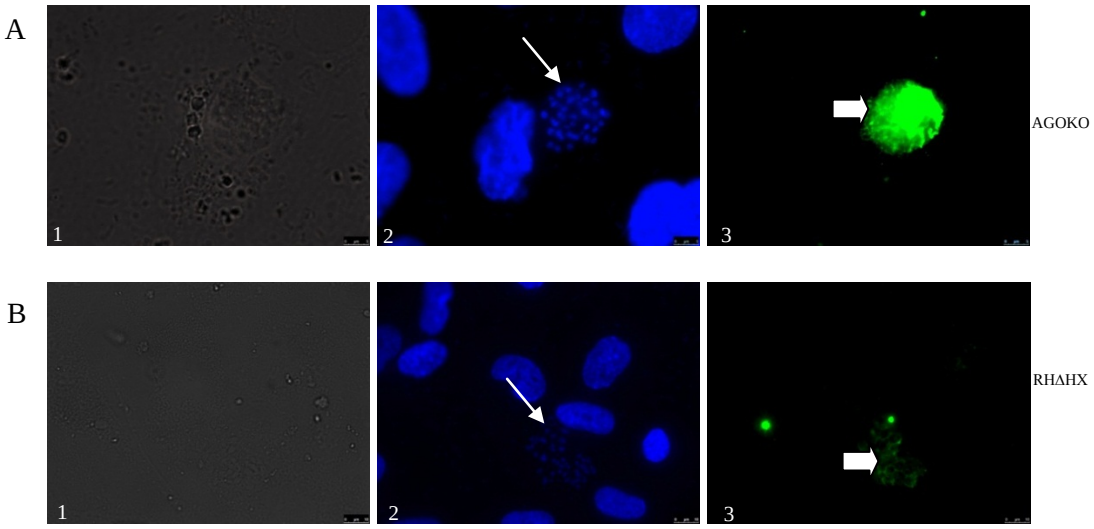


Figure 3.5: Growth analysis of AGOKO transgenic and parental clones at 24 and 36 hours. A) The nuclei were stained with Hoechst. The thick arrow indicates the nucleus of the HFF host cell, while the thin arrow indicates the nuclei of the *T. gondii* parasites in a vacuole. A1 shows 16 parasites in a vacuole and A2 shows 32 parasites in a vacuole. Parental (RHΔHX) and transgenic strains were allowed to infect HFF monolayers and the numbers of parasites were counted per vacuole 24 and 36 hrs post-infection (B and C). At 24 hours, vacuoles for both parental and transgenic clones were found to contain 8 and 16 parasites (B). At 36 hours post-infection, the majority of vacuoles contained 32 parasites (C). Analyses were performed in triplicate with standard deviation shown (error bars).

3.2.3.2 The role of argonaute in bradyzoite differentiation

A database (ToxoDB.org) search for *T. gondii* argonaute revealed that the interconversion ability was less in Type 1 parasites (RH) then type 2 (ME49), but that the expression of argonaute was higher in type 1 parasite compared to type 2. The information suggested that the level of argonaute expression may affect the ability of the parasite to differentiate into bradyzoites. This hypothesis was assessed by comparing the differentiation ability of AGOKO and wild type parasite strains (RHΔHX) under alkaline growth conditions. Cyst formation was quantified by counting the number of vacuoles that formed cysts under tested condition in both strains (Figure 3.6a). Percentages of vacuoles that formed cysts were ~25% from RHΔHX and ~62% from AGOKO, which supports the hypothesis that argonaute expression is correlated with bradyzoite differentiation capacity (Figure 3.6b).

a



b

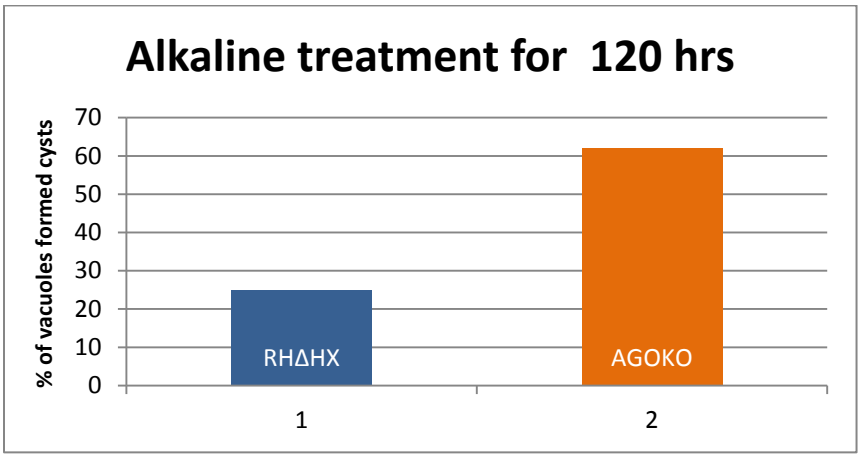


Figure 3.6: a) Differentiation assay of transgenic AGOKO parasites in comparison to parental parasites (RHΔHX). Parasites were cultured for 5 days (120 hrs) under alkaline condition (pH 8.3). *Dolichos biflorus*- FITC stain was used to reveal cyst structure, and Hoechst staining to visualize the nuclei. Image shows the Cyst structure after Dolichos-FITC by AGOKO parasites (A3). b) Quantification of cyst formation in AGOKO parasites compare to parental parasites (RHΔHX). Single analyses were performed.

Chapter 4
Discussion and conclusion

RNA interference (RNAi) is a naturally occurring, evolutionary conserved, highly efficient post transcriptional gene regulation pathway. Many components of the RNAi were discovered in *T. gondii*, two of which are miRNAs and the argonaute protein (Al Riyahi *et al*., 2006; Braun *et al*., 2010). However the molecular mechanisms of the interacting components in this pathway in *T. gondii* are still not fully understood. Argonaute protein is the core component in the RNAi mechanism (Fire *et al*., 1998; Rand *et al*., 2004; Hutvagner and Simard, 2008). The elucidation of the different components and their molecular mechanism may eventually lead to development of potential drug targets that would prevent the parasitic infection. The current study was aimed to investigate the variants of the argonaute protein expressed by *T. gondii* and the contribution of argonaute protein to post transcriptional gene silencing pathway as well as its role in parasite growth and differentiation.

Previous studies by several research groups reported different sizes of argonaute protein expressed in *T. gondii* (databases from ToxoDB.org and NCBI) (Figure 1.5). The research groups reported the same sequences at the 3’ end of the argonaute transcript but their observations varied at the 5’ end (Figure 1.5). FASTA format of amino acid sequences for argonaute protein shown the C-terminal end is same, but the length of the N-terminal end varies (Figure 1.6). For this reason, the domain reported for argonaute exhibits a similar content for the catalytic PIWI and MID domain (C-terminal end domain) whereas the PAZ domain and N- terminal shows the discrepancy (Figure 1.6). Therefore, it was of interest to study the possibility that *T. gondii* express multiple forms

of argonaute. It is possible that multiple forms could arise from multiple transcript variants from alternative splicing of the primary transcript, or multiple proteins from the use of different start codons on the transcript. Notably, a few studies have suggested that *T. gondii* may utilize alternative splicing mechanism as a means of post transcriptional gene regulation (Suvorova and White, 2014). However, others have reported 2.5 kb single transcript for argonaute after northern blot analysis by using DIG-labeled probe prepared from the sequences correspondent to the internal PIWI domain (Al Riyahi *et al*., 2006).

Type 1 parasite (RHΔHX) was used for this experiment because it is more virulent and optimal for rapid replication and efficiency of lysing host cells, which allows for isolation of large amounts of parasites (Saeij *et al*., 2005). This study is the first to describe a comprehensive analysis of argonaute transcript variants in *T. gondii*. Using multiple overlapping PCR primer sets, we identified single variant of transcript for TgAgo which correlates with the length reported from Barik research group (marked as 2 in the Figure 1.5). Additionally, no splice variants of the transcript were observed from the investigation, which also correlates with the northern blot report by Al Riyahi *et al*., (2006).

Moreover, we also demonstrated that *T. gondii* expresses single form of argonaute protein by immunoblot analysis (Figure 3.2), and the molecular weight of the argonaute is ~ 58.5 kDa, which is similar to the size to that reported by Al Riyahi *et al*., (2006). However, the length of the transcript detected correlates with Barik research group (GenBank accession no. DQ177874, marked as 2 in the Figure 1.5) but according to that sequence the calculated molecular weight supposed to be around ~83 kDa. So, it can be

assume that the 5' UTR might be longer than the length reported by Al Riyahi *et al*., (2006), which spans from exon 3-5 whereas the coding sequence is from exon 6 to 7. This shows that *T. gondii* expresses a single species of argonaute that is shorter than most other eukaryotes. In *vitro* RNA cleavage assay with TgAgo mutant with deleted different portions of N-terminus showed that the cleavage activity of argonaute was conserved as long as the putative PIWI domain was intact (Musiyenko *et al*., 2012). Furthermore, archeal argonaute protein from *Archaeoglobus fulgidus* lacks PAZ domain but crystal structure revealed that amino acid residues at the N- terminus can form a domain which can bind with siRNA and guide the RISC complex (Ma *et al*., 2005). Moreover, Song *et al.* (2004) mentioned that PAZ domain is poorly conserved and not easy to detect from the primary amino acid sequences. Thus, N-terminal amino acid sequences from short version of TgAgo could form a PAZ like fold. In that case, crystallographic analysis may disclose the argonaute folding structure during the silencing activity or if PAZ domain is absent further study will be fascinating to analyze TgAgo, that how TgAgo works in absence of PAZ domain?

Argonaute protein is an important component in RISC complex in RNAi pathway for post transcriptional gene regulation in many organisms including different protozoan parasites (Ullu *et al*., 2004; Hutvagner and Simard, 2008; Meister, 2013). MiRNAs are an important factor in gene silencing pathways where miRNAs guide the RISC complex towards mRNA and bind with mRNA by sequence complementarity (either perfect or near-perfect complementary binding) (Carthew, 2006; Zhang *et al*., 2007; Ghildiyal and Zamore, 2009; Braun *et al*., 2010). Furthermore, previous studies from our group demonstrated that the introduction of a luciferase plasmid construct with binding sites for

two highly expressed *Tg*-miRNA families, namely *Tg*-miR-60a (Rn60a) and *Tg*-miR-4a (Rn4a), exhibited silencing of the luciferase expression by endogenous *Tg*-miRNAs in *T. gondii* RHΔHX strains (Crater *et al*., 2015). My results revealed a silencing effect on luciferase expression in RHΔHX strains. However, in AGOKO strains the silencing ability was reduced (Figure 3.4). This observation suggests that argonaute has a major contribution in RNAi based gene silencing pathway in *T. gondii*. Further studies of the molecular mechanism and associated components of RNAi silencing pathways will undoubtedly reveal a great surprise.

Argonaute proteins are associated with developmental processes in diverse organisms (Carmell *et al*., 2002). Reduction of argonaute expression in Drosophila and *C. elegans* produces defective germlines (Cox *et al*., 2000; Tijsterman *et al*., 2002). Hence, it has been speculated that absence of argonaute expression in *T. gondii* may lead to effects on parasite growth. At 24 and 36 hours post infection an average doubling time is approximately 6-8 hours for both AGOKO and wild type (RHΔHX) strains (Figure 3.5), which is also similar to previously reported values for wild type strain (Black and Boothroyd, 2000), suggesting that reduction of argonaute expression does not affect *T. gondii* growth.

To monitor if knock out of argonaute expression has any influence on bradyzoite differentiation, my experimental findings demonstrated that AGOKO strains formed around 40% more cyst structures than the parental RHΔHX, which verified that argonaute expression inhibits bradyzoite conversion (Figure 3.6).

Taken together, the results presented in this thesis indicate that *T. gondii* expresses single species of argonaute protein which is shorter than most other eukaryotes. The knockout of TgAgo expression resulted in a reduction in silencing ability which suggests that argonaute has a major contribution in RNAi based gene silencing pathway in *T. gondii*. It was also observed that TgAgo plays a role in differentiation since the knockout of argonaute expression resulted in an increase in bradyzoite formation. Following this result, future work might be directed to determine the possible functional domain of TgAgo, by utilizing knockin studies of various argonaute variants truncated for specific domains in AGOKO strains and monitor if argonaute complements the function in the AGOKO strain.

References

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

Baek D, Villén J, Shin C, Camargo F. D, Gygi S. P, & Bartel D. P. (2008). The impact of microRNAs on protein output. *Nature* 455: 64-71.

Bartel DP. (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* 116(2):281-297.

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

Black, D. L. (2003). Mechanisms of alternative pre-messenger RNA splicing. *Annual Review of Biochemistry*, 72(1), 291-336.

Blader, I. J., Coleman, B. I., Chen, C. T., & Gubbels, M. J. (2015). Lytic cycle of *Toxoplasma gondii:* 15 years later. *Annual Review of Microbiology*, 69, 463-485.

Boothroyd, J. C., & Grigg, M. E. (2002). Population biology of *Toxoplasma gondii* and its relevance to human infection: do different strains cause different disease? *Current Opinion in Microbiology*, 5(4), 438-442.

Bohne, W., Heesemann, J., & Gross, U. (1993). Induction of bradyzoite-specific *Toxoplasma gondii* antigens in gamma interferon-treated mouse macrophages. *Infection and Immunity*, 61(3), 1141-1145.

Bohne, W., Heesemann, J., & Gross, U. (1994). Reduced replication of *Toxoplasma gondii* is necessary for induction of bradyzoite-specific antigens: a possible role for nitric oxide in triggering stage conversion. *Infection and Immunity*, 62(5), 1761-1767.

Bohne, W., Wirsing, A., & Gross, U. (1997). Bradyzoite-specific gene expression in *Toxoplasma gondii* requires minimal genomic elements. *Molecular and Biochemical Parasitology*, 85(1), 89-98.

Brennecke, J., Aravin, A. A., Stark, A., Dus, M., Kellis, M., Sachidanandam, R., & Hannon, G. J. (2007). Discrete small RNA-generating loci as master regulators of transposon activity in *Drosophila*. *Cell*, 128(6), 1089-1103.

Braun, L., Cannella, D., Ortet, P., Barakat, M., Sautel, C. F., Kieffer, S., & 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.

Cai, X., Hagedorn, C. H., & Cullen, B. R. (2004). Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs. *RNA*, 10(12), 1957-1966.

Carthew, R. W. (2006). Gene regulation by microRNAs. *Current Opinion in Genetics & Development*, 16(2), 203-208.

Carmell, M. A., Xuan, Z., Zhang, M. Q., & Hannon, G. J. (2002). The Argonaute family: tentacles that reach into RNAi, developmental control, stem cell maintenance, and tumorigenesis. *Genes & Development*, 16(21), 2733-2742.

Carmody, S. R., & Wente, S. R. (2009). mRNA nuclear export at a glance. *Journal of Cell Science*, 122(12), 1933-1937.

Cenci-Goga, B. T., Rossitto, P. V., Sechi, P., McCrindle, C. M., & Cullor, J. S. (2011). *Toxoplasma* in animals, food, and humans: an old parasite of new concern. *Foodborne Pathogens and Disease*, 8(7), 751-762.

Cowman, A. F., & Crabb, B. S. (2002). The *Plasmodium falciparum* genome--a blueprint for erythrocyte invasion. *Science*, 298(5591), 126-128.

Cox, D. N., Chao, A., & Lin, H. (2000). Piwi encodes a nucleoplasmic factor whose activity modulates the number and division rate of germline stem cells. *Development*, 127(3), 503-514.

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*. *American Journal of Biomedical Sciences*, 4(1), 1-13.

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

Cuellar, T. L., & McManus, M. T. (2005). MicroRNAs and endocrine biology. *Journal of Endocrinology*, 187(3), 327-332.

Czech, B., & Hannon, G. J. (2011). Small RNA sorting: matchmaking for Argonautes. *Nature Reviews Genetics*, 12(1), 19-31.

Darnell, J. E. (2013). Reflections on the history of pre-mRNA processing and highlights of current knowledge: a unified picture. *RNA*, 19(4), 443-460.

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.

Derouin, F., & Garin, Y. J. (1991). *Toxoplasma gondii*: blood and tissue kinetics during acute and chronic infections in mice. *Experimental Parasitology*, 73(4), 460-468.

Donald, R. G., Carter, D., Ullman, B., & Roos, D. S. (1996). Insertional tagging, cloning, and expression of the *Toxoplasma gondii* hypoxanthine-xanthine-guanine phosphoribosyltransferase gene Use as a selectable marker for stable transformation. *Journal of Biological Chemistry*, 271(24), 14010-14019.

Du, T., & Zamore, P. D. (2005). microPrimer: the biogenesis and function of microRNA. *Development*, 132(21), 4645-4652.

Dubey, J. P. (1998). Advances in the life cycle of *Toxoplasma gondii*. *International Journal for Parasitology*, 28(7), 1019-1024.

Dubey, J. P., Lindsay, D. S., & Speer, C. A. (1998). Structures of *Toxoplasma gondii* tachyzoites, bradyzoites, and sporozoites and biology and development of tissue cysts. *Clinical Microbiology Reviews*, 11(2), 267-299.

Dubey, J. P. (2006). Comparative infectivity of oocysts and bradyzoites of *Toxoplasma gondii* for intermediate (mice) and definitive (cats) hosts. *Veterinary Parasitology*, 140(1), 69-75.

Dubey, J. P. (2009). History of the discovery of the life cycle of *Toxoplasma gondii*. *International Journal for Parasitology*, 39(8), 877-882.

Dubey, J. P. (2016). Toxoplasmosis of animals and humans. CRC press.

Dzierszinski, F., Mortuaire, M., Dendouga, N., Popescu, O., & Tomavo, S. (2001). Differential expression of two plant-like enolases with distinct enzymatic and antigenic properties during stage conversion of the protozoan parasite *Toxoplasma gondii*. *Journal of Molecular Biology*, 309(5), 1017-1027.

Echeverria, P. C., Figueras, M. J., Vogler, M., Kriehuber, T., de Miguel, N., Deng, B., & Buchner, J. (2010). The Hsp90 co-chaperone p23 of *Toxoplasma gondii*: Identification, functional analysis and dynamic interactome determination. *Molecular and Biochemical Parasitology*, 172(2), 129-140.

Ekimler, S., & Sahin, K. (2014). Computational methods for microRNA target prediction. *Genes*, 5(3), 671-683.

Ender, C., & Meister, G. (2010). Argonaute proteins at a glance. *Journal of Cell Science*, 123(11), 1819-1823.

Engels, B. M., & Hutvagner, G. (2006). Principles and effects of microRNA-mediated post-transcriptional gene regulation. *Oncogene*, 25(46), 6163-6169.

Erickson, S. L., & Lykke-Andersen, J. (2011). Cytoplasmic mRNP granules at a glance. *Journal of Cell Science*, 124(3), 293-297.

Eulalio, A., Huntzinger, E., & Izaurralde, E. (2008). Getting to the root of miRNA-mediated gene silencing. *Cell*, 132(1), 9-14.

Ferguson, D. J. P. (2004). Use of molecular and ultrastructural markers to evaluate stage conversion of *Toxoplasma gondii* in both the intermediate and definitive host. *International Journal for Parasitology*, 34(3), 347-360.

Fire, A., Xu, S., Montgomery, M. K., Kostas, S. A., Driver, S. E., & Mello, C. C. (1998). Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature*, 391(6669), 806-811.

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.

Friedman, R. C., Farh, K. K. H., Burge, C. B., & Bartel, D. P. (2009). Most mammalian mRNAs are conserved targets of microRNAs. *Genome Research*, 19(1), 92-105.

Furuichi, Y., & Shatkin, A. J. (1989). Characterization of cap structures. *Methods in enzymology*, 180, 164-176.

Garcia-Ruiz, H., Carbonell, A., Hoyer, J. S., Fahlgren, N., Gilbert, K. B., Takeda, A., & Baladejo, M. T. M. (2015). Roles and programming of *Arabidopsis* ARGONAUTE proteins during Turnip mosaic virus infection. *PLoS Pathogens*, 11(3), e1004755.

Ghildiyal, M., & Zamore, P. D. (2009). Small silencing RNAs: an expanding universe. *Nature Reviews Genetics*, 10(2), 94-108.

Gregory, R. I., Yan, K. P., Amuthan, G., Chendrimada, T., Doratotaj, B., Cooch, N., & Shiekhattar, R. (2004). The Microprocessor complex mediates the genesis of microRNAs. *Nature*, 432(7014), 235-240.

Grishok, A., Pasquinelli, A. E., Conte, D., Li, N., Parrish, S., Ha, I., & Mello, C. C. (2001). Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control *C. elegans* developmental timing. *Cell*, 106(1), 23-34.

Gruber, A. R., Martin, G., Keller, W., & Zavolan, M. (2014). Means to an end: mechanisms of alternative polyadenylation of messenger RNA precursors. *Wiley Interdisciplinary Reviews: RNA*, 5(2), 183-196.

Gunawardane, L. S., Saito, K., Nishida, K. M., Miyoshi, K., Kawamura, Y., Nagami, T., & Siomi, M. C. (2007). A slicer-mediated mechanism for repeat-associated siRNA 5'end formation in *Drosophila*. *Science*, 315(5818), 1587-1590.

Haley, B., & Zamore, P. D. (2004). Kinetic analysis of the RNAi enzyme complex. *Nature Structural & Molecular Biology*, 11(7), 599-606.

Hall, N., Karras, M., Raine, J. D., Carlton, J. M., Kooij, T. W., Berriman, M., & James, K. (2005). A comprehensive survey of the *Plasmodium* life cycle by genomic, transcriptomic, and proteomic analyses. *Science*, 307(5706), 82-86.

Halbeisen, R. E., Galgano, A., Scherrer, T., & Gerber, A. P. (2008). Post-transcriptional gene regulation: from genome-wide studies to principles. *Cellular and Molecular Life Sciences*, 65(5), 798-813.

Hammond SM. (2005). Dicing and slicing: The core machinery of the RNA interference pathway. FEBS Lett 579:5822–5829.

Han, J., Lee, Y., Yeom, K. H., Kim, Y. K., Jin, H., & Kim, V. N. (2004). The Drosha-DGCR8 complex in primary microRNA processing. *Genes & development*, 18(24), 3016-3027.

Hassan, M. A., Melo, M. B., Haas, B., Jensen, K. D., & Saeij, J. P. (2012). De novo reconstruction of the *Toxoplasma gondii* transcriptome improves on the current genome annotation and reveals alternatively spliced transcripts and putative long non-coding RNAs. *BMC Genomics*, 13(1), 1.

Hoffmann, S., Batz, M. B., & Morris Jr, J. G. (2012). Annual cost of illness and quality-adjusted life year losses in the United States due to 14 foodborne pathogens. *Journal of Food Protection*, 75(7), 1292-1302.

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

Huynh, M. H., & Carruthers, V. B. (2009). Tagging of endogenous genes in a *Toxoplasma gondii* strain lacking Ku80. *Eukaryotic Cell*, 8(4), 530-539.

Hunter, C. A., & Sibley, L. D. (2012). Modulation of innate immunity by *Toxoplasma gondii* virulence effectors. *Nature Reviews Microbiology*, 10(11), 766-778.

Hutvágner, G., McLachlan, J., Pasquinelli, A. E., Bálint, É., Tuschl, T., & Zamore, P. D. (2001). A cellular function for the RNA-interference enzyme Dicer in the maturation of the *let-7* small temporal RNA. *Science*, 293(5531), 834-838.

Hutvagner, G., & Simard, M. J. (2008). Argonaute proteins: key players in RNA silencing. *Nature Reviews Molecular Cell Biology*, 9(1), 22-32.

Irvine, D. V., Zaratiegui, M., Tolia, N. H., Goto, D. B., Chitwood, D. H., Vaughn, M. W., & Martienssen, R. A. (2006). Argonaute slicing is required for heterochromatic silencing and spreading. *Science*, 313(5790), 1134-1137.

Iwasaki, S., & Tomari, Y. (2009). Argonaute-mediated translational repression (and activation). Fly, 3(3), 205-208.

Kawahara, Y., Zinshteyn, B., Sethupathy, P., Iizasa, H., Hatzigeorgiou, A. G., & Nishikura, K. (2007). Redirection of silencing targets by adenosine-to-inosine editing of miRNAs. *Science*, 315(5815), 1137-1140.

Ketting, R. F., Fischer, S. E., Bernstein, E., Sijen, T., Hannon, G. J., & Plasterk, R. H. (2001). Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in *C. elegans*. *Genes & Development*, 15(20), 2654-2659.

Khvorova, A., Reynolds, A., & Jayasena, S. D. (2003). Functional siRNAs and miRNAs exhibit strand bias. *Cell*, 115(2), 209-216.

Kim, K., & Weiss, L. M. (2004). *Toxoplasma gondii*: the model apicomplexan. *International Journal for Parasitology*, 34(3), 423-432.

Kiriakidou, M., Tan, G. S., Lamprinaki, S., De Planell-Saguer, M., Nelson, P. T., & Mourelatos, Z. (2007). An mRNA m 7 G cap binding-like motif within human Ago2 represses translation. *Cell*, 129(6), 1141-1151.

Kirkman, L. A., Weiss, L. M., & Kim, K. (2001). Cyclic nucleotide signaling in *Toxoplasma gondii* bradyzoite differentiation. *Infection and Immunity*, 69(1), 148-153.

Köhler, A., & Hurt, E. (2007). Exporting RNA from the nucleus to the cytoplasm. *Nature Reviews Molecular Cell Biology*, 8(10), 761-773.

Lagos-Quintana, M., Rauhut, R., Meyer, J., Borkhardt, A., & Tuschl, T. (2003). New microRNAs from mouse and human. *RNA*, 9(2), 175-179.

Landthaler, M., Yalcin, A., & Tuschl, T. (2004). The human DiGeorge syndrome critical region gene 8 and its *D. melanogaster* homolog are required for miRNA biogenesis. *Current Biology*, 14(23), 2162-2167.

Lewis, B. P., Green, R. E., & Brenner, S. E. (2003). Evidence for the widespread coupling of alternative splicing and nonsense-mediated mRNA decay in humans. *Proceedings of the National Academy of Sciences*, 100(1), 189-192.

Lewin, B. (2004). *Genes* VIII (pp. 119-122). Upper Saddle River, NJ: Pearson Prentice Hall.

Lee, R. C., Feinbaum, R. L. & Ambros, V. (1993). The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to lin-14. *Cell* 75, 843-854.

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.

Liu, D., Brockman, J. M., Dass, B., Hutchins, L. N., Singh, P., McCarrey, J. R., & Graber, J. H. (2007). Systematic variation in mRNA 3′-processing signals during mouse spermatogenesis. *Nucleic Acids Research*, 35(1), 234-246.

Lingel, A., Simon, B., Izaurralde, E., & Sattler, M. (2004). Nucleic acid 3'-end recognition by the Argonaute2 PAZ domain. *Nature Structural & Molecular Biology*, 11(6).

López-Lastra, M., Rivas, A., & Barría, M. I. (2005). Protein synthesis in eukaryotes: the growing biological relevance of cap-independent translation initiation. *Biological Research*, 38(2-3), 121-146.

Matranga, C., Tomari, Y., Shin, C., Bartel, D. P., & Zamore, P. D. (2005). Passenger-strand cleavage facilitates assembly of siRNA into Ago2-containing RNAi enzyme complexes. *Cell*, 123(4), 607-620.

Ma, J. B., Ye, K., & Patel, D. J. (2004). Structural basis for overhang-specific small interfering RNA recognition by the PAZ domain. *Nature*, 429(6989), 318-322.

Ma, J. B., Yuan, Y. R., Meister, G., Pei, Y., Tuschl, T., & Patel, D. J. (2005). Structural basis for 5′-end-specific recognition of guide RNA by the A. fulgidus Piwi protein. *Nature*, 434(7033), 666-670.

Malys, N., & McCarthy, J. E. (2011). Translation initiation: variations in the mechanism can be anticipated. *Cellular and Molecular Life Sciences*, 68(6), 991-1003.

Matlin, A. J., Clark, F., & Smith, C. W. (2005). Understanding alternative splicing: towards a cellular code. *Nature Reviews Molecular Cell Biology*, 6(5), 386-398.

Meister, G. (2013). Argonaute proteins: functional insights and emerging roles. *Nature Reviews Genetics*, 14(7), 447-459.

Moore, M. J. (2005). From birth to death: the complex lives of eukaryotic mRNAs. *Science*, 309(5740), 1514-1518.

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

Ngo, H., Tschudi, C., Gull, K., & Ullu, E. (1998). Double-stranded RNA induces mRNA degradation in *Trypanosoma brucei*. *Proceedings of the National Academy of Sciences*, 95(25), 14687-14692.

Pan, Q., Shai, O., Lee, L. J., Frey, B. J., & Blencowe, B. J. (2008). Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nature Genetics*, 40(12), 1413-1415.

Petsko, G. A., & Ringe, D. (2004). Protein structure and function. New Science Press.

Phillips, T. (2008). The role of methylation in gene expression. *Nature Education*, 1(1), 116.

Proudfoot, N. J., Furger, A., & Dye, M. J. (2002). Integrating mRNA processing with transcription. *Cell*, 108(4), 501-512.

Rand, T. A., Ginalski, K., Grishin, N. V., & Wang, X. (2004). Biochemical identification of Argonaute 2 as the sole protein required for RNA-induced silencing complex activity. *Proceedings of the National Academy of Sciences of the United States of America*, 101(40), 14385-14389.

Reinhart, B. J., Slack, F. J., Basson, M., Pasquinelli, A. E., Bettinger, J. C., Rougvie, A. E., & Ruvkun, G. (2000). The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans. *Nature*, 403(6772), 901-906.

Roos, D. S., Donald, R. G., Morrissette, N. S., & Moulton, A. L. C. (1995). Molecular tools for genetic dissection of the protozoan parasite *Toxoplasma gondii*. *Methods in Cell Biology*, 45, 27-63.

Rommereim, L. M., Triana, M. A. H., Falla, A., Sanders, K. L., Guevara, R. B., Bzik, D. J., & Fox, B. A. (2013). Genetic manipulation in Δku80 strains for functional genomic analysis of *Toxoplasma gondii*. *Journal of Visualized Experiments: JoVE*, (77).

Saeij, J. P., Boyle, J. P., & Boothroyd, J. C. (2005). Differences among the three major strains of *Toxoplasma gondii* and their specific interactions with the infected host. *Trends in Parasitology*, 21(10), 476-481.

Saurabh, S., Vidyarthi, A. S., & Prasad, D. (2014). RNA interference: concept to reality in crop improvement. *Planta*, 239(3), 543-564.

Sandberg, R., Neilson, J. R., Sarma, A., Sharp, P. A., & Burge, C. B. (2008). Proliferating cells express mRNAs with shortened 3'untranslated regions and fewer microRNA target sites. *Science*, 320(5883), 1643-1647.

Schwarz, D. S., Hutvágner, G., Du, T., Xu, Z., Aronin, N., & Zamore, P. D. (2003). Asymmetry in the assembly of the RNAi enzyme complex. *Cell*, 115(2), 199-208.

Shen, Y., Ji, G., Haas, B. J., Wu, X., Zheng, J., Reese, G. J., & Li, Q. Q. (2008). Genome level analysis of rice mRNA 3′-end processing signals and alternative polyadenylation. *Nucleic Acids Research*, 36(9), 3150-3161.

Skariah, S., McIntyre, M. K., & Mordue, D. G. (2010). *Toxoplasma gondii*: determinants of tachyzoite to bradyzoite conversion. *Parasitology Research*, 107(2), 253-260.

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.

Sonenberg, N., & Hinnebusch, A. G. (2009). Regulation of translation initiation in eukaryotes: mechanisms and biological targets. *Cell*, 136(4), 731-745.

Song, J. J., Smith, S. K., Hannon, G. J., & Joshua-Tor, L. (2004). Crystal structure of Argonaute and its implications for RISC slicer activity. *Science*, 305(5689), 1434-1437.

Starr, C., & Starr, L. (2014). Biology: Concepts and applications (9th ed.). New York.

Su, C., Evans, D., Cole, R. H., Kissinger, J. C., Ajioka, J. W., & Sibley, L. D. (2003). Recent expansion of *Toxoplasma* through enhanced oral transmission. *Science*, 299(5605), 414-416.

Suvorova, E. S., & White, M. W. (2014). Transcript maturation in apicomplexan parasites. *Current Opinion in Microbiology*, 20, 82-87.

Tenter, A. M., Heckeroth, A. R., & Weiss, L. M. (2000). *Toxoplasma gondii*: from animals to humans. *International Journal for Parasitology*, 30(12), 1217-1258.

Timothy W. Nilsen. (2007). Mechanisms of microRNA mediated gene regulation in animal cells. *Trends in Genetics* 23:243-249.

Tijsterman, M., Okihara, K. L., Thijssen, K., & Plasterk, R. H. (2002). PPW-1, a PAZ/PIWI protein required for efficient germline RNAi, is defective in a natural isolate of *C. elegans*. *Current Biology*, 12(17), 1535-1540.

Tomita, T., Bzik, D. J., Ma, Y. F., Fox, B. A., Markillie, L. M., Taylor, R. C., & Weiss, L. M. (2013). The *Toxoplasma gondii* cyst wall protein CST1 is critical for cyst wall integrity and promotes bradyzoite persistence. *PLoS Pathogens*, 9(12), e1003823.

Topisirovic, I., Svitkin, Y. V., Sonenberg, N., & Shatkin, A. J. (2011). Cap and cap-binding proteins in the control of gene expression. *Wiley Interdisciplinary Reviews: RNA*, 2(2), 277-298.

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

Vaquero, A., Loyola, A., & Reinberg, D. (2003). The constantly changing face of chromatin. *Science's SAGE KE*, 2003(14), 4.

Wells, S. E., Hillner, P. E., Vale, R. D., & Sachs, A. B. (1998). Circularization of mRNA by eukaryotic translation initiation factors. *Molecular Cell*, 2(1), 135-140.

Wienholds, E., Kloosterman, W. P., Miska, E., Alvarez-Saavedra, E., Berezikov, E., de Bruijn, E., & Plasterk, R. H. (2005). MicroRNA expression in zebrafish embryonic development. *Science*, 309(5732), 310-311.

Wightman, B., Ha, I., & Ruvkun, G. (1993). Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in *C. elegans*. *Cell*, 75(5), 855-862.

Xu, W., San Lucas, A., Wang, Z., & Liu, Y. (2014). Identifying microRNA targets in different gene regions. *BMC Bioinformatics*, 15(7), S4.

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

Yigit, E., Batista, P. J., Bei, Y., Pang, K. M., Chen, C. C. G., Tolia, N. H., & Mello, C. C. (2006). Analysis of the *C. elegans* Argonaute family reveals that distinct Argonautes act sequentially during RNAi. *Cell*, 127(4), 747-757.

Yi, R., Qin, Y., Macara, I. G., & Cullen, B. R. (2003). Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. *Genes & Development*, 17(24), 3011-3016.

Yuan, Y. R., Pei, Y., Ma, J. B., Kuryavyi, V., Zhadina, M., Meister, G., & Patel, D. J. (2005). Crystal structure of *A. aeolicus* argonaute, a site-specific DNA-guided endoribonuclease, provides insights into RISC-mediated mRNA cleavage. *Molecular Cell*, 19(3), 405-419.

Zhang, Y. W., Halonen, S. K., Ma, Y. F., Wittner, M., & Weiss, L. M. (2001). Initial characterization of CST1, a *Toxoplasma gondii* cyst wall glycoprotein. *Infection and Immunity*, 69(1), 501-507.

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.

Zhang, B., Wang, Q., & Pan, X. (2007). MicroRNAs and their regulatory roles in animals and plants. *Journal of Cellular Physiology*, 210(2), 279-289.

Appendix A

a) Primer sets used for TgAgo transcript analysis

Primer labels correspond to those in Figure 3.2: A and B	**Primer sets (5'→3')**
a : 860 bp	TgArg5'end: CCGCAGTATTTCGCAAATG
	TgArg3'end: GCTTCGCCACTGAGGTGAG
b : 816 bp	5'TgAgoNsiI: TTTATGCATATCATGCCTCCC CG
	GSP-Argo: CGCACACACAATGGGCGCTG
c : 298 bp	AGOexon4_fw: CCGTTTCGATCTGCA
	Ext_Ago5RACE: CGG GGC TGA ATG TCC CTG
d : 570 bp	Barik_TgAgo5': ATGGAGGGAT ACCAAATGTT CC
	AGOexon5_rev: GTATGCCGGGCGATCC
e : 738 bp	Barik_TgAgo5': ATGGAGGGAT ACCAAATGTT CC
	Ext_Ago5RACE: CGG GGC TGA ATG TCC CTG
	AGOexon5_rev: GTATGCCGGGCGATCC
f : 869 bp	RGG5Ago: GGGAATTCAATGAACGGAGGAGGCAGA
	AGO Exon 4_Rv: GACAAGTCCACCATTCG
g : 393 bp	RGG5Ago: GGGAATTCAATGAACGGAGGAGGCAGA
	AGO Exon 2_Rv: CGCGAGATGGCCCAAGG
h : 1179 bp	RGG5Ago: GGGAATTCAATGAACGGAGGAGGCAGA
	AGOexon5_rev: GTATGCCGGGCGATCC

Table: PCR Primer sets for analyzing the TgAgo transcript

Appendix B

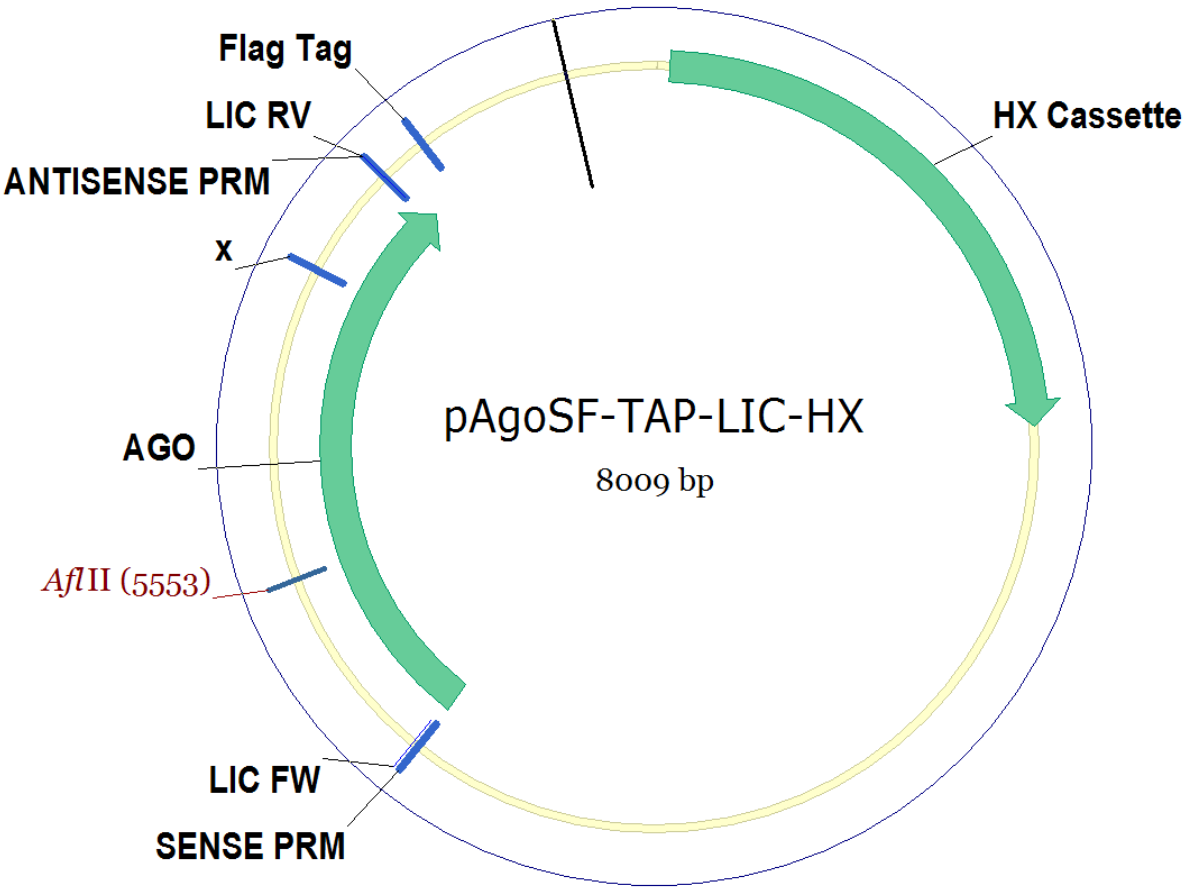


Figure B1: Illustration of pAGO-SF-TAP-LIC-HX transgenic plasmid. 3′ end genomic sequence fragment of argonaute gene (TGGT1_310160) was acquired by PCR analysis from *T. gondii* Type I strain. The location of strep and FLAG tags are shown. The HXGPRT cassette is also indicated which was used for transgenic parasite selection. Moreover, single cutting site of restriction enzyme is also marked, which was used to linearize the plasmid.

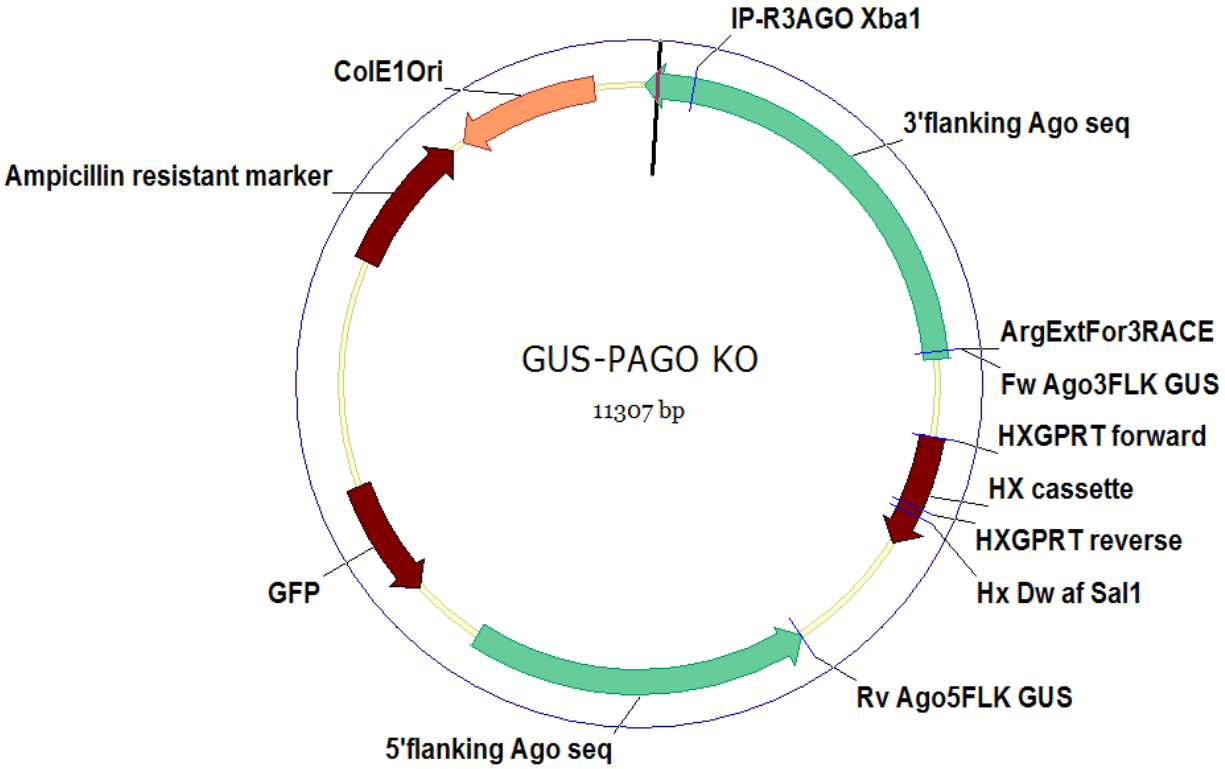


Figure B2: Illustration of pAGOKO transgenic plasmid. 5' end and 3' end flanking genomic sequence fragment of argonaute gene from *T. gondii* Type I strain (shown in green). The location of the HXGPRT cassette (backward orientation) is indicated, which was used for knock out of argonaute genomic sequence. The location of coding sequence for the GFP tag is also shown.

pGEM®-T Easy Vector Map and Sequence Reference Points

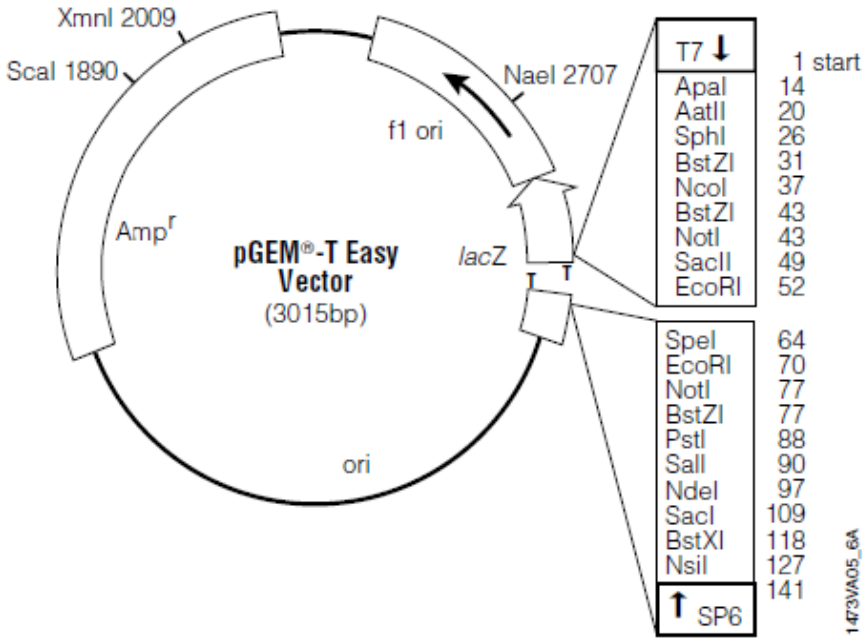


Figure B3: Illustration of pGEM-T easy cloning vector. Linearized vectors with a single 3´-terminal thymidine at both ends (shown in figure). The vector contains T7 and SP6 RNA polymerase promoters flanking a multiple cloning region within the α-peptide coding region of the enzyme β-galactosidase. The location of the coding sequence for ampicillin resistance is also shown, used for drug selection.

VITA AUCTORIS

Name	Md Amran Hossain
Place of Birth	Dhaka, Bangladesh
Year of Birth	1985
Education	Monipur High School, Dhaka, Bangladesh, 1998-2000
	Dhaka College, Dhaka, Bangladesh, 2000-2002
	University of Dhaka, Dhaka, Bangladesh, 2002-2006 (Exam held: 2009) BSc (Honors) in Laboratory Science
	University of Windsor, Windsor, Ontario 2015-Present MSc Biochemistry