

GCN5-B IS A NOVEL NUCLEAR HISTONE ACETYLTRANSFERASE
THAT IS CRUCIAL FOR VIABILITY IN THE PROTOZOAN PARASITE
TOXOPLASMA GONDII

Stacy E. Dixon

Submitted to the faculty of the University Graduate School
in partial fulfillment of the requirements
for the degree
Doctor of Philosophy
in the Department of Pharmacology and Toxicology,
Indiana University

December 2010

Accepted by the Faculty of Indiana University, in partial
fulfillment of the requirements for the degree of Doctor of Philosophy.

William J. Sullivan, Jr., Ph.D. – Chairman

Rebecca J. Chan, M.D./Ph.D

Barbara A. Hocevar, Ph.D.

Doctoral Committee

Sherry F. Queener, Ph.D.

June 17, 2010

Jian-Ting Zhang, Ph.D.

Acknowledgements

I first owe an enormous debt of gratitude to my mentor, Dr. Bill Sullivan, and Dr. Sherry Queener. In the summer of 2002, they allowed me to work in the lab as an undergraduate research fellow. The guidance and mentorship I received that summer changed my course and inspired me to pursue a research-related career. Over the last eight years, both Dr. Bill Sullivan and Dr. Sherry Queener have been incredibly supportive, have provided me with numerous opportunities, and are excellent instructors. I would not have reached this point without their assistance.

I must also thank the other members of my thesis committee: Dr. J.T. Zhang, Dr. Barbara Hocevar, and Dr. Rebecca Chan. Their advice, suggestions, and guidance have been very instrumental in the completion of my dissertation project. Likewise, I must extend my gratitude to all the faculty and staff of the Pharmacology and Toxicology department. Our department is comprised of friendly and helpful personnel, which has enhanced my experience as a student. I would especially like to thank Dr. Michael Vasko, our department chairman, for his support, advice, and many letters of recommendation. I would also like to thank Amy Lawson, Lisa King, and Miriam Barr for administrative assistance.

I also like to thank all of the collaborators that have assisted with portions of my dissertation project. This includes Drs. Keith Dunker and Vladimir Uversky (Indiana University School of Medicine) for their work on the intrinsic disorder of TgGCN5-B, Dr. Andy Tao (Purdue University) and his lab for their assistance with mass spectrometry, and Dr. Ali Hakimi (National Centre for Scientific Research in Grenoble, France) and his lab for the purification of the TgGCN5-B complex. In addition, I would like to thank Dr. Ron Wek (Indiana University School of Medicine), Dr. Kami Kim (Albert Einstein College of Medicine), and Dr. Michael White (University of South Florida) and Dr. Cynthia Hingtgen (Indiana University School of Medicine) for their insight and advice. I also wish to express gratitude to the PhRMA foundation for funding my research as well as the NIAID. Thank you to the American Society for Microbiology for travel assistance to their general meeting and for my attendance of the Kadner Institute.

I would not have been able to complete this work without excellent training from my mentor, Dr. Bill Sullivan, from Dr. Micah Bhatti, and from Pam Torkelson. Thank you for your instruction. I also worked with many summer and rotational students within the Sullivan lab and am grateful for the assistance each provided to my dissertation project.

I would also like to thank both the past and present members of the Sullivan lab especially Dr. Arunasalam Naguleswaran (Nagul) for all your advice.

Lastly, I sincerely thank my family for their love and support throughout this journey. My parents, Thom and Jerilynn Dixon, and my grandparents, Tom and Helen Dixon, have been influential in my life and have always been my pillar of support. To my mom especially, thank you for listening and providing encouragement. To Chad, thank you for joining me on this journey and for all the joy you have brought into my life.

ABSTRACT

Stacy E. Dixon

GCN5-B IS A NOVEL NUCLEAR HISTONE ACETYLTRANSFERASE THAT IS CRUCIAL FOR VIABILITY IN THE PROTOZOAN PARASITE *TOXOPLASMA GONDII*

Infection with the single-celled parasite *Toxoplasma gondii* (phylum Apicomplexa) is usually benign in normal healthy individuals, but can cause congenital birth defects, ocular disease, and also life-threatening infection in immunocompromised patients. Acute infection caused by tachyzoites is controlled by a healthy immune response, but the parasite differentiates into a latent cyst form (bradyzoite) leading to permanent infection and chronic disease. Current therapies are effective only against tachyzoites, are highly toxic to the patient, and do not eradicate the encysted bradyzoites, thus highlighting the need for novel therapeutics. Inhibitors of histone deacetylases have been shown to reduce parasite viability *in vitro* demonstrating that chromatin remodeling enzymes, key mediators in epigenetic regulation, might serve as potential drug targets. Furthermore, epigenetic regulation has been shown to contribute to gene expression and differentiation in *Toxoplasma*. This dissertation focused on investigating the physiological role of a *Toxoplasma* GCN5-family histone acetyltransferase (HAT), termed TgGCN5-B. It was hypothesized that TgGCN5-B is an essential HAT that resides within a unique, multi-subunit complex in the parasite nucleus. Studies of TgGCN5-B have revealed that this HAT possesses a unique nuclear localization signal (₃₁₁RPAENKKRGR₃₂₀) that is both necessary and sufficient to translocate the protein to the parasite nucleus. Although no other protein motifs have been identified in the N-terminal extension of TgGCN5-B, it is likely that this extension plays a role in protein-protein interactions. All GCN5 homologues function within large

multi-subunit complexes, many being conserved among species, but bioinformatic analysis of the *Toxoplasma* genome revealed a lack of many of these conserved components. Biochemical studies identified several potential TgGCN5-B associating proteins, including several novel apicomplexan transcription factors. Preliminary evidence suggested that TgGCN5-B was essential for tachyzoites; therefore, a dominant-negative approach was utilized to examine the role of TgGCN5-B in the physiology of *Toxoplasma*. When catalytically inactive TgGCN5-B protein was over-expressed in the parasites, there was a significant decrease in tachyzoite growth and viability, with initial observations suggesting defects in nuclear division and daughter cell budding. These results demonstrate that TgGCN5-B is important for tachyzoite development and indicate that therapeutic targeting of this HAT could be a novel approach to treat toxoplasmosis.

William J. Sullivan, Jr., Ph.D. – Chairman

Table of Contents

List of Tables	xi
List of Figures.....	xii
List of Abbreviations	xiv
Chapter 1: Introduction and Literature Review.....	1
I. <i>Toxoplasma</i> is a successful pathogen.....	1
A. <i>Toxoplasma</i> is a model apicomplexan parasite	1
B. The life cycle of <i>Toxoplasma</i>	3
C. Clinical toxoplasmosis	6
D. Treatment for toxoplasmosis	9
II. Epigenetics and chromatin remodeling.....	10
A. Definition of epigenetics.....	10
B. Histones and their code	11
III. GCN5 is a HAT involved in transcriptional activation	13
A. Definition, classification and summary of HATs	13
B. GCN5 is a conserved HAT	15
C. GCN5 functions within multi-subunit complexes	15
D. The role of GCN5 in cellular physiology.....	18
E. Nuclear import of HATs	19
IV. Gene expression and epigenetic modulation in <i>Toxoplasma</i>	20
A. Regulation of gene expression	20
B. The histones of <i>Toxoplasma</i> and their modifying enzymes.....	22
C. Two GCN5 homologues in <i>Toxoplasma</i>	25
V. Hypothesis and aims.....	28
Chapter 2: Materials and Methods.....	29
I. Tissue culture and parasite techniques	29
A. Host cell and parasite culture.....	29
B. Freezing and thawing of host cells and parasites.....	31
C. Parasite transfection.....	32
D. Cloning by limiting dilution	36
E. Harvesting parasites and generating lysates	37
F. Generating parasite nuclear-enriched lysates	38
II. Molecular biology techniques	39

A. General PCR protocol.....	39
B. Ligation, bacterial transformation, and plasmid preparation	43
C. Ligation independent cloning	47
D. Inverse PCR.....	48
E. Genomic PCR from parasites	49
F. Site-directed mutagenesis	50
III. Biochemical techniques.....	51
A. Immunofluorescence assay	51
B. Immunoprecipitation	54
C. SDS-PAGE and Western blotting	55
D. HAT assays.....	56
E. Bacterial inductions	57
F. Purification of MBP fusion proteins	58
G. Affinity chromatography.....	59
IV. <i>Toxoplasma</i> growth assays.....	60
A. Plaque assay	61
B. B1 gene detection assay to monitor <i>Toxoplasma</i> growth	61
C. Doubling assay	63
Chapter 3: Results.....	64
I. Aim 1: Determine how TgGCN5-B enters the parasite nucleus	64
A. TgGCN5-B requires its N-terminus to enter the parasite nucleus.....	64
B. TgGCN5-B contains a unique NLS within its N-terminus	65
C. The TgGCN5- B NLS is sufficient to localize β -gal to the nucleus.....	69
D. The NLS of TgGCN5-B has predictive value	71
E. Protein with a predicted analogous NLS localizes to nucleus	73
F. The nuclear chaperone utilized by TgGCN5-B remains to be resolved	75
II. Aim 2: Identify the proteins associating with TgGCN5-B.....	78
A. Bioinformatics reveal a dearth of common GCN5 complex members	78
B. TgGCN5-B is disordered in its N-terminal extension	79
C. Affinity chromatography to identify TgGCN5-B associating proteins	82
D. Expression of ectopic TgGCN5-B in <i>Toxoplasma</i>	88
E. Co-immunoprecipitation reveals novel TgGCN5-B associating proteins	91
F. Confirmation of TgGCN5-B associating proteins.....	95
III. Aim 3: Determine the role of TgGCN5-B in parasite physiology.....	99

A. Preliminary evidence suggests that TgGCN5-B is essential	99
B. Generation of a regulatable dominant-negative TgGCN5-B.....	100
C. Expression of dominant-negative TgGCN5-B reduces parasite viability	104
Chapter 4: Discussion and Future Studies.....	114
I. TgGCN5-B harbors a novel NLS within its N-terminal extension	114
A. Summary of Aim 1 results.....	114
B. The classification of the TgGCN5-B NLS.....	114
C. Utility of the TgGCN5-B NLS	116
D. Future studies	118
II. TgGCN5-B associates with novel Apicomplexa transcription factors	119
A. Summary of Aim 2 results.....	119
B. Comparison of affinity chromatography and co-IP techniques	119
C. The pursuit of Apicomplexa AP2s.....	121
D. The roles of the N-terminal extensions in Apicomplexa GCN5s.....	122
E. Future studies.....	123
III. TgGCN5-B dominant-negative phenotype decreases parasite viability.....	124
A. Summary of Aim 3 results.....	124
B. GCN5 regulates normal progression through the cell cycle	124
C. TgGCN5-B is a novel therapeutic candidate.....	126
D. Future studies	127
Chapter 5: Summary	129
Appendices	130
Appendix A: Parasite counting	130
Appendix B: Details regarding mass spectrometry for affinity chromatography ..	132
Appendix C: Additional affinity chromatography data.....	122
Appendix D: Monitoring parasite cell cycle stages via IFA.....	140
References.....	141
Curriculum Vitae	

List of Tables

Table I: Selection agents for <i>Toxoplasma</i>	34
Table II: <i>Toxoplasma</i> transfection vectors	35
Table III: Description of primers for PCR	41
Table IV: Plasmids.....	46
Table V: Antibodies used in IFA and Western blotting	53
Table VI: Predictive value of the TgGCN5-B NLS	73
Table VII: SAGA complex homologues in <i>Toxoplasma</i>	79
Table VIII: Proteins identified in both replicates of affinity chromatography.....	87
Table IX: Proteins co-immunoprecipitated with _{HA-MYC} GCN5-B _{FLAG}	94
Table A1: Proteins associating with MBP-GCN5-B from affinity chromatography #1	133
Table A2: Proteins associating with MBP-GCN5-B from affinity chromatography #2.....	135

List of Figures

Figure 1: Life cycle and transmission of <i>Toxoplasma</i>	5
Figure 2: Diverse roles of epigenetics in <i>Toxoplasma</i> biology.....	24
Figure 3: Comparison of apicomplexan GCN5s to other eukaryotic GCN5 homologues	27
Figure 4: Summary of guanine biosynthesis in <i>Toxoplasma</i> and the manipulation of the pathway for selection purposes	33
Figure 5: Inverse PCR to remove the nucleotides encoding the NLS of TgGCN5-B.....	49
Figure 6: Preliminary mapping of TgGCN5-B NLS.....	65
Figure 7: Mutation constructs of TgGCN5-B elucidate the complete NLS	67
Figure 8: Mutation constructs of TgGCN5-B reveal that specific residues in the upstream sequence are key to nuclear localization	68
Figure 9: The GCN5-B NLS is sufficient to localize <i>E. coli</i> β -gal to the parasite nucleus.....	70
Figure 10: AT-hook 054600 contains a predicted analogous NLS and localizes to the parasite nucleus	75
Figure 11: TgImp- α localizes to the parasite nucleus.....	77
Figure 12: TgGCN5-B contains several regions of intrinsic disorder within the N-terminus	81
Figure 13: Pictorial diagram representing affinity chromatography procedure.....	83
Figure 14: Purified recombinant MBP-GCN5-B is catalytically active	84
Figure 15: Purification of proteins associating with MBP-GCN5-B using affinity chromatography.....	85
Figure 16: Characterization of parasites expressing $_{HA-MYC}GCN5-B_{FLAG}$	89
Figure 17: Schematic of co-IP protocol used for detecting $_{HA-MYC}GCN5-B_{FLAG}$ associating proteins	92
Figure 18: Analysis from co-IP of $_{HA-MYC}GCN5-B_{FLAG}$ and associating proteins	93
Figure 19: Two TgGCN5-B associating proteins localize to the parasite nucleus..	97
Figure 20: Illustration of conditional protein expression through the destabilization domain.....	101
Figure 21: Addition of Shield-1 stabilizes expression of ectopic GCN5-B proteins	103
Figure 22: Addition of Shield-1 allows for expression of ectopic GCN5-B proteins	104

Figure 23: The presence of Shield-1 does not affect the growth of parental parasites	105
Figure 24: GCN5-B DN parasites cannot proliferate in the presence of Shield-1 ..	107
Figure 25: Normal progression through parasite cell cycle and division	110
Figure 26: GCN5-B DN parasites display abnormal nuclear division and daughter cell formation when treated with Shield-1	112
Figure A1: Counting grid of hemocytometer.....	130

List of Abbreviations

6TX	6-thioxanthine
AA	anacardic acid
ADA	alteration/deficiency in activation protein
AIDS	acquired immune deficiency syndrome
AP2	Apetala2
ATAC	ADA two a containing
ATc	anhydrotetracycline
ATM	ataxia telangiectasia mutated
β -Me	beta-mercaptoethanol
β -gal	<i>Escherichia coli</i> beta-galactosidase
Br	bromodomain
BSA	bovine serum albumin
CAM	chloramphenicol
CARM1	cofactor-associated arginine [R] methyltransferase 1
CAT	chloramphenicol acetyltransferase
CBP	CREB binding protein
cDNA	complementary deoxyribonucleic acid
ChIP	chromatin immunoprecipitation
ChIP-chip	chromatin immunoprecipitation coupled to microarray
CNS	central nervous system
Co-IP	co-immunoprecipitation
CREB	cAMP responsive element binding protein
DAPI	4',6'-diamino-2-phenylindole
DD	destabilization domain
ddH ₂ O	double distilled H ₂ O
DHFR	dihydrofolic acid reductase
DHFR-TS	dihydrofolic acid reductase-thymidylate synthase
DHPS	dihydropteroate synthase
DMEM	Dulbecco's Modified Eagle Medium
DMSO	dimethyl sulfoxide
DN	dominant-negative
DNA	deoxyribonucleic acid

Dub	deubiquitinase
ECL	enhanced luminol-based chemiluminescent
EST	expressed sequence tag
FITC	fluorescein isothiocyanate
GCN4	general control non-derepressible 4
GCN5	general control non-derepressible 5
gDNA	genomic DNA
GDP	guanosine diphosphate
GNAT	GCN5-related N-acetyltransferase
GST	glutathione-S-transferase
GTP	guanosine triphosphate
HA	hemagglutinin
HAART	highly active anti-retroviral therapy
HAT	histone acetyltransferase
HDAC	histone deacetylase
HFF	human foreskin fibroblast
HIV	human immunodeficiency virus
hN	host cell nucleus
HRP	horseradish peroxidase
Hsp	heat shock protein
hTERT	human telomerase reverse transcriptase
HXGPRT	hypoxanthine-xanthine-guanine phosphoribosyltransferase
ICM1	inner membrane complex 1
IFA	immunofluorescent assay
IMPDH	inosine-monophosphate dehydrogenase
IP	immunoprecipitation
IPTG	isopropyl β -D-1-thiogalactopyranoside
Imp- α	importin-alpha
Imp- β	importin-beta
kDa	kiloDalton
KO	knockout
LC-MS/MS	liquid chromatography coupled with tandem mass spectrometry
LIC	ligation independent cloning
MBP	maltose binding protein

MES	2-Morpholinoethanesulfonic acid
MOPS	3-(N-Morpholino) propanesulfonic acid
MoRF	molecular recognition feature
MPA	mycophenolic acid
MRA	<i>Mycoplasma</i> removal agent
mRNA	messenger ribonucleic acid
MYST	MOZ, Ybf2/Sas3, Sas2, and Tip60 members
NAD	nicotinamide adenine dinucleotide
NER	nucleotide excise repair
NHEJ	nonhomologous end-joining
NLS	nuclear localization signal
NPC	nuclear pore complex
OD ₆₀₀	optical density at 600 nm
PAGE	polyacrylamide gel electrophoresis
PBM	protein binding microarray
PCAF	p300/CBP associating factor
PCNA	proliferating cell nuclear antigen
PCR	polymerase chain reaction
PFA	paraformaldehyde
PHD	plant homeo domain
PIC	pre-initiation complex
PONDR	Predictor of Natural Disordered Regions
PYR	pyrimethamine
Ran	ras-related nuclear protein
RanGAP	Ran GTPase activating factor
RanGEF	Ran guanine nucleotide exchange factor
RCC1	regulator of chromosome condensation 1
RE	restriction enzyme or endonuclease
RNA	ribonucleic acid
SAGA	Spt-ADA-GCN5 acetyltransferase complex
SAGE	serial analysis of gene expression
SDS	sodium dodecyl sulfate
SET	Su(var)3-9, Enhancer-of-zeste, Trithorax
SGD	<i>Saccharomyces</i> Genome Database

SGF	SAGA associated factor
Shld	Shield-1
Spt	suppressors of Ty
TAF _{II}	TBP associated factors
TBP	TATA-binding protein
TFTC	TATA-binding protein-free TAF-containing complex
TgN	<i>Toxoplasma</i> nuclei
Tra1	transplantability associated gene 1
tRNA	transfer ribonucleic acid
TRRAP	transformation/transcription domain–associated protein
UAS	upstream activating sequence
UTR	untranslated region
UV	ultraviolet
WB	Western blot
wt	wild-type

Chapter 1: Introduction and Literature Review

The study of epigenetics has revolutionized the understanding of gene expression and regulation in a multitude of organisms including the medically relevant pathogen *Toxoplasma gondii*. Epigenetics, defined as changes to a genetic locus not encoded by the underlying DNA sequence, encompasses a number of coordinated cellular phenomena that impact the degree of gene expression [1]. These alterations allow for the same genome to give rise to a variety of phenotypes, which is particularly important for pathogens such as *Toxoplasma* that have complex life cycles involving multiple stages. The pharmacological exploitation of epigenetic mechanisms has arisen as a promising avenue for new drug discoveries towards several pathogens. This thesis will describe an enzyme involved in the epigenetic regulation of gene expression in the parasite *Toxoplasma gondii* and the potential of that enzyme as a therapeutic target.

I. *Toxoplasma* is a successful pathogen

A. *Toxoplasma* is a model apicomplexan parasite

Toxoplasma is one among several human and animal pathogens in the phylum Apicomplexa. The apicomplexans are a large and diverse group of unicellular protozoan parasites with an expansive environmental distribution. The majority of the parasites from this phylum are obligate intracellular parasites with complex life cycles that involve both asexually reproducing forms and sexual stages. The name of the phylum is derived from a group of unique organelles at the apical end that these parasites possess. The apicoplast is a multiple-membrane bound organelle acquired through a secondary endosymbiotic event and thought to be derived from chloroplasts [2,3]. Although no longer capable of photosynthesis, this organelle is essential to the apicomplexans, and it is believed that its primary function is the synthesis of fatty acids, isoprenoids, and heme [2,3]. The apical complex consists of multiple secretory organelles such as micronemes and rhoptries that are necessary for interaction and invasion of host cells by apicomplexans [4].

The apicomplexan parasites that commonly infect humans cause a variety of diseases and affect diverse populations. The parasites *Cystoisospora*, *Cyclospora*, and *Sarcosystis* are all intestinal pathogens that rarely cause disease in the United States. *Cryptosporidium* is another apicomplexan intestinal pathogen that usually infects

immunocompromised individuals such as AIDS patients. However, *Cryptosporidium* had its fifteen minutes of parasitic fame in 1993 when it contaminated the water supply of Milwaukee, Wisconsin, causing diarrheal disease in over 400,000 people [5]. *Babesia* is a parasite that is spread by certain species of ticks throughout the northeast United States. Infection is usually asymptomatic although in some cases, babesiosis can result in a hemolytic anemia that requires treatment. By far, the most infamous apicomplexan parasite is *Plasmodium*, the causative agent of malaria. This pathogen is also spread via an insect, but in this case it is the mosquito. About one-half of the world's population is at risk for infection with *Plasmodium*, and each year there are approximately 250 million cases of malaria worldwide, resulting in nearly one million deaths [6]. The countries prone to endemic malaria include many of the world's poorest countries, and those most vulnerable to death from the disease are children. Unfortunately, attempts to control the spread of malaria are complicated by the ease with which the parasite has become resistant to common treatment options. This situation has resulted in a substantial worldwide health economic burden and necessitates the development for new pharmacological agents to combat this pathogenic menace. Like its cousin *Plasmodium*, *Toxoplasma* is also present at high concentrations worldwide with approximately one-third of the world's population being infected [7]. Although this infection is life-long, most affected individuals are asymptomatic. *Toxoplasma* is mostly known as an opportunistic pathogen causing disease in immunocompromised individuals, although the parasite is also associated with congenital infections as well as ocular disease. The human pathogenic members of the phylum Apicomplexa represent a diverse array of clinical diseases and account for a significant worldwide health burden.

In addition, several apicomplexans are animal pathogens. *Toxoplasma* can infect virtually any warm-blooded vertebrate, making it a pathogen to livestock and domestic animals as well as humans. In particular, infection with *Toxoplasma* has a significant impact on sheep and is a common cause of abortion in ewes, necessitating a vaccine to protect herds [8]. *Neospora* and *Theileria* are both pathogens of cattle while *Eimeria* is a common infective agent of poultry. The prominent global prevalence of the apicomplexan parasites and their ability to infect a multitude of hosts necessitates the need to understand the biology of these ancient protozoans.

Toxoplasma is considered a model apicomplexan because the parasites are easy to propagate *in vitro*, there is an established mouse model of infection, and *Toxoplasma* is amenable to genetic manipulation [9,10]. Many of the initial

characterization studies of apicomplexan cellular biology such as invasion, motility, and function of the apicoplast were first investigated in *Toxoplasma* [10]. *Toxoplasma* is an experimentally tractable organism, and both classical and reverse genetic techniques have been implemented to study a variety of cellular processes. Additionally, *Toxoplasma* was one of the first apicomplexans to have its genome fully sequenced (12X coverage) with a database created to house the information and provide access to researchers (<http://ToxoDB.org>) [11]. Recently, many of the techniques initially established in *Toxoplasma* have been developed for other parasites, and genomic sequencing of other apicomplexans has been completed.

B. The life cycle of *Toxoplasma*

The life cycle of *Toxoplasma* is complex with sexual reproduction of the parasite taking place only within cats and other felids, its definitive hosts. However, the parasite can replicate asexually in a vast range of intermediate hosts, which include virtually all warm-blood vertebrates. With such a broad host range, *Toxoplasma* is a ubiquitous parasite with an expansive geographical distribution. The parasite is found in both temperate and tropical climates alike and does not demonstrate seasonal variation [12].

The three principal stages of *Toxoplasma* include the tachyzoite, bradyzoite, and sporozoite. Tachyzoites are the rapidly replicating, asexual form of the parasite, capable of infecting any nucleated cell. Tachyzoites divide via a specialized process termed endodyogeny, in which two identical daughter cells develop and then consume the mother cell after budding. As the proliferative form of the parasite, tachyzoites are responsible for tissue destruction and diseases associated with *Toxoplasma* infection. When under stress, such as from the immune response of the host, tachyzoites have the ability to differentiate into a quiescent form termed the bradyzoite. Bradyzoites are slowly developing tissue cysts that have a predilection for muscle tissue and the central nervous system. Bradyzoites are capable of evading the immune response and result in a latent, chronic lifelong infection of the host. The tachyzoite and bradyzoite stages, which are both haploid, are the only forms capable of replicating in intermediate hosts.

When a cat eats a mouse or any other animal that has a latent infection of *Toxoplasma* (contains bradyzoites), the parasite undergoes a series of differentiating events through several additional life stages within the epithelial cells in the small intestine of the cat. This process eventually results in the production of the sexually reproducing forms of *Toxoplasma*, the microgametocytes (males) or macrogametocytes

(females). Upon parasite sexual reproduction, oocysts are formed and shed into the environment through the feces of the cat. Oocysts are highly infectious and resistant to environmental insults, allowing them to persist for long periods of time. Initially, the oocysts are unsporulated. Sporulation, resulting in infectious sporozoites, occurs one to five days after excretion of the oocyst from the cat [13]. Once an intermediate host becomes infected with an oocyst, sporozoites are released into the intestinal lumen where they can then invade and differentiate into tachyzoites, which are able to disseminate throughout the body, thus completing the *Toxoplasma* life cycle. Figure 1 summarizes the life cycle and developmental stages of *Toxoplasma*.

The production of oocysts is critical for the propagation of the parasite and accounts for the coccidian classification of *Toxoplasma*. Typically, a cat will only shed oocysts once in its life, after its initial infection [14]. However, it is estimated that even after the ingestion of only a few bradyzoites, a cat is capable of shedding millions of oocysts, resulting in extensive environmental contamination [15]. Not only are oocysts able to withstand environmental conditions and remain viable for months, but they are also resistant to disinfectants. However, oocysts are destroyed at temperatures above 60°C [16]. Intermediate hosts such as livestock become infected with *Toxoplasma* when they ingest oocysts from the soil or those found on food sources. Humans become infected through the ingestion of oocysts from unwashed fruits, vegetables, or contaminated water. For humans, oocysts are the most infectious form of *Toxoplasma*, more so than bradyzoites or tachyzoites, and a single oocyst can result in an infection [17,18]. Humans can also become infected with *Toxoplasma* through the consumption of undercooked meat containing tissue cysts.

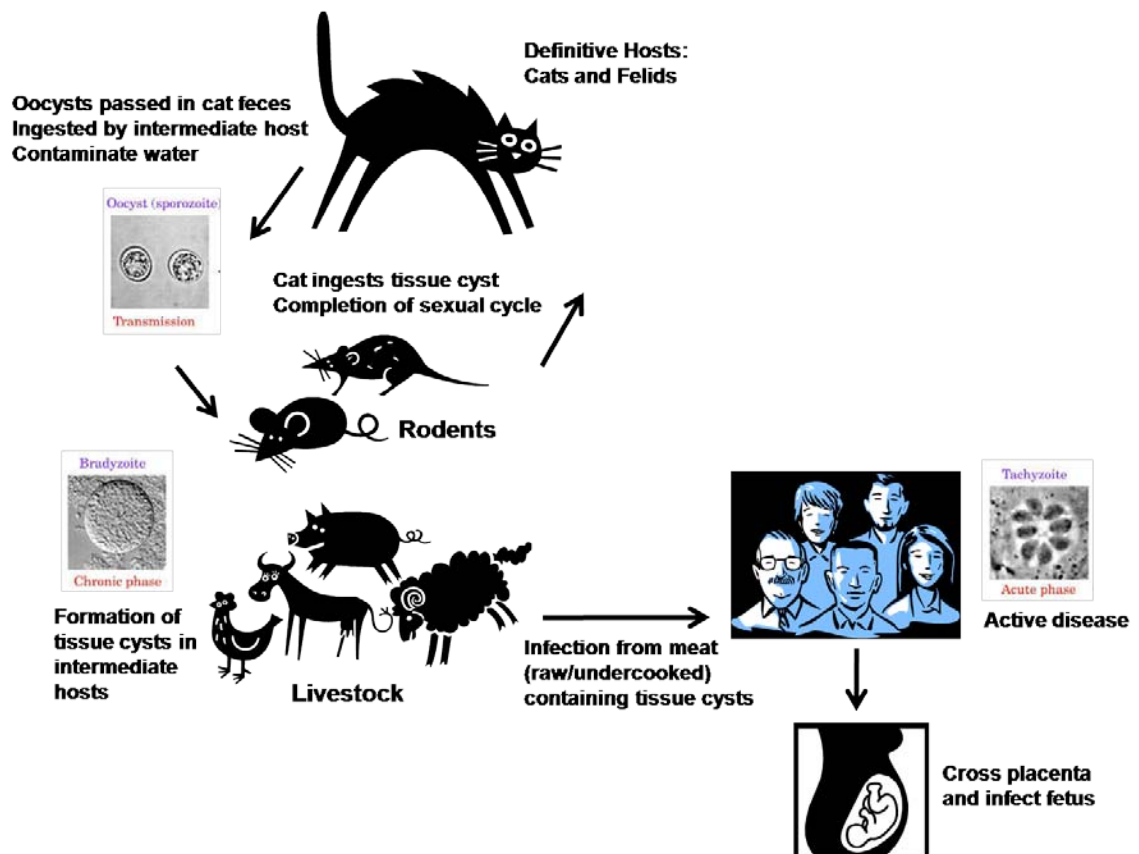


Figure 1: Life cycle and transmission of *Toxoplasma*. The definitive host of *Toxoplasma* is the cat (or other felids), and it is within the cat's gut that the sexual cycle of the parasites takes place, resulting in the formation of oocysts. An infected cat will shed highly infectious oocysts into the environment. Intermediate or dead-end hosts will consume oocysts, resulting in a chronic infection of *Toxoplasma* with bradyzoite cysts residing in the CNS and muscle tissue. When a cat eats a chronically infected rodent, the parasite's sexual cycle is completed. If humans consume oocysts from the environment (through contaminated water, on unwashed fruits and vegetables, or after eating undercooked meat containing tissue cysts), they become infected. Tachyzoites are the active form of the parasite responsible for tissue destruction and disease symptoms. Tachyzoites can cross the placenta and cause congenital infection of the fetus if a woman becomes infected for the first time during pregnancy. Figure adopted from Dubey et al. (1998) and http://www.toxomap.wustl.edu/life_c4.jpg [13].

C. Clinical toxoplasmosis

While the vast majority of the world's population is at risk for acquiring *Toxoplasma*, very few people will ever experience overt disease. The leading complications associated with infection by *Toxoplasma* include congenital birth defects and opportunistic disease although other complications exist, including ocular disease and an association with neurological and psychiatric disorders.

Many factors contribute to how *Toxoplasma* will affect its human host, one being the strain of the parasite causing the infection. *Toxoplasma* is highly clonal. The three strains that cause the vast majority of infections in humans differ from one another genetically by only 1% [19,20]. These strains, termed types I, II, or III, appear to have shared a common ancestor approximately 10,000 years ago [21]. Types I, II, and III strains of *Toxoplasma* have distinct phenotypic characteristics [19,20,22]. Type I strains are extremely virulent, grow faster than the other types in culture, and do not convert to bradyzoites *in vitro*. Types II and III are avirulent and are capable of *in vitro* and *in vivo* differentiation. The majority of human infections are attributed to the type II strain of *Toxoplasma* although any of the strains are capable of infecting humans. There can also be variations and atypical genotypes of *Toxoplasma* that diverge from the three main lineages [23]. The manifestations of toxoplasmosis in humans vary, and this may in part be attributed to the strain of *Toxoplasma* although others factors such as the health and genetic background of the host and environmental conditions may also contribute.

Roughly one-third of the world's population is infected with *Toxoplasma*; however, the prevalence varies region to region. For instance, in France, where more undercooked meat is eaten, approximately 80% of the population is seropositive for *Toxoplasma* while, in the United States between 16-40% are infected [7]. When humans acquire *Toxoplasma* postnatally, they are usually asymptomatic or develop mild disease. Common clinical signs are nonspecific, but can include cervical lymphadenopathy, fever, fatigue, and malaise. The disease is self-limiting and usually resolves without treatment. However, due to the life cycle of *Toxoplasma*, infection is life-long due to bradyzoite cysts residing permanently in host tissues.

Toxoplasmosis is a disease common to persons with immunosuppression. Patients under immunosuppression therapy after transplants and those persons who become immunocompromised due to certain malignancies are at risk. Toxoplasmosis is also a serious complication that has arisen with the advent of the HIV/AIDS pandemic.

In general, toxoplasmosis in the immunocompromised population results from reactivation of disease. Those persons who are chronically infected with the parasite and possess latent bradyzoite cysts are at risk for recrudescence of the bradyzoites into the actively destructive tachyzoites upon impairment of immunity. Because bradyzoites tend to reside in the CNS, the most common clinical manifestation for this form of toxoplasmosis is encephalitis. Patients often present with headache, confusion, motor weakness, and fever. Speech abnormalities and hemiparesis are the most common neurological findings [24]. If untreated, the disease can progress to seizures, stupor, coma, and even death. The predominant lesion of toxoplasmic encephalitis is cerebral necrosis, particularly of the thalamus [25]. Other clinical manifestations include chorioretinitis, pneumonia, myocarditis, and disseminated systemic disease.

Toxoplasma is capable of vertical transmission through its ability to cross the placenta of an infected female, resulting in congenital disease in the fetus. In the United States, 85% of women of childbearing age are susceptible to acute infection with *Toxoplasma* [26]. Congenital transmission occurs if a woman becomes infected with *Toxoplasma* for the first time during pregnancy. Risk of infection of the fetus is lowest during the first trimester and highest in the third trimester. Overall, the congenital infection risk is approximately 20 to 50% [27]. The severity of congenital disease is inversely related to the infection risk, with infection acquired during the first trimester being the most severe. If a pregnant woman is confirmed to be infected with *Toxoplasma* as determined by IgG and IgM antibody levels, the fetus can be tested by diagnostic PCR screening of the amniotic fluid [28]. Congenital infection can result in spontaneous abortion or still birth. In a live infant, congenital toxoplasmosis can manifest as the classic triad of symptoms, which includes chorioretinitis, hydrocephalus, and intracranial calcifications. However, clinical manifestations can be nonspecific and include convulsions, growth and mental retardation, learning disabilities, deafness, visual impairment, hepatosplenomegaly, and lymphadenopathy to name a few [26]. Most children are asymptomatic at birth but can develop sequelae such as ocular or neurological disease (learning disabilities) later in life [26,29,30]. It is estimated that 400 to 4,000 cases of congenital toxoplasmosis occur in the United States each year, resulting in a large healthcare economic burden due to the life-long sequelae associated with the disease [26].

Toxoplasma infection in humans can also present as ocular disease resulting from congenital disease, postnatal acquisition, or reactivation of latent infection.

Typically, ocular toxoplasmosis presents as posterior uveitis with the retina being the primary site of infection (retinitis). The choroid, the layer of blood vessels and connective tissue in the middle part of the eye, is often affected along with the retina, resulting in chorioretinitis. Other structures that can be infected include the vitreous, anterior chamber, or optic nerve [31,32]. Common symptoms include pain, redness, photophobia, and decreased vision. Congenital ocular toxoplasmosis tends to be bilateral, with multiple lesions located particularly in the macula, which can result in blindness [33]. Postnatally acquired disease is often unilateral and involves a focal area [33]. Most postnatally acquired ocular toxoplasmosis is identified in the second through fourth decades of life [34]. The major complication of ocular toxoplasmosis is recurrence, which can occur at any age [35]. Recurrence not only involves a similar pathology, as outlined above, but is also associated with additional complications such as neovascularization and retinal detachment [32]. Ocular toxoplasmosis may be more common than originally thought; a survey by Lum et al. (2005) estimated that in a 2-year period there were over 250,000 visits to ophthalmologists for active or inactive ocular toxoplasmosis in the United States [36].

For many years, latent toxoplasmosis was considered innocuous; however, recent studies suggest that this might not be the case. Due to its predilection to reside in the brain and CNS, *Toxoplasma* may alter the behavior of its host, a characteristic believed to enhance its transmission rate. This phenomenon has been termed the manipulation hypothesis [37]. Rodents chronically infected with *Toxoplasma* not only lose their innate fear of cat odors, but in effect became more attracted to them, thereby increasing the likelihood of being devoured by cats, resulting in *Toxoplasma* completing its sexual cycle [38,39,40,41]. In humans, infection with *Toxoplasma* has been linked to schizophrenia. Several studies, including a meta-analysis, demonstrate that individuals with schizophrenia and other psychoses have an increased prevalence of *Toxoplasma* infection based on seroconversion [42,43,44]. Interestingly, the genome of *Toxoplasma* was found to contain two homologues of tyrosine hydrolase, a key enzyme in the generation of dopamine [45]. Dopamine is a major neurotransmitter in the brain that is involved in many functions, including behavior and the reward system, and is also linked to schizophrenia [46]. It has been hypothesized that *Toxoplasma* might be able to manipulate behavior through modulation of dopamine within the brain of the host [45]. Additionally, serological evidence of *Toxoplasma* infection has been linked to increased tendency to have risk-taking behavior in humans. Studies show that infected humans

have an increased incidence of automobile accidents compared to controls [47,48]. The ability of *Toxoplasma* to manipulate its host is a fascinating subject that deserves further research in order to understand the implications of latent toxoplasmosis on humans.

In summary, *Toxoplasma* is a ubiquitous parasite that is responsible for a broad spectrum of clinical diseases in humans. Although approximately one-third of the world's population has been exposed to the parasite, most will not demonstrate clinical disease [49]. However, those that are congenitally infected, are immunosuppressed, or suffer from ocular toxoplasmic disease represent a significant health burden due to the reoccurring nature of the disease. Additionally, the implied roles of *Toxoplasma* in schizophrenia and host behavior manipulation extend the impact of the parasite on human health. Therefore, toxoplasmosis has a significant impact on the health and socio-economic state of the world.

D. Treatment for toxoplasmosis

The standard therapy for toxoplasmosis is a combination of pyrimethamine and sulfonamides (such as sulfadiazine or sulfadoxine), which are folate antagonists. Pyrimethamine, a competitive inhibitor of dihydrofolate reductase (DHFR), has a narrow therapeutic window and is only about six-fold more selective for *Toxoplasma* DHFR as compared to the mammalian enzyme [50]. Therefore, it is necessary to co-administer folinic acid as a supplement during treatment to minimize toxicity associated with pyrimethamine [51]. Folinic acid cannot be scavenged by the parasites; therefore, it will not interfere with therapy [52]. However, even with the administration of folinic acid, some patients will still experience adverse side effects of pyrimethamine treatment, which include thrombocytopenia and leucopenia [53]. Sulfonamides competitively inhibit the parasite enzyme dihydropteroate synthetase (DHPS), an enzyme not found in mammalian cells. An important difficulty with treatments involving sulfonamides is the propensity of patients to develop allergic reactions, resulting in rashes or, in extreme cases, Stevens-Johnson syndrome. Sulfonamides can also cause bone marrow suppression. In fact, 40 – 50% of patients treated with pyrimethamine and sulfonamides develop adverse effects, thus complicating the long-term prophylactic treatment needed for susceptible immunocompromised patients [54]. Alternative treatment options for toxoplasmosis include clindamycin, azithromycin, or atovaquone. Clindamycin and azithromycin both are antibiotics that inhibit protein synthesis and result in the loss of the apicoplast of *Toxoplasma* [55,56]. Atovaquone, a hydroxynaphthoquinone, is presumed

to inhibit the electron transport chain within the mitochondria. Often, corticosteroids are co-administered with *Toxoplasma* treatment regimes to reduce inflammation, particularly in ocular toxoplasmosis.

Special consideration for treatment of toxoplasmosis must be implemented in pregnant females. If the woman is infected but not the fetus, then spiramycin should be administered since this drug is unable to cross the placenta. If the fetus has acquired congenital toxoplasmosis, then the standard treatment of pyrimethamine, sulfadiazine, and folinic acid should be administered to the mother. However, pyrimethamine should not be given in the first 14 – 16 weeks of pregnancy due to the teratogenic nature of the drug during early fetal development [57].

Interestingly, drug resistance is rarely a complication of treating toxoplasmosis, because if resistance arises during clinical treatment, it cannot be readily transmitted. Humans are dead-end hosts for *Toxoplasma* and therefore do not contribute to the propagation of the parasite.

Aside from the mentioned side effects of each treatment regimen, the major caveat with current therapies is that none are capable of eliminating bradyzoite cysts, the latent form of toxoplasmosis. Bradyzoite cysts are quiescent and metabolically inactive; therefore, the current regimens that target active metabolic processes such as folate and protein synthesis are ineffective. This limitation is a significant disadvantage, making infection with *Toxoplasma* impossible to clear. Therefore, in susceptible populations such as immunocompromised human patients, reoccurrence of active disease is always a threat. This disadvantage necessitates the need to develop novel therapies that not only reduce toxicity but also eliminate or inhibit the recrudescence of latent bradyzoite cysts.

II. Epigenetics and chromatin remodeling

A. Definition of epigenetics

All nucleated cells in a human body contain the exact same genetic material. The DNA code or genetic sequence does not vary from hepatocyte to astrocyte, yet these two cell types are extremely different. How can an identical genetic code be read and interpreted in such variant manners as to produce such diverse cell types? The same question can be applied to protozoan parasites such as *Toxoplasma*. How can the same genetic code give rise to both a tachyzoite and a bradyzoite, two very distinct

life stages? Something beyond or above the genetic code must be contributing to the phenotype of individual cells, something epigenetic. Hence, the field of epigenetics was born.

Epigenetics is most basically defined as changes to a genetic locus not encoded by the underlying DNA sequence [1]. These changes account for a number of cellular phenomena that alter the degree of gene expression. By regulating the level of gene expression, the same genetic code can give rise to a multitude of phenotypes. In essence, the field of epigenetics investigates the cellular mechanisms that control how the genetic code is interpreted. Epigenetic control is critical not only for development and differentiation, but also for normal cellular function. The promise of the field of epigenetics is best represented by James D. Watson (2003), “You can inherit something beyond the DNA sequence. That’s where the real excitement in genetics is now.”

The molecular mechanisms underlying epigenetic modulation include a variety of cellular processes such as DNA methylation, RNA-directed gene silencing, and chromatin remodeling. DNA methylation is a well characterized covalent modification of cytosine residues of CpG dinucleotides. When these nucleotides are clustered together, they are referred to as CpG islands, and their associated DNA methylation results in transcriptional repression [58]. DNA methylation also plays a pivotal role in X chromosome inactivation and mammalian gene imprinting, two well characterized epigenetic phenomena [59]. RNA, in particular noncoding RNA, has also been implicated in epigenetic mediated gene silencing [60]. By far the most characterized epigenetic regulation involves alternations to chromatin. Chromatin is an intimate complex of DNA and proteins. One function of chromatin is to condense a cell’s DNA so that it can be packaged into chromosomes. Histone proteins, the primary constituents of chromatin, are key components to epigenetic regulation, as they can be both covalently modified and noncovalently remodeled to direct downstream cellular processes.

B. Histones and their code

The fundamental component of chromatin is the nucleosome, which is composed of an octamer of canonical (core) histone proteins (two each of histones H2A, H2B, H3, and H4) encircled by 147 bp of DNA. Histone 1 (H1) binds to the linker DNA connecting repetitive nucleosomes forming the classic pattern resembling beads on a string. Histones are small, basic proteins that are highly conserved throughout eukaryotes. Histone proteins consist of both a globular center core domain and a flexible tail domain.

The tails of histones extend from the surface of the nucleosome and associate with the DNA. Aside from the canonical histones, variants exist that possess key compositional differences. The presence of these histone variants adds an additional layer to the epigenetic regulation of chromatin as switching or replacing a core histone with a variant histone protein marks a given location for a specialized cellular function. For instance, H2AX, together with its associated post-translational modifications (i.e. phosphorylation), is a histone variant that associates with DNA damage and is involved in recruitment of the necessary DNA repair machinery to the site. Likewise, the histone variant H3.3 marks transcriptionally active regions, whereas another histone H3 variant, CENP-A, is centromere-specific and involved in chromosome segregation.

The condensed nature of chromatin and the role of histones were recognized to be repressive to transcription and gene expression as early as 1950 [61]. However, several years later, Allfrey observed that acetylation of histones correlated with gene activation [62]. This observation ignited interest in understanding the connection between histones and gene expression. Subsequently, other histone modifications such as phosphorylation and methylation were identified, but their significance remained unclear. The defining moment came when Brownell and Allis (1996) made the landmark observation that a *Tetrahymena* histone acetyltransferase (HAT) was homologous to the yeast transcriptional activator GCN5 (general control nonderepressible-5), thereby providing a direct connection between histone modification and gene expression [63]. Since this discovery, there has been an explosion of work to identify additional histone modifications, the residues modified, and the enzymes responsible.

It is now known that the N-terminal tails of histones are extensively modified, and to some extent even residues within the core globular domain of histones are subjected to modification. The various modifications include acetylation, phosphorylation, methylation, ADP-ribosylation, ubiquitination, and sumoylation. All of these modifications are reversible. A plethora of enzymes capable of catalyzing the incorporation or removal of these modifications have been identified. The majority of the chromatin-modifying enzymes are members of large multi-subunit complexes and exhibit remarkable specificity to target specific residues. Interestingly, many chromatin-modifying enzymes are also capable of targeting non-histone proteins [64,65]. Additionally, there are enzymes capable of remodeling chromatin and changing the nucleosomal composition in a non-covalent manner. These enzymes use the energy from ATP hydrolysis to alter the higher-order structure of nucleosomes, as well as

exchange canonical histones for variants. This multitude of alterations creates a dynamic chromatin environment for the regulation of gene expression.

The modifications of histones have both *cis*- and *trans*-effects. *Cis*-effects constitute changes in the physical properties of the histone such as structural modulation or alteration of electrostatic charge. For instance, histone acetylation on lysines neutralizes the positive charges of these basic residues allowing for expansion of the nucleosome structure. This grants the transcriptional machinery access to a particular region of DNA. Chromatin modulation can also result in *trans*-effects, in which there is recruitment or stabilization of specific binding partners or chromatin-associated proteins. These proteins act as readers of the modifications and recruit additional proteins or complexes to perform downstream chromatin alterations. For example, the bromodomain is a particular motif that recognizes acetylated lysine residues and is often found on HATs. It is believed that through the recognition of acetylated histone tails, bromodomain-containing proteins are able to direct additional modifications and thus participate in acetylation-dependent chromatin remodeling [66,67]. The complex language established by the modifications to chromatin resulted in the histone code hypothesis [68]. The histone code hypothesis proposes that there are writers and readers of the histone modifications. Together they constitute a cellular code that must be interpreted to produce specific downstream effects [68].

Although all the histone modifications are important and play a distinct role in cellular function, the focus of this thesis will be on histone acetylation by the enzyme GCN5.

III. GCN5 is a HAT involved in transcriptional activation

A. Definition, classification, and summary of HATs

As the bridge that linked histone acetylation to transcriptional activation, GCN5 as well as histone acetylation in general, became an extensively studied topic. Since the initial discovery by Brownell and Allis (1996), a multitude of HATs have been identified, many of which were previously categorized as transcriptional activators [63]. HATs can broadly be divided into two subcategories: type A HATs, which localize to the nucleus, and type B HATs, which have a cytoplasmic distribution. The general consensus is that type B HATs are involved in acetylating newly synthesized free histones prior to their nuclear import, whereas type A HATs acetylate nucleosomal histones within chromatin

and are linked with transcriptional activation. HATs are further characterized into families defined by sequence similarity and possession of conserved protein domains. Interestingly, each HAT family is very diverse in terms of the HAT domain structure and associated protein characteristics. For instance, the GNAT family (GCN5-related N-acetyltransferase) is characterized by a HAT domain of approximately 160 amino acids followed by a bromodomain, a motif that recognizes acetylated lysine residues [67,69]. On the other hand, MYST family HATs (named for the members MOZ, Ybf2/Sas3, Sas2, and Tip60) contain HAT domains of roughly 250 residues and possess N-terminal chromodomains, involved in recognition of methylated lysine residues [69,70]. Several other HAT families exist, each with their own characteristics.

Enzymatically, HATs transfer an acetyl group from acetyl-coenzyme A (acetyl CoA) to the ϵ -amino group of certain lysine side chains within histones [71]. This modification neutralizes the positive charge of lysine, which weakens the interaction between the histone and DNA, or between two nucleosomes, or results in a conformational change in the nucleosomal structure [72,73,74]. Such alterations to the chromatin structure allow various factors, such as the transcriptional machinery, to gain access to genetic loci.

The antagonists of HATs are histone deacetylases (HDACs) that catalyze the removal of the acetyl group from specific lysine residues, allowing the chromatin to resume a restricted state. Hypoacetylation of a given region and the activity of HDACs are considered repressive to transcription. The reciprocal activities of HATs and HDACs govern the steady-state histone acetylation balance within a cell and thereby regulate gene expression.

Very few HAT inhibitors exist. Some compounds, such as curcumin, garcinol, and anacardic acid are derived from naturally occurring products and show HAT inhibitory properties [75]; however, there are limitations to each. Curcumin has anti-oxidant properties and therefore might interfere with other cellular pathways, making it non-specific [75,76]. Garcinol has low potency, resulting in high cellular toxicity, including apoptosis [75,77]. Anacardic acid has low cell permeability, limiting its *in vivo* usage [75,78]. Other synthetically derived HAT inhibitors such as the bisubstrate inhibitors, histone peptide connected to CoA, are likewise unable to permeate cells [79,80]. In general, the currently available HAT inhibitors have limited *in vivo* utility due to lack of specificity, low potency, and decreased cell permeability [80]. The discovery or synthesis of additional HAT inhibitors would greatly facilitate the study of HATs and their

epigenetic roles in biological and medical processes. Several HDAC inhibitors are being examined in clinical trials for various human diseases, including various types of cancers [81].

B. GCN5 is a conserved HAT

GCN5 is a well conserved HAT with homologues in a multitude of diverse organisms from yeast to plants (*Arabidopsis*) to humans. As a member of the GNAT family, GCN5 shares similarities with other members including Elp3 (elongation protein 3), Hat1, and Hpa2 (histone and other protein acetyltransferase 2). The GCN5 homologues share common sequence motifs in their HAT domains, C-terminal bromodomains, and Ada2 interaction domains, a necessary co-activator (Figure 3, Chapter 1, section IV-C) [69]. Interestingly, vertebrates such as humans and mice possess a second GCN5 homologue termed PCAF (p300/CBP-associated factor) [82]. PCAF not only functions as a HAT but has intrinsic E3 ubiquitinase activity [83].

In vitro studies in *S. cerevisiae* have revealed that recombinant GCN5 strongly acetylates free histone H3 lysine 14 (K14), and to a lesser extent free histone H4 (K8 and K16) [84]. However, when GCN5 is incorporated into one of its multi-subunit complexes, its substrate specificity expands for histone H3 and includes histone H2B [85]. The acetylation of histone H3 K36, in yeast, has also been shown to occur in a GCN5-dependent manner [86]. Additionally, GCN5 acetylates free histone H3 K56 *in vitro* [87,88].

Mechanistically, GCN5 catalyzes the acetylation of histones by forming a ternary complex between the enzyme, histone, and acetyl-CoA. GCN5 deprotonates the histone substrate thereby, allowing the ϵ -amino group of the lysine to directly attack the bound acetyl-CoA. In *S. cerevisiae*, the glutamic acid residue 173 is essential for catalytic function and serves as a general base for deprotonation [89]. This residue is conserved throughout the GCN5 homologues.

C. GCN5 functions within multi-subunit complexes

In order for transcription to occur, a series of well coordinated events must take place. First gene-specific transcription factors bind to upstream activating sequences (UAS) in promoters and recruit additional transcriptional activation complexes and cofactors. These complexes include chromatin modifying enzymes that create an open chromatin confirmation, allowing general transcription factors to access the genetic loci.

These events lead to the recruitment and formation of the pre-initiation complex (PIC) and hence, the commencement of RNA polymerase II mediated transcription. As a HAT, GCN5 is the enzymatic component of several multi-subunit complexes that are integral for the activation and initiation of transcription. To date, the *S. cerevisiae* SAGA (Spt-Ada-GCN5 acetyltransferase) complex has been the most widely studied and best characterized GCN5 complex.

The *S. cerevisiae* SAGA complex is a 2 MDa complex that is composed of several distinct functional modules. Within SAGA, GCN5 is the enzymatic component of the acetyltransferase unit, along with the Ada (alteration/deficiency in activation) proteins [90]. Ada2, the co-activator of GCN5, along with Ada3, regulate the acetylation activities within SAGA [91,92]. Other members of SAGA include Spt proteins (suppressor of Ty) and TAFs (TBP-associated factors). Together, Spt3 and Spt8 form the TBP (TATA-binding protein) interaction unit with Spt3 being crucial for the recruitment of TBP and PIC formation [93]. The complex member Tra1, an ataxia telangiectasia mutated (ATM)-related protein, interacts with specific transcription factors such as GCN4 (general control nonderepressible 4) to recruit the entire complex to a given genetic locus [94,95]. Most of the other members of SAGA are integral for the architectural structure of the complex or have less defined functions. Recently, it was determined that SAGA contains a second enzymatic component, Ubp8, a deubiquitinase (Dub). Ubp8 is the enzyme responsible for the deubiquitination of H2B K123 within the context of its SAGA modular unit components Sus1, Sgf11, and Sgf73 [96,97,98]. Table VII (Chapter 3, section II-A) lists all 21 members of the *S. cerevisiae* SAGA complex.

The SAGA complex plays an integral role in the transcription of approximately 10% of the *S. cerevisiae* genome [99]. Most of the genes are stress induced, and include genes that regulate the response to environmental stress, starvation, DNA damage, and heat [99]. The SAGA complex members have also been implicated in a variety of other important cellular processes. For instance, SAGA can localize not only to gene promoters but also to the coding regions, where acetylation by GCN5 promotes nucleosome eviction and facilitates transcriptional elongation [100]. Additionally, the subunits Sus1, Sgf11, and Sgf74 link SAGA physically to the nuclear export machinery and are involved in the export of transcribed mRNA [101,102]. This study indicated the existence of coupling between transcriptional elongation and mRNA export [103].

There is a high level of homology between the *S. cerevisiae* SAGA complex and the GCN5 complexes purified from other species including *Drosophila* and humans.

Each of these species has SAGA complexes (also referred to as STAGA, Spt3-TAF9-GCN5 acetyltransferase in humans) that are conserved in protein subunit composition and structure. This similarity suggests that each complex serves a related function across species [104]. PCAF, the second GCN5 homologue in vertebrates, also functions within multi-subunit complexes that are analogous to the GCN5 complexes. Furthermore, metazoans possess two homologues of Ada2 (ADA2-A and ADA2-B) that play distinct roles in the various GCN5/PCAF complexes [104].

The other distinctive GCN5 complex in *S. cerevisiae* is the ADA complex. At approximately 700 kDa, this complex is considerably smaller than SAGA. The yeast ADA complex is composed of the SAGA members GCN5, Ada2, Ada3, and Sgf29, along with Ahc1 and Ahc2 (ADA histone acetyltransferase complex component), which are necessary for structural integrity. There are also several unidentified proteins that may be members the *S. cerevisiae* ADA complex [105,106,107]. Although not all functions of the *S. cerevisiae* ADA complex have been elucidated, the complex is able to acetylate H3 K14 and K18 [85].

Likewise, a smaller (700 kDa) GCN5-containing complex has been identified in both *Drosophila* and human cells [107,108,109]. The ATAC (ADA two a containing) complex was first purified in *Drosophila* due to its defining feature of possessing ADA2-A, one of GCN5's co-activator homologues [109]. Metazoans have two Ada2 homologues, and it has subsequently been determined that metazoan SAGA complexes selectively incorporate ADA2-B while ADA2-A is a member of the smaller GCN5 complex ATAC. The ATAC complex also contains proteins homologous to SAGA's Ada3 and Sgf29. Several other proteins have been identified from both *Drosophila* and human ATAC including an additional HAT enzyme, ATAC2 (*Drosophila*), which has specificity for histone H4 [108]. The human ATAC complex can contain either GCN5 or PCAF although the homologue of ATAC2 (CSRP2BP) does not appear to have *in vitro* H4 HAT activity, despite having a conserved GNAT domain [107]. The ATAC complex is involved in transcriptional activation but has additional functions including stimulation of nucleosome sliding (*Drosophila*), as well as a potential involvement in stress signaling pathways and even transcriptional repression (humans) [107,108].

In summary, GCN5 is a member of two distinctive complexes that are conserved throughout many species. Aside for the SAGA and ADA/ATAC complexes, there are several minor variations to each that further expand the GCN5 (or PCAF) complexes.

All the GCN5 complexes have been attributed to transcriptional activation; however, additional functions are being defined for each.

D. The role of GCN5 in cellular physiology

In *S. cerevisiae*, GCN5 is not essential and acts as a specific co-activator of only approximately 4% of genes, mostly genes involved in stress response pathways [110,111]. Although other HATs appear to compensate for the loss of GCN5 in *S. cerevisiae*, it has been determined that GCN5 has additional roles. For instance, it was determined that GCN5 functions as a global chromatin remodeling enzyme in *S. cerevisiae*, by setting a basal state of genome-wide acetylation upon which specific targeted acetylation is superimposed [112]. Additionally, more specific roles have been assigned to GCN5. The SAGA complex appears to be instrumental in coordinating transcription with RNA splicing because GCN5 activity was also shown to be required for recruitment of particular components of the splicing machinery [113]. GCN5 also plays a role in *S. cerevisiae* nucleotide excision repair, demonstrated by increased H3 K9 and K14 acetylation in response to ultraviolet (UV) irradiation [114]. This acetylation can result in increased transcription of certain nucleotide excision repair proteins and also appears to function in a transcription-independent manner [114,115]. Consequently, the loss of GCN5 results in deficient nucleotide excision repair [114].

In *Drosophila*, GCN5 is essential; deletion mutants fail to complete metamorphosis, dying at the end of the larval period [116]. Likewise, GCN5 is also essential for vertebrates, as GCN5-null mice die during embryogenesis due to a failure to form the dorsal mesoderm causing death by day 10.5 post coitum [117]. Interestingly, mice with HAT-dead or catalytically inactive GCN5 survive until day 16.5 post coitum and show defects in neural tube closure and exencephaly [118]. These data show the necessity of GCN5 activity for proper development. In contrast, PCAF-null mice develop normally, with no observable phenotype. However, increased levels of GCN5 expression were noted in specific tissues that normally expressed PCAF, suggesting a compensatory function [117,119].

Studies in chicken DT40 cells show that PCAF deletion does not affect cell growth, whereas cells lacking GCN5 have a delay in growth rate. Additionally, in the cells with deletion of GCN5, both PCAF and HDAC4 displayed increased expression, suggesting compensatory mechanisms exist [120]. These studies also revealed that GCN5 plays a role in the transcription of certain G₁/S phase transition-related genes as

well as apoptotic genes [120]. Studies in human and mouse cells indicate that GCN5 is also integral for telomere maintenance through regulation of components of the shelterin complex [121]. GCN5 and PCAF can also acetylate cellular proteins other than histones [122,123,124,125]. Collectively, these data indicate that GCN5 is an important mediator of transcriptional activation and plays a pivotal role in several other cellular processes.

E. Nuclear import of HATs

In order for proteins such as type A HATs to perform their role in transcriptional activation, they must enter the nucleus of the cell. The nuclear membrane forms a barrier that prevents the passive diffusion of macromolecules into the nuclear compartment. To traffic necessary nuclear proteins to their site of action, specific chaperones exist that facilitate nuclear import. These chaperones not only recognize their cargo via specific signals but navigate through the nuclear pore complex for delivery. In the classical nuclear localization model, the chaperones are termed karyopherins or importins and use the energy provided by Ran, a small Ras family GTPase, to direct nuclear traffic [126,127,128]. Ran exists as either a GTP- or GDP-bound nucleotide state, with compartmentalized regulatory proteins controlling the switch. Ran guanine nucleotide exchange factor (RanGEF) is the nuclear regulatory factor, while Ran GTPase-activating protein (RanGAP) is cytoplasmic. This separation provides an asymmetric distribution of the nucleotide states of Ran and allows for selective binding by karyopherins, thus facilitating nuclear trafficking [129,130,131].

The classical nuclear import pathway begins with the karyopherin importin- α recognizing and binding to a cargo protein via a signal (nuclear localization signal, NLS). Importin- β then binds importin- α forming a ternary complex that is translocated through the nuclear membrane via importin- β 's interactions with the nuclear pore complex. Once inside the nucleus, the binding of RanGTP to importin- β results in the dissociation of the complex and release of the cargo. Both importins are recycled back to the cytoplasm bound to RanGTP. In the cytoplasm, RanGTP is hydrolysed by RanGAP, resulting in release of the importins and formation of RanGDP. RanGDP is then shuttled back to the nucleus by a specialized transporter where it is converted to RanGTP by RanGEF [130,131,132]. Variations to this model exist and are represented by a variety of importins capable of performing the same role as importin- α . Alternatively, cargo can be bound directly to importin- β [133,134,135,136].

The first step in this process, the recognition of a cargo protein by an importin through a signal, is crucial in the transportation of proteins into the nucleus. These signals found on the cargo proteins are specific targeting sequences termed NLSs. The classical NLS is a short cluster of basic amino acids and is exemplified by the NLS of the SV40 large T antigen, PKKKRRV [137]. This type of NLS is termed monopartite because the entire signal exists in a single cluster of amino acids. Bipartite NLSs contain two clusters of basic residues separated by 10 – 12 amino acids. The NLS of nucleoplasmin, KRPAATKKAGQAKKKK, is a classical bipartite sequence [138]. Since the discovery of NLSs, a plethora have been discovered, some of which conform to the classical pattern while others are divergent, forming multiple classes and patterns [130,139,140].

The NLSs of HATs have been studied in order to understand the regulation of these important cellular modulators. PCAF contains an NLS at its C-terminus between amino acids 428 – 442 [141]. Five lysine residues within the NLS are autoacetylated. The acetylation of these lysines is crucial for nuclear localization as mutations to arginines or over-expression of HDAC3 (capable of deacetylating these residues) results in cytoplasmic accumulation [142]. The acetylation of HATs has also been shown to play an important role in regulating enzymatic function [143]. The only GCN5 NLS characterized to date is a GCN5 homologue from *Toxoplasma* that possesses a classic monopartite basic-rich cluster at its N-terminus and is capable to interacting with importin- α [144].

IV. Gene expression and epigenetic modulation in *Toxoplasma*

A. Regulation of gene expression

Toxoplasma is a haploid organism in both the tachyzoite and bradyzoite stages, and has a genome of approximately 63 Mb distributed among 14 chromosomes [145]. Sequencing and annotation of the *Toxoplasma* genome is available for strains from each type (I, II, and III), along with other genomic-wide data such as ESTs (expressed sequence tags), expression and proteomics data. This information is housed at *Toxoplasma* database (ToxoDB, <http://ToxoDB.org>) and is a valuable research tool [11].

The complexity of the life cycle of *Toxoplasma* as well as the ability to adapt to numerous host environments necessitates precise coordination and regulation of gene expression. Gene expression analysis has revealed that changes to the transcriptome

of the parasite are critical for survival and differentiation. The progression of gene expression through the cell cycle of *Toxoplasma* has been examined using microarrays. The parasites have a strict program of gene expression, with subsets of genes expressed at given points during the cell cycle [146]. Additionally, serial analysis of gene expression (SAGE) studies have revealed that *Toxoplasma* uses a just-in-time progression, where genes are expressed only as needed throughout the cell cycle [147]. Studies analyzing the transition from tachyzoites to bradyzoites using cDNA microarrays reveal a hierarchical gene expression program may direct the differentiation process [148]. Taken together, these studies indicate *Toxoplasma* utilizes a tightly controlled process, reliant on transcriptional regulation, to control gene expression. Although *Toxoplasma* possesses the basal transcriptional machinery, initial searches revealed a dearth of identifiable specific transcription factors found in other eukaryotes [149,150]. Interestingly, *Toxoplasma* was found to possess a large repertoire of chromatin remodeling enzymes, possibly implying an increased reliance on epigenetic modifications to regulate its gene expression [151,152,153]. Additionally, histone acetylation and methylation, as analyzed by chromatin immunoprecipitation (ChIP), are associated with modulation of gene expression pertinent to stage-specific gene expression in *Toxoplasma* [154,155,156].

Recently, *in silico* evidence suggested that apicomplexan parasites possess a lineage-specific expansion of proteins containing motifs similar to AP2 (Apetala2)-integrase DNA-binding domains [157]. AP2 proteins are transcription factors important for developmental transitions and the stress response in plants [158]. This information implies that apicomplexan parasites might utilize plant-like transcription factors as opposed to those that are prevalent in metazoans. Using protein binding microarrays (PBMs), De Silva et al. (2008) demonstrated that *Plasmodium* AP2-domains were able to bind specific DNA motifs, and these motifs were subsequently found upstream of coordinately regulated genes [159]. A study demonstrating that a specific *Plasmodium* AP2 protein (PF11_0442, AP2-O) activates ookinete stage-specific genes through binding of a specific DNA motif provided functional validation that AP2 proteins function as transcription factors within apicomplexans [160]. Additionally, select apicomplexan AP2 proteins were found to be associated with *Toxoplasma* HDAC3 and *Plasmodium* GCN5, thereby linking the AP2 proteins to epigenetic gene regulation [154,161]. These discoveries have opened new avenues for gene regulation research in Apicomplexa

parasites, with at least 60 AP2-domain proteins identified in *Toxoplasma*, compared to 27 in *Plasmodium*, and 17 in *Cryptosporidium* [157].

B. The histones of *Toxoplasma* and their modifying enzymes

As described for other eukaryotes, *Toxoplasma* possesses conserved homologues of the canonical histones H2A, H3, and H4, as well as, centromeric H3 (termed TgH2A1, TgH3, TgH4, and TgCenH3, respectively). H2B is also conserved, but the parasite contains two nearly identical proteins (TgH2Ba and TgH2Bb) with significant homology to canonical H2B [162]. Additionally, there is another lineage of H2B in *Toxoplasma* consisting of the variant histone TgH2Bv1. TgH2Bv1 is highly expressed and not differentially regulated; therefore, it is expected to be the main functioning H2B. Although TgH2Ba and TgH2Bb are very similar, only one, TgH2Ba, has been found to be expressed in tachyzoites and bradyzoites. TgH2Bb could not be amplified from cDNA or detected during histone purification. TgH2Ba may be differentially regulated since its expression was shown to be higher in tachyzoites compared to bradyzoites, indicating TgH2Ba may have a specialized role [162]. *Toxoplasma* also has variants of H2A (TgH2A1, TgH2AX, TgH2AZ) [163]. Due to its dimerization with TgH2Bv1 and association with acetylated histones at active genes, TgH2AZ acts as the main H2A variant. On the other hand, the variants TgH2AX and TgH2A1 appear to have specific roles in response to stress [163]. Two variants of H3 have been identified, TgH3 and TgH3.3, which function in chromatin structure and gene regulation, respectively [164]. Only four amino acids differ between the two H3 variant histones, but this slight change in sequence may result in distinct structural and functional differences [164]. To date, no variants of H4 have been identified. *Toxoplasma* lacks a histone H1 homologue. Due to the divergent nature of this protein among species and the evidence that another protozoa, *Tetrahymena*, grows normally when H1 is deficient, it is likely that *Toxoplasma* may not require H1 to function [165]. Alternatively, *Toxoplasma* may have a variant that is too divergent to detect.

All of *Toxoplasma*'s histones contain the prototypical N-terminal tails and other well characterized conserved residues that can be extensively modified for regulation of chromatin architecture [151,152]. Histone marks on TgH3 and TgH4 have been studied using chromatin immunoprecipitation (ChIP) and ChIP coupled to microarray (ChIP-to-chip) techniques [155]. Acetylation of H3K9 and H4 as well as trimethylation of H3K4, are specific concurrent histone marks that have been identified at promoters of actively

expressed genes [155]. Modifications of transcription repression have also been identified in *Toxoplasma* and include trimethylation of H3K9 and H4K20 [166].

Toxoplasma possesses a vast trove of proteins capable of modifying chromatin. This category includes enzymes able to write the code (acetyltransferases, methylases, kinases, etc.), erase the code (deacetylases, demethylases, phosphatases, etc.) and those able to re-organize the histones (ATP-dependent remodelers) [151,152,153]. To date, only a select handful of these enzymes have been studied.

The first chromatin modifying enzyme to be cloned and characterized in *Toxoplasma* was a homologue of GCN5 [167,168]. Importantly, several years later, a second GCN5 homologue in *Toxoplasma* (distinguished from the original TgGCN5-A as TgGCN5-B) was identified, which was remarkable given the fact that there was no precedence for any invertebrates to possess two GCN5 homologues [169]. Additionally, two more HATs were characterized that belong to the MYST family [170]. Enzymes including HDACs, arginine methyltransferases, SET-domain lysine methyltransferases, and a SWI/SNF family ATP-dependent chromatin remodeler have all been described [154,166,171,172]. Each of these enzymes plays a distinct role in *Toxoplasma*'s biology, and several have been shown to regulate gene expression pertinent to bradyzoite differentiation [151,152,153].

HATs and HDACs in *Toxoplasma* function within their prototypical roles of transcriptional activation and repression as demonstrated by the presence of TgGCN5-A at active promoters and TgHDAC3 at inactive promoters [154]. An additional role of the HAT TgMYST-B was discovered that links this enzyme to the DNA damage response through the regulation of ATM kinase [173]. Likewise, the SET-domain lysine methyltransferase KMTox (TgSET13) contributes to the oxidative stress response in *Toxoplasma* and is linked with regulation of genes involved in antioxidant defense [171]. Another SET-domain protein, TgSET8, is involved in transcriptional repression and heterochromatin assembly [166]. TgSET8 may also be involved in bradyzoite formation/persistence because high levels of monomethylation of H4K20, a mark written by this enzyme, were found in bradyzoites [166]. The ATP-dependent chromatin remodeler, TgSRCAP, is another epigenetic modulator implicated in differentiation because its mRNA levels increase during *in vitro* conversion to bradyzoites [154,172]. Additionally, pharmacological inhibition of either HDAC3 or the arginine methyltransferase TgCARM1 was found to induce bradyzoite differentiation [154,174]. Other HDAC inhibitors, such as apicidin and SAHA (suberoylanilide hydroxamic acid),

have anti-*Toxoplasma* activity [175,176,177]. Although this summary is just a sampling, these examples demonstrate the importance of epigenetic and chromatin modulations in *Toxoplasma* and are summarized in Figure 2.

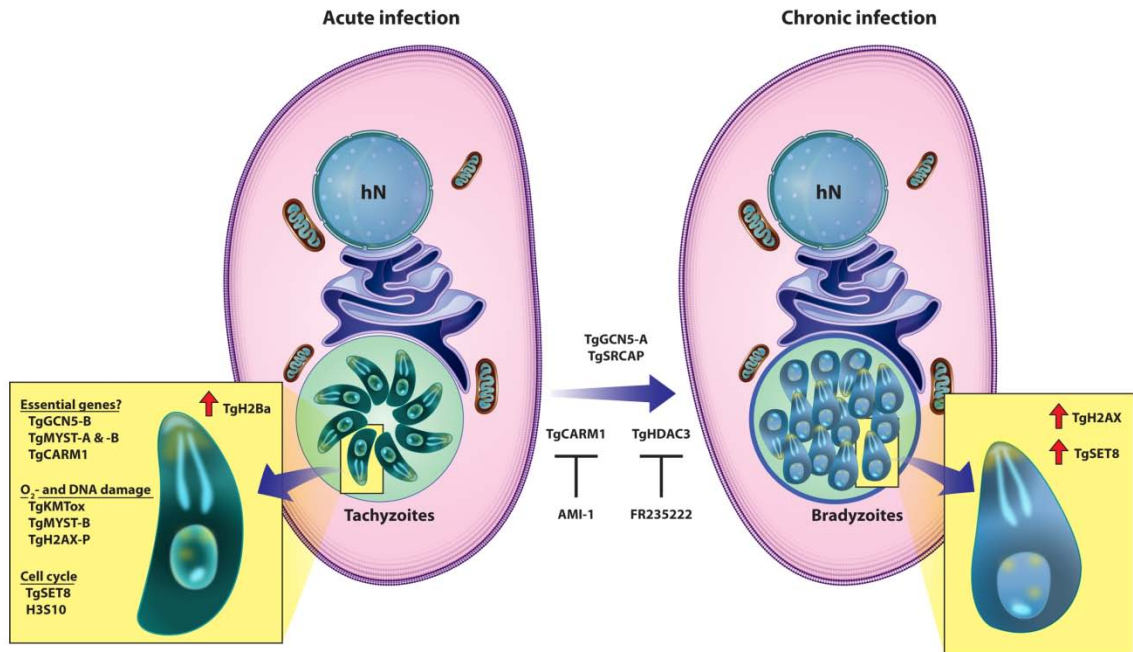


Figure 2: Diverse roles of epigenetics in *Toxoplasma* biology. Schematic diagram shows two host cells (hN = host cell nucleus) harboring a parasite vacuole containing either proliferative tachyzoites (acute infection, left) or a cyst containing slow-growing bradyzoites (chronic infection, right). The inserts depict a single *Toxoplasma* parasite in either the tachyzoite stage (left) or the bradyzoite stage (right) and list the known epigenetic characteristics associated with each life-cycle stage. In tachyzoites, TgGCN5-B, TgMYST-A and -B, and TgCARM1 are histone modifying enzymes that may be required for propagation, based on the inability to disrupt the genetic loci. KMTox and TgMYST-B have been linked to the response of *Toxoplasma* to reactive oxygen species (O₂⁻) and DNA damage, respectively. Additionally, the histone variant TgH2AX is phosphorylated in response to DNA damage in tachyzoites. TgSET8 has been implicated in *Toxoplasma* cell cycle regulation, as has the phosphorylation of histone H3 on Serine-10, which is delivered by an uncharacterized kinase. TgH2Ba appears to be a

variant histone exclusive to tachyzoites. TgGCN5-A and TgSRCAP have been implicated in the differentiation of tachyzoites into bradyzoites. The inhibition of TgCARM1 or TgHDAC3 by AMI-1 or FR235222, respectively, has been shown to induce cyst formation. The expression of the histone variant TgH2AX increases during bradyzoite differentiation. Presently, TgSET8 is the only chromatin remodeler identified to be associated in bradyzoite biology. Art in Figure 2 was done by Christopher Brown (2010) at the IUSM Office of Visual Media. Originally published in Dixon et al. (2010) [153].

C. Two GCN5 homologues in *Toxoplasma*

Two unique GCN5 homologues have been cloned and characterized from *Toxoplasma* [167,168,169]. Although no precedence exists for this occurrence in invertebrates, we have recently discovered that *Neospora*, another apicomplexan parasite and the nearest relative of *Toxoplasma*, also contains two GCN5 homologues. In contrast, *Plasmodium* only possesses a single GCN5 homologue [178]. The presence of two GCN5s in *Toxoplasma* suggests that these HAT homologues might have pivotal roles within the parasites.

TgGCN5-A (chromosome 3) encodes a protein of 1169 amino acids, while TgGCN5-B (chromosome 14) is a 1032 amino acid protein [168,169]. TgGCN5-A and –B protein sequences share 54% identity and 66% similarity and are nearly identical in their HAT catalytic domains and possess very similar bromodomains. The ADA2-binding regions are less conserved [169]. The N-terminal extensions (794 amino acids of TgGCN5-A and 625 amino acids of TgGCN5-B) exhibit the greatest divergence between the two. These N-terminal regions are unique in that they lack identifiable protein motifs and do not have homology to known protein sequences in other species. Interestingly, the N-terminal extensions of TgGCN5-A and –B bear no significant similarity to one another [169]. Most GCN5 homologues from lower eukaryotes have very short N-termini, so it is unusual that several apicomplexan parasites (*Toxoplasma*, *Neospora*, and *Plasmodium*) have GCN5 homologues with elongated N-terminal extensions (Figure 3).

The HAT domain of both TgGCN5s show strong similarity to yeast GCN5, suggesting that the *Toxoplasma* GCN5s might have a similar acetylation profile to other GCN5 homologues. Using purified recombinant full-length FLAG-tagged TgGCN5-A and –B from *Toxoplasma*, the enzymatic activities and substrate profiles of each HAT

were examined using *in vitro* HAT assays [154,169]. TgGCN5-B appeared to be the prototypical GCN5 due to its ability to acetylate K9, K14, and K18 of histone H3. However, TgGCN5-A had a narrow substrate specificity limited to acetylation of K18 on histone H3 [154,169]. To date, TgGCN5-A is the only GCN5 homologue with such limited substrate specificity.

Like their differences in *in vitro* substrate specificity, the *Toxoplasma* GCN5 homologues also have a differential interaction with the GCN5 co-activator ADA2. Interestingly, *Toxoplasma* possesses two distinct ADA2 homologues (termed TgADA2-A and TgADA2-B), in contrast to the single ADA2 homologue of *Plasmodium* [169,179]. Using directed yeast two-hybrid tests, it was determined that TgGCN5-A can only interact with TgADA2-B, while TgGCN5-B can interact with both TgADA2s [169]. These data are in agreement with the fact that the ADA2 binding domains of the *Toxoplasma* GCN5 homologues have notable differences and are less conserved than the catalytic domains and bromodomains [169].

Toxoplasma is able to survive without TgGCN5-A. Sullivan et al. (2006) constructed a knock-out TgGCN5-A (Δ GCN5-A) parasite strain [169]. Although parasites lacking TgGCN5-A have no obvious phenotype when cultured under normal growth conditions, they do not recover from certain stresses as robustly as parental wild-type parasites, and Δ GCN5-A parasites fail to express stress-induced bradyzoite-marker genes. Additionally, microarray data from this strain suggest that TgGCN5-A might be important for stress-induced bradyzoite differentiation [Sullivan and Nagulaswaran et al., unpublished]. In contrast, attempts to knock-out TgGCN5-B through an analogous approach have failed, suggesting that this GCN5 homologue may be essential for tachyzoite growth and development. Both anacardic acid and curcumin inhibit *Plasmodium* GCN5 *in vitro*, and both agents decrease proliferation of *Plasmodium* in culture [180,181].

The roles of the elongated N-terminal extensions in the apicomplexan GCN5 homologues remain to be elucidated. TgGCN5-A possesses a unique NLS (RKRVKR) within its N-terminus [144]. Although the N-terminus of TgGCN5-B is required for localization to the parasite nucleus, the NLS of TgGCN5-A is not conserved in TgGCN5-B [169]. Therefore, TgGCN5-B must use a novel signal to facilitate nuclear localization. Using a high-throughput yeast two-hybrid screen, LaCount et al. (2005) discovered that the N-terminus of *Plasmodium* GCN5 associates with a variety of other proteins [161]. The *Plasmodium* GCN5 associating proteins include other potential chromatin-modifying

proteins, as well as, some potential transcription factors. In this study, *Plasmodium* GCN5 associated with two AP2 domain proteins along with other proteins containing putative DNA-binding motifs (Myb DNA binding domain and AT-hook DNA binding domain). Interestingly, GCN5 was found to be the most highly connected protein in *Plasmodium* [161]. These studies suggest that the N-termini of the Apicomplexa GCN5 homologues are critical for the functions of these proteins.

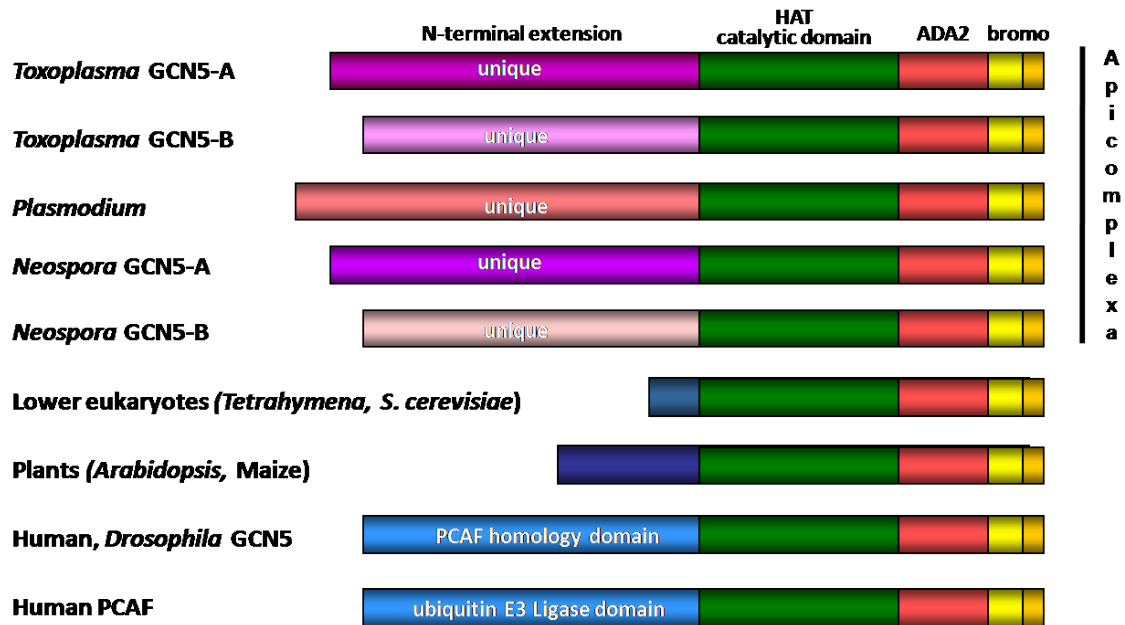


Figure 3: Comparison of apicomplexan GCN5s to other eukaryotic GCN5 homologues. Schematic diagrams compare *Toxoplasma* GCN5s with homologues from other species. Both *Toxoplasma* GCN5s share homology in the catalytic (green), ADA2 (red), and bromo- (yellow) domains. However, each possesses a unique N-terminal extension that does not share homology to protein sequences in other species. Other apicomplexans such as *Neospora* and *Plasmodium* also possess GCN5 homologues with elongated N-termini. *Drosophila* and human GCN5 homologues contain N-terminal extensions that have a high degree of similarity to the N-terminal region of PCAF. This PCAF homology domain is absent from the GCN5 homologues found in plants, fungi, and protozoa examined to date. Other invertebrates and protozoa have GCN5 homologues with short N-termini. Orange regions represent the C-termini of each homologue.

V. Hypothesis and aims

The preliminary studies characterizing the *Toxoplasma* GCN5 homologues suggest that these two HATs may have separate functions within the parasite. If this hypothesis is correct, the *Toxoplasma* GCN5 HATs would mimic the roles of GCN5s in higher eukaryotes such as mouse or human. Due to the differential interaction with ADA2 co-activator proteins and lack of homology between the two N-terminal extensions of each *Toxoplasma* GCN5 homologue, it is possible that the TgGCN5s form distinct complexes within the parasite. Interestingly, bioinformatic searches of the ToxoDB reveal a dearth of identifiable SAGA complex homologues. Therefore, the components of the *Toxoplasma* GCN5 complexes are likely to be novel or highly divergent. TgGCN5-A possess a unique NLS (RKRVKR) within its N-terminus that is not conserved in TgGCN5-B demonstrating that TgGCN5-B utilizes a different signal to enter the parasite nucleus. The fact that parasites are able to tolerate the deletion of TgGCN5-A, while TgGCN5-B appears essential, also suggests a distinct role for each homologue within the parasite. The initial characterization of the Δ GCN5-A parasite strain, suggests that TgGCN5-A is necessary for *Toxoplasma* differentiation and stress response, whereas TgGCN5-B is critical for parasite growth and survival. Collectively, this information led me to the thesis hypothesis that **TgGCN5-B is an essential HAT that resides within a unique, multi-subunit complex in the parasite nucleus**. This hypothesis was examined by performing the following specific aims: 1) Determine the nuclear localization signal of TgGCN5-B, 2) Identify the proteins associating with TgGCN5-B, and 3) Evaluate the impact of TgGCN5-B on parasite physiology.

Chapter 2: Materials and Methods

I. Tissue culture and parasite techniques

A. Host cell and parasite culture

Toxoplasma is a ubiquitous pathogen capable of infecting a multitude of host cells, including most mammalian cells frequently used in tissue culture research. For all studies presented in this work, human foreskin fibroblasts (HFF; ATCC #SCRC-1041) were used because they are large and flat, enabling an ease in monitoring parasite growth through microscopy. HFF cells also have strong contact inhibition allowing for confluent monolayers to be prepared and stored prior to passage of parasites. Additionally, since HFF cells have no or minimal growth after reaching confluency, they are resistant to metabolic inhibitors, allowing for drug selection of parasites [182]. The disadvantage of HFF cells is that they are a primary cell line, so they will senesce at higher passage numbers. Recently, another cell line has been developed, HTERT-HFF (human telomerase reverse transcriptase; ATCC #CRL-4001), which retain the characteristics of HFFs but are immortal [183]. For routine passage of parasites and generating parasite lysates, HTERT cells were used as they produced higher parasite yields as compared to HFF cells (personal observation). For all assays, HFF cells were utilized as these are the established cell line in the field.

Standard tissue-culture techniques were followed, and all work was performed under sterile conditions within a laminar flow hood. For maintenance of HFF and HTERT cells both lines were grown in Dulbecco's Modified Eagle's Medium (DMEM, Gibco #19625-126) containing 10% heat inactivated fetal bovin serum (FBS, Gibco #16000). To split cells, a confluent monolayer from a large flask (T75 or T150) was washed once in sterile phosphate buffer saline (PBS, Gibco #10010-049) to remove residual media. Next, the monolayer was disrupted using 0.25% Trypsin-EDTA (Gibco #25200) digestion for approximately 30 seconds at 37°C, followed by manual agitation. The trypsin was deactivated by resuspension of cells in host cell media. Cells were then seeded into the appropriate number of flasks or plates depending on downstream applications. All host cells were grown at 37°C in a humidified incubator containing 5% CO₂. Cells typically reached confluency in 4 – 10 days depending on cell type, inoculation conditions, and passage number. HFF cells grew slower than HTERT cells

(personal observation). HFF cells were maintained in culture to approximately passage 15, whereas HTERT could be kept through passage 30.

For all the work presented in this thesis, RH strain (type I lineage) derived *Toxoplasma* tachyzoites were used. Parasites were maintained via serial passage onto confluent host cell monolayers in *Toxoplasma* media, DMEM supplemented with 1% heat-inactivated FBS. A lower amount of serum must be used with parasite culture in order to minimize the exposure of parasites to antibodies or complement [182]. Upon host cell lysis, parasite lines were serially passed onto another confluent monolayer of host cells. Typically, parasites double inside the host cells every 7.5 hours. After lysis, parasites must be passed onto the next monolayer fairly quickly because the parasites have a limited extracellular half-life of approximately 10 hours [182]. Parasite lysis occurred every 2 – 4 days depending on inoculation. Parasites were grown at 37°C in a humidified incubator containing 5% CO₂. Both parasite and host cell growth was monitored using an inverted microscope with phase-contrast optics. All tissue culture and parasitic waste was decontaminated with either bleach or 70% alcohol and autoclaved prior to disposal.

Contamination of tissue culture is a serious complication that can destroy cell or parasitic lines or alter experimental results. Therefore, cautions were taken to ensure sterile tissue culture techniques were carried out. This included filtration of reagents through 0.2 µm filters and use of sterile equipment. Antimicrobials are typically included in tissue culture media as a preventative to ward off potential contaminants. However, continual use of antimicrobials could inadvertently select for resistant contaminants and thereby complicate treatment options if contamination arises. Therefore, antimicrobials were not routinely used in either host cell or *Toxoplasma* media. When contamination was suspected or arose in another area of the lab, gentamicin (5 µg/ml, Gibco #15710-072) was included in the media until the issue was resolved. If parasites or host cells had overt contamination, they were immediately destroyed and new stocks initiated.

An inconspicuous contaminate of tissue culture is *Mycoplasma*. *Mycoplasma* is a free-living bacterium that lacks a cell wall and is extremely small, making it undetectable using basic microscopy. It is estimated that 15 – 80% of long-term tissue culture cell lines are contaminated by this nuisance [184]. *Mycoplasma* contamination has been documented in cultures of *Plasmodium*, and its presence can result in false or blemished data interruption [185]. Therefore, both host cells and parasite cultures were

routinely tested for the presence of *Mycoplasma* using the MycoAlert *Mycoplasma* Detection kit (Lonza #LT07-318) per the manufacturer's protocol.

Mycoplasma was never detected in my culture samples; however, there are a few treatment options if it were to arise. First, contaminated parasites can be passed through mice and then re-isolated and put back into tissue culture. For this technique, the mouse's immune system will eradicate the *Mycoplasma* infection, and the parasites can be re-isolated from the peritoneal fluid. *Toxoplasma* cultures contaminated with *Mycoplasma* can also be treated with *Mycoplasma* Removal Agent (MRA, MP Biomedicals #093050044), an inhibitor of bacterial gyrase, per the manufacturer's guidelines [183,185].

B. Freezing and thawing of host cells and parasites

In order to preserve stocks of both low passage host cells and all parasite lines, cryopreservation in liquid nitrogen was used for long-term storage. In both cases, DMSO (dimethyl sulfoxide) was used as a cryoprotectant in a freeze mix that included 25% DMSO and 20% FBS in DMEM. To freeze host cells, a large flask (T150 or T75) was trypsinized, as described above, to dislodge the cells from the monolayer. Cells were next resuspended in host cell media and centrifuged at 400 x *g* for 10 minutes at 4°C to pellet the cells. Following centrifugation, the supernatant was discarded and the cellular pellet was resuspended in an equal volume of cold host cell media and freeze mix (2.0 ml each for T150 and 1.0 ml each for T75). The cellular mixture was immediately aliquoted into cryovials (Simport #T311-1), 0.5 ml each, and placed at -80°C. Following at least an overnight incubation at -80°C, the cryovials were transferred to liquid nitrogen for long-term storage. In a similar fashion, parasites that had lysed approximately 80% of the host cell monolayer from a T25 flask were scraped and centrifuged at 400 x *g* for 10 minutes at 4°C. For freezing *Toxoplasma*, it is best to harvest the parasites prior to complete host cell lysis, because the host cells provide an added protection during the freeze/thaw process. Following centrifugation, the supernatant was discarded, and the parasite pellet was resuspended in 0.5 ml each of cold *Toxoplasma* media and freeze mix. The mixture was then aliquoted (0.5 ml each) between two cryovials (Simport #T311-1), which were placed at -80°C overnight prior to extended storage in liquid nitrogen. To thaw both host cells and parasites, cryovials were quickly removed from liquid nitrogen and thawed in a 37°C water bath. Host cells were inoculated into a T75 flask containing 30 ml host cell media, whereas parasites

were inoculated into a T25 flask with a confluent host cell monolayer containing *Toxoplasma* media. Following inoculation, flasks were incubated as described earlier for standard culture. Following overnight incubation, the media of the flask was replaced to remove any residual DMSO, which could reduce viability of both host cells and parasites.

C. Parasite transfection

Toxoplasma is highly amenable to genetic manipulation, and a variety of tools have been developed to assist in both forward and reverse genetics in the haploid tachyzoites. *Toxoplasma* has a high frequency of non-homologous random integration of vectors, allowing for stable expression of transgenes within the parasites. Additionally, mutant strains of *Toxoplasma* have been engineered to assist in genetic manipulation. For instance, the RH Δ Ku80 strain has been engineered for high efficiency in homologous recombination through disruption of the gene Ku80, which is an integral component of the nonhomologous end-joining (NHEJ) pathway. This strain is ideal for disruption of genomic loci and also allows genes to be endogenously tagged [186,187].

In this thesis, two methods were used to genetically manipulate the parasites: expression of transgenes or the tagging of an endogeneous locus. In general, to express transgenes, the parasite strain RH Δ HX was utilized [188]. RH Δ HX parasites lack the gene HXGPRT (hypoxanthine-xanthine-guanine phosphoribosyltransferase; HX) which is involved in purine synthesis (Figure 4). Therefore, the parasites are resistant to 6-thioxanthine (6TX) since they are unable to scavenge xanthine. To produce purines, RH Δ HX parasites are reliant on *de novo* synthesis through the enzyme IMPDH (inosine monophosphate dehydrogenase), which is inhibited by mycophenolic acid (MPA). When used for transfection purposes, RH Δ HX parasites are given back the gene HXGPRT via the transfection vector. Addition of MPA will inhibit IMPDH and select for parasites that incorporated the transfected DNA, which included HXGPRT. Xanthine (XAN) is also given as a supplement, since it is the cofactor scavenged by the parasites [188]. Although this is the most common selection scheme used, other selection agents exist. Parasites can be given the gene chloramphenicol acetyltransferase (CAT) allowing for resistance to chloramphenicol (CAM) [189]. Additionally, a mutant version of the enzyme DHFR-TS (dihydrofolate reductase-thymidylate synthetase) can be used to confer resistance to pyrimethamine (PYR) [190]. Table I summarizes the above

selection schemes and lists the concentrations of each selection agent included in the parasite media.

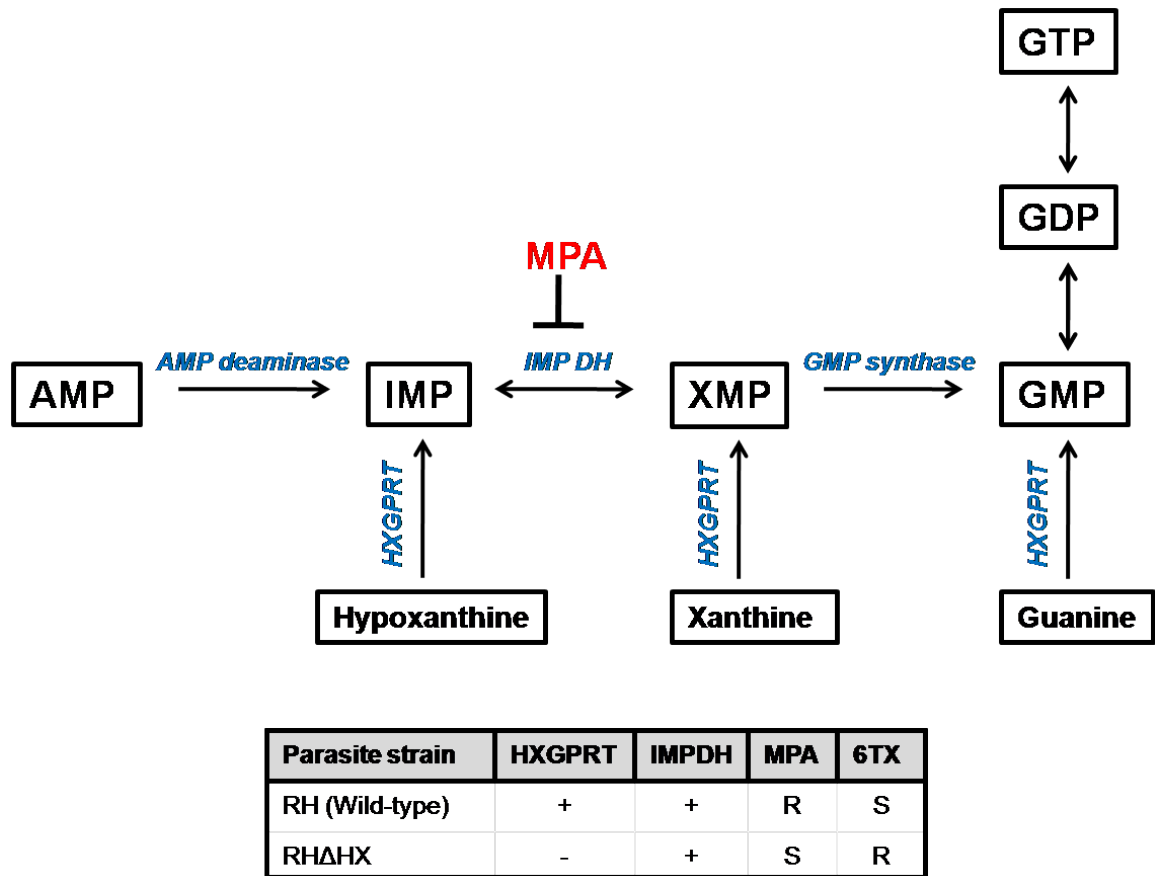


Figure 4: Summary of guanine biosynthesis in *Toxoplasma* and the manipulation of the pathway for selection purposes. The enzyme HXGPRT (hypoxanthine-xanthine-guanine phosphoribosyltransferase) is not essential for *Toxoplasma* as parasites lacking this enzyme are able to synthesize guanine nucleotides through AMP deaminase and IMPDH (inosine monophosphate dehydrogenase). Inhibition of IMPDH by addition of MPA (mycophenolic acid) results in the parasites being dependent on HXGPRT and xanthine for the production of guanine nucleotides. Therefore, the addition of MPA to RHΔHX parasites results in death, unless the parasites incorporated HXGPRT through transfection. Inversely, RHΔHX parasites are resistant to 6TX (6-thioxanthine), due to the lack of HXGPRT. This scheme allows for both positive selection with MPA and negative selection through 6TX. The table summarizes these points for both wild-type (RH) and RHΔHX parasites with + or – indicating the presence or absence of a given gene respectively [188]. R signifies resistance and S represents sensitivity.

Table I: Selection agents for *Toxoplasma*

Selection Agent	Concentration	Dilutant/Vehicle	Additional Information
Mycophenolic acid (MPA)	25 µg/ml	100% ethanol	For RHΔHX parasites when HXGPRT gene is utilized for selection Must include Xan as a supplement
Xanthine (XAN)	50 µg/ml	fresh 0.5M KOH	Used in combination with MPA
6-thioxanthine (6TX)	320 µg/ml	fresh 0.5M KOH	For RHΔHX parasites prior to transfection to verify strain
Chloramphenicol (CAM)	20 µM	100% ethanol	CAT gene utilized for selection
Pyrimethamine (PYR)	1 µM	70% ethanol	DHFR* gene utilized for selection
Shield-1 (Shld)	250 nM - 1 µM	100% ethanol	Used to stabilize any protein tagged with the DD

* indicates a mutant form of the enzyme DHFR that is resistant to pyrimethamine [190].

DD, destabilization domain.

To endogenously tag a genomic locus, the parasite strain RHΔKu80 was utilized. As mentioned earlier, this strain has been engineered to have a high efficiency in homologous recombination as compared to normal *Toxoplasma* strains [186,187]. RHΔKu80 parasites were a gift from the lab of Dr. Vern Carruthers (University of Michigan).

For simplicity, this section will generally describe the method of parasite transfection to introduce foreign DNA. All vectors designed for parasite transfection have a Bluescript pKS plasmid backbone (Stratagene) that contains an ampicillin resistance gene for selection in *E. coli* bacteria (used for construction and propagation). Other essential elements of the parasite expression vectors include a selectable marker for parasite expression and 5' and 3' flanking UTR (untranslated region) sequences for expression of the gene of interest. Table II lists all parasite vectors used and includes their properties and purpose.

Table II: *Toxoplasma* transfection vectors

Plasmid name	Strain / Selection	5'UTR / 3'UTR	Description
<i>ptubX_{FLAG}::HX</i>	Must be in RHΔHXGPRT parasites Gives back HXGPRT gene conferring resistance to MPA	Tubulin / DHFR	For over-expression of gene of interest (X). Contains a C-terminal FLAG epitope sequence NotI to linearize
<i>ptubX_{FLAG}::CAT</i>	For any parasite strain Confers resistance to chloramphenicol through CAT	Tubulin / DHFR	For over-expression of gene of interest (X). Contains a C-terminal FLAG epitope sequence NotI to linearize
<i>pLIC.3xHA::DHFR</i>	For use in RHΔKu80 parasites Confers resistance to PYR by providing the parasites with a resistant form of DHFR*	N/A	Contains the sequence necessary for LIC. Used to endogenously tag a genomic locus with 3X HA epitope at the C-terminus Insert must contain unique RE site for linearization
<i>pLIC.2xHA-DD::DHFR</i>	For use in RHΔKu80 parasites Confers resistance to PYR by providing the parasites with a resistant form of DHFR*	N/A	Contains the sequence necessary for LIC. Used to endogenously tag a genomic locus with 2X HA epitope followed by the DD at the C-terminus Insert must contain unique RE site for linearization

All *Toxoplasma* transfection vectors contain a Bluescript pKS⁺ backbone (Stratagene) conferring ampicillin resistance for selection in bacteria.

* The *Toxoplasma* minigene DHFR (dihydrofolate reductase) contains two point mutations, Ser(36) → Arg and Thr(83) → Asn, that confer resistance to pyrimethamine [190]. RE, restriction enzyme; LIC, ligation independent cloning; DD, destabilization domain; UTR, untranslated region; HX, hypoxanthine-xanthine-guanine phosphoribosyltransferase; CAT, chloramphenicol acetyltransferase; PRY, pyrimethamine.

Prior to transfection, vector DNA was linearized via restriction digestion overnight to facilitate recombination. Following digestion, DNA was sterilized by ethanol precipitation in which 2.5X cold 100% ethanol and 0.1X 3M NaOAc, pH 5.2 was added to the digested vector (X = volume of digestion), followed by mixing and incubation at -20°C for at least 30 minutes. The DNA was then pelleted by centrifugation at 16,000 x g at 4°C for 10 minutes, washed once in 70% ethanol, and then centrifuged again under the same conditions. The 70% ethanol wash was discarded while in the laminar flow hood to ensure DNA sterility. The DNA pellet was allowed to dry for approximately 2 hours to remove residual ethanol. The DNA was then used either immediately for parasite transfection or stored at -20°C until needed.

For transfection, parasites from a freshly lysed T25 flask were purified through a sterile 3 μ M polycarbonate 25 mm filter (Whatman #110612) and harvested via centrifugation at 400 x *g* for 10 minutes at room temperature. The supernatant was removed, and the parasite pellet was resuspended in 10 ml of cytomix, an electroporation buffer designed to mimic the parasites' intracellular environment. Cytomix contains 120 mM KCl, 0.15 mM CaCl₂, 10 mM K₂HPO₄/KH₂PO₄ (pH 7.6), 25 mM HEPES (pH 7.6), 2 mM EDTA, 5 mM MgCl₂ [191]. Cytomix was made and sterilized prior to transfection and stored at 4°C. Parasites were counted using a hemocytometer (Appendix A). Meanwhile, supplemented cytomix was prepared fresh by adding 15 mg ATP and 18 mg glutathione to 12.5 ml cytomix (final concentration of 2 mM ATP and 5 mM glutathione), following by sterilization with a 0.22 μ M filter (Fisher #09-7194A). The DNA pellet was resuspended in 100 μ l of supplemented cytomix and allowed to sit at room temperature until needed. Following counting, the parasites were centrifuged again under the same conditions, and the pellet was resuspended to 2.0×10^7 parasites per 0.3 ml in supplemented cytomix. Next, 0.3 ml of parasites was mixed with the DNA for a final volume of 0.4 ml that was placed in a sterile 2 mM gap electroporation cuvette (Fisher #BTX620). Parasites were electroporated in a BTX ECM 630 with a single 1.5 kV pulse with a resistance setting of 25 ohms and a capacitor setting of 25 μ F. Following electroporation, the parasites were left undisturbed for 15 minutes at room temperature, then inoculated into 2 – 4 T25 flasks of confluent monolayers containing *Toxoplasma* media, and incubated as described earlier. Parasites were put on drug selection 24 hours after transfection or following the first lysis. The concentration of each selection agent used is listed in Table I.

D. Cloning by limiting dilution

If a clonal, homogenous parasite population was required following transfection, the parasites were seeded into 96-well plates containing confluent host cell monolayers at a concentration of approximately a single parasite per well. This allowed for a single parasite to generate a clonal population that could be used for down-stream applications. Prior to cloning, transfected parasites were passed under drug selection at least twice. Freshly lysed parasites from a T25 flask were harvested by scraping and purified through a 3 μ M polycarbonate 25 mm filter (Whatman #110612). The concentration of parasites was determined by counting 4 times using a hemocytometer (Appendix A). Next, the parasites were diluted 1:1000 (10 μ l into 10 ml *Toxoplasma*

media), and based on the parasite count, and factoring in the 1:1000 dilution, the parasites were further diluted into 20 ml *Toxoplasma* media with drug to obtain approximately 1 parasite per well (usually 10 – 30 μ l of 1:1000 dilution was added to 20 ml *Toxoplasma* media). Subsequently, the latter dilution was distributed into a 96-well plate containing confluent HFF cells. The plates were incubated as described earlier and not disturbed for 7 days. Typically, two 96-well plates were set-up for each flask of transfected parasites. After incubation, the plates were inspected for single parasite plaques, representing a clonal population. If multiple plaques were seen in a well, the well was discarded. Once wells containing single plaques were identified, they were allowed to grow until approximately 50% of the well was lysed (about 14 days). At this point, the well was scraped with a pipet tip, and parasites were inoculated into a 12-well plate for expansion. Typically 10 – 20 clones were selected between all the 96-well plates (3 – 6 per plate). Each clone was expanded and evaluated (via immunofluorescence assay or genomic PCR) for desired phenotype (i.e. expression of transgene).

E. Harvesting parasites and generating lysates

Typically, parasites were harvested from T150 flasks of hTERT-HFF host cells because approximately 10^9 to 10^{10} parasites could be obtained from a single flask, about 10-fold greater than from a T150 of parasites in HFF host cells (personal observation). To inoculate a T150 flask for harvesting, 4 – 5 ml lysed parasites were inoculated into the flask and allowed to incubate for approximately 36 – 40 hours as previously described. When the parasites began to lyse the host cell monolayer, the flask was scraped, and the parasites were purified from host cell debris by filtration through a 3 μ M polycarbonate 47 mm filter (Whatman #111112) into a conical tube. The parasites were pelleted by centrifugation at 400 x g for 10 minutes at 4°C. The media was aspirated away from the parasite pellet, which was subsequently washed in PBS and centrifuged again under the same conditions. Next, the PBS was aspirated and the parasite pellet was resuspended in 1 ml PBS, transferred to an eppendorf tube, and centrifuged for 10 minutes at 1.0 x g at 4°C. After centrifugation, this final PBS wash was aspirated away and the parasite pellet was stored at -80°C until needed. The freezing and thawing of parasite pellets assisted in the lysing process.

Alternatively, some experiments required harvesting of intracellular parasites, in order to preserve parasitic processes such as protein complex formation. For these

experiments, a slightly modified harvest was utilized. When parasite vacuoles were large (greater than 100 parasites) and engorged, the flask was scraped and the parasites were purified from host cell debris by filtration through a 3 μ M polycarbonate filter into a conical tube. The parasites were pelleted by centrifugation at 400 x *g* for 10 minutes at 4°C. The media was aspirated away from the parasite pellet, which was subsequently washed in PBS and centrifuged again under the same conditions. Next, the PBS was aspirated and the parasite pellet was resuspended in 1 ml 1% paraformaldehyde (PFA), diluted in PBS, and allowed to incubate at room temperature for 8 minutes. The addition of 1% PFA served to cross-link any protein-protein interactions. Next, 125 mM glycine was added (100 μ l from 1.25 M stock glycine solution) to quench the PFA reaction. The mixture was incubated at room temperature for an additional 5 minutes, followed by centrifugation at 1.0 x *g* at room temperature to pellet the parasites. The parasite pellet was then washed two times in PBS with similar centrifugations at 4°C prior to being stored at -80°C.

The preserved parasite pellets were used to generate lysates that could be used for immunoprecipitation or analyzed via SDS-PAGE and western blotting. The parasite pellet was first thawed on ice for approximately 10 minutes. Next, the pellet was resuspended in 500 μ l Lysis Buffer (50 mM Tris, pH 7.5, 150 mM NaCl, 0.1% Nonidet P40) supplemented with mammalian protease inhibitor cocktail (10 μ l per 1 ml Lysis Buffer; Sigma-Aldrich #P8340). The mixture was then incubated at 4°C with rocking for 15 minutes. Next, the lysate was sonicated 3 times for 15 seconds at 30% power with a 30 second recovery on ice between each round using an Ultrasonic processor sonicator. Sonication was used to lyse the parasites and to shear the genomic DNA. Next, all insoluble material was separated from soluble lysate by centrifugation at 16,000 x *g* for 10 minutes at 4°C. The soluble lysate was removed from the insoluble pellet into a new, pre-chilled tube and quantified for protein concentration using Bio-Rads' *DC* protein assay kit (#500-0111), per the manufacturer's instructions. To determine protein concentration, a standard curve was established using a range of known concentrations of albumin (albumin standard, Thermo Scientific #23210) diluted in the Lysis Buffer.

F. Generating parasite nuclear-enriched lysates

To prepare nuclear-enriched lysates, parasites were harvested from a T150 as described previously (Chapter 2, Section I-E). The parasite pellet was resuspended in 2 ml cytoplasmic lysis buffer (10 mM HEPES, pH 7.9, 10 mM KCl, 2 mM MgCl₂, 1 mM

DTT) supplemented with 20 μ l mammalian protease inhibitor cocktail (Sigma #P8340). The mixture was incubated on ice for 15 minutes followed by the addition of NP-40 to a final concentration of 0.6%. The mixture was vortexed vigorously 3 times for 15 seconds with a 30 second recovery on ice after each round. Next, the mixture was centrifuged at 2,800 x *g* for 10 minutes at 4°C. The supernatant representing the cytoplasmic fraction was removed. Next, the pellet (nuclear fraction) was resuspended in 2 ml nuclear lysis buffer (50 mM HEPES, pH 7.9, 50 mM KCl, 300 mM NaCl, 1 mM DTT, 10% glycerol) supplemented with 20 μ l mammalian protease inhibitor cocktail (Sigma #P8340). The mixture was incubated at 4°C with rocking for 20 minutes followed by sonication (3 times for 15 seconds at 30% power with 30 second recovery on ice after each round). Following sonication, the mixture was centrifuged for 10 minutes at 2,800 x *g* at 4°C. The supernatant was the nuclear-enriched fraction. The nuclear-enriched fraction was dialyzed into an appropriate solution for downstream applications. For use in affinity chromatography, the nuclear fraction was dialyzed in MBP wash buffer (50 mM Tris (pH 7.5), 150 mM NaCl, 10% glycerol).

II. Molecular biology techniques

A. General PCR protocol

All PCRs were assembled in a designated area using PCR-only specific supplies and equipment to avoid contaminating DNA. All reactions were assembled on ice. Primers were ordered from either Invitrogen or IDT (Integrated DNA Technologies) and resuspended in ddH₂O to a concentration of 100 μ M. Table III lists all the primers used, as well as a description for each. PCRs were performed with Phusion™ High-fidelity DNA polymerase (NEB #F530S), a proof-reading enzyme that produces blunt-ended products, per the manufacturer's guidelines. A 50 μ l reaction was assembled containing 10 μ l 5X GC Buffer (preferred due to the GC-rich nature of the *Toxoplasma* genome), 1.5 μ l DMSO, 1 μ l 10 mM dNTPs, 10 μ M both forward and reverse primers, 10 – 20 ng template DNA (added after all the above reagents) and 0.5 μ l Phusion™ DNA polymerase (always added last). The reaction was mixed gently by tapping followed by a brief centrifugation. After assembly, the reaction was placed in a thermocycler (Eppendorf Mastercycler) with a pre-heated lid temperature of 104°C.

The general thermocycler protocol is as follows:

1. 98°C for 30 seconds
2. 98°C for 10 seconds
3. T_m for 30 seconds

To calculate the T_m , the following formula was used based on the nucleotide sequence of the primers: $[2*(A + T)] + [4*(G + C)]$

Once the T_m for each primer was determined, the lowest T_m minus 5°C was used

4. 72°C for 30 seconds per kilobase of amplicon
5. Repeat steps 2 – 4 for a total of 30 cycles
6. 72°C for 10 minutes
7. 4°C indefinitely

Upon completion of the thermocycler program, the PCR reaction products were analyzed via agarose gel electrophoresis using a 0.8% agarose gel containing 1 µg/ml ethidium bromide. For agarose gel electrophoresis, the running buffer was 1X TBE diluted from a 10X TBE stock (89 mM Tris-base, pH 8.0, 89 mM Boric acid, 20 mM EDTA, pH 8.0). The gel was run at approximately 117V for about 30 minutes. To visualize the PCR products, the gel was viewed under low UV light (365 nm), and the desired band was excised using a clean razor blade. The gel slice containing the appropriate DNA was purified using a gel extraction kit (Invitrogen #K2100-12) per the manufacturer's protocol. The purified PCR product had one of two fates. It was either 1) cut directly with restriction enzymes and ligated into a destination vector or 2) cloned into pCR®4Blunt-TOPO® (Invitrogen #K2875-J10) following the manufacturer's instructions. In general, if the amplicon was to clone a new fragment from genomic or cDNA, then the PCR was sub-cloned into pCR®4Blunt-TOPO® so the product could be confirmed by sequencing at ACGT, Inc. (Wheeling, IL). Primers used to sequence the amplicon in pCR®4Blunt-TOPO® were M13 forward and M13 reverse. Alternatively, if the PCR product was amplified from plasmid DNA, then usually it was directly ligated into its destination vector.

Table III: Description of primers for PCR

#	Description	Orientation	Primer Sequence (5' → 3')
1	Ndel-FLAG-GCN5-Bstart	Sense	CATATGAAAATG(F)GCGCCTTCAGAGTGTCCCAGC
2	GCN5-Bend-AvrII-stop	Anti-sense	CCTAGGCTAGAAAAATGTCGGATGCTTCGCGCCCAAGCC CCTCGTCTCC
3	Ndel-HA-myc-GCN5-B-start	Sense	CATATGAAAATG(H)(M)GCGCCTTCAGAGTGTCCCAGCGAC
4	GCN5-Bend-AvrII (no stop)	Anti-sense	CCTAGGGAAAAATGTCGGATGCTTCGCGCC
5	Ndel-FLAG-Δ310-GCN5-B	Sense	CATATGAAAATG(F)AGGCCCGCGGAGAACAAGAAGCGCGGC AGAGACGCCTGCGAGGGC
6	Ndel-FLAG-Δ313-GCN5-B	Sense	CATATGAAAATG(F)GAGAACAAGAAGCGCGGCAGAGACGCC TGCGAGGGCCCTCAGGCG
7	Ndel-FLAG-Δ310-R311A-GCN5-B	Sense	CATATGAAAATG(F)GCGCCCGCGGAGAACAAGAAGCGCGGC AGAGACGCCTGCGAGGGC
8	Ndel-FLAG-Δ310-P312A-GCN5-B	Sense	CATATGAAAATG(F)AGGGCGGCGGAGAACAAGAAGCGCGGC AGAGACGCCTGCGAGGGC
9	Ndel-FLAG-Δ310-RP311/312AA-GCN5-B	Sense	CATATGAAAATG(F)GCGGCGGCGGAGAACAAGAAGCGCGGC AGAGACGCCTGCGAGGGC
10	Inverse PCR to create ΔNLS	Sense	GCTAGCGACGCCTGCGAGGGCCCTCAGGCG
11	Inverse PCR to create ΔNLS	Anti-sense	GCTAGCCGCCTGGGCGAGCGGCGGTGC
12	BglII-β-gal-start	Sense	AGATCTAAAATGGCGGTCGTTTTACAACGTCGTGACTGGGAA
13	β-gal-end-AvrII (no stop)	Anti-sense	CCTAGGTTTTTGACACCAGACCAACTGGTAATG
14	β-gal-end-NLS-AvrII (no stop)	Anti-sense	CCTAGGTCTGCCGCGCTTCTTGTCTCCGCGGGCCTTTTTTG ACACCAGACCAACTGGTAATG
15	β-gal-end-NLS-HA-AvrII	Anti-sense	CCTAGGCTACGCGTAGTCCGGGACGTCGTACGGGTATCTGC CGCGCTTCTGTCTCCGCGGGCCTTTTTTGACACCAGACCA ACTGGTAATG
16	BglII-NLS-β-gal-start	Sense	AGATCTAAAATGGCGAGGCCCGCGGAGAACAAGAAGCGCGG CAGAGTCGTTTTACAACGTCGTGACTGGGAA
17	Ndel-HA-Imp-α-start	Sense	CATATGAAAATG(H)ATGGAGCGCAAGTTGGCCGATCGTCGAT CG
18	Imp-α-end-stop-AvrII	Anti-sense	CCTAGGCTACTGGCCGAAGTTGAAGCCTCCCTGAGGCGGCG CTGC
19	BamHI-GCN5-Bstart (MBP fusion)	Sense	GGATCCGCGCCTTCAGAGTGTCCCAGCGAC
20	GCN5-Bend-stop-HindIII	Anti-sense	AAGCTTTCCTAGAAAAATGTCGGATGCTTCGC
21	LIC-AHook-056400 fwd	Sense	<u>TACTTCCAATCCAATTTAATG</u> CTTACCACCCACGGTGAACG CTTTGATCTGC
22	LIC-AHook-056400 rev	Anti-sense	<u>TCCTCCACTTCCAATTTTAGC</u> ACCGCGAGCAGGCACCACGAA ACCGTCCACTGGACC
23	LIC-AP2-3816 fwd	Sense	<u>TACTTCCAATCCAATTTAATG</u> CGAACTCTGCTGTACCGTGTC GGGAGATACG
24	LIC-AP2-3816 rev	Anti-sense	<u>TCCTCCACTTCCAATTTTAGC</u> GGCAGTGGCGGGTTCTCCAC CTCAAGACATGATGAGAGC
25	LIC-AP2-3948 fwd	Sense	<u>TACTTCCAATCCAATTTAATG</u> CAGAGTACGGCCGAATGCTTCA AAAGTGG

26	LIC-AP2-3948 rev	Anti-sense	<u>TCCTCCACTTCCAATTTTAGC</u> AAAGTCTTCGTCAACAACGAAC TTGCGAGTGC
27	SDM GCN5-B E703G fwd	Sense	CAGCAGAAATTCGCC GGC ATCGCTTTCCTCGCG
28	SDM GCN5-B E703G rev	Anti-sense	CGCGAGGAAAGCGAT GCC GGCGAATTTCTGCTG
29	NdeI-DD	Sense	CATATGAAAATGGCGGGAGTGCAGGTGGAAACCATCTCC
30	DD-Nsi-AvrII	Anti-sense	CCTAGGATCGATATGCATTTCCGGTTTTAGAAGCTCCACATC GAAGACGAGAGTGGC
31	Nsi-HA-GCN5-Bstart	Sense	ATGCAT(H)GCGCCTTCAGAGTGTCCCAGCGACGCG
32	LIC-GCN5-B fwd	Sense	<u>TACTTCCAATCCAATTTAATG</u> CCGTTCTCAAACCTGTGAGTC G
33	LIC-GCN5-B rev	Anti-sense	<u>TCCTCCACTTCCAATTTTAGC</u> GAAAAATGTCGGATGCTTCGC GCCCACAAGC
34	B1 assay fwd	Sense	GGAGGACTGGCAACCTGGTGTCTG
35	B1 assay rev	Anti-sense	TTGTTTCACCCGGACCGTTTAGCAG
36	pLIC vectors sequencing	Sense	ACCTCACTCATTAGGCACCCAGG
37	ptubX _{FLAG} ::HX sequencing	Sense	TCAGGACGCTTGCGCTCATCGC
38	pMAL-c2X sequencing	Sense	AACCTCGGGATCGAGGGAAGG
39	GCN5-B 1928 sequencing	Sense	AGATGCCAGAGAGTACATTGTCCGTCTCG
40	GCN5-B 656 sequencing	Sense	CAACGCGAGCATTGGAGACATTGC
41	GCN5-B 2828 sequencing	Sense	ACACGACTGCGCAGATGTTTGCGGACG
42	GCN5-B 2852 sequencing	Sense	ACGAAGTTCAGTTGATGTTCAAGAA
43	DD sequencing	Anti-sense	TGGGTGCCCAGTGGCACCAT
44	AP2-3816 sequencing	Sense	AGGTTTCGCGAGGCCTTAATGC
45	AP2-3948 sequencing	Sense	ACCTGCTGCTCCTCAAGTGCAGG

(F) = FLAG epitope sequence GACTACAAGGACGACGACGACAAG

(H) = HA epitope sequence TACCCGTACGACGTCCCGGACTACGCG

(M) = c-myc epitope sequence GAGCAGAAGCTCATCTCTGAGGAGGACCTC

LIC for sense primer TACTTCCAATCCAATTTAATGC

LIC for anti-sense primer TCCTCCACTTCCAATTTTAGC

Shading = nucleotides mutated for site-directed mutagenesis

B. Ligation, bacterial transformation, and plasmid preparation

To ligate inserts into destination vectors, two approaches were used. The first involved cloning the insert into pCR®4Blunt-TOPO® followed by sequence verification. The insert was then removed from pCR®4Blunt-TOPO® through restriction enzyme digestion. The products of the digestion were resolved via agarose gel electrophoresis as previously described. The desired band was excised and purified using a gel extraction kit (Invitrogen #K2100-12). The purified DNA was then ready for ligation. Alternatively, a PCR could be digested with restriction enzymes directly after gel purification. Following this digestion, the enzymes and buffers could be removed from the DNA by either agarose gel electrophoresis or phenol/chloroform extraction followed by ethanol precipitation. The DNA was then ready to ligate into its destination vector. The destination vector was also digested with restriction enzymes and purified in the same manner. All restriction enzymes and appropriate buffers were from NEB (New England Biolabs). Restriction digestion reactions were performed at recommended temperatures, usually 37°C for at least one hour.

To extract DNA via phenol/chloroform treatment, an equal volume of phenol:chloroform:isoamyl alcohol (25:24:1, USB #75831) was added, mixed, and centrifuged at 16,000 x *g* for 3 minutes. The top, aqueous layer contains the DNA and was carefully removed to a new tube. Next, an equal volume of pure chloroform was added, mixed, and centrifuged at 16,000 x *g* for 2 minutes. Again, the top layer (containing the DNA) was removed to a new tube and mixed with 2.5X 100% cold ethanol, 0.1X 3M NaOAc, pH 5.2, and 0.01X glycerol, where X is the volume of starting material (from phenol/chloroform extraction). The solution was mixed and incubated at -20°C for at least 30 minutes to precipitate the DNA. Following incubation, the mixture was centrifuged at 16,000 x *g* for 10 minutes at 4°C to pellet the DNA. The DNA pellet was washed once in 70% ethanol and centrifuged again at 16,000 x *g* for 5 minutes. This wash was removed and the pellet was allowed to dry before resuspension in 10 mM Tris, pH 8.0.

To ligate the purified insert into the purified vector, a quick ligation protocol employing T4 DNA ligase was utilized (NEB #M2200S). First, both the vector and insert DNA were quantified with a NanoDrop® spectrophotometer (ND-1000). For the ligation reaction, approximately 10 ng vector was combined with 50 ng insert in a 10 µl volume completed with ddH₂O. To this mixture, 10 µl of 2X reaction buffer was added along with 1 µl of T4 DNA ligase. The reaction was mixed and allowed to incubate at room

temperature for 10 minutes. Following incubation, the ligated vector was transformed into *E. coli*.

Bacterial transformation into *E. coli* was performed following cloning into pCR®4Blunt-TOPO® or after ligation of an insert into its destination vector. Basically 2.5 µl of reaction mixture was gently added to 50 µl One Shot® TOP10 chemically competent *E. coli* (Invitrogen #C4040-03). The cells were gently mixed by tapping and incubated on ice for 20 minutes. Following the incubation, the cells were heat shocked at 42°C for 30 seconds, followed by 2 minutes recovery on ice. Next, 250 µl S.O.C. media (Invitrogen #15544034) was added to the cells, followed by a one hour incubation at 37°C with shaking at 250 rpms. Next, the transformed *E. coli* bacteria were plated on LB agar plates containing appropriate antibiotic selection. Plates were incubated at 37°C overnight. All pCR®4Blunt-TOPO® reactions required kanamycin selection (50 µg/ml) while the destination vectors used ampicillin selection (50 µg/ml). Table IV lists all vectors constructed and the primers used to amplify each insert.

To test for bacteria containing the desired plasmid DNA, between 4 – 8 bacterial colonies were selected from each plate and inoculated into 2 ml liquid LB with appropriate antibiotic selection. The cultures were allowed to grow between 8 and 16 hours at 37°C with shaking at 250 rpms. After sufficient culture growth, 1.5 ml of each culture was used to isolate the plasmid DNA with 5 Prime's FastPlasmid Mini Kit (#2300000) per the guidelines of the manufacturer. Following purification, the plasmid DNA was screened via restriction digestion, followed by agarose gel electrophoresis. Any plasmid displaying the predicted bands following restriction digestion was sent for sequence verification at ACGT, Inc. (Wheeling, IL). The primers used to sequence all destination vectors are listed in Table III. A glycerol stock from the original bacterial culture was also made for all positives plasmids. To make the glycerol stock, 250 µl bacterial culture was added to a sterile eppendorf tube and a large drop of sterile glycerol was added (approximately 25 µl or 10%). The glycerol stock was stored at -80°C and was used to inoculate additional cultures for propagation of the plasmid.

For a large scale preparation of plasmids, a starter culture of 2 ml was inoculated from the glycerol stock and grown at 37°C with shaking at 250 rpms for 8 hours. Next, 100 µl of the starter culture was used to inoculate a 50 ml culture that was subsequently incubated overnight at 37°C. This large culture was used for plasmid isolation with GenElute™HP Plasmid Maxiprep Kit (Sigma-Aldrich #NA0310) per the manufacturer's guidelines. Upon completion, the elutant was centrifuged at 16,000 x *g* for 7 minutes to

remove residual resin from the purified plasmid DNA. After centrifugation, the supernatant was ethanol precipitated as previously described, and the DNA was resuspend in 0.5 ml 10 mM Tris, pH 8.0. The DNA was then quantified with a NanoDrop® spectrophotometer (ND-1000) and saved at -20°C until needed for downstream applications.

Table IV: Plasmids

#	Plasmid Name	Parental vector	Insert	Primers
1	TOPO®-GCN5-B	pCR®4Blunt-TOPO®	FL GCN5-B	1 & 2
2	FLAG GCN5-B	<i>ptubX</i> _{FLAG} ::HX	FL GCN5-B	1 & 2
3	HA-myc GCN5-B _{FLAG}	<i>ptubX</i> _{FLAG} ::HX	FL GCN5-B	3 & 4
4	FLAG Δ310GCN5-B	<i>ptubX</i> _{FLAG} ::HX	Δ310GCN5-B	5 & 2
5	FLAG Δ313GCN5-B	<i>ptubX</i> _{FLAG} ::HX	Δ313GCN5-B	6 & 2
6	TOPO®-GCN5-BΔNLS [†]	pCR®4Blunt-TOPO®	GCN5-BΔNLS	10 & 11
7	FLAG GCN5-BΔNLS	<i>ptubX</i> _{FLAG} ::HX	GCN5-BΔNLS	1 & 2 *
8	FLAG Δ310-R311A-GCN5-B	<i>ptubX</i> _{FLAG} ::HX	Δ310-R311A-GCN5-B	7 & 2
9	FLAG Δ310-P312A-GCN5-B	<i>ptubX</i> _{FLAG} ::HX	Δ310-P312A-GCN5-B	8 & 2
10	FLAG Δ310-RP-AA-GCN5-B	<i>ptubX</i> _{FLAG} ::HX	Δ310-RP-AA-GCN5-B	9 & 2
11	β-gal _{FLAG}	<i>ptubX</i> _{FLAG} ::HX	FL β-gal	12 & 13
12	β-gal-NLS _{FLAG}	<i>ptubX</i> _{FLAG} ::HX	FL β-gal + NLS	12 & 14
13	β-gal-NLS _{HA}	<i>ptubX</i> _{FLAG} ::HX	FL β-gal + NLS	12 & 15
14	NLS-β-gal _{FLAG}	<i>ptubX</i> _{FLAG} ::HX	NLS + FL β-gal	16 & 13
15	HA Imp-α	<i>ptubX</i> _{FLAG} ::HX	FL Imp-α	17 & 18
16	AT-hook-056400 _{HA}	<i>pLIC</i> .3xHA::DHFR	last 1,130 bp w/o TGA	21 & 22
17	MBP-GCN5-B	<i>pMAL</i> -c2X (NEB)	FL GCN5-B	19 & 20
18	AP2-3816 _{HA}	<i>pLIC</i> .3xHA::DHFR	last 1,530 bp w/o TGA	23 & 24
19	AP2-3948 _{HA-DD}	<i>pLIC</i> .2xHA-DD::DHFR	last 1,014 bp w/o TAA	25 & 26
20	TOPO®-DD-linker [♦]	pCR®4Blunt-TOPO®	DD	29 & 30
21	TOPO®-DD- _{HA} GCN5-B	pCR®4Blunt-TOPO®	DD & FL GCN5-B	31 & 2
22	TOPO®-DD- _{HA} GCN5-B DN [‡]	pCR®4Blunt-TOPO®	DD & FL GCN5-B; SDM for DN	27 & 28
23	DD- _{HA} GCN5-B-WT	<i>ptubX</i> _{FLAG} ::CAT	DD & FL GCN5-B	29 & 2
24	DD- _{HA} GCN5-B-DN	<i>ptubX</i> _{FLAG} ::CAT	DD & FL GCN5-B DN	29 & 2

FL = full-length; bp = base pairs; w/o = without; TGA, TAA = stop codons; DD = destabilization domain; WT = wild-type; SDM = site-directed mutagenesis; DN = dominant-negative allele

† Refer to Chapter 2, Section II-D for details on construction of this plasmid.

*To generate the insert for the construct *ptubX*_{FLAG} GCN5-BΔNLS::HX, the plasmid TOPO®-GCN5-BΔNLS was used as the template DNA along with primers 1 & 2 to amplify GCN5-B lacking only the 10 residue NLS (aa 311-320).

♦ The linker is an NsiI restriction site followed by 6 random nucleotides and an AvrII restriction site. This allows a gene to be cloned downstream and in-frame with the DD using the NsiI and AvrII restriction sites.

‡Site-directed mutagenesis was performed on plasmid #21 (TOPO®-DD-_{HA} GCN5-B) with primers 27 & 28 to mutate the glutamic acid at position 703 to glycine (E703G), thus creating the dominant-negative allele.

Note: Plasmids #21 and #22 were used as template to amplify inserts for construction of plasmids #23 and #24, respectively.

C. Ligation independent cloning

Ligation independent cloning (LIC) was an alternative method used to ligate inserts into certain destination vectors [192]. This method of cloning was utilized to construct all parasite transfection vectors used for endogenously tagging genes in *Toxoplasma*. These vectors all contain LIC in their name (Table II) and were a gift from Dr. Michael White (University of South Florida). LIC eliminates the need for traditional restriction enzyme digestion followed by ligation. Instead, the LIC method involved generating insert PCR products that incorporate conjoining sequences at the termini. After treatment with T4 DNA polymerase in the presence of a single deoxyribonucleotide triphosphate, overhangs were generated that were complementary to overhangs produced after vector treatment. The complementary overhangs allowed for sufficient annealing prior to transformation into bacteria. After transformation, repair enzymes within the bacteria completed the ligation of the plasmid [192]. This process is highly efficient and allowed a single insert to be shuttled between several vectors without the need to switch restriction enzyme sites.

For LIC, the method developed and modified by Huynh et al. (2009) was utilized [187]. PCR inserts were generated as previously described using Phusion™ High-fidelity DNA polymerase (NEB #F530S) with sense and anti-sense primers listed in Table III. The LIC complementary conjoining sequences are underlined for each primer sequence. The DNA template was *Toxoplasma* genomic DNA (gDNA) prepared as described in the next section (Chapter 2, Section II-D) and diluted in ddH₂O. To endogenously tag a gene, approximately 1 – 2 kB genomic fragment immediately upstream of the stop codon (but not including the stop codon) was amplified. This fragment must contain a unique restriction site at least 350 bp from either end, in order to linearize the vector prior to parasite transfection. Following amplification of the desired fragment by PCR, the insert was gel purified as previously described and quantified with a NanoDrop® spectrophotometer (ND-1000). Next, 0.2 pmol insert was treated with T4 DNA polymerase (Invitrogen #18005017) in the following 20 µl reaction: 4 µl 5X buffer, 1 µl 100 mM DTT, 0.8 µl 100 mM dCTP (Invitrogen #10217016), 0.5 µl T4 DNA polymerase, 8.7 µl ddH₂O, 5 µl insert. The reaction was mixed on ice and then placed in a thermocycler (Eppendorf Mastercycler) for the following program:

1. 22°C for 30 minutes
2. 75°C for 20 minutes
3. 4°C indefinitely

The insert was either used immediately or stored at -20°C until needed.

To prepare the vector, approximately 5 µg vector DNA was digested with PacI overnight at 37°C. After digestion, the vector DNA was subjected to phenol:chloroform extraction, followed by ethanol precipitation as described previously. Next, the linearized vector was treated with T4 DNA polymerase (Invitrogen #18005017) in the following 60 µl reaction: 12 µl 5X buffer, 3 µl 100 mM DTT, 2.4 µl 100 mM dGTP (Invitrogen #10218014), 1.5 µl T4 DNA polymerase, 35.1 µl ddH₂O, 6 µl vector (200 – 300 ng/µl). The reaction was mixed on ice and then placed in a thermocycler. The same program described above for the insert was used for the vector DNA as well. At the completion of the thermocycler program, the vector DNA was either used immediately or stored at -20°C until needed.

To anneal the T4-treated vector and insert DNA, 1 µl treated vector was combined with 2 µl treated insert. The mixture was incubated at room temperature for 10 minutes followed by the addition of 1 µl 25 mM EDTA with an additional incubation of 5 minutes at room temperature. Next, 1 µl of the annealing reaction was transformed into 50 µl One Shot® TOP10 chemically competent *E. coli* (Invitrogen #C4040-03) as previously described. The remainder of the LIC construct preparation is the same as the plasmid preparations described earlier (Chapter 2, Section II-B).

D. Inverse PCR

Inverse PCR was the method used to delete the NLS (30 nucleotides) from TgGCN5-B cDNA. The inverse PCR is shown in Figure 5. Template DNA was pCR®4Blunt-TOPO® containing full-length TgGCN5-B cDNA. Primers #10 and #11 were used to amplify the entire plasmid excluding the 30 nucleotides containing the NLS. The PCR reaction utilized Phusion™ High-fidelity DNA polymerase (NEB #F530S) and was set-up as previously described, except an extended elongation time of 4 minutes was used to ensure the entire plasmid was amplified. Each primer incorporated a NheI site at the termini. Therefore, following PCR amplification, the purified DNA was digested with NheI overnight at 37°C. The digested vector was next subjected to a rapid ligation reaction with T4 DNA ligase (NEB #M2200S) as described earlier. The remainder of the plasmid preparation was as previously described (Chapter 2,

Section II-B). The 30 nucleotides encoding the NLS of TgGCN5-B were replaced with a NheI restriction site (6 nucleotides that encode for the amino acids Ala and Ser).

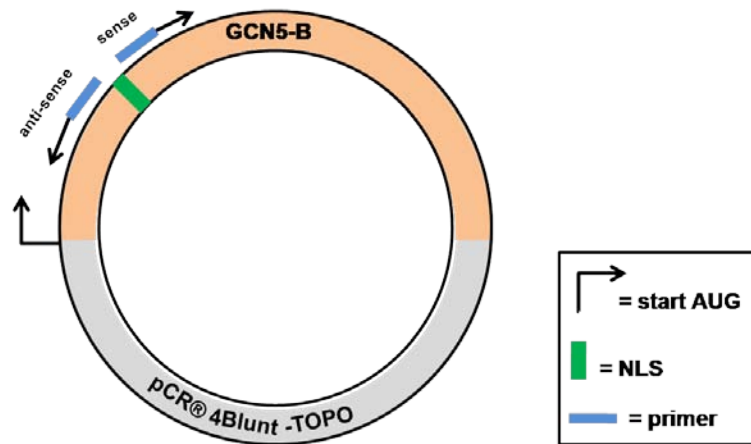


Figure 5: Inverse PCR to remove the nucleotides encoding the NLS of TgGCN5-B.

Primers #10 (sense) #11 (anti-sense) were used to amplify the entire plasmid pCR®4Blunt-TOPO® containing full-length TgGCN5-B cDNA excluding the 30 nucleotides encoding the NLS. The plasmid was re-ligated after digestion with NheI, a restriction site incorporated into each primer. This procedure replaced the NLS nucleotides of TgGCN5-B with a NheI site.

E. Genomic PCR from parasites

The method described below utilized parasite DNA as a template for genomic PCR. This protocol was used to amplify genomic fragments for vector construction or to screen parasite clones for the incorporation of genetic information into the genome. In general, only a small amount (approximately $1 \times 10^{5-6}$) of parasites were needed to have enough material for PCR. For this protocol, 1 ml of lysed parasites was transferred to a sterile eppendorf tube. The parasites were pelleted by centrifugation at $3.3 \times g$ for 5 minutes at room temperature. The media was aspirated away from the parasite pellet, and these pellets were either frozen until needed at -80°C or processed directly for genomic PCR.

To process parasite pellets for genomic PCR, the pellets were thawed on ice (if frozen) and then resuspended in 49 μl Genomic PCR Lysis Buffer (50 mM KCl, 10 mM Tris pH 8.3, 2.5 mM MgCl_2). Next, 0.5 μl each of 1% SDS and ProteinaseK (20 mg/ml) was added to each lysate. All lysates were mixed and incubated at 55°C overnight in a hybridization oven. The following day, the lysates were heated at 95°C for 10 minutes to

deactivate the ProteinaseK. The lysates were now ready to be used as templates in the PCR reactions.

General PCR procedures were used to amplify inserts from parasite gDNA and were performed as previously described (Chapter 2, Section II-A). All parasite genomic PCRs for screening purposes were performed with *Taq* DNA polymerase (Invitrogen #10342-053). The following formula is a master mix for four reactions. If there were more reactions, this recipe could be appropriately manipulated. Master Mix for *Taq* parasite genomic PCR included 79.5 µl H₂O, 10.0 µl 10X *Taq* PCR buffer, 1.0 µl each primer (100 µM), 1.0 µl dNTPs (10 mM), 3.0 µl MgCl₂, and 0.5 µl *Taq* DNA polymerase. All reactions were assembled on ice. A 25 µl volume total PCR reaction was obtained by adding 24 µl master mix with 1 µl of treated lysate. PCRs were then placed in a thermocycler (Eppendorf Mastercycler) with the lid temperature pre-incubated to 104°C. The following thermocycler protocol was performed:

1. 95°C for 2 minutes
2. 95°C for 1 minute
3. T_m for 2 minutes (See T_m calculation in Chapter 2, Section II-A)
4. 72°C for 2 minutes
5. Repeat steps 2 – 4 for a total of 35 cycles
6. 72°C for 10 minutes
7. 4°C hold

After conclusion of the thermocycler protocol, all reactions were analyzed via agarose gel electrophoresis as previously described. To view the PCR products, the gel was visualized under UV light.

F. Site-directed mutagenesis

Site-directed mutagenesis was utilized to mutate a single amino acid (3 nucleotides) in the HAT domain of TgGCN5-B that is critical for enzymatic activity. The mutation of the glutamic acid at residue 703 to glycine (E703G) results in a catalytically inactive enzyme or HAT-dead TgGCN5-B [89]. To perform the mutagenesis of this residue, Stratagene's QuikChange II XL Site-Directed Mutagenesis kit (#200521) was utilized and the manufacturer's recommendations were followed. The first step was to use PCR to synthesize the mutated strand. The template DNA was pCR®4Blunt-TOPO® containing full-length TgGCN5-B cDNA (plasmid #21). The reaction included 10 ng of template DNA, 1 µl 1:10 dilution of primers #27 and #28, 5 µl 10X reaction buffer,

1 μ l dNTP mix, 3 μ l QuikSolution, 1 μ l *PfuUltra* DNA polymerase, and ddH₂O to complete the volume to 50 μ l. The reaction was mixed on ice and then placed in a thermocycler (Eppendorf Mastercycler; lid preheated to 104°C) for the following program:

1. 95°C for 1 minute
2. 95°C for 50 seconds
3. 60°C for 50 seconds
4. 68°C for 8 minutes
5. Repeat steps 2 – 4 for 18 cycles
6. 68°C for 7 minutes
7. 4°C indefinitely

Following completion of the PCR thermocycler program, 1 μ l DpnI was added to the PCR mixture, and the reaction was placed at 37°C overnight. DpnI is an endonuclease that specifically cuts methylated and hemimethylated DNA. The purpose of this step is to digest the parental DNA template leaving only the newly synthesized mutated DNA. After DpnI digestion, 2 μ l of the digested PCR reaction was transformed into 50 μ l One Shot® TOP10 chemically competent *E. coli* (Invitrogen #C4040-03) as previously described. The remainder of the site-directed mutagenesis preparation is the same for the plasmid preparations described earlier. To confirm the mutated residues, plasmids were sequenced by ACGT, Inc. (Wheeling, IL).

III. Biochemical techniques

A. Immunofluorescence assay

Immunofluorescence assays (IFA) were used to monitor parasite expression of epitope-tagged transgenic proteins using commercial antibodies as well as native proteins with custom-made antibodies. Parasites (50 – 100 μ l from lysed T25 flask or approximately $1 \times 10^{4-5}$) were inoculated into a 12-well plate with coverslips containing confluent HFF cell monolayers. The parasites were incubated at previously described conditions (Chapter 2, Section I-A) for 18 – 24 hours. Following incubation, the parasite media was removed and the coverslips were washed 3 times in PBS to remove any extracellular parasites or debris. Next, the cells were fixed and permeabilized using 100% cold methanol for 10 minutes at -20°C, followed by 3 washes in PBS. After the PBS washes, the coverslips were blocked in 3% bovine serum albumin (BSA, Sigma-Aldrich #A3059) diluted in PBS for 2 hours at room temperature or overnight at 4°C.

Following blocking, the primary antibody was diluted in 3% BSA in PBS and incubated for 1 – 2 hours at room temperature, followed by 3 washes in PBS. Table V provides descriptions, including the dilutions, of primary and secondary antibodies used for IFA. The secondary antibody, which is conjugated to a fluorescent dye, was diluted in 3% BSA in PBS and incubated on the coverslips for 1 hour at room temperature in the dark. For the remainder of the IFA protocol everything was done in the dark to prevent photobleaching. The coverslips were washed 3 times in PBS followed by application of 4',6-diamidino-2-phenylindole (DAPI; Invitrogen #D3571) diluted to 0.2 μ M in PBS for 5 minutes at room temperature. After 3 final washes in PBS, coverslips were mounted onto a glass slide using 8 μ l mounting reagent containing 50% glycerol with Mowiol 4-88 (Calbiochem #81381) and DABCO (Sigma #10981) to reduce photobleaching. Slides were allowed to dry prior to being viewed with a Leica DMLB scope with a 100X HCX Plan Apo oil immersion objective. All images were captured with a monochrome SPOT-RTSE (model 12) camera and Spot Diagnostic software (version 7.0) and pseudocolored using Adobe Photoshop 7.0.

The majority of the IFAs presented in this manuscript were fixed and permeabilized with cold 100% methanol at -20°C for 10 minutes. However, this method of fixation/permeabilization can distort the parasite morphology. For the purpose of visualization of nuclear versus cytoplasmic protein distribution, a slight distortion of parasite morphology does not alter the results, which is why methanol fixation was still used. Alternatively, another method for fixation and permeabilization exists that preserves the parasite morphology. This method uses paraformaldehyde (PFA) to fix the parasites and Triton-X to permeabilize. When analyzing parasites for cell-cycle and division, this alternative was utilized. In general, after washing infected coverslips in PBS, cells were fixed in 3% PFA diluted in PBS for 15 minutes at room temperature. After 3 washes in PBS, the cells were quenched by adding 1 ml 0.1M glycine in PBS for 5 minutes at room temperature. Cells were again washed in PBS 3 times, followed by permeabilization with 0.2% Triton X-100 diluted in 3% BSA in PBS for 10 minutes at room temperature. After 3 washes in PBS, the cells were blocked in 3% BSA and the remainder of the IFA is as described above.

Table V: Antibodies used in IFA and Western blotting

Antibody Description	Source	Dilution
Primary Antibodies		
Anti-HA (clone 3F10) rat, monoclonal	Roche #11867423001 #11815016001 (affinity resin)	WB or IFA: 1:1000 – 2000
Anti-FLAG® M2 mouse, monoclonal	Sigma #F1804	IFA: 1:2000
Anti-FLAG® rabbit, polyclonal	Sigma #F7425 #A4596 (M1 affinity gel)	WB: 1:1000
Anti-c-myc (clone 9E10) mouse, monoclonal	Roche #11667149001	WB: 1:1000
Anti-acetyl histone H3 rabbit, polyclonal	Upstate #06-755	WB: 1:1000
Anti-histone H3 rabbit, polyclonal	Upstate #06-755	WB: 1:1000
Anti-acetyl lysine rabbit, polyclonal	Stressgen #KAP-TF120	WB: 1:500*
Anti-Sag1 <i>Toxoplasma</i> mouse, monoclonal	Meridian Life Science, Inc. #C86319M	IFA: 1:4000
Anti-TgGCN5-B (peptide) rabbit, polyclonal from terminal bleed (not affinity purified)	QBC, custom antibody designed to last 30 amino acids	WB or IFA: 1:1000
Anti-TgIMC1 <i>Toxoplasma</i> mouse, gift from Dr. Michael White (University of South Florida)	Custom antibody	IFA: 1:2000
Anti-TgPCNA <i>Toxoplasma</i> rabbit polyclonal, gift from Dr. Michael White (University of South Florida)	Custom antibody	IFA: 1:5000
Anti-MBP mouse, monoclonal	NEB #E80325	WB: 1:5000
Secondary Antibodies		
ECL rat IgG, HRP-linked (from goat)	GE Healthcare # NA935	WB: 1:5000
ECL mouse IgG, HRP-linked (from goat)	GE Healthcare #NA931	WB: 1:5000
ECL rabbit IgG HRP-linked (from goat)	GE Healthcare # NA934	WB: 1:5000
Alexa Fluor® 488 anti-rat, IgG (from goat)	Molecular Probes # A-11006	IFA: 1:2000
Alexa Fluor® 488 anti-mouse, IgG (from goat)	Molecular Probes # A-11017	IFA: 1:2000
Alexa Fluor® 488 anti-rabbit, IgG (from goat)	Molecular Probes # A-11070	IFA: 1:2000
Alexa Fluor® 594 anti-mouse, IgG (from goat)	Molecular Probes # A-11072	IFA: 1:2000

WB, Western blot; HA, hemagglutinin; MBP, maltose binding protein; ECL, enhanced luminol-based chemiluminescent; HRP, horseradish peroxidase

The dilutant for WBs was either 5% non-fat dry milk or 3% BSA in TBST. Most antibodies were diluted in milk although occasionally BSA was required (denoted by *). The dilutant for all IFAs was 3% BSA in PBS.

B. Immunoprecipitations

Immunoprecipitations were typically performed using parasite lysates to isolate epitope-tagged proteins. Parasite lysates were generated and quantified as described in Chapter 2, Section I-E. Typically 300 – 500 µg of lysate was used for each immunoprecipitation reaction. The majority of proteins were expressed with an HA epitope tag, although a few had a FLAG tag. A resin that contains antibody coupled to beads exists for both tags (Roche anti-HA affinity matrix #11815016001 and Sigma Aldrich Anti-FLAG® M1 Agarose Affinity Gel #A4596). Prior to being added to lysates, the resin was washed 3 times in Lysis buffer (50 mM Tris, pH 7.5, 150 mM NaCl, 0.1% Nonidet P40) and then approximately 50 µl of slurry (resin/Lysis buffer mixture) was added to each lysate. The mixtures were incubated overnight at 4°C with rocking to immunoprecipitate the epitope-tagged protein. The next day, the resins were washed 3 – 5 times in Lysis buffer for 10 minutes each. At this point, the resin/immunoprecipitated material was ready for down-stream applications such as HAT assays or analysis via SDS-PAGE and western blotting. To prepare the resin/immunoprecipitated material for SDS-PAGE, the resin was resuspended in 50 – 100 µl of 1X SDS load dye. The master mix for the 1X SDS load dye includes 300 µl ddH₂O, 100 µl 4X NuPAGE Loading Dye (Invitrogen #NP0007), 20 µl beta-mercaptoethanol (β-Me; Sigma #M7154) and was made fresh each time. Next, the samples were heated at 95°C for 10 minutes. The combination of the SDS load dye and the heating disassociates the immunoprecipitated material from the resin. The samples were then ready for analysis via SDS-PAGE.

If a resin (antibody coupled to beads) did not exist for a given antibody to be used in an immunoprecipitation reaction, then usually 5 – 10 µl of antibody (depending on concentration, if available) was added directly to lysate (see Table V). This was incubated overnight at 4°C with rocking. The next day, 50 µl agarose beads coupled to either Protein A (Roche #11719408001) or Protein G (Roche #11719416001) were added to each immunoprecipitation reaction and allowed to incubate for an additional hour at 4°C with rocking. The beads were washed 3 times in Lysis buffer prior to being added. Protein A has the ability to bind rabbit antibodies, whereas Protein G binds best to mouse antibodies. The addition of the beads coupled to either Protein A or G allows for the antibody/immunoprecipitated protein complex to be captured and purified from the rest of the lysate. The remainder of the immunoprecipitation protocol is the same as described above.

Another modification to immunoprecipitation reactions involves the stringency of the Lysis buffer and/or wash buffer. The concentration of NP-40 in the Lysis/wash buffer can be lowered to 0.05% or switched to another detergent such as Triton-X, which is less stringent.

C. SDS-PAGE and Western blotting

Prior to analyzing protein samples via gel electrophoresis, all samples were mixed with 4X NuPAGE Loading Dye (Invitrogen #NP0007) and 5% beta-mercaptoethanol (β -Me; Sigma #M7154). Typically, 30 μ l of sample was mixed with 10 μ l 4X loading dye and 2 μ l β -Me. Samples were then heated to greater than 70°C for 10 minutes prior to being resolved on 4 – 12% NuPAGE Bis-Tris SDS gels (Invitrogen #NP0335BOX). Each gel included 10 μ l of SeeBlue® Plus2 Protein Standard (Invitrogen #LC5925) to estimate molecular weight. The running buffers for electrophoresis contained either MES (2-Morpholinoethanesulfonic acid) or MOPS (3-(N-Morpholino) propanesulfonic acid). Each chemical has a different pKa accounting for a variance in running time and separation range of proteins. MES runs faster than MOPS, although MOPS provided better resolution of higher molecular weight proteins [193]. Therefore, MOPS was the preferred buffer. All buffers for SDS-PAGE and Western blotting were diluted from 20X stocks: 20X MOPS Buffer (50 mM MOPS, 50 mM Tris-base, 0.1% SDS, 1 mM EDTA, pH 7.7) or 20X MES Buffer (50 mM MES, 50 mM Tris-base, 0.1% SDS, 1 mM EDTA, pH 7.3). Following electrophoresis, gels could be processed for Western blotting (see below) or stained with Simply Blue™ Safe Stain (Invitrogen #LC6060) or silver-stained to visualize the proteins.

To silver-stain, the gel was fixed twice for 15 minutes in a solution of 30% ethanol and 10% acetic acid. The fixing reagent was removed, and the gel was next incubated for 30 minutes in Reducer, a thiosulfate buffer containing 10 mM sodium acetate, 30% ethanol, and 1 mg/ml sodium thiosulfate (pH 6.0 with acetic acid). The Reducer was made without the addition of sodium thiosulfate and stored at 4°C until needed. Prior to use, the Reducer was warmed to room temperature, and sodium thiosulfate was added. Following the incubation in the Reducer, the gel was washed 3 times in ddH₂O. Next, 50 mg silver nitrate was diluted in 50 ml ddH₂O. To this solution, 12.5 μ l of 37% formaldehyde was also added. The gel was incubated in the silver nitrate solution for 30 minutes. After this incubation, the gel was rinsed briefly in ddH₂O, prior to the addition of the Developer solution (2.5 % Na₂CO₃ + 25 μ l of 37% formaldehyde per 50 ml). After

approximately 2 minutes of incubation in Developer, the protein bands of the gel would appear. At this point, the reaction was quenched by the addition of 1% acetic acid. The gel was allowed to incubate an additional 5 – 10 minutes prior to visualization and recording.

For Western blotting, proteins from the gel were transferred to a nitrocellulose membrane (Invitrogen #LC2001) with a 1X dilution of Transfer Buffer from a 20X stock solution (25 mM Bicine, 25 mM Bis-Tris, 1 mM EDTA, pH 7.2). To ensure proper transfer, the membrane was stained with Ponceau S (Sigma-Aldrich #P7170-1L) to visualize proteins. Next, the membrane blot was blocked in either 5% non-fat dry milk or 3% BSA diluted in TBST (20mM Tris, pH 7.4, 150mM NaCl, 0.1% v/v Tween-20). Blocking was either overnight at 4°C or for at least 2 hours at room temperature. The preferred blocking reagent was non-fat dry milk, although some antibodies required BSA (where indicated, see Table V). BSA was a more sensitive blocking reagent, although more background occurred with its use. Immediately following blocking the primary antibody was applied, diluted in either 5% milk or 3% BSA in TBST. The primary antibody was incubated on the blot for at least one hour at room temperature, followed by 3 washes in TBST. Next, the secondary antibody diluted in either 5% milk or 3% BSA in TBST was applied to the blot and allowed to incubate for at least one hour at room temperature followed by 3 washes in TBST. Refer to Table V for antibody specifics including dilutions. All secondary antibodies were coupled to horseradish peroxidase (HRP). To visualize the Western, the blot was treated in the dark with Amersham's ECL™ detection reagent (#RPN2209) for one minute. Excess reagent was removed and the blot was wrapped in cellophane and placed in a metal film tray. In a photography dark room, the blot was exposed to High Sensitive Blue photographic film (RPS imaging #33-0810) for various times (30 seconds to 5 minutes). The film was processed through an automated developer.

D. HAT assays

In vitro HAT assays were used to evaluate the enzymatic activity of TgGCN5-B. For all work presented in this thesis, a non-radioactive method was utilized. Typically, HAT assays were performed “on-bead,” meaning with enzymes purified from immunoprecipitation reactions that were still attached to the resin. Prior to starting the assay, a 5X HAT buffer was prepared and included 250 mM Tris HCl, pH 8.0, 25% glycerol, 0.5 mM EDTA, 250 mM KCl, 5 mM DTT, 5 mM PMSF, and 50 mM sodium

butyrate. This buffer was aliquoted, stored at -80°C, thawed, and diluted when needed. For on-bead HAT assays, the IP resin was first washed once in 1X HAT buffer prior to being resuspended in 28 µl 1X HAT buffer. To this mixture, 1 µl recombinant histone H3 (1 µg/µl, Upstate #14-411) and 1 µl acetyl CoA (1 mM, Sigma-Aldrich #A2056) was added. The mixture was then incubated at 37°C for one hour with manual mixing every 10 minutes. After the completion of the incubation, 10 µl 4X NuPAGE Loading Dye (Invitrogen #NP0007) and 2 µl β-Me were added. The reactions were heated for 10 minutes at greater than 70°C and resolved using SDS-PAGE gel electrophoresis, followed by processing for Western blotting as previously described (Chapter 2, Section III-C). Anti-acetyl histone H3 antibody (Table V) was used for detection.

Alternatively, a similar reaction could be assembled using recombinant enzyme. In this case, purified enzyme was combined with 6 µl 5X HAT buffer, 1 µl recombinant histone H3, 1 µl acetyl CoA, and completed to a total volume of 30 µl with ddH₂O. The remainder of the procedure is the same as described above.

E. Bacterial Inductions

Bacterial inductions were utilized to generate proteins, such as the MBP-TgGCN5-B fusion protein used in affinity chromatography. This method also can be applied to the production of both GST- and His-tagged proteins. The methods described below represent a general protocol for bacterial inductions. For each individual protein that was purified, slight modifications might be needed to obtain optimal results.

After completion of vector construction for any plasmid destined for bacterial induction, the plasmid must be transformed into bacteria capable of inducing proteins. The preferred bacterial strain (and the one used for all MBP fusion constructs) was One Shot® BL21 Star™ (DE3) chemically competent *E. coli* (Invitrogen #C601003). The plasmid(s) were transformed into this bacterial strain as previously described. From the transformation plate, a single colony was picked and inoculated into a 2 ml LB culture (containing appropriate antibiotic) for growth at 37°C overnight. Part of this culture was used to make a glycerol stock (as previously described; Chapter 2, Section II-B) while the rest was used for a pilot induction. The pilot induction followed the same protocol as will be outlined for a large scale induction, except the induced culture was 10 ml instead of 100 ml. Pilot inductions were performed to ensure the bacteria were able to produce the protein of interest, test the conditions, and ensure the protein was soluble.

For the large scale induction, a 5 ml starter culture was inoculated from the glycerol stock and allowed to grow overnight at 37°C. The next day, 1 ml of the starter culture was used to inoculate a 100 ml culture. The 100 ml culture was grown at 37°C until an OD₆₀₀ of approximately 0.6 was reached. Next, the 100 ml culture was moved to a shaking incubator (250 rpms) at 15°C (in cold room) and allowed to acclimate to the new temperature for approximately 30 minutes. After acclimation, 0.5 mM IPTG (isopropyl β-D-1-thiogalactopyranoside) was added to the culture, which continued to incubate overnight at 15°C. Before the additional of IPTG, a 1 ml sample was removed from the culture that would serve as an uninduced control when testing the induction. The next day, another 1 ml sample was removed (induced sample), while the rest of the culture was pelleted by centrifugation (2,800 x g for 30 minutes). The media was decanted away from the bacterial pellets, which were subsequently stored at -20°C until needed.

The 1 ml samples were used to test the induction and ensure protein was produced. The bacteria from each sample were pelleted by centrifugation (16,000 x g for 2 minutes) followed by removal of the media away from the bacterial pellet. Next, the bacterial pellets were resuspended in 100 µl PBS and lysed by sonication (3 times for 15 seconds at 30% power following by 30 second recovery on ice after each round using an Ultrasonic processor sonicator). The insoluble protein was separated by centrifugation at 16,000 x g for 10 minutes at 4°C. The soluble fraction was removed to a new, pre-chilled tube on ice, while the insoluble pellet was resuspended in 100 µl 1X SDS load dye and sonicated again under the same conditions. At this point, both the soluble and insoluble fractions were ready to be analyzed via SDS-PAGE followed by staining with Simply Blue™ Safe Stain (Invitrogen #LC6060) to visualize the proteins. A Western blot could also be performed if necessary. If the induction was successful, then the larger bacterial pellets would be processed for harvesting and purifying the induced protein. The next section will outline this protocol for MBP-tagged proteins. Similar methods exist for other protein tags.

F. Purification of MBP fusion proteins

To purify MBP-fused proteins expressed in bacteria, the bacterial pellets were thawed on ice for approximately 20 minutes, followed by resuspension in 5 ml MBP column buffer (20 mM Tris-HCl, 200 mM NaCl, 1 mM EDTA). Protease inhibitors were added to each bacterial suspension and included 1 uM PMSF and 50 µl of Sigma's

protease inhibitor cocktail (#P8849). Next, the bacterial solution was sonicated 3 times for 30 seconds at 45% power with a 30 second recovery on ice after each round using an Ultrasonic processor sonicator. Following sonication, the solution was centrifuged at $2,800 \times g$ for 20 minutes to separate the soluble and insoluble fractions. The soluble supernatant was removed from the insoluble pellet. A 100 μ l sample from the soluble portion was saved for analysis (termed crude extract). To prepare for purification, the remaining soluble fraction was diluted with an equal volume column buffer.

To purify the MBP-fusion proteins, a column containing 1 ml amylose resin (NEB #E80215) was assembled with glass wool in a 10 ml syringe. The column was washed 3 times with 3 ml column buffer. Next, the diluted soluble fraction was passed over the column. The flow through from this step was saved for analysis. The column was next washed 4 times in 3 ml column buffer. The washes were also saved for analysis. To elute the bound protein from the column, an elution buffer consisting of column buffer supplemented with 10 mM maltose was utilized. The elution buffer was added to the column in 1 ml aliquots with each fraction collected individually. After completion of the purification, all samples including crude extract, flow through, washes, and elutions were analyzed via SDS-PAGE followed by staining with Simply Blue™ Safe Stain (Invitrogen #LC6060) to visualize the proteins. If the purification was successful, the desired elution fractions were combined and dialyzed in coupling buffer (0.1 M phosphate buffer, pH 7.0).

G. Affinity chromatography

The first step in affinity chromatography was to couple the desired protein to the Affi-Gel 15 resin (Bio-Rad #153-6051). To determine TgGCN5-B associating proteins, two Affi-Gel resins were prepared, one containing MBP alone and the other with MBP-GCN5-B. To prepare each resin, 1 ml of Affi-gel was utilized. Prior to coupling the Affi-Gel was washed once in 30 ml ice cold ddH₂O followed by a wash in 10 ml of cold coupling buffer (0.1 M phosphate buffer, pH 7.0). Next, the dialyzed protein (prepared as described in previous section) was added to the Affi-Gel resin and allowed to incubate overnight with rocking at 4°C. The next day, the unbound portion was removed and the Affi-Gel resin was washed in 10 ml 10 mM Tris, pH 7.5. This was followed by a single wash each in 10 ml 100 mM glycine, pH 3.0 and 10 ml 10 mM Tris, pH 8.8. The final 2 washes were in 10 ml 10 mM Tris, pH 7.5. The coupled Affi-Gel resin was stored in 5 ml 10 mM Tris, pH 7.5 containing 0.01% azide at 4°C until needed.

For affinity chromatography, the storage solution was removed from the coupled Affi-Gel resins by centrifugation at $2,800 \times g$ for 1 minute. The MBP resin was then washed 3 times in 5 ml MBP wash buffer (50 mM Tris, pH 7.5, 150 mM NaCl, 10% glycerol). Next, the dialyzed nuclear-enriched parasite lysate (Chapter 2, Section I-F) was added to the MBP resin and allowed to incubate at 4°C for 4 hours with rocking. This step was to pre-clear the nuclear-enriched lysate of any protein that would bind non-specifically to MBP. After incubation, the pre-cleared nuclear-enriched lysate was separated from the MBP-Affi-Gel resin by centrifugation and added to the MBP-GCN5-B-Affi-gel resin. The MBP-GCN5-B-Affi-Gel resin was washed 3 times in 5 ml MBP wash buffer prior to the addition of the pre-cleared nuclear-enriched lysate. The pre-cleared nuclear-enriched lysate was incubated with the MBP-GCN5-B-Affi-gel resin overnight at 4°C with rocking. After incubation with the nuclear-enriched lysate, each column was washed 3X in 5 ml MBP wash buffer prior to elution of the bound proteins.

To elute the proteins that bound each individual resin (MBP bound, and MBP-GCN5-B bound), 2 ml 100 mM glycine pH 3.0 was added to each resin, followed by incubation for 5 minutes at 4°C . The eluted protein fraction was separated from the resin by centrifugation and removed to a new tube containing 1.0 M Tris, pH 8.0, (10% of the eluted volume). The eluted portion was saved for analysis and downstream applications.

After eluting the bound proteins, each Affi-gel resin was washed once in 10 ml 10 mM Tris, pH 8.8 and 10 mM Tris, pH 7.5. The resin was stored in 5 ml 10 mM Tris, pH 7.5 containing 0.01% azide and stored at 4°C . The resins could be used multiple times.

Samples of both the eluted proteins from the MBP and MBP-GCN5-B resins were analyzed via SDS-PAGE followed by silver staining to visualize the proteins. The eluted fractions were also analyzed via mass spectrometry (in collaboration with Dr. W. Andy Tao, Purdue University) for identification (Appendix B). Protein identified in the MBP-GCN5-B sample but not the MBP sample was a potential TgGCN5-B interacting protein.

IV. *Toxoplasma* growth assays

Several assays were utilized to monitor the replication rate and survival of *Toxoplasma* under various conditions. When evaluating the growth of different mutants of *Toxoplasma*, multiple types of growth assays were used to validate the results.

A. Plaque assay

Toxoplasma plaque assays were performed in 12-well culture plates containing confluent monolayers of HFF cells. Prior to infection with the parasites, the host cell media in each well was replaced with *Toxoplasma* media. Freshly lysed parasites from each strain/mutant to be examined were counted 4 times using a hemocytometer (Appendix A). Based on the counts, the parasites were diluted in *Toxoplasma* media to a concentration of 1.0×10^5 parasites per 1 ml. From these dilutions, 5 μ l (corresponding to 500 parasites) was inoculated into each well in duplicate or triplicate. The parasites were incubated for 2 hours at previously described conditions (Chapter 2, Section I-A) to allow for invasion. Following incubation, the media and any parasites that did not invade were removed and replaced with *Toxoplasma* media containing treatment or vehicle control. Plates were incubated as previously described (Chapter 2, Section I-A) for 5 – 6 days. During the incubation period, it was important to not disturb the plates, as movement could distribute the parasites and cause the formation of secondary plaques. After 6 days, plates were examined using light microscopy to ensure that control parasites had formed visible plaques. Next, the media was removed, wells were washed once in PBS, and each well was fixed with 100% cold methanol at -20°C for 10 minutes. Following fixation and removal of the methanol, the plates were allowed to dry overnight at room temperature under ambient conditions. After drying, the host cell monolayer became opaque and the parasite plaques could be visualized as cleared, transparent areas among the nontranslucent background. Plaques could be counted and scored at this time, although if a difficulty in visualization occurred, the wells could be stained with crystal violet to enhance the contrast between the cleared parasite plaques and the host cell monolayer. Each individual well was counted for number of plaques three independent times, blind to treatment and strains. The data was recorded in Microsoft Office Excel, where subsequent statistical analysis was performed. To determine significant values, the student's t-test was utilized with p values < 0.01 being significant.

B. B1 gene detection assay to monitor *Toxoplasma* growth

To prepare for the B1 gene detection growth assay [194,195], freshly lysed parasites from each strain/mutant to be examined were counted 4 times using a hemocytometer (Appendix A). Parasites were diluted in *Toxoplasma* media to a concentration of 1.0×10^5 parasites per 1 ml. From these dilutions, 10 μ l per well

(corresponding to 1000 parasites) was inoculated into an entire row (6 wells) of a 24-well culture plate containing confluent HFF monolayers. Prior to parasite inoculation, the host cell media was replaced with *Toxoplasma* media. The parasites were incubated for 2 hours at previously described conditions (Chapter 2, Section I-A) to allow for invasion. Following incubation, the media and any parasites that did not invade was removed and replaced with *Toxoplasma* media containing treatment or vehicle control. Plates were incubated as previously described (Chapter 2, Section I-A). Every 24 hours a well (containing host cells and parasites) for each strain and treatment was harvested. The B1 gene is specific for *Toxoplasma*, so the presence of host cells will not affect the assay.

To harvest and subsequently purify the genomic DNA, Qiagen's DNeasy Blood and Tissue kit (#69506) was used. The lysis solution for harvesting the cells/parasites included 200 μ l PBS, 200 μ l AL buffer, and 20 μ l Proteinase K (the latter two items were provided in the kit) per well. After addition of lysis solution, each well was scraped with the end of a pipet tip to obtain all material, and the contents were stored in an eppendorf tube. Following harvesting, each well was viewed under a light microscope to ensure all material was obtained. Samples were stored at -20°C until completion of the assay. Harvests were performed through day 5 when control parasite strains typically began to lyse the host cell monolayer.

Upon completion of the assay, the genomic DNA was purified from each sample using Qiagen's DNeasy Blood and Tissue kit (#69506) per the manufacturer's protocol. Included in the genomic DNA purification was a parasite pellet containing a known quantity of parasites, which served as a standard for quantification. Following DNA purification, samples were heated at 95°C for 5 minutes to heat inactivate residual Proteinase K.

Quantitative real-time PCR was utilized to amplify the *Toxoplasma* B1 gene from each sample [194]. A 25 μ l reaction was assembled in triplicate for each sample. The reaction included 1 μ l 1:10 dilution of sample DNA, 12.5 μ l SYBR green mix (Applied Biosystems #4309155), 1.0 μ l of each primer (#34 and #35, diluted to 12.5 μ M), and 9.5 μ l ddH₂O. In addition to each sample, triplicate reactions were assembled for 1 μ l of DNA from 4 standards of known parasite quantities (10^2 , 10^3 , 10^4 , 10^5). The quantitative PCR was performed using the StepOnePlus™ Real-Time PCR System (Applied Biosystems) with data analysis provided by the system's software. To determine significant values, the student's t-test was utilized with p values < 0.01 being significant.

C. Doubling assay

To directly assess the difference between various *Toxoplasma* strains and treatments with respect to parasite replication rates, a doubling assay was performed. This assay monitors the number of parasite doublings as a function of time and thereby provides a means to monitor parasite growth at an individual level [182]. The doubling assay was performed in T25 flasks containing confluent monolayers of HFF host cells. Prior to infection with the parasites, the host cell media in each flask was replaced with *Toxoplasma* media. Freshly lysed parasites from each strain/mutant to be examined were counted 4 times using a hemocytometer (Appendix A). Parasites were diluted in *Toxoplasma* media to a concentration of 1.0×10^5 parasites per 1 ml. Then, 1 ml containing 1.0×10^5 parasites was inoculated into each T25 flask for each strain and each treatment. The parasites were incubated for 2 hours at previously described conditions (Chapter 2, Section I-A) to allow for invasion. Following incubation, the media, along with parasites that did not invade was removed from each flask and replaced with *Toxoplasma* media containing treatment or vehicle control. The flasks were incubated as previously described (Chapter 2, Section I-A). Every 12 hours, flasks were removed and examined via light microscopy. The number of individual parasites within 50 randomly selected vacuoles was counted for each time point. The assay was continued with counting every 12 hours until parasite vacuoles began to lyse. The doubling assay counts were performed blind to both strain and treatment. Data was recorded in Microsoft Office Excel. Since the parasite vacuoles were scored using a discontinuous numerical set, the data could not be analyzed by standard parametric statistics.

Chapter 3: Results

I. Aim 1: Determine how TgGCN5-B enters the parasite nucleus

A. TgGCN5-B requires its N-terminus to enter the parasite nucleus

Bhatti et al. (2006) initially discovered that *Toxoplasma* possesses two GCN5 homologues, which at the time was thought to be unique among protozoans [169]. In the initial characterization of these HATs, it was determined that parasites expressing a FLAG-tagged truncated TgGCN5-B lacking its N-terminal extension (_{FLAG}GCN5-BΔ528) excluded the protein from the parasite nucleus [169]. Scanning of the primary amino acid sequence of N-terminal region of TgGCN5-B revealed that this protein does not possess the same nuclear localization signal that is required for TgGCN5-A to enter the parasites nucleus (RKRVKR) [144]. However a basic-rich region, a hallmark characteristic of nuclear localization signals (NLSs), was identified from amino acid 316 – 319 (KKRGR). Bhatti et al. began to determine if this was the NLS of TgGCN5-B by designing various FLAG-tagged truncation mutants of TgGCN5-B, expressing these proteins in the parasite, and determining their cellular distribution via IFA [Sullivan and Bhatti et al., unpublished data]. The results are summarized in Figure 6. TgGCN5-B lacking the first 320 amino acids (_{FLAG}GCN5-BΔ320) showed cytoplasmic localization, whereas TgGCN5-B lacking only the first 304 amino acids (_{FLAG}GCN5-BΔ304) was able to enter the parasite nucleus. Unexpectedly, the truncation mutant that lacked the first 315 amino acids but still included the basic-rich stretch (_{FLAG}GCN5-BΔ315) was excluded from the parasite nucleus. Therefore, it was concluded that the five amino acid basic-rich stretch by itself was not sufficient to localize TgGCN5-B to the parasite nucleus but required additional upstream residues.

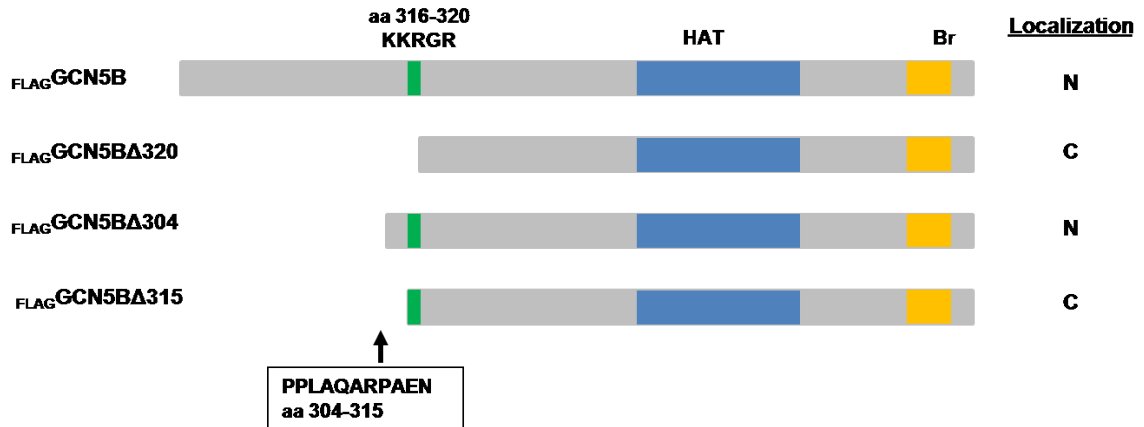


Figure 6: Preliminary mapping of TgGCN5-B NLS. The diagrams represent different truncation mutants of TgGCN5-B that were tagged N-terminally with the FLAG epitope and over-expressed in the parasites. The localization of each protein was determined via IFA and results are listed in the right-hand column, where N = nuclear and C = cytoplasmic. For each TgGCN5-B protein diagram, the blue region represents the HAT catalytic domain whereas the orange box signifies the bromodomain (Br). The basic rich stretch (KKRGR) from amino acids (aa) 316 – 320 is depicted by a green box. From these experiments it was concluded additional residues between amino acids 304 to 315 (in box) are necessary for proper nuclear localization [Sullivan and Bhatti et al., unpublished data].

B. TgGCN5-B contains a unique NLS within its N-terminus

To determine the full and necessary NLS of TgGCN5-B, I designed two additional truncation mutants to delineate which residues in the 11 amino acids upstream of the basic-rich stretch were required for TgGCN5-B to enter the parasite nucleus (Figure 7). Similar to the previous constructs, these two truncation mutants were FLAG-tagged N-terminally, expressed in the parasites, and visualized via IFA to determine their cellular distribution. The construct $\text{FLAGGCN5-B}\Delta 313$ (Plasmid #5), which begins with a glutamic acid residue at position 314 of TgGCN5-B (ENKKRGR), displayed a largely cytoplasmic distribution (Figure 7A). However, the construct $\text{FLAGGCN5-B}\Delta 310$ (Plasmid #4), which begins with the dipeptide RP at position 311 of TgGCN5-B (RPAENKKRGR), was able to enter the parasite nucleus (Figure 7B). These data revealed that the NLS of TgGCN5-B requires the 5 residues immediately upstream of the basic-rich stretch. Therefore, the complete NLS of TgGCN5-B is the 10 amino acids, RPAENKKRGR, from positions 311 to 320. When these 10 amino acids were

excised from full-length TgGCN5-B ($\text{FLAG GCN5-B}\Delta\text{NLS}$; Plasmid #7), the mutated protein was predominantly cytoplasmic (Figure 7C), validating that these 10 residues are necessary for nuclear localization.

Although the ectopic protein $\text{FLAG GCN5-B}\Delta 310$ was largely nuclear, the staining pattern did not entirely correspond to the DAPI staining pattern, which represents the DNA. Either a portion of $\text{FLAG GCN5-B}\Delta 310$ is located in the nuclear periphery or the nuclear membrane (areas not stained by DAPI), or something in the first 310 amino acids facilitates nuclear localization. TgGCN5-B is acetylated (Figure 16 D). It is possible that the acetylation of TgGCN5-B might facilitate the nuclear localization of the protein.

Next, I wanted to investigate the nature of the additional five amino acids upstream of the basic-rich stretch. I hypothesized that this upstream sequence could 1) be non-specific and just served as a “landing pad” for chaperone binding, or 2) the arginine and the proline, both common residues to NLSs, were necessary for the proper localization [140,196]. To examine these hypotheses, I designed mutated versions of the construct $\text{FLAG GCN5-B}\Delta 310$ in which I mutated select residues to alanines. As done previously, I expressed these mutated proteins in the parasites and determined the localization via IFA. Mutating the arginine at position 311 or the proline at position 312 individually to alanine in the constructs $\text{FLAG GCN5-B}\Delta 310\text{-R311A}$ (Plasmid #8) and $\text{FLAG GCN5-B}\Delta 310\text{P312A}$ (Plasmid #9), respectively, did not exclude these mutant proteins from the parasite nucleus (Figure 8A and B). However, it appears as if these mutant proteins linger in the nuclear periphery, particularly $\text{FLAG GCN5-B}\Delta 310\text{-R311A}$. It is possible that although the mutation of a single residue does not exclusively hinder nuclear localization, it could slow the process, resulting in the mutant proteins concentrating around the nuclear pore. The double mutation of both the arginine and proline together in the construct $\text{FLAG GCN5-B}\Delta 310\text{-RP311/312AA}$ (Plasmid #10) was excluded from the parasite nucleus (Figure 8C) indicating that the presence of either the arginine or the proline are required for nuclear localization. Therefore, it can be concluded that an additional basic residue (the arginine) and/or helix breaking residue (the proline) in the upstream sequence preceding the basic-rich stretch are necessary and contribute to the nuclear localization signal of TgGCN5-B.

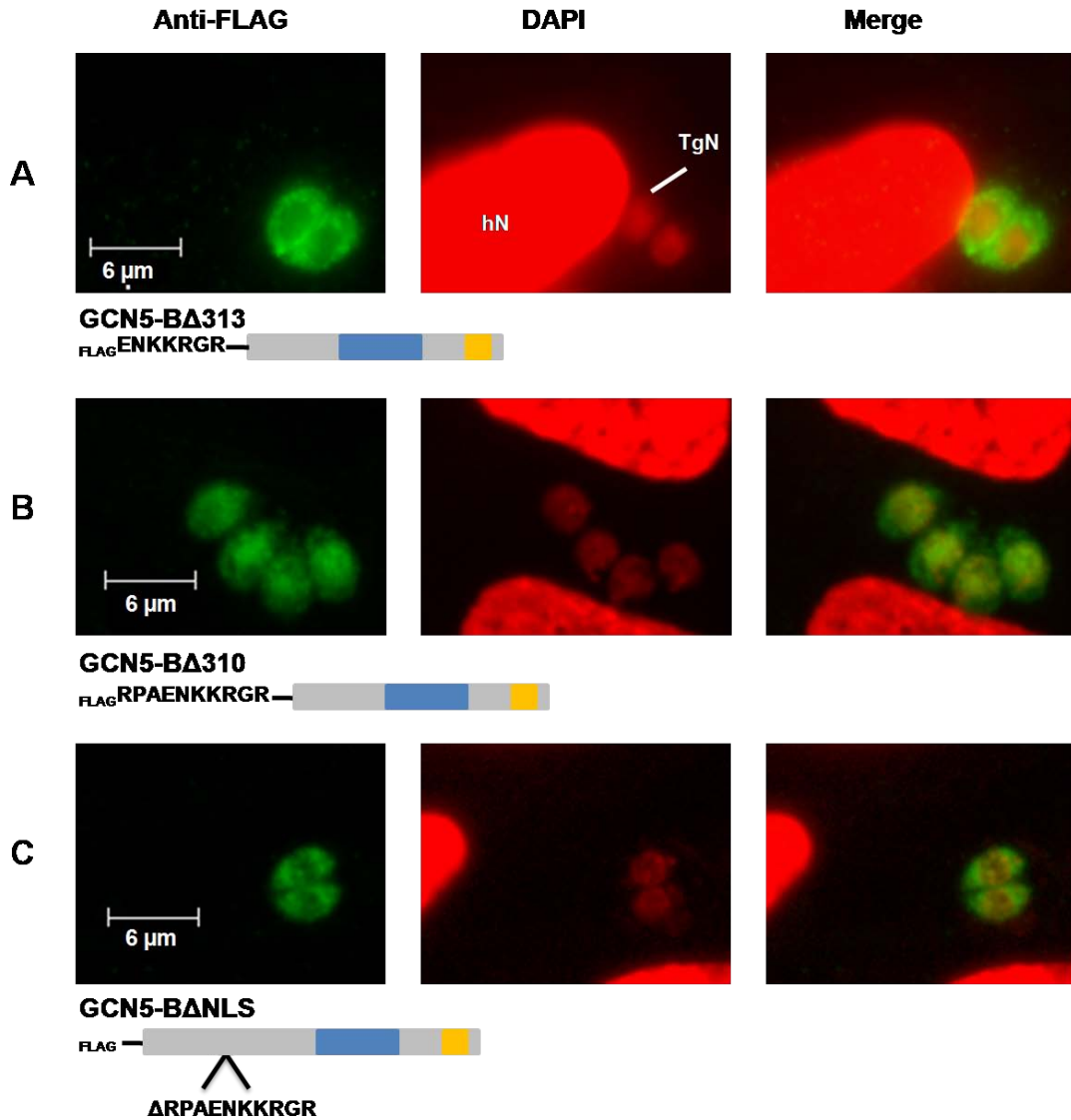


Figure 7: Mutation constructs of TgGCN5-B elucidate the complete NLS. All constructs above contain an N-terminal FLAG tag followed by various mutant versions of TgGCN5-B. Each was expressed within the parasites with localization determined via IFA. Below each construct name is a diagram of the protein. The blue box of each protein diagram represents the HAT catalytic domain of TgGCN5-B, whereas the orange box depicts the bromodomain. Panel A demonstrates the cytoplasmic distribution of TgGCN5-B after removal of the first 313 amino acids ($\Delta 313$). Nuclear localization is restored when only the first 310 residues are removed from TgGCN5-B ($\Delta 310$, Panel B). Internal deletion of just the 10 residue NLS (Δ NLS, Panel C) excluded TgGCN5-B from the nucleus. TgN, *Toxoplasma* nuclei; hN, host cell nucleus; Green = Anti-FLAG; Red = DAPI, 4',6-diamidino-2-phenylindole.

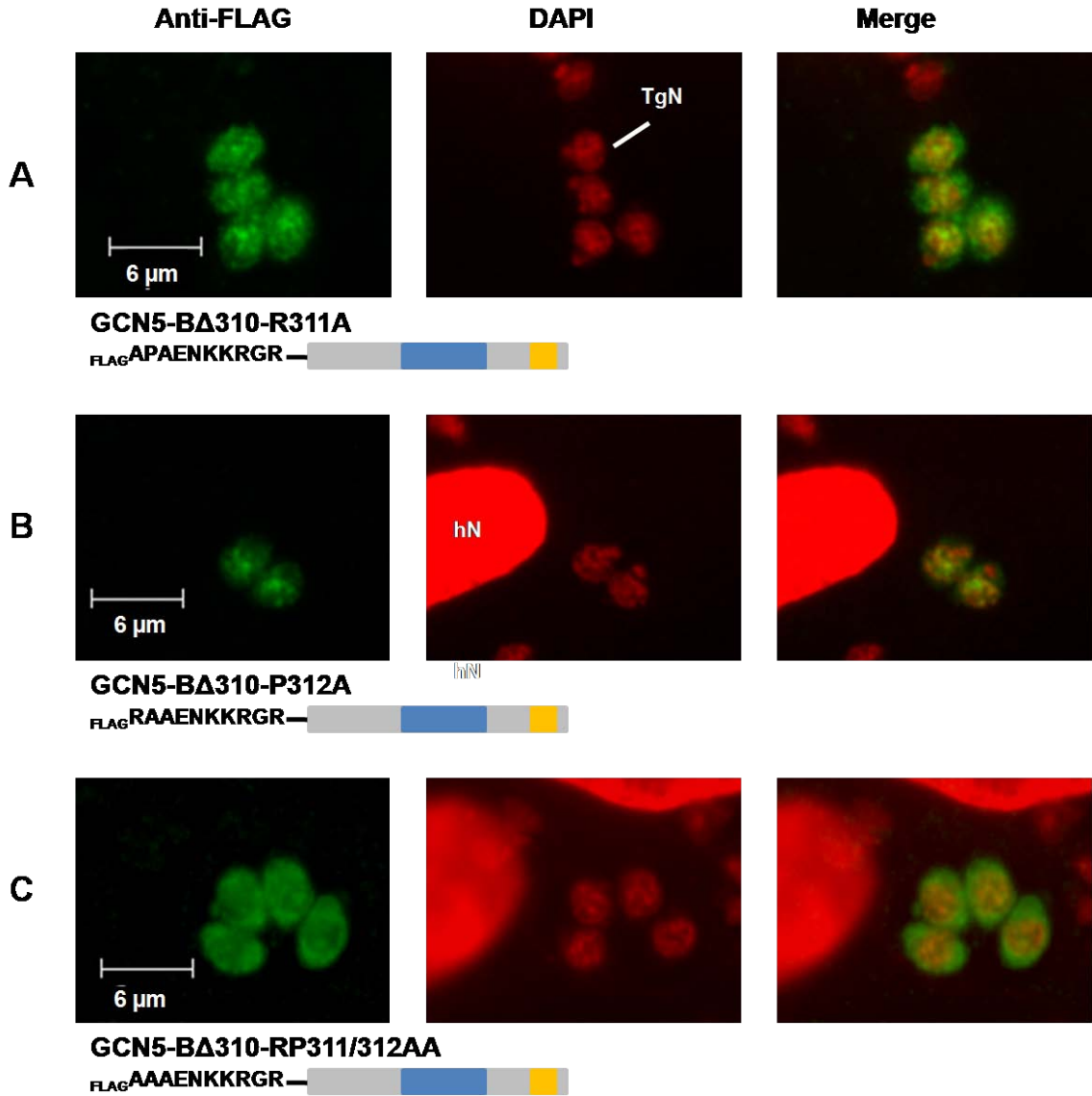


Figure 8: Mutation constructs of TgGCN5-B reveal that specific residues in the upstream sequence are key to nuclear localization. All constructs expressed above contain an N-terminal FLAG tag followed by various mutant versions of TgGCN5-B. Below each construct name is a diagram of the protein. The blue box of each protein diagram represents the HAT catalytic domain of TgGCN5-B, whereas the orange box depicts the bromodomain. Panels A and B demonstrate that a single point mutation of either the arginine (residue 311) or proline (residue 312) still allow nuclear entry. Nuclear localization is hindered when both the upstream arginine and proline are mutated to alanines (Panel C). TgN, *Toxoplasma* nuclei; hN, host cell nucleus; Green = Anti-FLAG; Red = DAPI, 4',6-diamidino-2-phenylindole.

C. The TgGCN5-B NLS is sufficient to localize β -gal to the nucleus

In order to demonstrate that a putative NLS is sufficient and is the minimal required sequence, it must be capable of localizing a non-nuclear protein to the nucleus. To determine the sufficiency of the TgGCN5-B NLS, I followed a similar approach used by Bhatti et al. (2005) for TgGCN5-A [144]. When the *E. coli* protein β -galactosidase (β -gal) was expressed in the parasites, it was distributed throughout the cytoplasm (Figure 9A) and excluded from the parasite nucleus. However, the addition of the TgGCN5-A NLS allowed partial localization of β -gal to the parasite nucleus [144]. When the NLS of TgGCN5-B is fused to the C-terminal end of β -gal followed by a FLAG tag (β -gal-NLS_{FLAG}) the protein becomes predominantly nuclear (Figure 9B). To exclude the possibility of the combination of the NLS-FLAG forming a unique sequence that was responsible for the nuclear localization, the FLAG tag was replaced with an HA tag (β -gal-NLS_{HA}). The latter construct also demonstrated chiefly nuclear distribution (Figure 9C), thereby eliminating the possibility of the tag affecting the localization. The localization of both the above ectopic β -gal-NLS fusion proteins corresponded nearly exclusively with the DAPI staining pattern of the nuclear DNA. Therefore, it can be assumed that the NLS of TgGCN5-B is sufficient and does not require any additional upstream residues or post-translational modifications. Interestingly, when the NLS of TgGCN5-B was moved to the N-terminal portion of β -gal, the protein was no longer able to localize to the parasite nucleus (Figure 9D). It is possible that in this circumstance, the NLS is being cleaved from the rest of the protein; therefore, it is unable to result in nuclear localization.

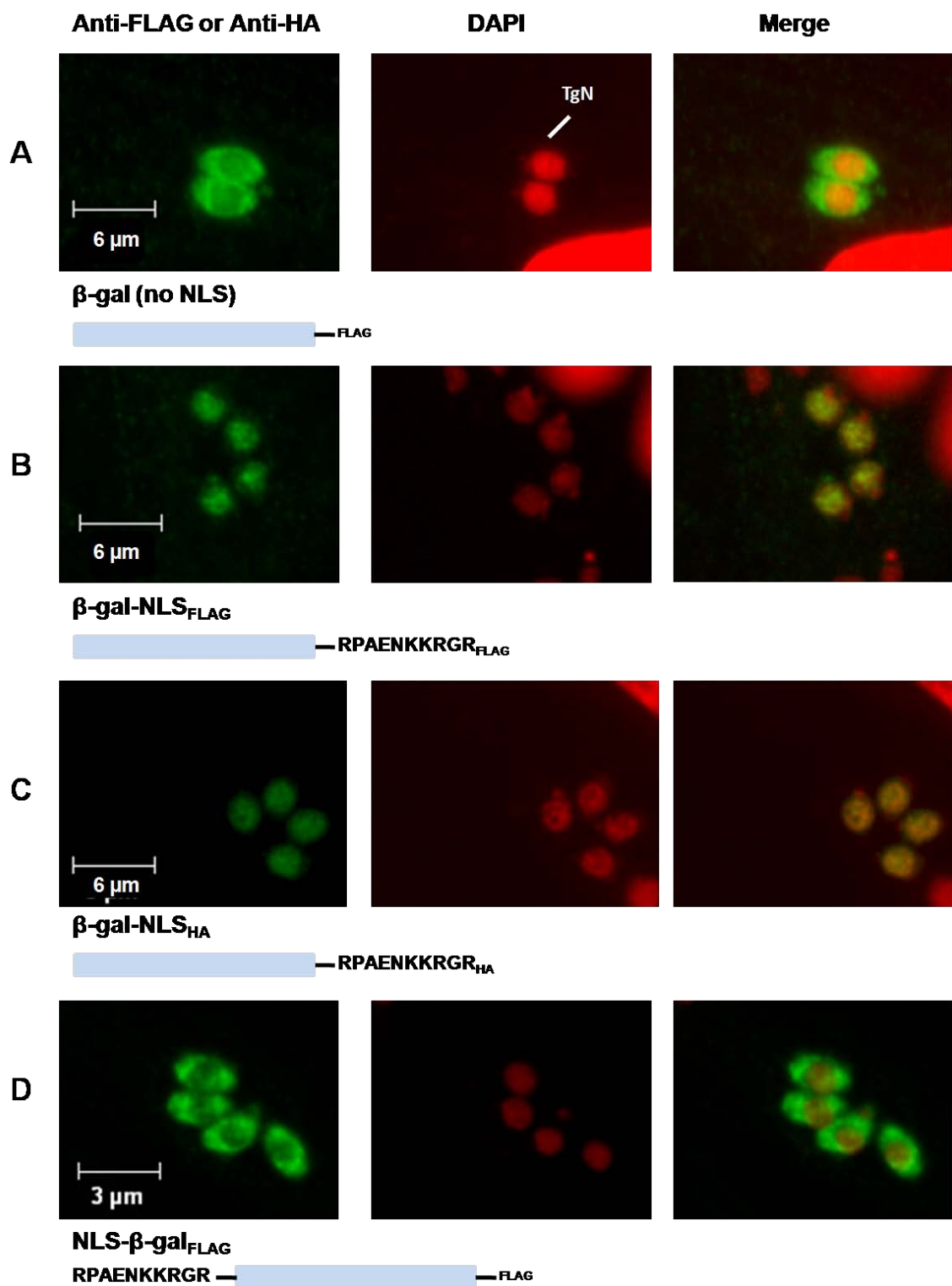


Figure 9: The GCN5-B NLS is sufficient to localize *E. coli* β -gal to the parasite nucleus. Panel A demonstrates that over-expression of bacterial β -gal in the parasites with a C-terminal FLAG tag results in cytoplasmic distribution. However, when the 10 residue NLS of TgGCN5-B is added to the C-terminus of the protein followed by a FLAG-tag (Panel B) or HA-tag (Panel C) the protein localizes to the parasite nucleus. If the NLS is attached to the N-terminus of β -gal, it no longer is able to localize the protein to the parasite nucleus (Panel D), possibly due to cleavage. The diagrams on the right of each panel depict each construct. TgN, *Toxoplasma* nuclei; Green = Anti-FLAG or Anti-HA; Red = DAPI, 4',6-diamidino-2-phenylindole.

D. The NLS of TgGCN5-B has predictive value

Although the genome of *Toxoplasma* has been sequenced to 12X coverage (<http://ToxoDB.org>), many predicted proteins remain uncharacterized and have unknown functions [11]. Identifying potential motifs such as NLSs in an uncharacterized protein can contribute to determining the function of the protein. To determine the utility of the TgGCN5-B NLS as a predictor for nuclear localization of other *Toxoplasma* proteins, I searched the ToxoDB (<http://ToxoDB.org>) for predicted proteins harboring a similar motif (Table VI). TgGCN5-B was the only protein to possess the exact 10 residue NLS. However, when permutations were allowed for residues that were not basic or proline (RP . . . KKR . R, with "." being any amino acid), then three unique proteins were identified. This dataset included two uncharacterized hypothetical proteins and another protein possessing a PHD-finger domain, commonly found in chromatin remodeling enzymes. By far, the largest dataset was obtained when just the basic cluster (KKR . R) was searched, resulting in over 800 proteins being identified. Many of the proteins in this dataset are hypothetical proteins, and only a select few were listed in Table VI. Indeed, some of the proteins identified in our search are likely nuclear proteins (i.e. DNA polymerases or DNA repair proteins) or possess domains that might interact with DNA (AT-hook domains or zinc finger motifs). This bioinformatic survey demonstrates that the TgGCN5-B NLS might be helpful in predicting additional nuclear proteins within *Toxoplasma*.

Additionally, a similar search could be expanded to include other Apicomplexa parasites, such as *Neospora*, since many proteins in other phylum members remain uncharacterized. *Neospora*, *Toxoplasma*'s closest apicomplexan cousin, is currently the only other protozoan known to possess two GCN5 homologues. With these two

parasites being so closely related, I sought to determine if the NLSs between the GCN5 homologues were conserved. Interestingly, the GCN5-B homologues in both parasites contain 82% identity, and the N-terminal extensions of both are well conserved including the NLS regions. *Neospora*'s GCN5-B contains an identical basic-rich cluster (KKRGR) that is likely to function as part of its NLS. The upstream sequence flanking *Neospora*'s basic-cluster (RPVPES) is similar to *Toxoplasma*'s additional upstream activating residues (RPAEN). Conversely, the GCN5-A homologue of *Neospora* appears to have either a divergent or novel NLS as only 3 of the 6 residues from *Toxoplasma*'s TgGCN5-A NLS are identical (*Neospora*: SKRLKM, *Toxoplasma*: RKRVKR. Identical residues are underlined.)

Table VI: Predictive value of the TgGCN5-B NLS

Accession Numbers	Search Results	Predicted homology/motifs	GO Terms
RPAENKKRGR			
TGGT1_046420	RPAENKKRGR	TgGCN5-B	acetyltransferase
RP . . . KKR . R			
TGGT1_071200	RPSDAKKRER	PHD-finger domain containing protein	zinc ion binding
TGGT1_113380	RPKKGKKRKKR	conserved hypothetical protein	none
TGME49_091900	RPEGRKKRLR	conserved hypothetical protein	nuclear
RP . . KKR . R			
None			
RP . KKR . R			
TGGT1_071910	RPKKKRSR	GIY-YIG catalytic domain protein	nuclease activity
TGGT1_056400	RPKKKRRR	AT-hook motif containing protein	DNA binding
TGME49_085520	RPQKKRSR	prip interacting protein, pimt putative	methyltransferase
RPKKR . R			
TGGT1_068070	RPKKRSR	conserved hypothetical protein	none
KKR . R	Total of 801 <i>Toxoplasma</i> genes		
TGGT1_117800	KKRGR	DNA polymerase theta	nucleic acid binding
TGGT1_082080	KKRGR & KKRQR	SET domain containing protein (TgSET1)	nuclear
TGGT1_034540	KKRQR	DNA polymerase lambda	DNA replication
TGGT1_116960	KKRRR	PHD-finger domain containing protein	zinc ion binding
TGGT1_010370	KKRQR	zinc finger (C2H2 type) protein	none
TGGT1_121530	KKRGR	DNA polymerase epsilon	nuclear
TGGT1_032500	KKRQR	DNA mismatch repair protein	Mismatch repair

The predictive value of the TgGCN5-B NLS was determined by screening the ToxoDB (<http://ToxoDB.org>, release 6.0) using the protein motif pattern tool. The shaded rows represent the amino acid patterns used in each search with each “.” corresponding to any amino acid. For all search results, only the TGGT1 accession number is listed. In the case that there was not a TGGT1 homologue, the TGME49 accession number is listed. The last search, which included only the basic cluster (KKR.R), yielded 801 genes, many of which were hypothetical proteins. Listed in the table for this search are only a select few of the top hits, the ones most likely to be nuclear proteins based on their homology assignment. GO terms represents the gene ontology (www.geneontology.org; version 1.803) predictions associated with each gene. These predictions include either the predicted cellular compartment (nuclear) or a predicted cellular function.

E. Protein with a predicted analogous NLS localizes to the nucleus

To test the predictive value of the TgGCN5-B NLS, I decided to examine the localization of the AT-hook domain containing protein TGGT1_056400 (will subsequently be referred to as AT-hook 056400). This protein was identified in the bioinformatic search (Table VI) and contains an analogous predicted NLS at its C-terminal end represented by a proline interrupting a basic cluster of amino acids, RPKKRRR (aa 2,515 to 2,522). AT-hook domains are small DNA binding motifs with a preference for A/T rich DNA regions. These motifs are often found in proteins associated with

chromatin interactions or transcriptional regulation [197]. Based on their functions, AT-hook proteins are believed to be nuclear. In *Theileria*, several closely related AT-hook domain containing proteins are able to localize to the host cell nucleus [198,199]. However, localization to the parasite nucleus has yet to be demonstrated for any Apicomplexa AT-hook motif protein. AT-hook 056400 is predicted to be a large protein of 3,768 amino acids with 3 putative AT-hook domains, predicted by SMART (Simple Modular Architecture Research Tool, <http://smart.embl-heidelberg.de/>) [200,201]. These domains exist between amino acids 1,003 to 1,015 (E-value = 5.08e+01), 1,305 to 1,317 (E-value = 2.50e+02) and 2,528 to 2,540 (E-value = 1.46e+00).

To examine the localization of AT-hook 056400, I followed the methods established by Huynh et al. (2008) for endogenously tagging a protein using *Toxoplasma* RHΔKU80 [186,187]. Through the incorporation of a 3xHA epitope tag at the C-terminus of AT-hook 056400 (Plasmid #16), I was able to determine the localization of the native protein via IFA. As depicted in Figure 10, AT-hook 056400 is distributed throughout the parasite nucleus. Although further testing would be needed to demonstrate that this basic cluster is the exact NLS of AT-hook 056400, the bioinformatic search assisted in identifying a potential region for NLS location, and also correctly predicted a nuclear protein. Incidentally, unlike *Theileria*, this *Toxoplasma* AT-hook protein localizes to the parasite nucleus rather than that of the host cell.

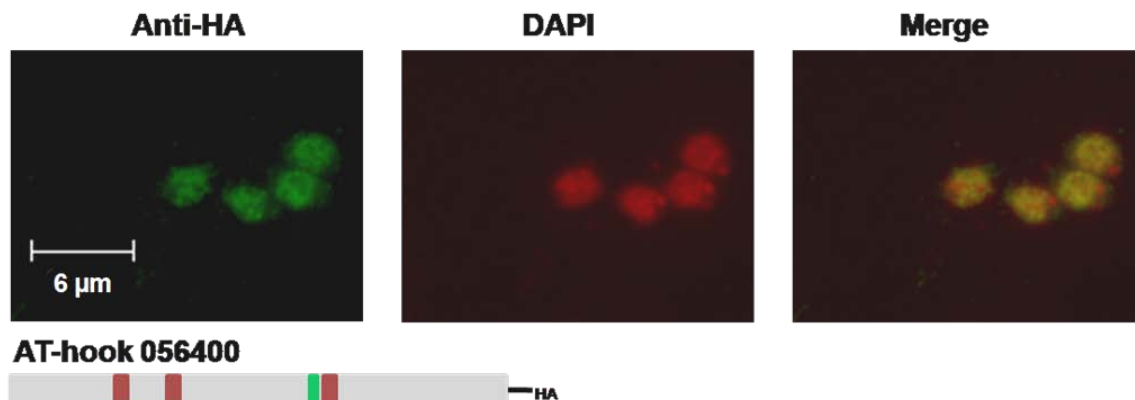


Figure 10: AT-hook 056400 contains a predicted analogous NLS and localizes to the parasite nucleus. The *Toxoplasma* protein AT-hook 056400 (TGGT1_056400) was tagged endogenously with a C-terminal 3xHA epitope tag. This protein localizes to the parasite nucleus as determined by IFA with staining for anti-HA. The diagram at the right depicts the protein and its domains. Red boxes indicate locations of putative AT-hook domains whereas the green box represents the predicted analogous NLS. TgN, *Toxoplasma* nuclei; Green = Anti-HA; Red = DAPI, 4',6-diamidino-2-phenylindole.

F. The nuclear chaperone utilized by TgGCN5-B remains to be resolved

As discussed in the introduction section, the classical pattern of nuclear localization involves a ternary complex consisting of importin- β (Imp- β) binding importin- α (Imp- α), which associates with a nuclear protein via its NLS. This ternary complex is then shuttled through the nuclear pore via the interactions of Imp- β with the nuclear pore complex (NPC). Using *in vitro* protein synthesis and ^{35}S -labeling followed by co-immunoprecipitation, Bhatti et al. (2005) demonstrated that TgGCN5-A is able to associate with TgImp- α via its NLS [144]. Following a similar approach, I set out to determine if TgGCN5-B also was able to interact with Imp- α via its NLS. This approach failed to demonstrate a definitive interaction between TgGCN5-B and TgImp- α . As an alternative approach, I co-expressed epitope tagged versions of both TgGCN5-B and TgImp- α within the same parasite strain to determine if I could establish an *in vivo* interaction. Both proteins were over-expressed in the parasites under the tubulin promoter. A stable, clonal parasite population (resistant to MPA) expressing N-terminally FLAG-tagged TgGCN5-B ($_{\text{FLAG}}$ GCN5-B; Plasmid #2) was transfected with Plasmid #15 (conferring resistance to CAM) to express N-terminally HA-tagged TgImp- α ($_{\text{HA}}$ Imp- α). IFA analysis of this double expressing parasite strain demonstrates that $_{\text{HA}}$ Imp- α localizes to the parasite nucleus as expected (Figure 11). However, when co-

immunoprecipitations were performed, an interaction between $_{HA}$ Imp- α and $_{FLAG}$ GCN5-B was not detected despite strong over-expression of each protein (data not shown). Co-immunoprecipitations were performed with either anti-HA affinity matrix to pull down $_{HA}$ Imp- α and then immunoprecipitated material was probed in Western blot analysis with anti-FLAG to determine if $_{FLAG}$ GCN5-B was concurrently pulled down. Reciprocal reactions were also performed, but in both cases an interaction between the two proteins was not established.

Recently, it has been established that several nuclear proteins bypass Imp- α and associate directly with Imp- β to gain access to the nucleus [133,134,135,136]. Since I could not detect an interaction between TgGCN5-B and TgImp- α , I tested if TgGCN5-B could interact with TgImp- β . Nuclear proteins that associate directly with Imp- β do not bind to Imp- β in a uniform location. Therefore, I cloned two overlapping fragments of *Toxoplasma*'s Imp- β homologue from parasite cDNA. In total, these two fragments represent the entire protein and have an overlapping portion of 42 residues. I then repeated the interaction studies using *in vitro* protein synthesis with ^{35}S -labeling followed by co-immunoprecipitation with each individual fragment of TgImp- β . No interaction was detected between TgGCN5-B and either portion of TgImp- β (data not shown).

In conclusion, using multiple approaches, I was unable to elucidate the chaperone TgGCN5-B utilizes to enter the parasite nucleus. I used *in vitro* protein synthesis with ^{35}S -labeling followed by co-immunoprecipitation to test both TgImp- α and TgImp- β and could not establish an interaction for either within the experimental parameters. One disadvantage of this approach is if post-translational modifications (PTMs) are needed for an interaction, then in this system the PTMs would likely be absent. Alternatively, I also over-expressed tagged versions of both TgImp- α (HA-tagged) and TgGCN5-B (FLAG-tagged) within the parasites and performed co-immunoprecipitation experiments. If PTMs were necessary for an interaction, then these proteins should have the proper modifications since they were harvested from the parasite lysate. However, this method did not show an interaction between TgImp- α and TgGCN5-B. This method was not repeated for TgImp- β as the whole protein was not cloned, and preliminary studies suggest that TgGCN5-B does not interact with TgImp- β . *Toxoplasma* possesses other nuclear chaperone homologues, including several transportins and importins 4, 5, and 7 [131]. Additionally, BLAST searches of the ToxoDB have identified several other proteins with similarity to Imp- α [131]. It is possible that TgGCN5-B utilizes another protein to gain access to the parasite nucleus.

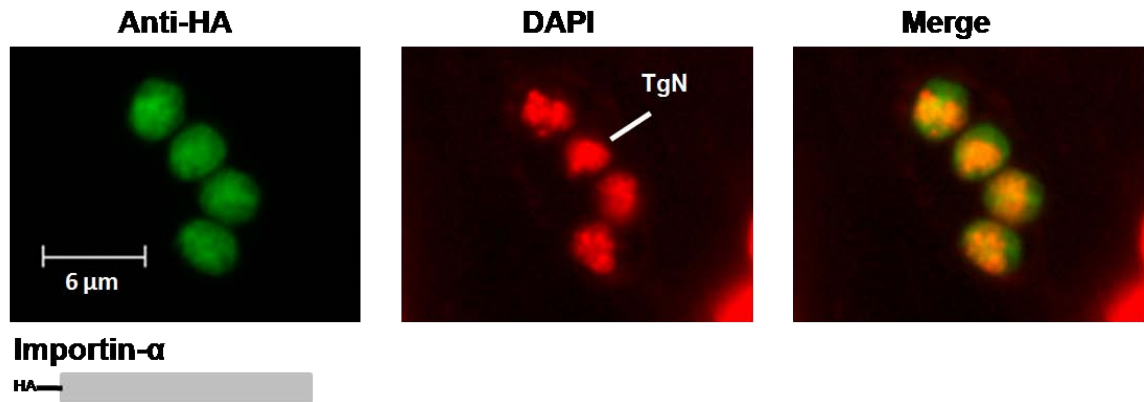


Figure 11: TgImp-α localizes to the parasite nucleus. *Toxoplasma*'s homologue of Imp-α was tagged with an HA epitope at its N-terminus and over-expressed in parasites. IFA with staining for anti-HA was used to determine that TgImp-α localizes to parasite nucleus. TgN, *Toxoplasma* nuclei; Green = Anti-HA; Red = DAPI, 4',6-diamidino-2-phenylindole.

II. Aim 2: Identify the proteins associating with TgGCN5-B

A. Bioinformatics reveal a dearth of common GCN5 complex members

GCN5 and other chromatin remodeling enzymes operate as members of large multi-subunit complexes [104,202]. For GCN5, the most characterized complex is the *S. cerevisiae* SAGA complex. The SAGA complex consists of 21 proteins and includes two enzymatic components, GCN5 and Ubp8, a deubiquitinase [202]. In *S. cerevisiae*, the SAGA complex regulates approximately 10% of genes but is also involved in several other cellular processes, including transcriptional elongation and mRNA export [99,100,101]. SAGA complex components are well conserved in metazoans and vertebrates, including *Drosophila* and humans [104].

Given the conserved nature of the SAGA complex, I attempted to determine if this complex is also conserved in *Toxoplasma*. The ToxoDB was searched using the protein-protein BLAST (basic local alignment search tool; BLASTp), with the amino acid sequences of the *S. cerevisiae* SAGA homologues as the query subjects. The results of this bioinformatics search are listed in Table VII. Aside from the two known GCN5 and two ADA2 homologues, there is an unusual lack of other conserved SAGA members in *Toxoplasma*. Searches with the amino acids sequences of Taf5/Taf90, Tra1, Ubp8, and Chd1 identified potential homologues with notable E-values, where the closer the E-value is to zero, the more reliable the prediction. However, these 4 proteins did not provide much insight into the possible GCN5 complex(s) in *Toxoplasma*. For instance, the identified potential Ubp8 homologue could be any deubiquitinase enzyme, with the homology based on the catalytic domain only. This argument can also be made for the Chd1 homologue. Since yeast Chd1 is a chromodomain-containing protein, the analogous protein in the search may only be conserved within this region [202]. The potential Tra1 homologue would be an interesting protein to pursue given that in *S. cerevisiae*, Tra1 interacts with specific transcription factors to recruit the SAGA complex to given promoters [202]. However, the identified *Toxoplasma* protein resembling Tra1 is an enormous size, over 8,000 amino acids, that can be very difficult to study. The lack of other identifiable homologues suggests that if a SAGA complex does indeed exist in *Toxoplasma*, then the members are either highly divergent or novel.

Table VII: SAGA complex homologues in *Toxoplasma*

Protein	Used in BLASTp	Top Hit	E-value
GCN5	YGR252W	GCN5-A TGGT1_004130	2.3 e-79
		GCN5-B TGGT1_046420	2.0 e-82
ADA2	YDR448W	ADA2-A TGGT1_097980	9.9 e-47
		ADA2-B TGGT1_007690	6.8 e-33
ADA1	YPL254W	NONE	
ADA3	YDR176W	NONE	
Spt20	YOL148C	NONE	
Spt3	YDR392W	NONE	
Spt7	YBR081C	TGGT1_004130	2.4 e-6
Spt8	YLR055C	TGME49_016880	3.2e-7
Taf5/Taf90	YBR198C	TGME49_016880	1.6 e-26
Taf6/Taf60	YGL122C	NONE	
Taf9/Taf17	YMR236W	NONE	
Taf10/Taf25	YDR167W	NONE	
Taf12/Taf68	YDR145W	NONE	
Tra1	YHR099W	TGGT1_106430	1.1 e-66
SGF29	YCL010C	NONE	
SGF11	YPL047W	NONE	
SGF73	YGL066W	NONE	
Sus1	YBR111W-A	NONE	
Ubp8	YMR223W	TGGT1_043540	2.1 e-26
Chd1	YER164W	TGGT1_011870	1.2 e-135
Rtg2	YGL252C	NONE	

The table depicts bioinformatic search results for the presence of conserved *S. cerevisiae* SAGA complex homologues in *Toxoplasma*. The SAGA members are listed in the left-hand column followed by the systematic name assigned to each protein in the *Saccharomyces* Genome Database (SGD, www.yeastgenome.org). The amino acid sequence of each protein was retrieved from the SGD and then used in a BLASTp search on the ToxoDB (<http://ToxoDB.org>, release 6.0). The top homologous *Toxoplasma* protein identified in each search is listed via its accession number. The right-hand column represents the E-value for each identified homologue.

B. TgGCN5-B is disordered in its N-terminal extension

An unusual aspect of each *Toxoplasma* GCN5 homologue is a lengthy N-terminal extension. Other protozoa and invertebrates, such as *S. cerevisiae* and *Tetrahymena*, have a single GCN5 homologue with a very short N-terminus. On the other hand, GCN5 homologues from higher-order metazoans and vertebrates have long N-terminal extensions [104]. Therefore, *Toxoplasma*'s GCN5 homologues and the GCN5s from other Apicomplexa parasites resemble those of higher eukaryotes as opposed to their closer kin (Figure 3; Chapter 1, Section IV-C).

Interestingly, the N-terminal extensions of *Toxoplasma*'s GCN5 homologues do not have identifiable protein motifs or domains other than the NLSs discussed in Aim 1 [144]. These N-terminal extensions also do not share homology with other characterized proteins including each other [169]. Thus, the functions of the elongated N-termini in *Toxoplasma*'s GCN5s remain elusive. To analyze the amino acid sequence of TgGCN5-B's N-terminus, I collaborated with Drs. Vladimir N. Uversky and A. Keith Dunker (Indiana University School of Medicine) to determine if this region was intrinsically disordered.

Regions of a protein that lack fixed structures are referred to as having intrinsic disorder. A multitude of proteins, especially eukaryotic proteins, are unstructured or contain regions that lack a distinct three-dimensional configuration [203]. Intrinsically disordered proteins can have a variety of important biological functions including signal transduction, molecular recognition, and regulation [204,205,206]. A protein's primary amino acid sequence can be used to predict the degree of disorder. Disordered regions are characterized by low amino acid sequence complexity, which includes few bulky, hydrophobic amino acids and an enrichment of polar and charged amino acids. For instance, regions of intrinsic disorder are likely to contain amino acids such as Ala, Arg, Gly, Gln, Ser, Glu, Lys, and Pro and are also depleted in Trp, Tyr, Phe, Ile, Leu, Val, Cys, and Asn residues [204,205,206,207]. To date, many computational predictors of disorder have been developed to identify regions or entire proteins characterized by intrinsic disorder [205]. Determining the degree of disorder in a protein can assist in predicting the biological relevance of a given domain, as many regions of disorder map to areas of protein-protein interactions or post-translational modifications. It has previously been determined that the genomes of early-branching eukaryotic protozoa, including *Toxoplasma*, contain a large proportion of predicted proteins containing intrinsic disorder [208]. Therefore, insight into the degree of intrinsic disorder in TgGCN5-B might help provide clues as to the function(s) of its N-terminal extension.

Drs. Vladimir N. Uversky and A. Keith Dunker analyzed the primary amino acid sequence of TgGCN5-B using several PONDR® (Predictor of Natural Disordered Regions) prediction versions as well as other computational techniques [205]. Figure 12 demonstrates that TgGCN5-B has multiple regions of intrinsic disorder in its N-terminus, including a region that corresponds to the sequence of the NLS. This is an intriguing correlation since intrinsically disordered regions have been shown to map to regions of protein-protein interactions, and the NLS region of TgGCN5-B is very likely to bind to a

chaperone protein for nuclear import [204,205,207]. Likewise, the ADA2 binding domain of TgGCN5-B also corresponds to a region of intrinsic disorder. From this observation, it can be hypothesized that other regions of intrinsic disorder in TgGCN5-B might predict regions of protein-protein interaction. As expected, the HAT catalytic domain is not disordered, thereby following the prediction that regions of structure or distinct three-dimensional characteristics lack disorder. In summary, the analysis of TgGCN5-B's amino acid sequence suggests that certain regions, particularly in the N-terminal extension, are intrinsically disordered, which may suggest functions such as protein-protein interaction domains or protein modification sites.

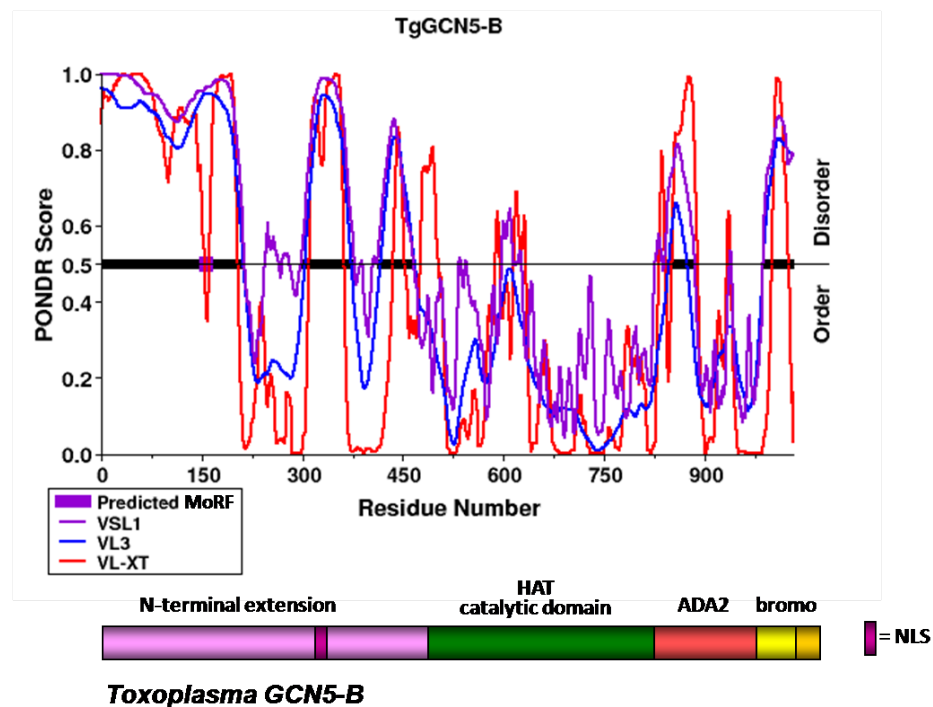


Figure 12: TgGCN5-B contains several regions of intrinsic disorder within the N-terminus. Three computational PONDR® algorithms (VSL1, VL3, VL-XT) were utilized to examine the primary amino acid sequence of TgGCN5-B. PONDR® scores above 0.5 are considered disordered, whereas those below are ordered. Large contiguous regions of disorder represent a predicted MoRF (molecular recognition feature). The diagram of TgGCN5-B depicting its molecular domains indicates that certain regions of disorder map to both the NLS and the ADA2-binding domain, indicating that these MoRF regions might predict protein-protein interaction sites.

Based on preliminary evidence, it is hypothesized that TgGCN5-B associates with unique proteins within a multi-subunit complex in *Toxoplasma*. The data leading to this hypothesis include: 1) Very few SAGA complex members are conserved in *Toxoplasma*, indicating TgGCN5-B associating proteins are likely to be novel or highly divergent, and 2) TgGCN5-B has an elongated N-terminal extension lacking identifiable protein motifs but containing several regions of intrinsic disorder, which may be an indicator of protein-protein interaction regions. To address this hypothesis, biochemical techniques including affinity chromatography and co-immunoprecipitations were utilized to identify proteins associating with TgGCN5-B.

C. Affinity chromatography to identify TgGCN5-B associating proteins

To determine TgGCN5-B associating proteins through affinity chromatography, recombinant TgGCN5-B protein was produced in bacteria. Next, the recombinant protein was coupled to a column through which *Toxoplasma* nuclear-enriched lysate was passed. Proteins able to interact with TgGCN5-B bound the column, whereas other proteins flowed through the column and were washed away. Then, all interacting proteins were eluted from the column, analyzed by SDS-PAGE, and identified by mass spectrometry. Figure 13 summarizes the affinity chromatography procedure.

To produce recombinant TgGCN5-B, the pMAL™ Protein Fusion & Purification System (NEB, #E8000S) was utilized. With this system, full-length TgGCN5-B was incorporated downstream of the *malE* gene of *E. coli* in the expression vector pMAL-c2X (Plasmid #17). The *malE* gene encodes maltose-binding protein (MBP). When induced, bacteria expressed the fusion protein MBP-GCN5-B. In parallel, empty pMAL-c2X vector was also used to produce MBP, for use as a control. Panel A of Figure 14 shows both recombinant MBP and MBP-GCN5-B following purification from bacteria lysate over amylose resin. The samples were resolved on SDS-PAGE and stained with Simply Blue®. For both MBP and MBP-GCN5-B there seems to be a breakdown product running slightly smaller than the expected size (MBP ~ 42 kDa and MBP-GCN5-B ~ 142 kDa). The presence of a breakdown product was not anticipated to interfere with the affinity purification of associating proteins.

Interestingly, MBP-GCN5-B was shown to be catalytically active (Figure 14, Panel B), as evaluated by a non-radioactive HAT assay using 1.25 µg of recombinant protein. Additionally, the HAT activity of MBP-GCN5-B was inhibited by the addition of 20 µM anacardic acid (DMSO = vehicle control). This was the first time that either

Toxoplasma GCN5 homologue was produced recombinantly with HAT activity. Previously, recombinant TgGCN5s tagged with either six histidine residues (His-tag) or GST (glutathione-S-transferase) never demonstrated *in vitro* HAT activity [Sullivan and Bhatti, unpublished data].

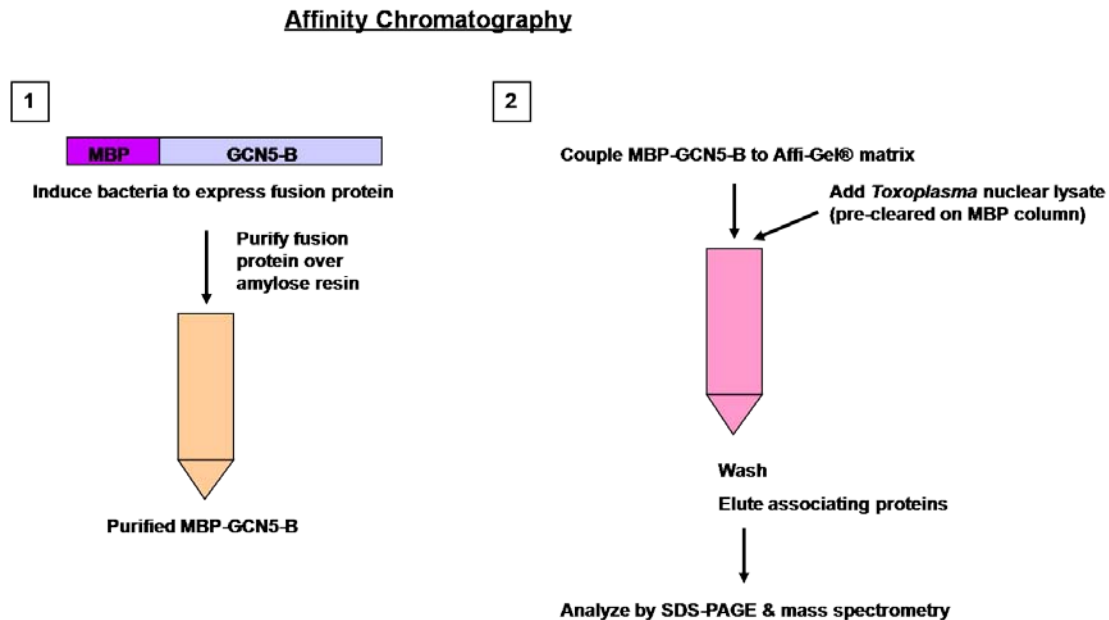


Figure 13: Pictorial diagram representing affinity chromatography procedure. The first step (1) was to produce the fusion protein MBP-GCN5-B, the maltose-binding protein (MBP) coupled to full-length TgGCN5-B. Bacteria were induced to express the fusion protein that was subsequently purified from bacterial lysate over a column containing amylose resin. The fusion protein bound the resin through the MBP domain and was then removed from the column by the addition of maltose. Next (2), the purified MBP-GCN5-B fusion protein was coupled to Affi-gel® matrix for the assembly of an affinity column. *Toxoplasma* nuclear-enriched lysate was first pre-cleared on an MBP affinity column prior to being used in the MBP-GCN5-B affinity column. Nuclear proteins capable of associating with TgGCN5-B should be bound to MBP-GCN5-B, whereas non-specific proteins would pass through the column. After several washes, all associating proteins were eluted from the column and analyzed via SDS-PAGE and mass spectrometry. Any protein identified from the MBP-GCN5-B column but not in the eluted proteins from the MBP column are potential TgGCN5-B associating proteins.

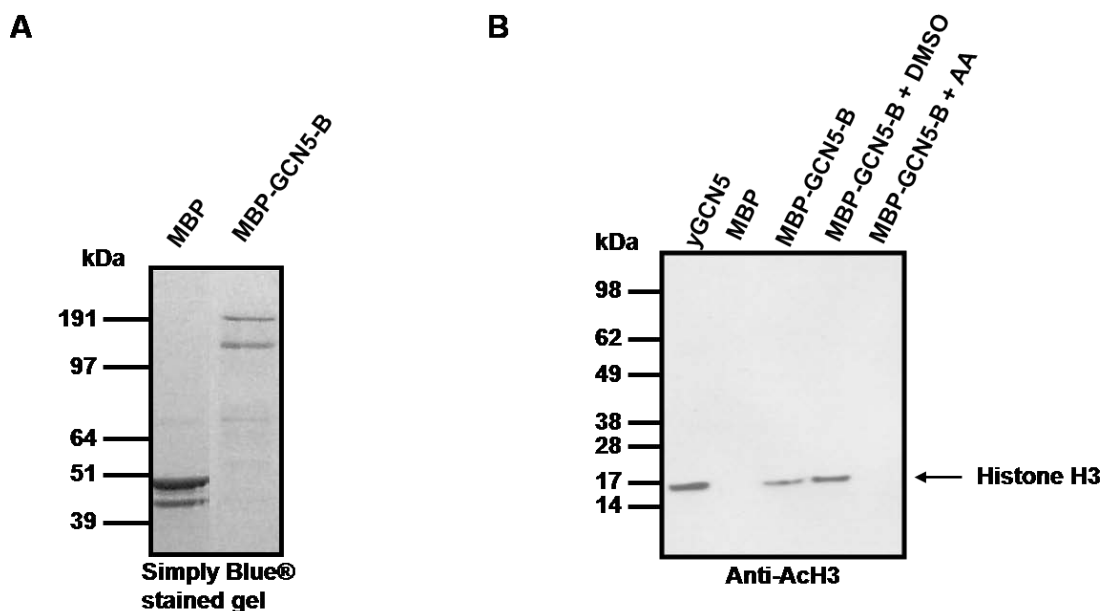


Figure 14: Purified recombinant MBP-GCN5-B is catalytically active. Panel A represents purified recombinant proteins of MBP (~42 kDa) and MBP-GCN5-B (~142 kDa) resolved on SDS-PAGE and stained with Simply Blue®. There appears to be a breakdown product for both MBP and MBP-GCN5-B; however, the presence of this product should not affect downstream applications. Purified samples such as these were used when making the affinity columns. Panel B shows purified recombinant MBP-GCN5-B to be catalytically active. An *in vitro* HAT assay was performed using *S. cerevisiae* GCN5 (yGCN5; positive control), MBP, or MBP-GCN5-B. All HAT reactions were analyzed by Western blot probed with anti-AcH3 (anti-acetyl histone H3). MBP-GCN5-B catalytic activity was inhibited by the addition of 20 μ M anacardic acid (AA) but was unaffected by the addition of vehicle (DMSO).

After purifying both MBP and MBP-GCN5-B, each protein was coupled separately to Affi-Gel® 15 resin (Bio-Rad #153-6051) to generate an affinity column for each. Next, *Toxoplasma* nuclear-enriched lysate was prepared from RH strain tachyzoites. Nuclear-enriched lysate was used to reduce cytoplasmic, contaminating proteins. Since TgGCN5-B is an HAT that localizes to the parasite nucleus, it can be assumed that its associating proteins must also reside within the parasite nucleus. The nuclear-enriched lysate was first pre-cleared on the MBP affinity column. The purpose of this step was to remove proteins that bind non-specifically to MBP. Next, the pre-cleared nuclear-enriched lysate was added to the MBP-GCN5-B affinity column.

Proteins able to adhere to TgGCN5-B or those proteins associate indirectly through the interaction of with another interacting protein should bind to the column, whereas other proteins would pass through the column. The bound proteins from each affinity column, MBP and MBP-GCN5-B, were eluted by acid wash, analyzed on SDS-PAGE followed by silver staining, and identified by mass spectrometry. The affinity chromatography protocol was repeated in two independent trials. For each trial, the procedure was done in duplicate, with one sample set analyzed by SDS-PAGE and the other parallel set sent to the laboratory of Dr. W. Andy Tao (Purdue University) for mass spectrometry analysis (Appendix B). Figure 15 shows a representative SDS-PAGE gel analyzing each sample after affinity chromatography. Protein bands identified in the MBP-GCN5-B eluted sample not found in the MBP eluted sample were considered possible TgGCN5-B associating proteins. The sampled labeled “Unbound” contained proteins from the nuclear lysate that did not bind to either column.

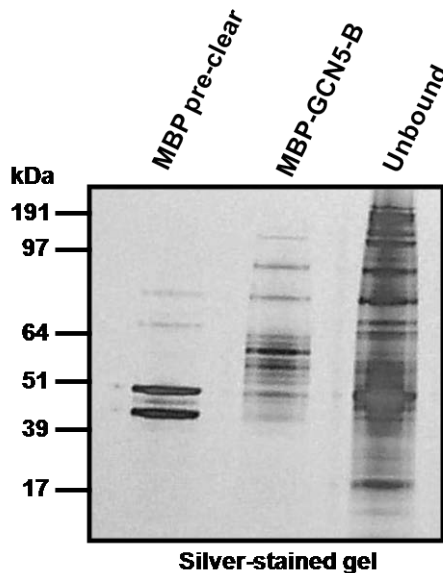


Figure 15: Purification of proteins associating with MBP-GCN5-B using affinity chromatography. Nuclear lysate from tachyzoites was pre-cleared through MBP affinity column prior to incubation with MBP-GCN5-B affinity column. After washing, bound proteins from both columns were eluted, concentrated, and resolved on SDS-PAGE with silver staining. Proteins present in the MBP-GCN5-B lane that are not represented in the MBP lane are candidates for identification by mass spectrometry. The last lane shows the unbound proteins.

For both replicates of the affinity chromatography experiment that were analyzed by mass spectrometry, a large number of candidate TgGCN5-B associating proteins were identified. The data sets for the individual experiments are included in Appendix C. For the second replicate of the affinity chromatography, very few proteins were bound to MBP; therefore, data set #2 is quite large. It is likely that there was not efficient elution of the MBP column for affinity chromatography #2. If this hypothesis is true, then some of the potential Tg-GCN5-B associating proteins for this trial might be contaminants or non-specifically binding. Additionally, some of the proteins identified are likely cytoplasmic, indicating that the nuclear-enriched fractions had some residual cytoplasmic protein contents. Table VIII lists the proteins that were identified in both replicates. The major caveat of the affinity chromatography experiment was that neither of the *Toxoplasma* ADA2 homologues was identified. As known interacting proteins and co-activators of TgGCN5-B, either TgADA2-A or -B was expected to be identified in the affinity chromatography experiment [169]. It is also important to note that many of the candidate proteins are highly abundant proteins in *Toxoplasma* (ribosomal and cytoskeletal proteins) and might represent non-specific binding to the column.

Since neither of the TgADA2 homologues was identified, I was a bit skeptical regarding the rest of the data obtained from the affinity chromatography. It is possible that the majority of the proteins might be binding the recombinant TgGCN5-B in a non-specific manner. One reason for this non-specific binding could be the improper folding of the recombinant protein. Although MBP-GCN5-B is catalytically active, this does not exclude the possibility that other portions of the protein, the N-terminal extension for example, might not be in the correct conformation for properly associating with other proteins. Additionally, the recombinant TgGCN5-B will likely not contain any post-translation modifications that might be needed for protein association.

The drawbacks aside, a few interesting proteins were identified in the affinity chromatography, such as another chromatin remodeling enzyme, the arginine N-methyltransferase 33.m01376. However, this protein was identified only in the first data set. Several histone proteins were also identified in the affinity chromatography experiment, indicating that some of the proteins (histone proteins) may be substrates of TgGCN5-B. However, since I had little confidence in the results from the affinity chromatography, I proceeded to identify TgGCN5-B associating proteins through another method, co-immunoprecipitation.

Table VIII: Proteins identified in both replicates of affinity chromatography

Entry #	Protein Description	Accession Number
1	small heat shock protein, \ bradyzoite-specific protein	44.m02755
2	caltractin (centrin), putative	50.m03356
3	60S ribosomal protein L18a, putative	55.m05004
4	ribosomal protein L21, putative	50.m00012
5	hypothetical protein	41.m01274
6	28 kDa antigen	42.m00015
7	membrane skeletal protein IMC1, putative	44.m00031
8	prohibitin, putative	49.m00051
9	ATP synthase, putative	42.m00065
10	elongation factor 1-beta, putative	42.m00069
11	ATP synthase, putative	76.m01572
12	40s ribosomal protein S6, putative	27.m00119
13	60S acidic ribosomal protein P2, putative	583.m00610
14	fibrillarin, putative	583.m00637
15	hypothetical protein	44.m06355
16	calmodulin, putative	541.m01151
17	histone H2A, putative	55.m04926
18	lysophospholipase, putative	76.m01665
19	hypothetical protein	31.m00869
20	hypothetical protein	583.m05696
21	hypothetical protein	583.m11414
22	conserved hypothetical protein	55.m00224
23	ubiquinol-cytochrome C reductase complex 14 kDa	80.m00018
24	eukaryotic translation initiation factor 3 delta subunit, putative	80.m02245
25	tubulin beta chain	57.m00003
26	malate:quinone oxidoreductase, putative	80.m00006
27	60s ribosomal protein L31, putative	57.m01771
28	adenylate kinase, putative	42.m00116
29	hypothetical protein	55.m10265
30	eukaryotic translation initiation factor 3, putative	83.m01278
31	histone H2A	145.m00002
32	histone H4, putative	49.m03134
33	hypothetical protein	583.m05686
34	hypothetical protein	83.m00011
35	thioredoxin, putative	50.m00069
36	u1 small nuclear ribonucleoprotein 70 kda-related protein	20.m03892
37	membrane skeletal protein IMC1	44.m00004
38	gbp1p protein (RNA binding protein), putative	55.m00241
39	conserved hypothetical protein	583.m00682
40	hypothetical protein	583.m00707
41	40S ribosomal protein S26, putative	49.m03356
42	60S acidic ribosomal protein P0	38.m00002
43	U1 small nuclear ribonucleoprotein, putative	541.m01233
44	ribosomal protein L5, putative	641.m00186
45	40S ribosomal protein S24, putative	33.m01367
46	articulon 4	41.m00021
47	inner membrane complex protein (IMC3)	35.m01595
48	hypothetical protein	44.m02644
49	hypothetical protein	55.m05032
50	protease-related	59.m03479
51	surface protein rhoptry, putative	583.m00003
52	conserved hypothetical protein	33.m01321
53	myosin light chain TgMLC1-related	583.m05420

For each protein, the description provided on the ToxoDB (<http://ToxoDB.org>) as well as the model ID accession number is listed.

D. Expression of ectopic TgGCN5-B in *Toxoplasma*

To determine the TgGCN5-B associating proteins through co-immunoprecipitation, a parasite line over-expressing a triple-tagged version of TgGCN5-B was generated. In this strain, TgGCN5-B was tagged with HA and c-myc epitopes at the N-terminus and a FLAG epitope at the C-terminus ($_{\text{HA-MYC}}\text{GCN5-B}_{\text{FLAG}}$; Plasmid #3). The expression of $_{\text{HA-MYC}}\text{GCN5-B}_{\text{FLAG}}$ was under the regulation of the *Toxoplasma* tubulin promoter. Figure 16 shows both IFA and Western blot data characterizing the $_{\text{HA-MYC}}\text{GCN5-B}_{\text{FLAG}}$ expressing parasites. As expected, $_{\text{HA-MYC}}\text{GCN5-B}_{\text{FLAG}}$ localizes to the parasite nucleus (Figure 16, Panel A). Not only can $_{\text{HA-MYC}}\text{GCN5-B}_{\text{FLAG}}$ be immunoprecipitated from parasite lysate (Figure 16, Panel B), but this tagged version of TgGCN5-B remains catalytically active, as determined by an *in vitro* HAT assay (Figure 16, Panel C). Additionally, $_{\text{HA-MYC}}\text{GCN5-B}_{\text{FLAG}}$ is acetylated (Figure 16, Panel D). Acetylation of HATs has been reported in the literature for both PCAF and p300/CBP and is important for localization and regulation of each HAT, respectively [141,142,143].

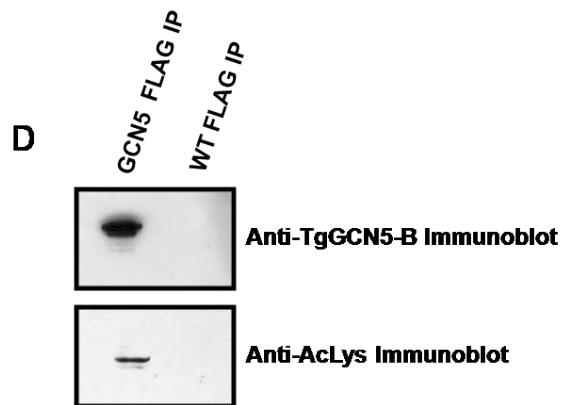
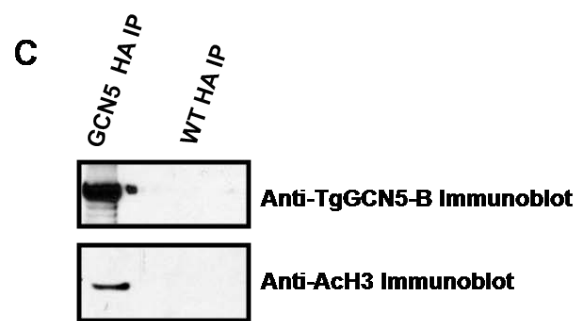
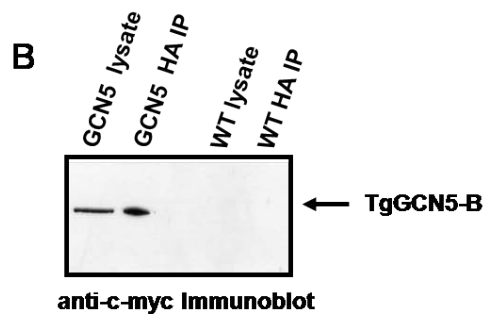
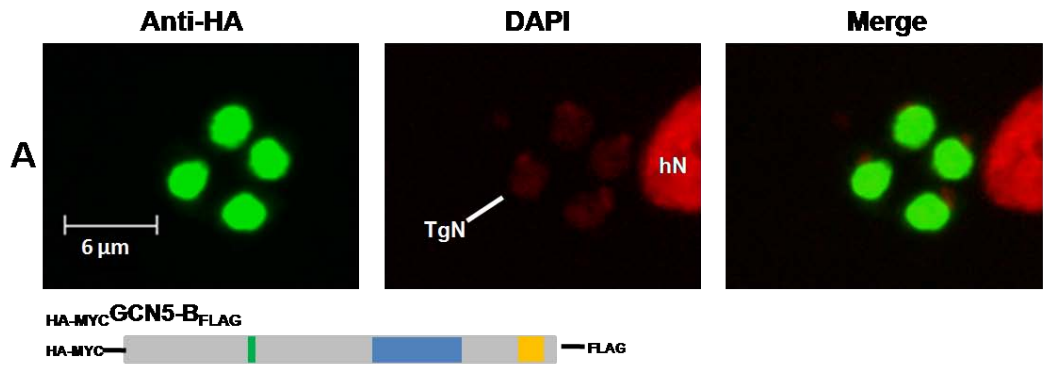


Figure 16: Characterization of parasites expressing $_{HA-MYC}GCN5-B_{FLAG}$. Parasites were transfected with Plasmid #3, selected, and cloned to obtain a homogenous population. Panel A: IFA using anti-HA demonstrating that $_{HA-MYC}GCN5-B_{FLAG}$ localized to the parasite nucleus. A diagram of the $_{HA-MYC}GCN5-B_{FLAG}$ protein is next to IFA data (green box, NLS; blue box, catalytic domain; orange box, bromodomain). TgN, *Toxoplasma* nuclei; hN, host cell nucleus; Green = Anti-HA; Red = DAPI, 4',6-diamidino-2-phenylindole. Panel B: Immunoblot showing $_{HA-MYC}GCN5-B_{FLAG}$ could be immunoprecipitated (IP) from 300 μ g parasite lysate with anti-HA affinity resin. The blot was probed with anti-c-myc, and $_{HA-MYC}GCN5-B_{FLAG}$ parasite lysate (30 μ g) as well as lysate and IP material from parental wild-type (WT) parasites were included as controls. Panel C: After an immunoprecipitation (IP) of 400 μ g $_{HA-MYC}GCN5-B_{FLAG}$ parasite lysate with anti-HA affinity resin, an on-bead HAT assay was performed. The IP material was divided for two immunoblots: one with anti-TgGCN5-B and the other with anti-AcH3 (acetyl-histone H3). Top immunoblot with anti-TgGCN5-B demonstrates that the IP was successful, whereas the bottom immunoblot with anti-AcH3 represents catalytic activity. Parental wild-type (WT) parasites were included as a negative control. Panel D: After an immunoprecipitation (IP) of 300 μ g $_{HA-MYC}GCN5-B_{FLAG}$ parasite lysate with anti-FLAG affinity gel, the IP material was divided for two immunoblots: one with anti-TgGCN5-B and the other with anti-AcLys (acetyl-lysine). This result demonstrates that $_{HA-MYC}GCN5-B_{FLAG}$ was acetylated. Parental wild-type (WT) parasites were included as a negative control.

E. Co-immunoprecipitation reveals novel TgGCN5-B associating proteins

We collaborated with Dr. Ali Hakimi (National Centre for Scientific Research in Grenoble, France) to perform the co-immunoprecipitation (co-IP) experiment to identify the TgGCN5-B complex. Dr. Hakimi's laboratory has had previous success in purifying the associating proteins of TgHDAC3 [154,209]. Dr. Hakimi's laboratory has the capability to grow hundreds of large flasks (T150) of parasites for purification and processing for co-IP experiments. The large amount of parasite material is preferable to obtain optimal results. Figure 17 is a representation of the purification scheme used to co-IP TgGCN5-B associating proteins. Once $_{HA-MYC}GCN5-B_{FLAG}$ and its associating proteins were purified, fractions were resolved by SDS-PAGE with silver-staining and Western blot analysis. Figure 18 depicts both a silver-stained gel and an immunoblot, demonstrating the success of the $_{HA-MYC}GCN5-B_{FLAG}$ co-IP. The silver-stained gel shows that several additional proteins of varying molecular weights were pulled down with $_{HA-MYC}GCN5-B_{FLAG}$ (Figure 18, Panel A).

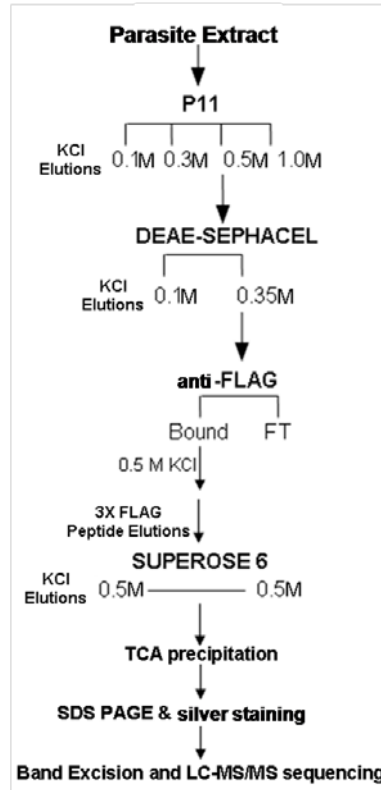


Figure 17: Schematic of co-IP protocol used for detecting $_{HA-MYC}GCN5-B_{FLAG}$ associating proteins. This schematic represents the protocol followed by Dr. Hakimi and colleagues for purification of $_{HA-MYC}GCN5-B_{FLAG}$ and its associating proteins. Parasite extract was generated from 250 large flasks of intracellular parasites harvested 18 hours post infection. The parasite extract was fractionated by chromatography through both phosphocellulose (P11) and DEAE-Sepharcel columns. Next, the co-IP was performed using anti-FLAG affinity gel. The bound proteins were eluted with 3X FLAG peptide in KCl. The elutants were subsequently fractionated using SUPEROSE 6 gel filtration, precipitated with trichloroacetic acid (TCA), and resolved on SDS-PAGE. Protein bands were visualized by silver staining. To identify the associating co-IP proteins, bands were excised and analyzed by nanocapillary liquid chromatography coupled with tandem mass spectrometry (LC-MS/MS sequencing). Figure adopted from Saksouk et al. (2005) and Bhatti (2006) [154,209].

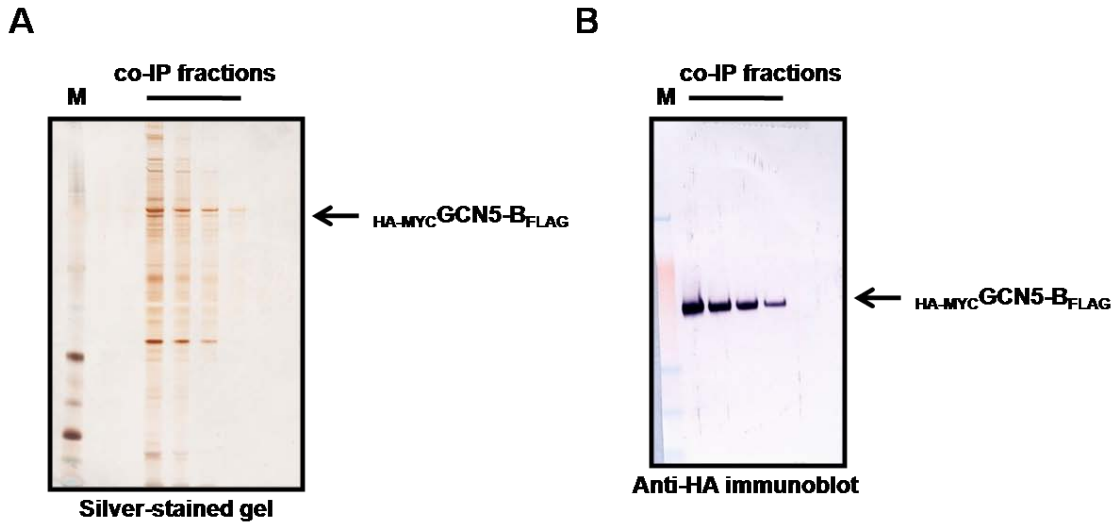


Figure 18: Analysis from co-IP of $HA-MYC GCN5-B_{FLAG}$ and associating proteins.

Following the co-IP protocol outlined in Figure 17, fractions were analyzed by SDS-PAGE (Panel A, silver-stained gel) and Western blot (Panel B, anti-HA immunoblot). Several proteins with a variety of molecular weights were coimmunoprecipitated with $HA-MYC GCN5-B_{FLAG}$ (Panel A). The arrow shows the protein band likely to be $HA-MYC GCN5-B_{FLAG}$ (~ 100 kDA). Panel B is an immunoblot with anti-HA of the same four co-IP fractions. This demonstrates that $HA-MYC GCN5-B_{FLAG}$ was present in each fraction. Various protein bands were excised from the silver-stained gel and identified by LC-MS/MS sequencing. Data is courtesy of Dr. Ali Hakimi (National Centre for Scientific Research in Grenoble, France). M, molecular marker.

The mass spectrometry data of candidate TgGCN5-B associating proteins are listed in Table IX. TgADA2-A was co-immunoprecipitated with $HA-MYC GCN5-B_{FLAG}$, increasing the confidence in this method. Interestingly, four proteins containing an AP2 domain (50.m03194, 20.m03816, 80.m03948, 33.m01324) were also identified as potential associating proteins of $HA-MYC GCN5-B_{FLAG}$. As discussed in the introduction (Chapter 1, Section IV-A), AP2 domain proteins are the recently discovered lineage of plant-like transcription factors, conserved in Apicomplexa parasites. These AP2 proteins may be transcription factors that recruit TgGCN5-B to specific promoters for regulation of gene expression. Another potential transcription factor identified was the AT-hook motif-containing protein (583.m05282). AT-hook motifs bind to AT-rich regions of DNA and are often found in proteins associated with chromatin interactions or transcriptional regulation [197]. Two associating proteins (42.m03344 and 46.m01622) contain PHD-

finger domains, a motif that can bind to tri-methylated lysine residues, and commonly found on proteins associated with chromatin modulation and gene regulation [210]. There were few proteins identified to be associated with TgGCN5-B that have an unknown function and are labeled as hypothetical. Several other proteins, such as the heat shock proteins, are highly expressed *Toxoplasma* proteins and could be potential contaminating proteins. None of the proteins identified in the HA-MYC GCN5-B_{FLAG} co-IP experiment corresponded to proteins found in both replicates of the affinity chromatography experiments.

Table IX: Proteins co-immunoprecipitated with HA-MYC GCN5-B_{FLAG}

Entry #	Protein Description	Accession Number
1	GCN5-B	49.m03346
2	PHD-finger domain-containing protein	42.m03344
3	hypothetical protein	49.m03263
4	PHD-finger domain-containing protein	46.m01622
5	AP2 domain transcription factor XII-4 (AP2XII-4)	50.m03194
6	AT-hook motif-containing protein	583.m05282
7	ADA2-A	35.m00936
8	AP2 domain transcription factor VIIa-5 (AP2VIIa-5)	20.m03816
9	AP2 domain transcription factor IX-7 (AP2IX-7)	80.m03948
10	AP2 domain transcription factor X-8 (AP2X-8)	33.m01324
11	hypothetical protein	72.m00394
12	hypothetical protein	42.m03515
13	alanyl-tRNA synthetase, putative	38.m01067
14	DnaK family protein	42.m03533
15	heat shock protein 90, putative	49.m00060
16	myosin A, putative	46.m00001
17	cell division protein 48, putative	59.m03661
18	aconitate hydratase, putative	42.m03524
19	elongation factor 2	20.m03912
20	heat shock protein, putative	38.m01113
21	heat shock protein 70, putative	59.m00003
22	heat shock protein 70, putative	583.m00009
23	heat shock protein 70, putative	50.m00085
24	tryptophanyl-tRNA synthetase, putative	80.m00063
25	eukaryotic translation initiation factor 3 subunit 6 interacting protein, putative	59.m00055
26	elongation factor 1-alpha, putative	76.m00016
27	protein disulfide isomerase	27.m00003
28	ATP synthase beta chain, putative	55.m00168
29	DnaJ domain-containing protein	583.m05418
30	hypothetical protein	80.m02161

For each protein, the description provided on the ToxoDB (<http://ToxoDB.org>) as well as the model ID accession number is listed.

F. Confirmation of TgGCN5-B associating proteins

The next step is to confirm through a second approach that TgGCN5-B does indeed interact with one or more of the proteins identified in the co-IP experiment. From the data set, I was most interested in confirming an interaction with one of the AP2-domain proteins or the AT-hook protein because these proteins likely function as gene-specific transcription factors within *Toxoplasma*. Additionally, proteins with these domains have been associated with *Plasmodium* GCN5 [161]. Unfortunately, the PHD-finger proteins are very large (greater than 4,000 amino acids), indicating they could be more difficult to genetically manipulate. As for the other proteins in the data set, there were no precedents for interactions with GCN5s from other species, and some appeared to be possible contaminants, so currently the other proteins are not being pursued.

To begin the confirmation process, I used the method of endogenously tagging a genetic locus in RHΔKu80 parasites [187]. I utilized this method because all the AP2-domain proteins and the AT-hook protein were large proteins (greater than 2,000 amino acids), meaning they would be difficult to clone, tag, and ectopically-express. The first protein that was successfully endogenously-tagged with a 3xHA epitope at its C-terminus was the AT-hook motif-containing protein (583.m05282 – Plasmid #16). This protein was also the same AT-hook protein identified in Aim 1 as containing an analogous NLS to TgGCN5-B (hereafter referred to as AT-hook 056400, its latest accession number). The PHD-finger domain containing protein 46.m01622 (TGGT1_071200) was also identified in the bioinformatics search of Aim 1 and contains a potential analogous NLS to TgGCN5-B. The second protein that was endogenously-tagged with a 3xHA epitope was the AP2-domain protein 20.m03816 (Plasmid #18). Hereafter, this protein will be referred to as AP2-3816. Figure 19 depicts IFA data (Panels A and B) demonstrating that each of these proteins, AT-hook 054600 and AP2-3816, are localized to the parasite nucleus. Western blot data from each parasite strain is shown in Figure 19, Panel C. As expected, AP2-3816 is approximately 260 kDa. Interestingly, the predominant band for AT-hook 054600 is ~80 kDa; however, the expected protein size is over 400 kDa. Either this band is a breakdown product, the protein is truncated, or the protein is predicted incorrectly and is actually smaller than expected.

Currently, reciprocal directed co-IPs are underway to determine if AP2-3816 interacts with TgGCN5-B. For these experiments, parasites are treated with paraformaldehyde prior to harvesting to cross-link proteins and preserve their

associations, and then AP2-3816 is immunoprecipitated from the parasite lysate with anti-HA affinity resin. The IP material is then analyzed by Western blotting with both anti-HA (to confirm the pull-down of AP2-3816) and anti-TgGCN5-B to determine if there is an association. This experiment will also be performed with AT-hook 056400. Additionally, another AP2-domain protein (80.m03948, hereafter referred to as AP2-3948) was recently endogenously tagged in the parasites at the C-terminus with a 2xHA epitope followed by the DD (destabilization domain) [211,212]. Since all these proteins have low basal levels of expression, I wanted to determine if I could increase the protein expression level by exploiting the dynamics of the DD and its ligand Shield-1 (Aim 3). This parasite line is currently being characterized.

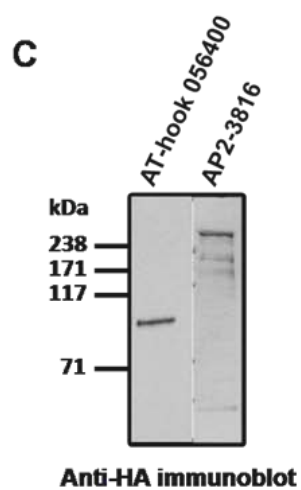
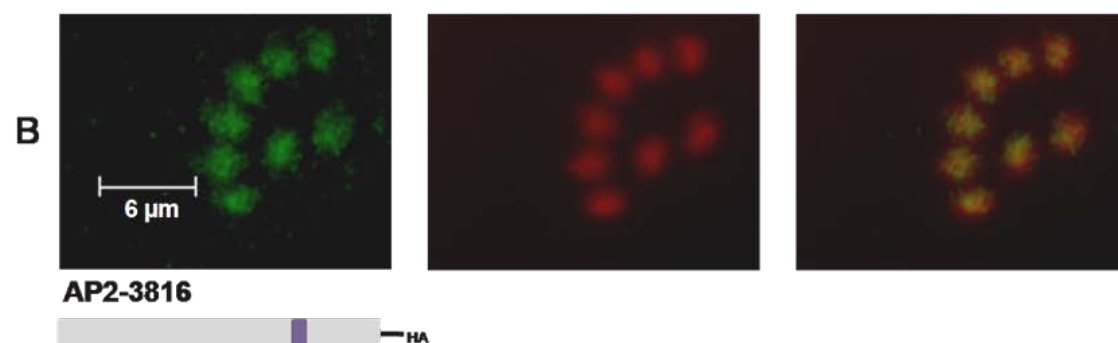
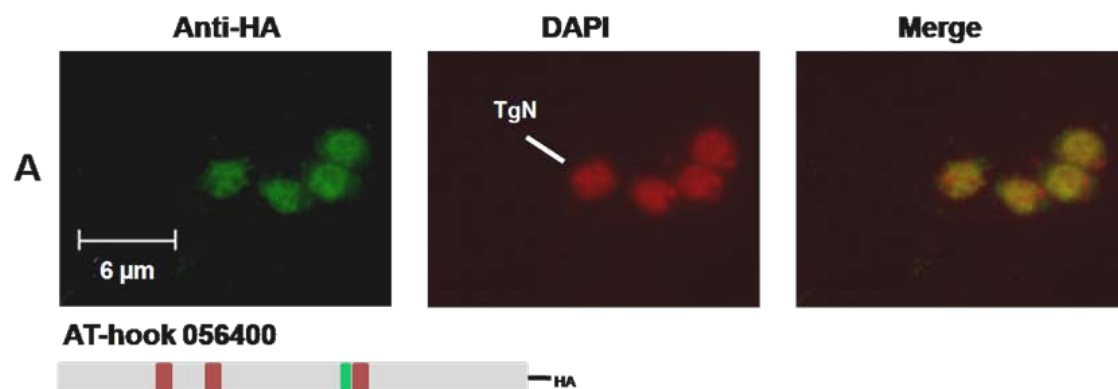


Figure 19: Two TgGCN5-B associating proteins localize to the parasite nucleus.

AT-hook 056400 and AP2-3816 proteins were each tagged endogenously with a C-terminal 3xHA epitope in separate RHΔKu80 parasites. Both proteins are localized to the parasite nucleus (Panels A and B) as determined by IFA with anti-HA. The diagrams to the right of the IFA pictures are schematic representations of the proteins. For AT-hook 056400, red boxes indicate the AT-hook domains, and the green box identifies the location of the putative NLS (Aim 1). In the diagram of AP2-3816, the purple box signifies the AP2 domain. TgN, *Toxoplasma* nuclei; Green = Anti-HA; Red = DAPI, 4',6-diamidino-2-phenylindole. Panel C is an anti-HA immunoblot of 50 µg whole parasite lysate from the strains possessing endogenously tagged AT-hook 056400 or AP2-3816. The predominant band for AT-hook 056400 is considerably smaller than the expected size of ~400 kDa, whereas AP2-3816 is at the expected size of ~260 kDa.

III. Aim 3: Determine the role of TgGCN5-B in *Toxoplasma* physiology

A. Preliminary evidence suggests that TgGCN5-B is essential

Toxoplasma tachyzoites are haploid in the asexual stage. Therefore, to generate a knockout (KO) only a single allelic replacement is needed. Sullivan et al. (2006) have generated a type I parasite strain that lacks TgGCN5-A (Δ GCN5-A) by replacing the genomic locus with a selectable marker through homologous recombination [169]. This strain does not demonstrate any observable difference from wild-type under normal culture conditions, indicating that TgGCN5-A is not an essential gene. However, in media at alkaline pH (8.1) the Δ GCN5-A parasites show deficiencies in responding to stress [Sullivan and Naguleswaran, unpublished data]. After generating the Δ GCN5-A parasite strain, Sullivan and other members of the lab, including myself, attempted to generate a KO of TgGCN5-B using the same method of allelic replacement through homologous recombination, but a KO of TgGCN5-B was never obtained. Since *Toxoplasma* is haploid, disruption of an essential gene will result in non-viable parasites. Therefore, since we were never able to generate a TgGCN5-B KO, we hypothesized that this GCN5 homologue might be crucial to *Toxoplasma* and likely essential. If TgGCN5-B is essential, it could be exploited as a new therapeutic target, although studying an essential gene can be byzantine.

Meissner et al. (2002) developed a conditional KO system for *Toxoplasma* [213], based on the *E. coli* tetracycline-repressor system that allows for gene regulation at the transcriptional level [183]. Meissner et al. (2002) modified this system for *Toxoplasma* by performing a genetic screen using random insertion to identify a transcriptional activating domain for establishment of a tetracycline transactivator-based inducible system [213]. The efficiency of this system has been demonstrated for several genes [213,214,215,216,217,218]. I attempted to use this conditional KO system for TgGCN5-B.

The first step to producing a conditional KO involved generating a parasite line expressing an exogenous copy of TgGCN5-B under the tet-regulatable promoter. Once this line was established, the endogenous locus of TgGCN5-B must then be disrupted through the insertion of a selectable marker via homologous recombination. After completion of both steps, the addition of anhydrotetracycline (ATc) will reduce the level of exogenous protein expression and in effect, knock out the gene. I was able to generate a clonal parasite line expressing a regulatable exogenous copy of TgGCN5-B.

Unfortunately, after several attempts, I was never able to disrupt the endogenous locus of TgGCN5-B. While troubleshooting this system, another approach became available for the study of essential genes.

B. Generation of a regulatable dominant-negative TgGCN5-B

Herm-Gotz et al. (2007) adapted a conditional protein expression system from mammalian cells for use in *Toxoplasma* [219]. This system allows for modulation of the stability of a target protein through coupling with the destabilization domain (DD). The DD is a 12 kDa mutant version of the human rapamycin binding protein FKBP12 [211,212]. When the DD is expressed as a tag on the target protein, this causes rapid degradation of the target protein, likely through the proteasome. However, the target protein can be rapidly stabilized, protected from proteasomal degradation, through the addition of the small membrane-permeable (750 Da) ligand of the DD, Shield-1. This system is reversible and allows the stabilized protein to be tuned to the correct level by varying the concentration of Shield-1. Therefore, the control is at the protein level rather than at the promoter level, as was the case for the previous system. This system has subsequently been marketed by Clontech Laboratories, Inc. as the ProteoTuner™ Systems. The regulation of exogenous protein expression through the DD has been demonstrated for both *Toxoplasma* and *Plasmodium* [219,220]. In particular, the use of the DD to control the expression of dominant-negative alleles was presented as an alternative approach to study essential genes [219]. I took advantage of this approach to study the effects of TgGCN5-B on parasite viability.

A catalytically inactive TgGCN5-B mutant was linked to the DD and expressed within RHΔHX parasites to generate a regulatable dominant-negative TgGCN5-B parasite strain. The mutant TgGCN5-B protein was only expressed in the parasites in the presence of the ligand Shield-1 (Shld), as illustrated in Figure 20. When expressed, the mutant TgGCN5-B would compete with the native protein for protein complex formation and substrates, thus hindering the native protein from functioning properly. As a control, a wild-type version of TgGCN5-B (no mutation) was also expressed under regulation of the DD in a separate RHΔHX parasite line.

To generate the catalytically inactive TgGCN5-B, site-directed mutagenesis was used to mutate the glutamic acid residue 703 to glycine (E703G). This residue is essential for the catalytic mechanism of GCN5 homologues [89]. Additionally, a similar conserved glutamic acid residue is key to the catalytic function of MYST HAT

homologues. Smith et al. (2005) demonstrated that mutation of the homologous TgMYST-A glutamic acid residue 279 to glycine (E279G) hindered catalytic activity of this HAT [170]. Once the mutant cDNA of TgGCN5-B was generated, it was linked to the DD at the N-terminus along with an HA epitope tag (Figure 20). Likewise, a wild-type (non-mutant) TgGCN5-B was constructed in the same manner to serve as a control. Each DD-linked TgGCN5 was inserted into a parasite expression vector for expression under the tubulin promoter (Plasmids #23 and #24).

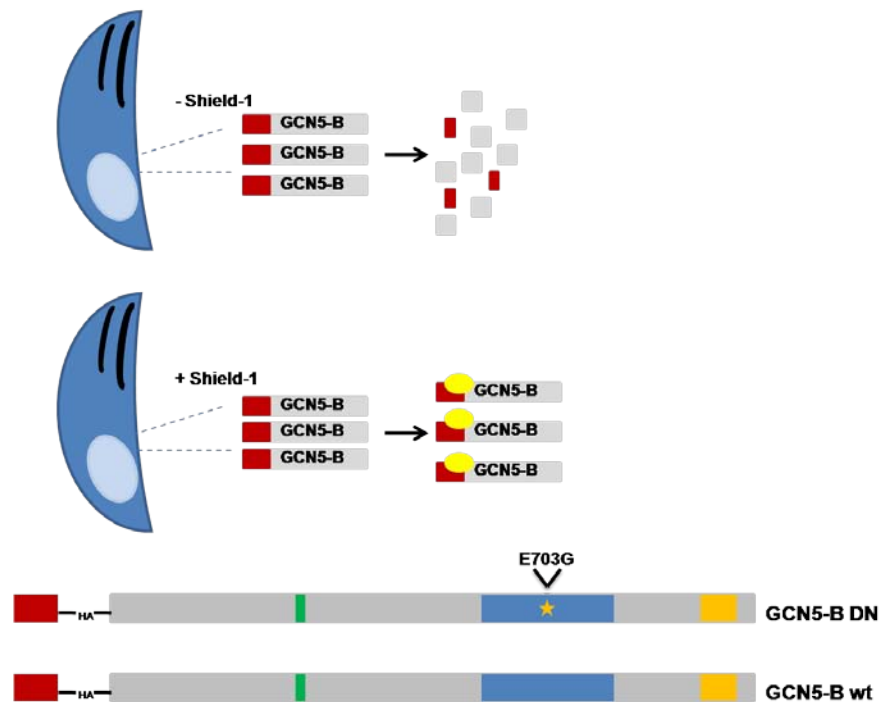


Figure 20: Illustration of conditional protein expression through the destabilization domain. Pictorial representations of *Toxoplasma* tachyzoites transfected with various forms of GCN5-B fused at the N-terminus with the destabilization domain (DD, in red). Without the stabilizing ligand Shield-1 (top), ectopic GCN5-B is degraded, whereas in the presence of Shield-1 (bottom) ectopic GCN5-B is stabilized. The diagrams represent the various GCN5-B constructs, which include the N-terminal DD domain (red) and HA epitope tag. For each GCN5-B protein diagram, the green region signifies the NLS, the blue region represents the HAT catalytic domain, and the orange box signifies the bromodomain. In the top diagram, the star (★) symbolizes the mutation of glutamic acid residue 703 to glycine (E703G), representing the dominant-negative allele (DN). The bottom diagram represents the wild-type enzyme (wt). Figure adopted from Striepen et al. (2007) [221].

Both GCN5-B DN (Plasmid #24) and GCN5-B wt (Plasmid #23) were transfected into RHΔHX parasites so that a clonal line of each could be isolated. After transfection, each parasitic population was under drug selection (20 μM CAM) and subsequently cloned to obtain a homogenous parasite population for each. The selection and cloning of the parasites was not done in the presence of Shld. Once clones for each population were obtained, they were screened for expression and regulation of ectopic GCN5-B protein (DN or wt) using IFA after the addition of Shld (1 μM) for 4 hours. Figure 21 is an IFA from the clonal parasite strains for ectopic GCN5-B (DN or wt) used for all subsequent studies. Shld (1 μM) rapidly stabilized each protein with expression present 2 hours after the addition of the ligand. There was no noticeable difference in protein expression (as viewed via IFA) between 2 and 4 hours following Shld application (data not shown). A single application of 1 μM Shld would stabilize the protein for at least 3 days (data not shown). Samples lacking Shld still included the equivalent amount of vehicle (100% ethanol). In the vehicle treated samples, a slight amount of ectopic GCN5-B protein could be detected for each parasite strain (Figure 21). The presence of this protein likely represented protein being targeted for degradation or partially degraded. As expected, both ectopic GCN5-B proteins localized to the parasite nucleus (Figure 21). Interestingly, parasites expressing GCN5-B DN also appeared to have protein in the nuclear periphery or in the cytoplasm surrounding the nucleus. This was not seen for parasites expressing GCN5-B wt protein, which localized exclusively with the nuclear DNA. Assuming equivalent expression of each ectopic GCN5-B protein, it is possible that the acetylation status of TgGCN5-B might facilitate nuclear localization. Therefore, if the GCN5-B DN cannot be autoacetylated, then its nuclear localization might not be complete or might occur at a slower rate. Figure 22 is a Western blot verifying the IFA data from Figure 21.

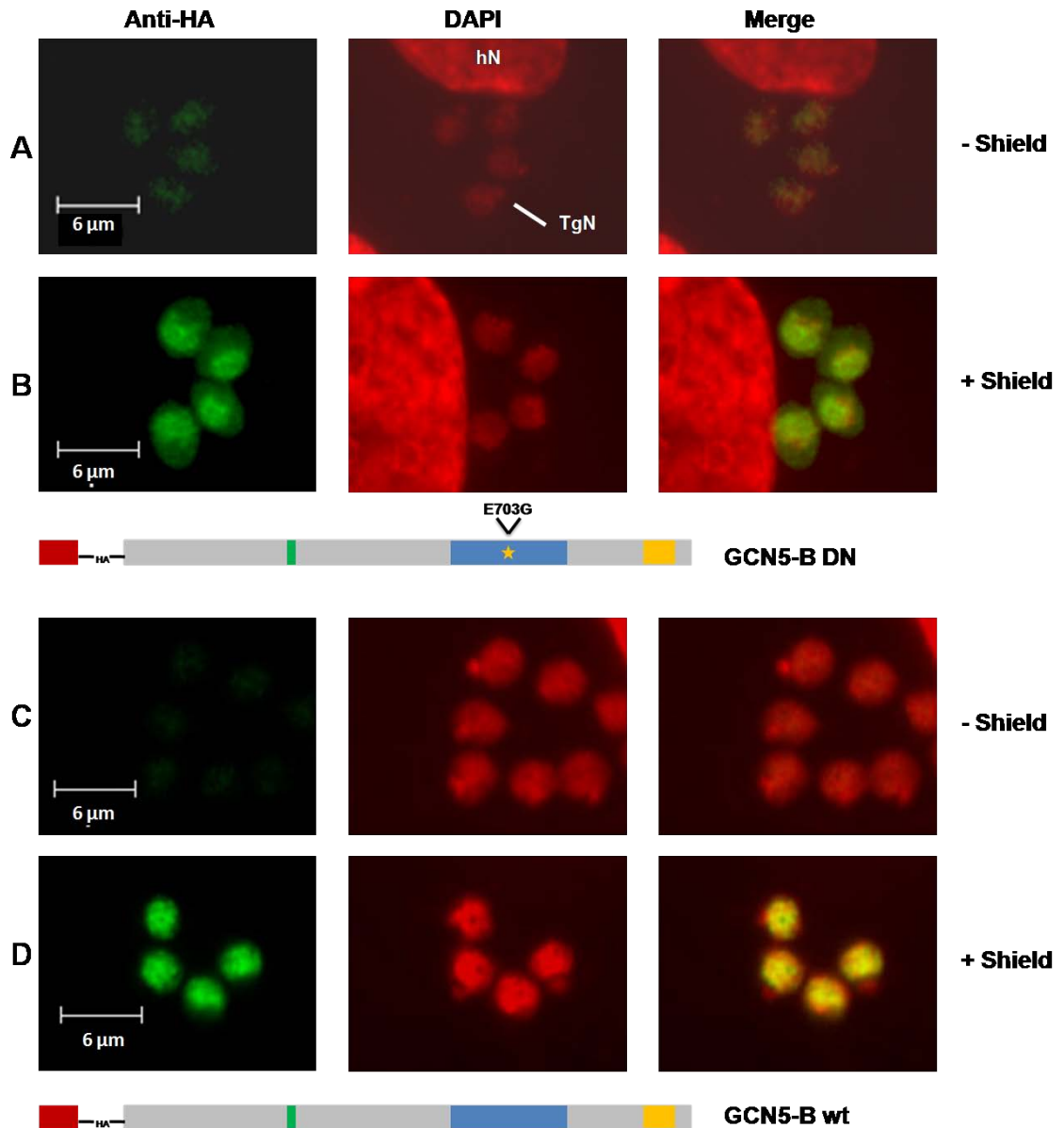


Figure 21: Addition of Shield-1 stabilizes expression of ectopic GCN5-B proteins. IFA data reveals nuclear localization of ectopic TgGCN5-B proteins from clonal parasite populations that were engineered to express either a dominant-negative allele of TgGCN5-B with the point mutation E703G (GCN5-B DN; Panel B) or a wild-type allele (GCN5-B wt; Panel D). The expression of each protein was dependent on regulation via the destabilization domain (DD, red box). Shield-1 (1 μM) or vehicle (100% ethanol) was added to cultures 4 hours prior to processing for IFA. Panels A and C depict a low basal level of expression of ectopic proteins in vehicle-treated samples. TgN, *Toxoplasma* nuclei; hN, host cell nucleus; Green = Anti-HA; Red = DAPI, 4',6-diamidino-2-phenylindole

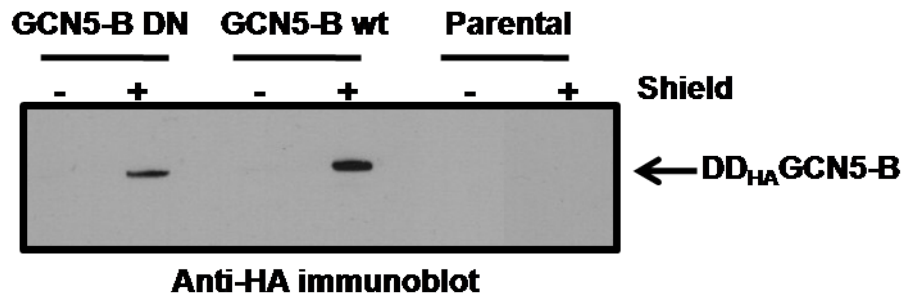


Figure 22: Addition of Shield-1 allows for expression of ectopic GCN5-B proteins.

Shield-1 (+; 1 μ M) or vehicle (-; 100% ethanol) was added to parasite cultures 4 hours before harvest. The whole cell parasite lysate (350 μ g) was immunoprecipitated with anti-HA affinity resin, followed by analysis by Western blotting with anti-HA. Distinct protein bands for each ectopic GCN5-B protein are evident in Shield-1-treated samples (+), whereas background is undetectable in vehicle-treated (-) or parental samples.

C. Expression of dominant-negative TgGCN5-B reduces parasite viability

After generating parasites expressing the ectopic GCN5-B proteins (DN or wt) under regulation by the DD, I wanted to determine if the expression of either ectopic GCN5-B protein altered the growth of *Toxoplasma* tachyzoites. To evaluate parasite growth and viability, I used three different assays. For doubling assays, the number of parasites per vacuole was counted at certain time points post infection. This assay provided insight into the replication rate of the parasites [182]. The B1 growth assay and plaque assay both provided an estimate of the number of parasites per well based on either quantitative real-time PCR for a parasite-specific gene or the number of plaques per well, respectively [182,194]. The data sets gathered from the B1 growth assays and the plaque assays were amendable to standard parametric statistical analysis; however, due to the discontinuous integer scoring of the doubling assay, the data was not analyzable by standard parametric statistics.

Even though it has been reported that the addition of Shld does not alter the growth of *Toxoplasma* tachyzoites, I first repeated these studies with my parental (RH Δ HX) parasites at 1 μ M Shld [219]. I did not observe a significant difference in the growth of parental parasites in the presence of Shld during any of the assays (Figure 23). The data in Figure 23 depicts a single experiment for each assay; however, all assays were repeated with at least three independent replicates each with similar results.

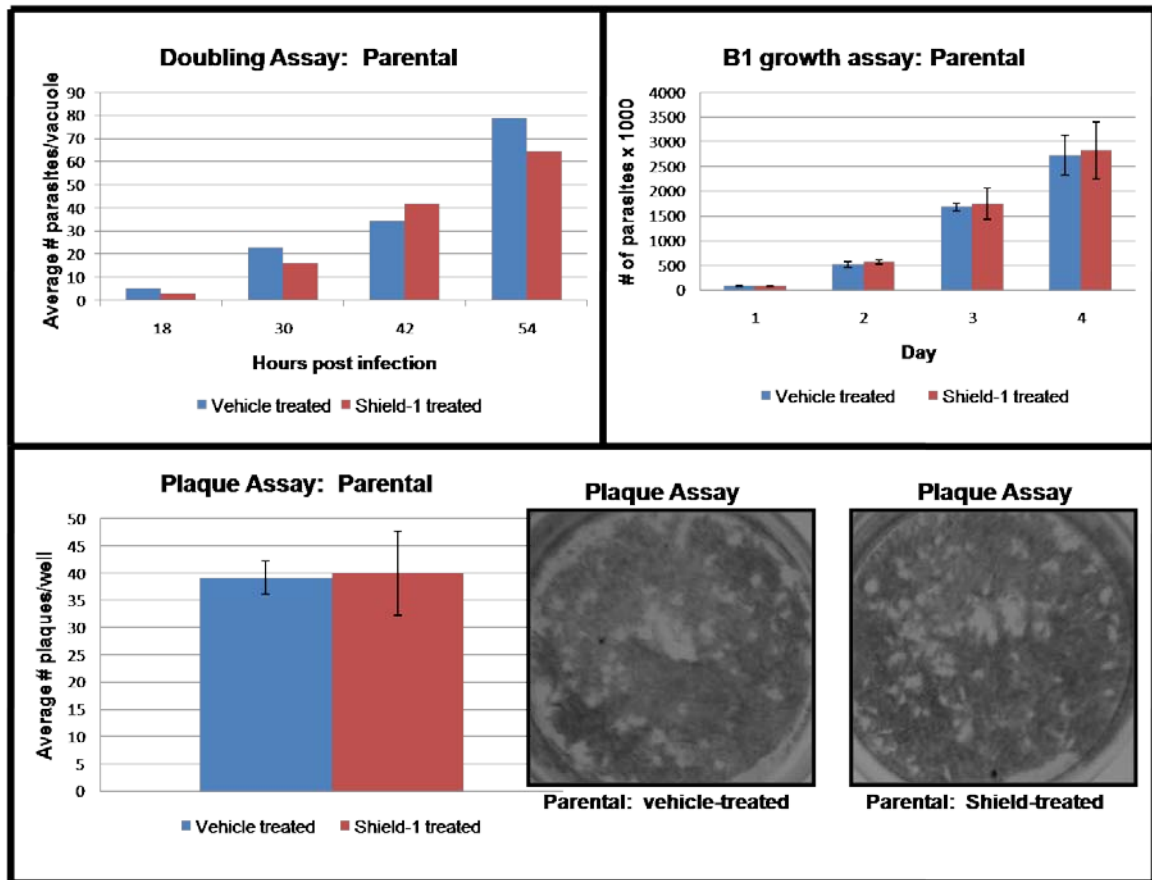


Figure 23: The presence of Shield-1 does not affect the growth of parental parasites. Parental (RHΔHX) parasites were evaluated in three different growth assays. All assays demonstrated that the parasites had similar growth and viability in the presence of vehicle (100% ethanol) and Shield-1 (1 μ M). The data from each growth assay represents one experiment; however, all assays were repeated three times independently with similar results. Error bars represent the standard deviation for each data set. There was not a significant growth difference observed when parental (RHΔHX) parasites were treated with Shield-1 (1 μ M) compared to vehicle control (100% ethanol) as assessed by student's t-test analysis. Pictures representing individual wells stained with crystal violet from the plaque assay are also included. Plaques are visible as light colored areas against the dark background.

Since the presence of Shld does not affect the growth or viability of parental parasites, I proceeded to evaluate the parasite strains expressing the ectopic GCN5-B proteins (DN or wt) using the same growth assays. There were not significant growth differences between GCN5-B wt and GCN5-B DN parasites for vehicle treatment (Figure

24). When compared to the parental parasites, a slight slowing of growth was noted for some time points during some of the assays for both GCN5-B wt and GCN5-B DN parasites. It is likely that this minor growth difference may be due to the presence of the DD and increased proteosomal activity.

Importantly, in the presence of Shld (1 μ M) parasites expressing the GCN5-B wt protein do not have a significant change in viability or growth from the vehicle treated controls. However, the addition of Shld (1 μ M) significantly hindered the growth of parasites expressing the dominant-negative allele (GCN5-B DN) compared to vehicle controls. The growth differences between vehicle and Shld treated GCN5-B DN parasites were quite dramatic. Data from both the B1 and plaque assays indicated that GCN5-B DN parasites stop proliferating in the presence of Shld (Figure 24). This is best illustrated by the pictures of individual wells from the plaque assay (Figure 24), which revealed the near absence of plaques in the GCN5-B DN Shld-treated sample.

The doubling assay data provided insight into the mechanism for the decreased viability of GCN5-B DN parasites under Shld treatment. The GCN5-B DN parasites were able to replicate until approximately 30 hours, albeit at a slightly slower rate. However, after this time point, parasite growth ceased, with the average vacuole size containing 16 – 32 parasites per vacuole for at least another 24 hours. The GCN5-B DN Shld-treated parasites were followed for up to 72 hours, and they never progressed beyond this point (data not shown).

It is important to note that for all growth assays, Shld or vehicle were not added until 2 hours post infection allowing parasites time to attach and invade the host cells. Therefore, the growth defect observed with the expression of GCN5-B DN was independent of adhesion and invasion. Additionally, a dominant-negative allele under DD regulation was also generated for TgGCN5-A (GCN5-A DN). As expected, these parasites did not have an observable growth defect in the presence of Shld (data not shown). This result indicated that the reduced growth phenotype was exclusive to the TgGCN5-B dominant-negative allele and cannot be attributed to the DD system.

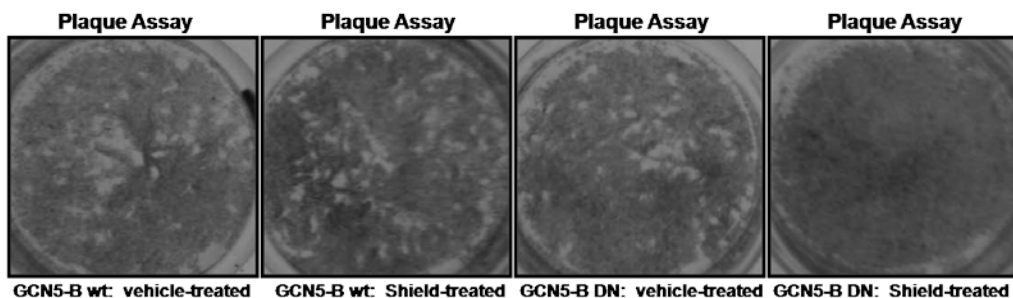
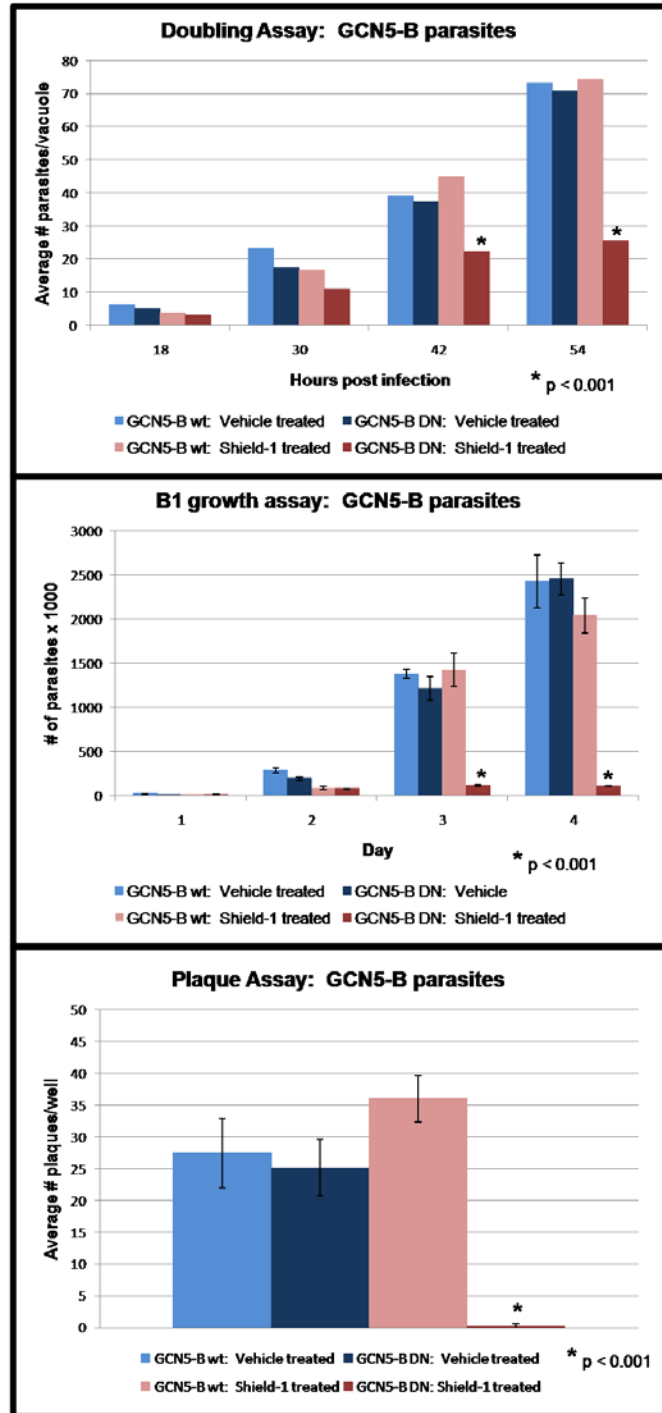


Figure 24: GCN5-B DN parasites cannot proliferate in the presence of Shield-1.

The growth of GCN5-B wt and GCN5-B DN parasites with or without Shield-1 was evaluated in three different assays. All assays demonstrated that there was not a significant difference between GCN5-B wt and GCN5-B DN parasites in the presence of vehicle (100% ethanol). There was also no growth difference between vehicle and Shield-1 treated (1 μ M) GCN5-B wt parasites. On the contrary, the addition of Shield-1 significantly decreased proliferation of GCN5-B DN parasites in all three assays. The data from each growth assay represents one experiment; assays were repeated three times independently with similar results. Error bars represent the standard deviation for each data set. * represents p values less than 0.001 as evaluated by the student's t-test. Pictures representing individual wells stained with crystal violet from the plaque assay are also included. Plaques are visible as light colored areas against the dark background.

Since the growth defect observed for GCN5-B DN parasites in the presence of Shld is independent of adhesion and invasion of host cells, I decided to further investigate the nature of this defect by monitoring common cellular markers in an IFA. As tachyzoite parasites progress through the cell cycle and replicate, distinct stages can be distinguished based on the staining patterns of PCNA (proliferating cell nuclear antigen), IMC1 (inner membrane complex 1), and DAPI (4',6-diamidino-2-phenylindole) [222,223]. Staining with PCNA and DAPI make it possible to monitor changes in nuclear morphology, while staining with IMC1 delineates the perimeter of the parasites, including daughter cell formation with the mother parasite. Typically parasites in G1 phase have a small nucleus and represent approximately 60% of an asynchronous population, whereas parasites in S phase have an expanded and enlarged nucleus and account for 30% of parasites in an asynchronous population. In S phase, the first indication of daughter formation is detectable. Parasites undergoing mitosis and cytokinesis, 10% of parasites in an asynchronous population, have U-shaped nuclei and the developing daughter cells can be identified (Appendix D) [222]. Therefore, IFAs were performed with staining for *Toxoplasma* specific PCNA and IMC1 (Table V) to monitor parasite progression through the cell cycle and division. For IFAs, parasites were allowed to attach and invade host cells for 2 hours prior to the addition of vehicle (100% ethanol) or Shld (1 μ M). Parasites were allowed to grow for either 2 or 3 days before being processed for IFA.

Figure 25 shows representative IFA data from parental and GCN5-B wt parasites treated with Shld as well as GCN5-B DN parasites treated with vehicle. These images are from the Day 3 time point and represent an asynchronous parasite population. The parasites in these images were all normal, although the parasites were at different points in the cell cycle or division (Appendix D). Even parasites within the same vacuole were not synchronized and showed different cellular morphology. Although only a single image is shown, there was not a noticeable difference among any of the parasites on the slide. Additionally, these images are representative of data from parental and GCN5-B wt vehicle-treated parasites on Day 3 and of the corresponding strains/treatments on Day 2.

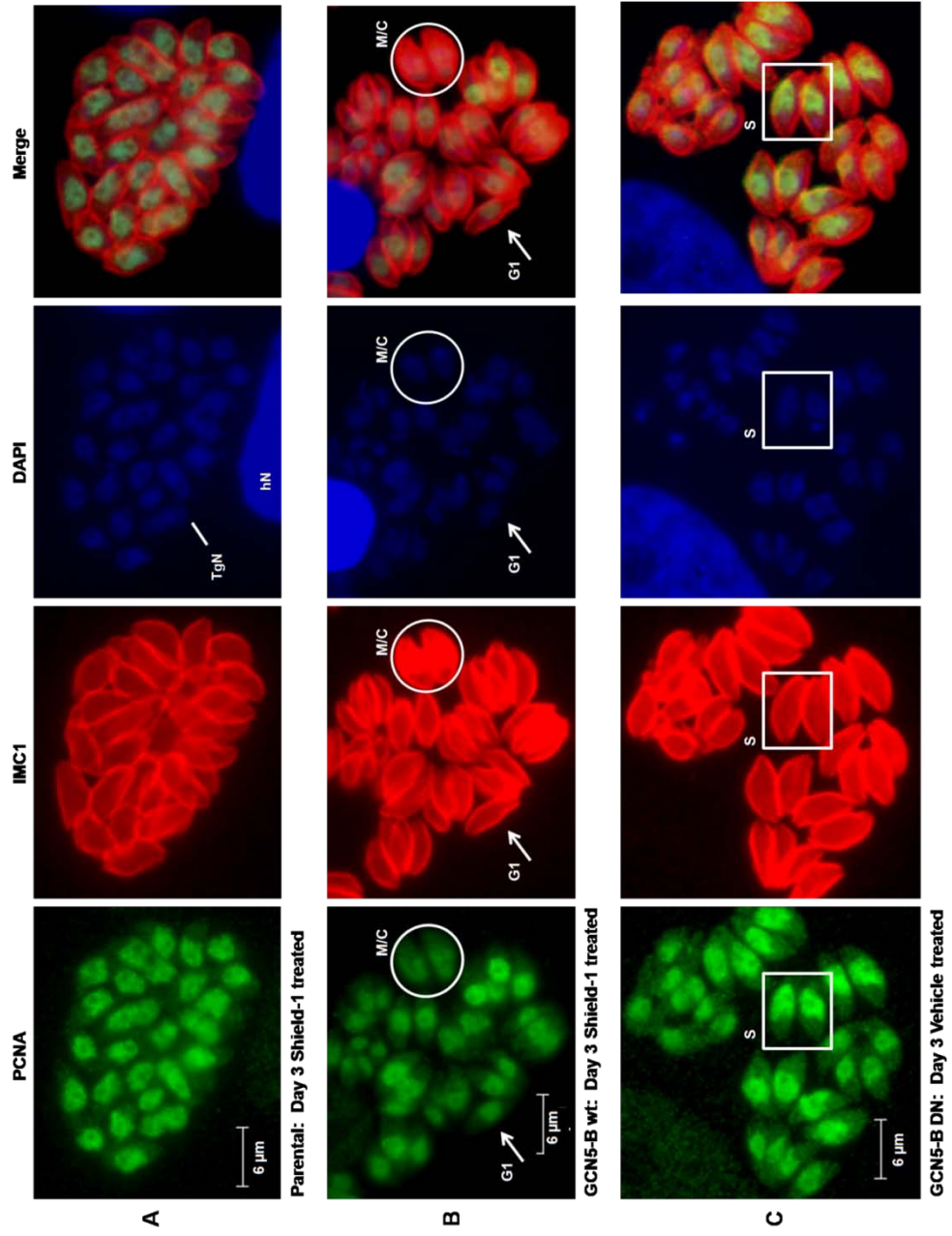


Figure 25: Normal progression through parasite cell cycle and division. IFA images from parasites demonstrated normal cellular morphology after 3 days of Shield-1 treatment (Panel A: Parental; Panel B: GCN5-B wt) or vehicle treatment (Panel C: GCN5-B DN). PCNA and DAPI monitor nuclear morphology through the cell cycle, while IMC1 stains the inner membrane complex, thus revealing the development of daughter cells. TgN, *Toxoplasma* nuclei; hN, host cell nucleus; PCNA, proliferating cell nuclear antigen (green); IMC1, inner membrane complex 1 (red); DAPI, 4',6-diamidino-2-phenylindole (blue); G1, Gap 1 cell cycle phase; S, synthesis cell cycle phase; M/C, mitosis and cytokinesis.

Importantly, IFA data from both Day 2 and Day 3 of the GCN5-B DN Shld treated parasites revealed that the parasites were unable to replicate normally. Figure 26, Panel A shows that after 2 days of Shld treatment, some GCN5-B DN parasites began to display unusual characteristics such as elongation of nuclei, extra-nuclear PCNA staining, and parasites without any appreciable PCNA staining. However, other parasites in this sample showed normal morphology. After 3 days of Shld treatment, these defects became more pronounced, and more parasites were affected. For instance, Panel B depicts parasites with defective daughter cell formation as determined by IMC1 staining as well as the unusual characteristics listed above. Significantly, Panel C depicts a vacuole in which half the parasites lacked PCNA staining and appeared dead, whereas the other half had atypical morphologies and were also likely not viable. These data are preliminary, and the abnormal features and morphologies suggest defects in nuclear division and daughter cell budding in parasites expressing the dominant-negative allele of TgGCN5-B.

The deficient phenotype of the GCN5-B DN parasites was not immediate as some parasites appeared normal even after 2 or 3 days of Shld treatment. This is to be expected with a dominant-negative mutant because it takes time for the mutant allele to be expressed, associate with complex members, and replace native protein. However, the parasites ultimately succumbed to the effects of the dominant-negative allele and were not able to continue normal cellular growth. Collectively, the growth assay and IFA data indicated that TgGCN5-B is critical for *Toxoplasma* tachyzoite survival.

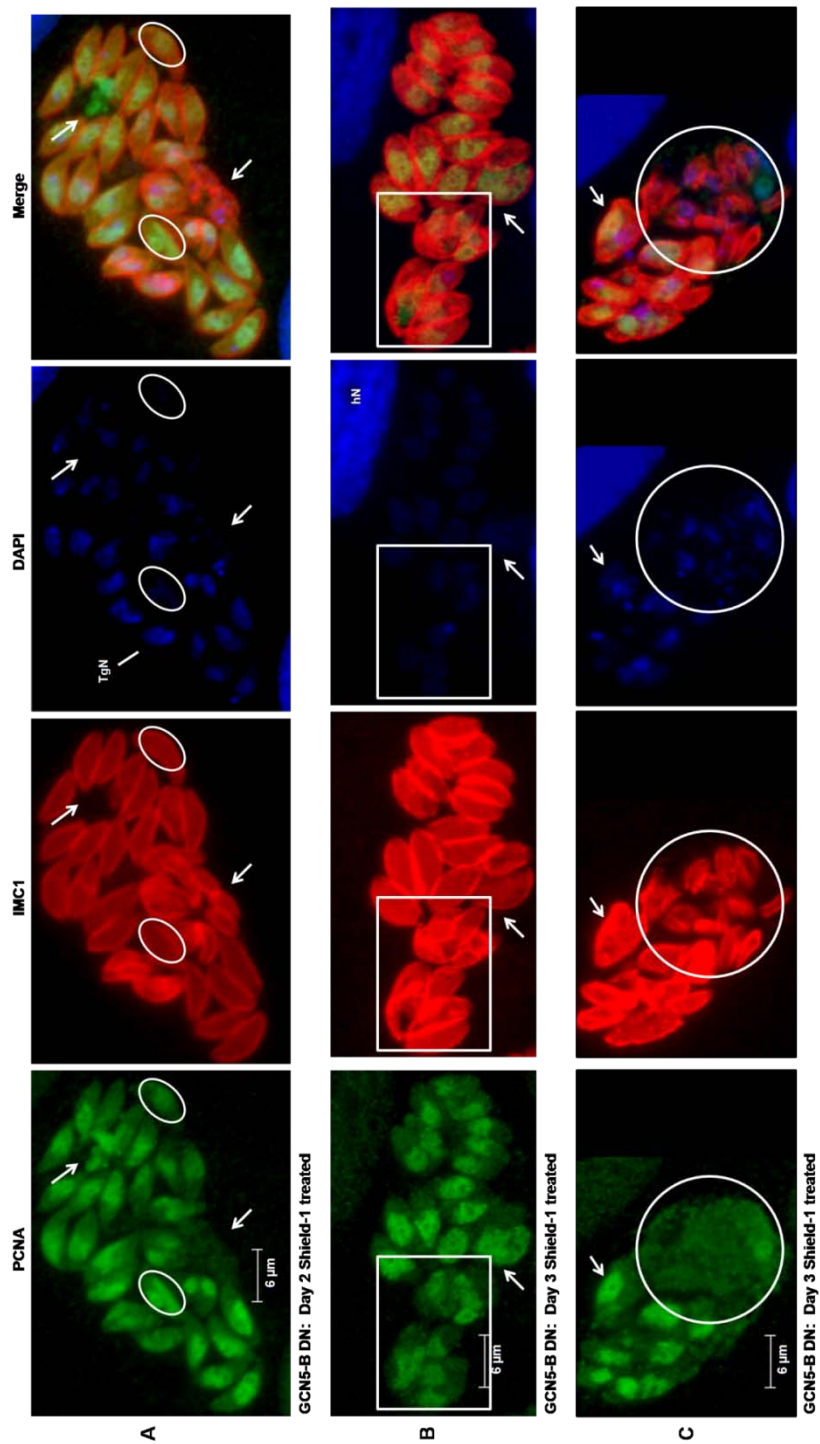


Figure 26: GCN5-B DN parasites display abnormal nuclear division and daughter cell formation when treated with Shield-1. IFA images from GCN5-B DN parasites demonstrated the progression of unusual morphological characteristics (highlighted with arrows, circles, or squares) after treated with Shield-1. Panel A shows a GCN5-B DN parasite vacuole 2 days after addition of Shield-1. Top arrow shows exogenous PCNA staining outside of the parasite, whereas the bottom arrow shows parasites that lack prominent PCNA and DAPI staining. The left circle highlights a parasite with unusual nuclear morphology, while the right circle shows a parasite that lacks detectable DAPI staining. Panel B shows GCN5-B DN parasites 3 days after the addition of Shield-1. Boxed parasites display defective daughter cell formation, as well as nuclear abnormalities. The arrow highlights a large and engorged parasite with excess PCNA and DNA while its sister parasite just to the right lacks both these markers. In Panel C (3 days of Shield treatment), the lower half of the parasite vacuole contains atypical parasites lacking PCNA and having defective membrane formation, suggesting these parasites are dead. Parasites in the top half of this vacuole have abnormal nuclear morphology and atypical daughter cell formation as highlighted by the arrow. TgN, *Toxoplasma* nuclei; hN, host cell nucleus; PCNA, proliferating cell nuclear antigen (green); IMC1, inner membrane complex 1 (red); DAPI, 4',6-diamidino-2-phenylindole (blue).

Chapter 4: Discussion and Future Studies

I. TgGCN5-B harbors a novel NLS within its N-terminal extension

A. Summary of Aim 1 results

It was discovered that TgGCN5-B possesses a novel NLS, ³¹¹RPAENKKRGR₃₂₀, within its N-terminal extension. These 10 amino acids were required for nuclear localization, as the deletion of just these residues excluded the protein from the parasite nucleus. The NLS of TgGCN5-B was also sufficient to localize bacterial β -galactosidase to the parasite nucleus. Despite several attempts, an interaction between TgImp- α or TgImp- β and TgGCN5-B could never be established. Therefore, the chaperone that assists in the transportation of TgGCN5-B to the parasite nucleus remains undefined. It was demonstrated that the NLS of TgGCN5-B has predictive value, as several proteins identified in a bioinformatics search of the ToxoDB contain similar but not identical basic-rich amino acid clusters. One protein identified in the bioinformatics search, AT-hook 056400, was shown to the parasite nucleus.

B. The classification of the TgGCN5-B NLS

As determined by the bioinformatics search, no other known *Toxoplasma* protein contains a motif identical to the NLS of TgGCN5-B. The GCN5-B homologue in *Neospora*, another Apicomplexa parasite, has a similar motif (³⁵²RPVPESKKRGR₃₆₂), but it is not identical. Database searches of other Apicomplexa parasites and related protozoa also failed to identify another protein harboring the exact TgGCN5-B NLS motif. Interestingly, searches of the NLS database (<http://cubic.bioc.columbia.edu/db/NLSdb/>) did not reveal any other NLSs with this exact sequence, thus rendering the NLS of TgGCN5-B unique [139].

NLSs can be classified into two broad sub-categories: monopartite and bipartite. Monopartite NLSs consist of a short stretch of basic amino acids, whereas bipartite NLSs are comprised of two basic clusters separated by 10 – 12 amino acids. The NLS of TgGCN5-B can be classified as a monopartite motif because only 4 amino acids separate the beginning Arg residue from the core basic cluster (KKRGK).

Kosugi et al. (2009) have further categorized NLSs into six different classes, including five classes of monopartite signals, based on sequence homology and binding specificities to importin- α [140]. Using their classification system, the TgGCN5-B NLS

most resembles the Class 2 NLSs exemplified by the sequence algorithm (P/R)XXKR(^DE)(K/R), where X is any amino acid, and (^DE) represents any residue except Asp or Glu [140]. The defining member of Class 2 is the well characterized c-myc NLS, ₃₂₀PAAKRVKLD₃₂₈ [140,224]. Similar to the algorithm and the c-myc NLS, the NLS of TgGCN5-B has a basic-core cluster of both Lys and Arg residues split by a neutral amino acid (Gly). Furthermore, the TgGCN5-B basic cluster is flanked by proceeding residues, which begin with both an Arg and a Pro.

However, the major caveat of suggesting that the TgGCN5-B NLS should be categorized within Class 2 is that the TgGCN5-B NLS does not appear to associate with *Toxoplasma*'s homologue of Imp- α . All the classes of NLSs described by Kosugi et al. (2009) are defined by both sequence similarities and Imp- α binding specificity [140]. Although several *Toxoplasma* proteins have similarities to Imp- α , the homologue identified by Bhatti et al. (2005) is the only protein to have a conserved importin β binding domain [131,144]. It is possible that additional divergent Imp- α homologues exist in *Toxoplasma* because other eukaryotes such as humans possess multiple Imp- α isoforms [130,131].

Kosugi et al. (2009) performed numerous mutational analyses of each NLS class and determined that both the flanking residues and the basic core pattern could influence overall NLS activity [140]. Likewise, Makkerh et al. (1996) have characterized the upstream PAA residues and downstream LD residues that flank the basic-cluster of the c-myc NLS and found both flanks contribute to nuclear import activity [196]. Furthermore, they provided evidence supporting the importance of both neutral and acidic flanking residues, and their ability to rescue single inactive basic-rich clusters [196].

Likewise, the NLS of TgGCN5-B illustrates the significance of additional residues since its basic cluster was not able to result in nuclear localization without the addition of several upstream residues. The nature of the residues is also critical as mutations of both the Arg and Pro, at positions 311 and 312, hindered activity of the TgGCN5-B NLS. This result demonstrates that an upstream Pro or additional basic residue (Arg or Lys) can serve to activate certain NLSs. Interestingly, the residue immediately following the TgGCN5-B NLS is an Asp (D), a residue that augments the activity of the c-myc NLS [196]. When the TgGCN5-B NLS was attached to the C-terminus of β -gal followed by the FLAG epitope tag, this Asp residue was inadvertently included because the FLAG epitope begins with Asp. However, the Asp was determined not to be necessary for the

function of the TgGCN5-B NLS, because replacement of the FLAG epitope with an HA epitope (begins with Tyr) did not alter the nuclear localization.

As discussed in the introduction (Chapter 1, Section III-E), specific Lys residues within the NLS of PCAF are autoacetylated, and these post-translational modifications are critical for proper nuclear localization [141,142]. TgGCN5-B is acetylated (Figure 16). However, acetylation does not appear to be required for the nuclear localization of TgGCN5-B. The mutant protein GCN5-B Δ NLS, which lacks the 10 residue NLS, retained its acetylation status, indicating that residues outside the NLS of TgGCN5-B are acetylated (data not shown). However, this does not conclusively exclude the possibility that the TgGCN5-B NLS is acetylated. The catalytically inactive TgGCN5-B dominant-negative protein maintained its ability to localize to the parasite nucleus; however, localization is not exclusively nuclear as some ectopic GCN5-B DN protein appears to be in the cytoplasm. This distribution pattern is similar to the pattern observed when ectopic TgGCN5-B Δ 310 was expressed in the parasites. The majority of TgGCN5-B Δ 310 localized to the nucleus, but some remained in the cytoplasm. It is possible that a post-translational modification, such as acetylation, could facilitate or retain the nuclear localization of both these TgGCN5-B proteins. Alternatively, the presence of GCN5-B DN beyond the parasite nucleus could be an artifact of over-expression.

C. Utility of the TgGCN5-B NLS

The NLS of TgGCN5-B was shown to have predictive value as assessed by a bioinformatics search of the ToxoDB for predicted proteins harboring a similar amino acid pattern. Various permutations of the TgGCN5-B NLS were identified in several predicted proteins (Table VI), some being known nuclear proteins such as DNA polymerases. To further evaluate the predictive value of the TgGCN5-B NLS, an identified protein, AT-hook 056400, was shown to localize to the parasite nucleus. However, the analogous basic-rich cluster of AT-hook 056400 (_{2,515}RPKKRRR_{2,522}) would need to be further evaluated in order to confirm that the motif functions as an NLS.

Since many *Toxoplasma* proteins remain uncharacterized and have unknown functions, identifying potential motifs such as NLSs in an uncharacterized protein can contribute to determining the function of the protein. Likewise, the predictive value of the TgGCN5-B NLS can be utilized to search other Apicomplexa and related protozoa

databases for the recognition of similar motif patterns, thereby identifying other potential nuclear proteins.

Intriguingly, the NLSs of both *Toxoplasma* GCN5s fall within regions of the proteins that are considered intrinsically disordered (Figure 12 and data not shown). Although unstructured and lacking a three-dimensional confirmation, regions of intrinsic disorder can be attributed to important biological functions, such as protein-protein interactions [204,205,207]. Combining the predictive capabilities of the TgGCN5-B NLS with computational predictors of disorder could further enhance the recognition of signal motifs within uncharacterized proteins.

As mentioned in the introduction, HAT inhibitors are notorious for their low cellular permeability. However, conjugation of an oligo-Arg peptide to a coenzymeA analogue HAT inhibitor resulted in cell permeability [225]. Since NLSs are small peptide motifs rich in Arg and Lys residues, could they also serve a similar function and enhance molecular inhibitor design? Additionally, coupling of the SV40 large T antigen NLS with halogenated monocyclic aromatic compounds increases the cellular uptake of the peptide, suggesting that manipulation of NLSs could enhance their cellular permeability [226]. Taken together, it is conceivable that NLS peptides might be useful tools for not only localizing molecules to the nucleus, but also assisting in cellular permeability and could be utilized in future pharmacological development.

Furthermore, Kosugi et al. (2008) have described a novel method of designing peptide inhibitors based on an activity-based profile representing the functional contribution of individual amino acids within a peptide sequence [227]. Using this method, they developed two peptide inhibitors that bind specifically to importin- α and disrupt classical nuclear import [227]. Likewise, peptide aptamer technology has been utilized to inhibit the NLS of a human cytomegalovirus protein resulting in decreased viral replication [228]. UL84 is an essential human cytomegalovirus replication factor, which harbors a nonconventional NLS of 282 amino acids. This NLS interacts with Imp- α in a mechanistically distinct manner. Exploiting these unique characteristics, Kaiser et al. (2009) isolated several peptide aptamers (small random peptide sequences attached to inactive scaffolds) that inhibit the UL84 NLS and Imp- α association and have antiviral activity [228,229]. Peptide aptamers act similar to antibodies, binding specifically to target sequences, and thereby blocking the interaction of other molecules [229]. Targeted inhibition of protein-protein interaction with peptide aptamers is an approach that should be considered as means to exploit the unique properties of other NLSs.

D. Future studies

It would be beneficial to determine the chaperone that binds the NLS of TgGCN5-B and facilitates its localization into the parasites nucleus. However, *Toxoplasma* possesses several potential proteins with similarities to Imp- α , and other importin-family homologues; therefore, many possibilities exist and there is no obvious insight to suggest one protein over the others [131,144]. It would be interesting to determine if TgGCN5-B could interact *in vitro* with yeast Imp- α . The SV40 large T antigen NLS is capable of interacting with Imp- α homologues from other species [130]; therefore, it is plausible that the NLS of TgGCN5-B could interact with another Imp- α from other species. If such an interaction were to occur, it could help to determine if another Imp- α homologue in *Toxoplasma* is likely to be the TgGCN5-B nuclear chaperone, or if one of the other classes of chaperones (Imp- β or transportins) should be further investigated [131]. Additionally, if one of the inhibitors of Imp- α identified by Kosugi et al. (2008) were expressed in the parasites and inhibited localization of FLAG-tagged TgGCN5-B, then this would indicate an Imp- α type chaperone [227]. To exclude the possibility that TgGCN5-B does not interact with TgImp- β , the entire TgImp- β should be cloned and utilized in co-IP studies. If it is discovered that TgGCN5-B does indeed interact with a unique or novel chaperone to enter the parasite nucleus, then this interaction might be an ideal target for inhibition by protein aptamers.

To complete the characterization of the TgGCN5-B NLS, it should be determined if any of the lysines within the NLS are acetylated. Although the acetylation status of TgGCN5-B does not appear to regulate nuclear localization, this should be confirmed. Mass spectrometry can be used to determine the lysine residue(s) that are acetylated on TgGCN5-B.

Another future study should determine if the analogous motif identified in AT-hook 056400 functions as an NLS. This can simply be assessed by adding these residues to β -gal and determining its localization via IFA.

II. TgGCN5-B associates with novel Apicomplexa transcription factors

A. Summary of Aim 2 results

The N-terminal extension of TgGCN5-B is devoid of identifiable motifs, other than the NLS described in Aim 1. Additionally, the protein sequence of the N-terminal extension of TgGCN5-B is not homologous to other characterized GCN5 homologues; hence, the function of this region is yet undetermined [169]. Bioinformatics reveal that the N-terminal extension of TgGCN5-B contains several distinct regions of intrinsic disorder, suggesting the extension might have a notable biological function, possibly as a region for protein-protein interactions [204,205,207]. Classically, GCN5 homologues function within large multi-subunit complexes that are conserved among species; however, *Toxoplasma* lacks many of these conserved components [104,202]. However, La Count et al. (2005) demonstrated that the N-terminus of *Plasmodium* GCN5 associated with a variety of proteins and was at the center of the largest protein interaction network within *Plasmodium* [161]. To determine the proteins that associate with TgGCN5-B two biochemical techniques, affinity chromatography and co-immunoprecipitation were utilized. The results indicate that GCN5-B associates with several proteins including novel Apicomplexa transcription factors.

B. Comparison of affinity chromatography and co-IP techniques

The two biochemical techniques used to determine TgGCN5-B associating proteins yielded very different results, with no proteins being identified in all three experiments (2 affinity chromatography experiments and a single co-IP experiment). Based on several observations, I feel that the co-IP experiment and its results are more likely to reflect actual protein associations with TgGCN5-B compared to the affinity chromatography experiments.

First, a recombinant TgGCN5-B produced in bacteria was used for the affinity chromatography. Although this protein was catalytically active, it is possible that when produced in this manner, parts of the protein could exist in a non-native conformation or be mis-folded. If structural confirmation is important for protein interaction, then the conditions of the affinity chromatography may not have been ideal for native protein-protein interactions to occur. Likewise, any post-translational modifications of TgGCN5-B that facilitate protein-protein interactions would likely not be present on the recombinant protein, and therefore not contribute to protein associations. As

demonstrated in Aim 2, TgGCN5-B is acetylated; however, the implication of this mark has yet to be determined.

The affinity chromatography results included a large amount of ribosomal proteins and other obvious cytoplasmic contaminants, whereas the data from the co-IP of TgGCN5-B had less cytoplasmic contaminants. The multiple column fractionations incorporated into the co-IP experimental protocol facilitated the removal of these contaminants and reduced the likelihood of non-specific associations.

Finally, since the affinity chromatography experiment failed to identify either TgADA2-A or -B as associating proteins of TgGCN5-B, I had little confidence in the data generated from this experiment. It has been shown that TgGCN5-B can interact with both TgADA2-A and -B by yeast two-hybrid, so the absence of either from the affinity chromatography data is unusual [169]. The presence of TgADA2-A in the co-IP data set further increases the confidence in this approach.

However, there were still some interesting proteins identified in the affinity chromatography experiment. For instance, TgPRMT5 (protein arginine methyltransferase 5; 33.m01376) was identified in the first replicate of the affinity chromatography experiment. It has been reported that TgPRMT5 modifies histone H3R2 [152]. Since both TgGCN5-B and TgPRMT5 modify lysines on histone H3, it is possible that these proteins interact in a collaborative manner to regulation gene expression. Perhaps the methylation of histone H3R2 modulates the acetylation of histone H3 or vice versa. Additionally, a putative 14-3-3 protein (55.m00015) was also identified in the first replicate of the affinity chromatography experiment. 14-3-3 proteins are important eukaryotic regulatory molecules capable of binding phosphoserine and phosphothreonine in variety of proteins, including kinases and phosphatases, and thereby regulate a multitude of cellular events [230,231]. Intriguingly, it has been shown that a 14-3-3 protein binds and prevents the nuclear localization of HDAC4, consequently hindering the function of HDAC4 [232]. It is possible that the putative 14-3-3 protein 55.m00015 could bind and regulate the function of TgGCN5-B. Additionally, 14-3-3 proteins were shown to be acetylated, thereby modulating their binding affinities for other substrates [233]. Therefore, the putative 14-3-3 protein 55.m00015 might be a substrate of TgGCN5-B. Unfortunately, since both of these proteins, TgPRMT5 (33.m01376) and the 14-3-3 protein 55.m00015, were only identified in the first replicate of affinity chromatography and not in the second replicate, I did not select them for further analysis.

A simple method to refine the affinity chromatography protocol to reduce the number of contaminating proteins would be to excise bands from the SDS-PAGE gel rather than analyzing the entire sample. The excision of prominent protein bands was done in the co-IP experiment. Although each excised band is likely to contain multiple proteins, this technique would reduce the number of non-specific associating proteins.

C. The pursuit of Apicomplexa AP2s

Toxoplasma and other apicomplexans are without the conventional eukaryotic specific transcription factors involved in regulating gene expression. However, it was discovered that Apicomplexa harbor proteins containing the AP2 DNA-binding domain commonly found in plant transcription factors [157]. Furthermore, studies in *Plasmodium* demonstrate that select AP2 domains, do indeed bind to specific DNA motifs, and one AP2 in particular (PF11_0442, AP2-O) has been attributed to the activation of stage-specific genes [159,160]. Interestingly, after their discovery, AP2 proteins were shown to be associated with chromatin remodeling enzymes. One of the proteins identified in the purification of the HDAC3 complex in *Toxoplasma* was subsequently found to be an AP2 protein (TgCRC350, TGME49_0727100) [152,154]. In *Plasmodium*, proteins capable of associating with N-terminal extension of PfGCN5 were discerned via a high-throughput yeast two-hybrid screen [161]. From this screen, it was later discovered that 2 AP2 proteins associated with PfGCN5 (PF10_0075 and MAL8P1.153.) [152,161]. The AP2 protein PF10_0075 also bears a putative AT-hook DNA binding motif [161]. Importantly, the high-throughput yeast two-hybrid screen revealed that many protein-protein interaction networks exist in *Plasmodium*, and that PfGCN5 is at the center of the most highly connected network; therefore, PfGCN5 is the most interconnected protein in *Plasmodium* [161]. Six *Plasmodium* AP2 proteins have been mapped to interaction networks consisting of several other proteins involved in chromatin remodeling and transcription regulation [152]. Interestingly, an AP2 factor from *Arabidopsis* (CBF1) associated within a GCN5 complex, thereby providing another connection between GCN5 and AP2 domains [234].

From this evidence, I decided to pursue the AP2 proteins that were identified in the co-IP experiment for further confirmation. I also included the AT-hook motif protein since this domain was also found on one of the *Plasmodium* AP2 proteins. The AP2 protein 33.m01324 was only recently annotated as an AP2 protein, and it is currently being pursued. I was never able to amplify gDNA from the C-terminus of AP2

50.m03194 (required to endogenously tag the protein), hindering the examination of this protein. However, the AP2 proteins 20.m03816 and 80.m03948 as well as AT-hook 583.m05282 (AT-hook 056400) were endogenously tagged within separate parasites strains and are currently being evaluated to confirm interactions with TgGCN5-B.

It should be noted that the *Toxoplasma* homologues of the PfGCN5 associating proteins were not identified in the co-IP experiment. However, some of the proteins in these two data sets have similar protein motifs. Apart from the AP2 and AT-hook domains, PFF1440w is a PfGCN5 interacting protein that contains a PHD domain and two TgGCN5-B associating proteins (42.m03344 and 46.m01622) also contain this domain [161]. It is plausible that in each organism, GCN5 forms specific complexes that mediate definitive cellular processes. In order to perform these specialized roles, the associating proteins differ; however, functional domains, such as the AP2 or PHD domains, might remain conserved. Likewise, identification of the TgGCN5-A complex is expected to yield a distinct set of associating proteins with conserved functional domains. Since *Toxoplasma* possesses two GCN5 homologues, an unusual feature for a lower eukaryote, it is predicted that each TgGCN5 forms discrete complexes in order to regulate distinct and independent processes within *Toxoplasma*.

D. The roles of the N-terminal extensions in Apicomplexa GCN5s

The major differences among GCN5 homologues exist at their N-terminal extensions (Figure 3). Yeast GCN5, with a short N-terminus must associate within a complex in order to acetylate nucleosomal histone, whereas mammalian GCN5s with longer N-terminal extension readily acetylate nucleosomal substrates [235]. This analysis suggests that the N-terminal extension facilitates the binding affinity of GCN5s for their substrates. The metazoan GCN5 homologues possess a PCAF homology domain within their N-terminal extensions [104]. In PCAF, this region contains the ubiquitin E3 ligase domain, although ubiquitinase activity has not been demonstrated for GCN5 homologues other than PCAF [83]. The remaining functions of the N-terminal extensions have yet to be defined.

The *Plasmodium* high-throughput yeast two-hybrid study provided direct evidence that the elongated N-terminal extensions of Apicomplexa GCN5s are involved in protein-protein interactions [161]. Therefore, it is very likely that some of the proteins associating with TgGCN5-B do so via the N-terminus. This is in agreement with the bioinformatics data demonstrating that TgGCN5-B N-terminus is intrinsically disordered.

It is also plausible that the Apicomplexa GCN5 N-terminal extensions also function as the mammalian extensions and enhance the activity and substrate recognition of these enzymes. However, the N-termini of the *Plasmodium* GCN5 and that of the two *Toxoplasma* GCN5s do not share similar sequence homology [169,178]. These results could indicate that each homologue associates with distinct proteins to perform different cellular functions. The Apicomplexa GCN5s do not contain the PCAF homology domain, and currently ubiquitinase activity has not been demonstrated for any of these enzymes.

E. Future studies

The first study, which is currently underway, is to confirm through an independent experiment that TgGCN5-B does indeed associate with one or more of the proteins identified in Aim 2. As discussed previously, the AP2 proteins and AT-hook motif protein seemed to be the most logical proteins with which to begin this study. It would also be interesting to determine if TgGCN5-B and these proteins regulate similar groups of genes. Currently, ChIP-to-chip (chromatin immunoprecipitation coupled to microarray analysis) has been shown to determine the gene networks regulated by certain chromatin modifying enzymes [166]. Applying such a strategy to both TgGCN5-B and an associating AP2 protein would reveal the patterns of genes regulated by both proteins and could reveal commonalities between the two gene networks.

Another interesting experiment would be to determine if the TgGCN5-B associating proteins differed during stress conditions or between type I and type II parasite strains. It is possible that other proteins interact with TgGCN5-B only under certain environmental conditions in order to regulate a different set of genes. Likewise, the determination of the TgGCN5-A associating proteins will likely reveal a different set of associating proteins, as we hypothesize that these two HATs form distinct complexes within the parasites. The elucidation of additional TgGCN5 complex members will further enhance how these HATs regulate *Toxoplasma* gene expression.

Although it was noted that TgGCN5-B is acetylated, the mechanistic details and implications of this post-translational modification have yet to be determined. Additionally, *S. cerevisiae* GCN5 is a substrate for sumoylation, and although this modification does not alter catalytic activity, it appears to have a regulatory role [236]. The proteins required for sumoylation are conserved in *Toxoplasma*, making this post-translation modification of TgGCN5-B a possibility [237]. Moreover, the phosphorylation

of human GCN5 inhibits its catalytic function [238]. It should be determined if these post-translational modifications also occur on TgGCN5-B.

Finally, if a unique and specific interaction is determined to occur between TgGCN5-B and one of the associating proteins, then this interaction might be ideal for targeting with inhibitor peptide aptamers (Chapter 4, Section I-C). However, it should be determined if TgGCN5-B and its associating protein interaction is critical for propagation of the parasites or differentiation prior to testing aptamer inhibitors.

III. TgGCN5-B dominant-negative phenotype decreases parasite viability

A. Summary of Aim 3 results

Traditional attempts to KO TgGCN5-B in haploid *Toxoplasma* tachyzoites failed, suggesting this gene might be essential. In order to study the role of TgGCN5-B within the parasites, a dominant-negative mutant was generated. A catalytically inactive TgGCN5-B mutant protein was expressed in the parasites under the regulation of the DD. When induced to express the mutant TgGCN5-B protein, parasites demonstrated a reduction in viability as assessed in three different growth assays. IFA data revealed that parasites expressing the mutant TgGCN5-B protein had defective nuclear division and daughter cell formation. A recombinant wild-type version of TgGCN5-B protein and a catalytically inactive TgGCN5-A mutant protein, expressed in the same manner, did not show an observable phenotype, indicating the growth defects are exclusive to the dominant-negative TgGCN5-B.

B. GCN5 regulates normal progression through the cell cycle

GCN5 appears to be intimately involved in cell-cycle regulation. *S. cerevisiae* mutants that are deficient in GCN5 accumulated in the G2/M phase, whereas the lack of GCN5 in DT40 mammalian cells caused suppression at the G1/S phase transition [120,239]. The latter study not only demonstrated that GCN5 regulates the transcription of key cell-cycle factors, but also that GCN5 influenced the transcription of apoptosis-related genes [120].

GCN5 is critical for proper mitotic progression [240,241]. *S. cerevisiae* mutants lacking GCN5 demonstrated improper nuclear segregation as well as a delay in mitotic spindle elongation [240]. Additionally, this study demonstrated that GCN5 associated with centromeres, was able to regulate the variant centromeric nucleosomes, and is

important for kinetochore function [240]. Subsequently, examination of the knock-down of ADA2a and ADA3 in mammalian cells, which causes GCN5 to dissociate from the ATAC (Ada Two A Containing) complex, resulted in mitotic dysfunction manifested by defects in nuclear division, spindle formation, and centrosome multiplication [241]. The ATAC complex was shown to localize to the mitotic spindle and control cell cycle progression through the GCN5-mediated acetylation of cyclin A, resulting in the degradation of this protein. In the absence of GCN5 acetylation, cyclin A phosphorylates the HDAC SIRT2 thereby decreasing its deacetylase activity and ultimately leading to accumulation of hyperacetylated α -tubulin [125,241]. Additionally, GCN5 was found to acetylate the cell-division cycle (CDC)-6, a protein essential for initiation of DNA replication and promoting cell-cycle progression, with this modification contributing to the regulation of the CDC6 [124]. Importantly, a global analysis of protein acetylation status revealed that a plethora of cell-cycle related genes are acetylated, further suggesting that HATs are key regulators of the cell cycle [233].

Given the results of these studies, it is not surprising that the dominant-negative TgGCN5-B mutant exhibited dramatic growth reduction. Interestingly, the defective nuclear division and deficient daughter cell budding morphologies seen in the GCN5-B dominant-negative mutant resemble the phenotype of *Toxoplasma* tachyzoites exposed to oryzalin, a microtubule-disrupting agent [242]. In the presence of 2.5 μ M oryzalin, parasites demonstrated defective nuclear division through disruption of the spindle microtubules [242]. Given this similarity and the previous findings that GCN5 mutants have defects in mitotic spindles, it is likely that TgGCN5-B plays a distinct role in regulating *Toxoplasma* nuclear division through modulation of mitotic spindle microtubules. Although *Toxoplasma* does not appear to have a homologue of cyclin A, two divergent cyclins have been identified in *Toxoplasma* [243]. It is possible that TgGCN5-B is able to acetylate and regulate one or more of these cyclins in a manner similar to the regulation demonstrated for cyclin A. *Toxoplasma* appears to have a homologue of CDC6 (TGGT1_035930). It would be interesting to determine if this protein is acetylated by TgGCN5-B. Alternatively, TgGCN5-B might act as a global regulator of *Toxoplasma* cell-cycle progression. It is possible that the dominant-negative TgGCN5-B mutant could mimic GCN5-null DT40 cells, which demonstrates that a broad range of cell-cycle genes are under the transcriptional regulation of GCN5 [120].

Parasites lacking TgGCN5-A do not have a growth defect when grown in normal culture conditions [169]. Also, a dominant negative TgGCN5-A mutant regulated

through the DD did not display growth defects, indicating that the phenotype of the TgGCN5-B dominant-negative is specific for the GCN5-B HAT. It is likely that the *Toxoplasma* GCN5s form distinct complexes, and I propose that the TgGCN5-B complex is exclusively responsible for regulating parasite nuclear division and daughter cell budding through modulation of mitotic spindle microtubules. Therefore, TgGCN5-A cannot compensate for the loss of TgGCN5-B.

C. TgGCN5-B is a novel therapeutic candidate

Given the severe growth phenotype exhibited by the TgGCN5-B dominant negative mutants, it is likely that specific pharmacological inhibition of TgGCN5-B will hinder *Toxoplasma* tachyzoite growth and viability. However, there are still many important considerations that must be taken into account when attempting to therapeutically target TgGCN5-B. First, the human host cells will also possess GCN5 homologues; therefore, it is critical that an inhibitor be specific to TgGCN5-B and has little or no effect on human GCN5 and PCAF. Second, it has not been determined if TgGCN5-B is crucial to bradyzoite development or maintenance. The generation of the dominant-negative TgGCN5-B mutant should be recapitulated in a type II strain, one capable of converting to bradyzoites *in vitro*, in order to determine the effect of the loss of TgGCN5-B on this form of the parasites. Ideally, a novel anti-toxoplasmosis therapy should not only target tachyzoites but inhibit the development and maintenance of bradyzoites to clear the infection. Finally, the current problems with HAT inhibitors such as low cell permeability and low specificity should be addressed.

If a small molecular inhibitor screen is initiated to find inhibitors of TgGCN5-B, a high-throughput assay must be developed in order to screen the inhibitors. A standard HAT assay utilizing recombinant enzyme would be ideal for this application. Previously, attempts to produce active recombinant TgGCN5-B from bacteria have failed [Sullivan and Bhatti, unpublished data]. However, I have shown that expression of recombinant TgGCN5-B fused at the N-terminus to maltose-binding protein (MBP) is active when produced in *E. coli* (Figure 14B). Large-scale production of recombinant MBP-GCN5-B would be ideal to use in inhibitor screening assay.

As an alternative approach, the associating proteins of TgGCN5-B could be novel targets for inhibitor design. The AP2-domain proteins that associate with TgGCN5-B represent a class of proteins unique to Apicomplexa parasites that are not found in humans. An inhibitor capable of blocking the interaction of one of the AP2-

domain proteins with TgGCN5-B would be specific to *Toxoplasma*, minimizing the effect on the human host cells.

D. Future studies

The first studies should further characterize the TgGCN5-B dominant-negative phenotype. For instance, studies are currently underway to confirm that histone H3 acetylation is decreased when the dominant-negative mutant TgGCN5-B protein is expressed. An overall decrease in histone H3 acetylation is expected since histone H3 is a major target of GCN5 proteins. However, when the GCN5-containing ATAC complex was disrupted, it was noted that histone H4 lysine 16 (K16) was hyperacetylated, an effect attributed to the dysregulation of the HDAC SIRT2 [241]; therefore, if hypoacetylation of histone H3 is not observed, it could be due to another underlying mechanism. Additionally, it has been suggested that IFA staining with an antibody to examine the centrosomes would likely reveal evidence of mis-segregation, which could lead to chromosome fragmentation and account for the abnormal nuclear morphologies depicted in the TgGCN5-B dominant-negative parasites. Since the dominant-negative TgGCN5-B expressing parasites resemble the phenotype of parasites exposed to the spindle microtubule disrupting agent oryzalin, the α -tubulin dynamics should be examined in the dominant-negative parasites [242]. PCNA has been shown to be acetylated [233]. It should be determined if *Toxoplasma*'s PCNA is also acetylated, and if there is a change in this modification with the expression of dominant-negative TgGCN5-B protein.

As suggested previously, the TgGCN5-B dominant-negative approach should be repeated in a type II strain of the parasite to determine the role of TgGCN5-B in bradyzoite development and maintenance. For this experiment, the GCN5-B DN protein should be turned on after the parasites have converted to bradyzoites cysts. Then not only can the viability of the bradyzoites be examined, but it can also be determined if bradyzoites expressing GCN5-B DN protein are able to convert back into tachyzoites. If the expression of the GCN5-B DN protein hinders either process, this further validates the utility of TgGCN5-B as a novel therapeutic target.

Although microarray analysis of the TgGCN5-B dominant-negative mutant would provide insight into the genes regulated by this HAT, the utility of a large study must be considered. The cost-benefit ratio of performing this microarray analysis might not be ideal since it is likely that a large number of transcripts are going to be dysregulated,

given the severe phenotype. A more cost-effective approach would be to study the transcript levels of select cell-cycle and cell division genes at given time-points after the addition of Shld through real-time PCR analysis. This is a direct approach and might initially be more beneficial than a global analysis. Additionally, to obtain information of the gene network regulated by TgGCN5-B, chromatin immunoprecipitation followed by either microarray analysis or high-throughput sequencing (ChIP-chip or ChIP-seq) can be performed using parasites expressing $_{HA-MYC}GCN5-B_{FLAG}$.

Additionally, studies are currently being initiated to examine the loss of native endogenous TgGCN5-B. This effect is being examined by tagging the C-terminus of the genomic locus of TgGCN5-B with a 2xHA epitope tag followed by the DD domain [187]. When endogenous TgGCN5-B is tagged in this manner, the parasites must be kept on Shld in order for TgGCN5-B to be expressed. The removal of Shld should cause degradation of TgGCN5-B and should mimic a KO. Repetition of the growth assays and IFAs with cell-cycle markers should produce a similar phenotype as the dominant-negative TgGCN5-B mutant. The confirmation of this phenotype through an independent method will further validate that TgGCN5-B is critical for *Toxoplasma* tachyzoite proliferation.

Finally, given the severe phenotype of the TgGCN5-B dominant-negative mutants, a small molecule inhibitor screen should be initiated. To screen the inhibitors, a high-throughput assay could be designed utilizing the recombinant and catalytically active MBP-GCN5-B protein (Aim 2). Even if an inhibitor is not clinically applicable, it could prove to be a useful laboratory tool.

Chapter 5: Summary

This dissertation has characterized TgGCN5-B, one of two GCN5-family homologues in *Toxoplasma*. It was discovered that TgGCN5-B harbors a unique NLS within its N-terminal extension, associates with several proteins including novel Apicomplexa transcription factors, and is critical for tachyzoite viability because parasites expressing a dominant-negative TgGCN5-B displayed a severe growth phenotype. The decreased viability seen in the dominant-negative TgGCN5-B mutants suggests that TgGCN5-B is a novel therapeutic target. Pharmacological inhibition of this enzyme should hinder *Toxoplasma* tachyzoite proliferation. However, since inhibition of the catalytic activity of TgGCN5-B could also affect GCN5 homologues in the host, it was suggested that peptide aptamer inhibitors also be considered. Specifically designed peptide aptamer inhibitors could disrupt the interaction of TgGCN5-B with its associating proteins (AP2-domain proteins) or block its NLS. However, little is known about the role of TgGCN5-B in *Toxoplasma* bradyzoite differentiation and development. Therefore, studies of the associating proteins of TgGCN5-B within bradyzoites as well as the affect of the dominant-negative on bradyzoites should be commenced. In conclusion, the histone acetyltransferase TgGCN5-B plays an important role in *Toxoplasma* proliferation *in vitro*.

Appendices

Appendix A: Parasite counting

To count parasites, 10 μ l freshly lysed or filter-purified parasites was loaded into each side of a hemocytometer. For each individual side of the hemocytometer, all the parasites within 5 squares of the 25 squares of the main grid were counted (see diagram below). Two replicate counts could be obtained from one hemocytometer. Typically, four replicate counts were required, so the hemocytometer was cleaned, and fresh parasites were loaded to obtain an additional set of replicate counts.

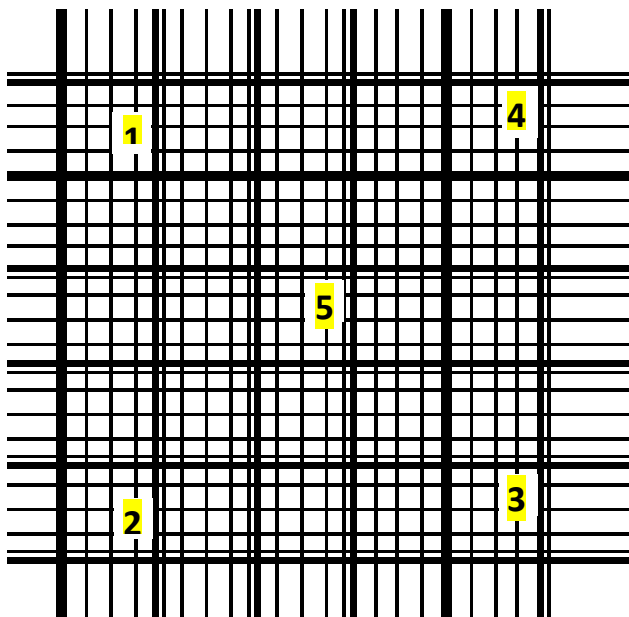


Figure A1: Counting grid of hemocytometer. The highlighted numbers represent the 5 individual squares counted for each replicate.

After obtaining all replicates of parasite counts, calculation could be performed. In general, the total mean was obtained (Ave. of Averages), representing the average number of parasites per square. The total mean was multiplied by 25 since the counting grid has 25 squares. The product is always multiplied by 1×10^4 due to the amount placed in the hemocytometer. This gives the concentration of parasites/ml.

EXAMPLE

	A	B	C	D
Box 1	30	35	37	19
Box 2	24	21	29	25
Box 3	26	26	15	27
Box 4	35	32	24	20
Box 5	20	27	21	31
Totals	164	141	126	122
Average	32.8	28.2	25.2	24.4

1) Concentration of parasites

Ave. of Averages = 27.65 X 25 = 691.25 X10 ⁴ parasite/ml OR 6.9 X10 ⁶ parasite/ml
--

2) For cloning by limiting dilution into 96-well plate

Ave. of Averages = 27.65 X 25 = 691.25 X10 ⁴ parasite/ml OR 6.9 X10 ⁶ parasite/ml
--

6.9 X10 ⁶ parasites/ml / 1000 = 6.9 X10 ³ parasites/ml OR 6.9 parasites/ μ l

100 wells/ 6.9 parasites/ μ l = 14.5 uL X 2 = 29 μ l + 20ml Media
--

After determining the concentration of parasites, a 1:1000 dilution is ALWAYS made (second line of calculation). The concentration of this dilution can also be written to correspond to parasites/ μ l. The third line of the calculation determines the approximately amount of the 1:1000 dilution needed to have roughly 1 parasite per well in the 96-well plates. For this calculation, 100 wells is use for simplification. The answer is double to account for any errors in counting. The final answer is the volume from the 1:1000 that should be added to 20 ml media for dispersion into the plate.

3) For growth assays

Ave. of Averages = 27.65 X 25 = 691.25 X10 ⁴ parasite/ml OR 6.9 X10 ⁶ parasite/ml
--

(1.0X10 ⁵ parasite/ml)(1.0 ml) = (6.9 X10 ⁶ parasite/ml)V; solve for V
--

V = 0.0145 ml or 14.5 μ l + 985.5 μ l media for concentration of 1.0X10 ⁵ parasite/ml
--

At a concentration of 1.0X10⁵ parasite/ml, 10 μ l corresponds to 1000 parasites, and 5 μ l corresponds to 500 parasites.

Appendix B: Details regarding mass spectrometry for affinity chromatography

Mass spectrometry analysis performed in collaboration with Dr. W. Andy Tao (Purdue University)

- Samples from affinity chromatography were denatured, reduced, and alkylated
- Mixture was then injected into an Agilent 1100 HPLC using micro capillaries to a C18 pre-packed column with a 0.1% formic acid solvent
- Elution of peptide was with an acetonitrile gradient running from 0 – 80% for 90 minutes.
- Peptides were ionized using a nano spray tip and delivered into the mass spectrometry machine, an Orbitrap linear iontrap hybrid
- Mass spectrometry data was searched using Sorcerer Software and ToxoDB

Mass spectrometry details from personal communication with Jacob Galan, researcher in the lab of Dr. W. Andy Tao.

Appendix C: Additional affinity chromatography data

Table A1: Proteins associating with MBP-GCN5-B from affinity chromatography #1

Entry #	Protein Description	Accession Number
1	protein arginine N-methyltransferase, putative	33.m01376
2	small heat shock protein, putative \ bradyzoite-specific protein, putative	44.m02755
3	caltractin (centrin), putative	50.m03356
4	eukaryotic translation initiation factor 3 subunit 9, putative	41.m00006
5	60S ribosomal protein L18a, putative	55.m05004
6	TCP-1\cpn60 family chaperonin, putative	49.m00030
7	chorismate synthase, putative	20.m00001
8	mitotic checkpoint protein BUB3, putative	50.m03067
9	ribosomal protein L21, putative	50.m00012
10	hydroxymethylhydropterin pyrophosphokinase-dihydropteroate synthase	55.m00011
11	hypothetical protein	41.m01274
12	28 kDa antigen	42.m00015
13	membrane skeletal protein IMC1, putative	44.m00031
14	prohibitin, putative	49.m00051
15	nucleoside-triphosphatase I	65.m00001
16	ATP synthase, putative	42.m00065
17	elongation factor 1-beta, putative	42.m00069
18	ATP synthase, putative	76.m01572
19	40s ribosomal protein S6, putative	27.m00119
20	60S acidic ribosomal protein P2, putative	583.m00610
21	fibrillarin, putative	583.m00637
22	hypothetical protein	44.m06355
23	protein transport protein Sec23, putative	80.m00011
24	calmodulin, putative	541.m01151
25	histone H2A, putative	55.m04926
26	lysophospholipase, putative	76.m01665
27	hypothetical protein	31.m00869
28	hypothetical protein	583.m05696
29	hypothetical protein	583.m11414
30	conserved hypothetical protein	55.m00224
31	eukaryotic translation initiation factor 3 subunit 3, putative	50.m03396
32	ubiquinol-cytochrome C reductase complex 14 kDa protein-related	80.m00018
33	hypothetical protein	583.m05375
34	hypothetical protein	42.m03493
35	eukaryotic translation initiation factor 3 delta subunit, putative	80.m02245
36	tubulin beta chain	57.m00003
37	malate:quinone oxidoreductase, putative	80.m00006
38	GTP-binding nuclear protein RAN\TC4, putative	50.m00042
39	transketolase, putative	59.m03618
40	sec23\Sec24 domain-containing protein	65.m01096
41	60s ribosomal protein L31, putative	57.m01771
42	adenylate kinase, putative	42.m00116
43	hypothetical protein	55.m10265
44	eukaryotic translation initiation factor 3, putative	83.m01278
45	histone H2A	145.m00002
46	histone H4, putative	49.m03134
47	hypothetical protein	583.m05686
48	hypothetical protein	37.m00747
49	hypothetical protein	83.m00011
50	thioredoxin, putative	50.m00069
51	u1 small nuclear ribonucleoprotein 70 kda-related protein	20.m03892
52	conserved hypothetical protein	113.m00798

53	membrane skeletal protein IMC1	44.m00004
54	40S ribosomal protein S12, putative	20.m03903
55	zinc finger (C2H2 type) protein, putative	76.m00006
56	gbp1p protein (RNA binding protein), putative	55.m00241
57	conserved hypothetical protein	583.m00682
58	ubiquinol-cytochrome c reductase, putative \ Rieske iron-sulfur protein, putative	641.m00178
59	hypothetical protein	55.m00144
60	hypothetical protein	583.m00707
61	40S ribosomal protein S26, putative	49.m03356
62	60S acidic ribosomal protein P0	38.m00002
63	Sec61beta family protein	27.m01477
64	hypothetical protein	42.m03397
65	subtilase family serine protease, putative	20.m00387
66	U1 small nuclear ribonucleoprotein, putative	541.m01233
67	ribosomal protein L5, putative	641.m00186
68	hypothetical protein	57.m03130
69	60s ribosomal protein L36, putative	49.m03096
70	40S ribosomal protein S24, putative	33.m01367
71	articulin 4	41.m00021
72	hypothetical protein	23.m00237
73	serine/threonine protein phosphatase, putative	42.m00006
74	inner membrane complex protein (IMC3)	35.m01595
75	hypothetical protein	44.m02644
76	outer membrane protein roma	641.m01580
77	myosin regulatory light chain, putative	583.m05709
78	hypothetical protein	55.m05032
79	histone H2A, putative	55.m04942
80	40S ribosomal protein S3, putative	44.m04669
81	protease-related	59.m03479
82	14-3-3 protein, putative	55.m00015
83	KH domain containing protein	72.m00843
84	surface protein rhoptry, putative	583.m00003
85	hypothetical protein	541.m01172
86	heat shock protein 60	50.m00006
87	hypothetical protein	42.m07426
88	hypothetical protein	583.m00606
89	chaperone protein dnaK, putative \ heat shock protein 70, putative	50.m00085
90	developmentally regulated GTP-binding protein 1, putative	49.m00057
91	hypothetical protein	50.m05687
92	conserved hypothetical protein	33.m01321
93	ROP8, putative	80.m02343
94	myosin light chain TgMLC1-related	583.m05420
95	hypothetical protein	55.m05095
96	60S ribosomal protein L7a, putative	46.m00028
97	ADP/ATP carrier, putative	50.m00009
98	transcription regulatory protein SNF2, putative	641.m01573
99	vacuolar ATP synthase subunit B, putative	38.m01889

Table A2: Proteins associating with MBP-GCN5-B from affinity chromatography #2

Entry #	Protein Description	Accession Number
1	conserved hypothetical protein	50.m03254
2	60s ribosomal protein L4, putative	583.m00619
3	small heat shock protein, putative \ bradyzoite-specific protein, putative	44.m02755
4	membrane skeletal protein IMC1	44.m00004
5	heat shock protein 70, putative	59.m00003
6	ATP synthase beta chain, putative	55.m00168
7	dense granule protein 3	42.m00013
8	gbp1p protein (RNA binding protein), putative	55.m00241
9	60S ribosomal protein L18a, putative	55.m05004
10	RNA recognition motif. (a.k.a. RRM, RBD, or RNP domain)-containing protein	46.m01699
11	histone H2B, putative	50.m03422
12	conserved hypothetical protein	583.m00682
13	WD-40 repeat protein, putative	641.m01564
14	ribosomal protein L21, putative	50.m00012
15	nucleolar protein family A, putative	59.m00071
16	hypothetical protein	44.m02723
17	ribosomal protein L11, putative	583.m00014
18	hypothetical protein	583.m00707
19	60s ribosomal protein L13a, putative	80.m00062
20	hypothetical protein	41.m01274
21	calmodulin	50.m03285
22	28 kDa antigen	42.m00015
23	ATP synthase alpha chain, putative	20.m00382
24	membrane skeletal protein IMC1, putative	44.m00031
25	prohibitin, putative	49.m00051
26	ATP synthase, putative	42.m00065
27	hypothetical protein	76.m01605
28	rhoptry antigen, putative	551.m00238
29	conserved hypothetical protein	33.m01321
30	elongation factor 1-beta, putative	42.m00069
31	splicing factor 3b subunit 10, putative	583.m05642
32	ATP synthase, putative	76.m01572
33	major surface antigen p30	44.m00009
34	translation initiation factor 2 beta, putative	46.m00016
35	40s ribosomal protein S6, putative	27.m00119
36	histone H2B variant 1	25.m00008
37	ribosomal protein L30\L7e, putative \ wx protein	583.m00012
38	60S acidic ribosomal protein P2, putative	583.m00610
39	fibrillarin, putative	583.m00637
40	60S ribosomal protein L22, putative	49.m05661
41	histone acetyltransferase GCN5, putative	49.m03346
42	hypothetical protein	44.m06355
43	thioredoxin, putative	25.m00203
44	nucleolar RNA-binding domain-containing protein	65.m02513
45	40s ribosomal protein S14, putative	55.m00221
46	calmodulin, putative	541.m01151
47	histone H2A, putative	55.m04926
48	conserved hypothetical protein	113.m00780
49	40S ribosomal protein S17, putative	25.m00216
50	60S ribosomal protein L30, putative	44.m00044

51	40S ribosomal protein S7, putative	49.m00013
52	microneme protein 4 (MIC4)	25.m00006
53	hypothetical protein	583.m05696
54	40S ribosomal protein S13, putative	59.m03516
55	hypothetical protein	583.m11414
56	40s ribosomal protein S20, putative	41.m01387
57	conserved hypothetical protein	55.m00224
58	tubulin alpha chain	583.m00022
59	myosin A, putative	46.m00001
60	L12 ribosomal protein	52.m00028
61	rhoptry antigen, putative	83.m02145
62	eukaryotic translation initiation factor 3 delta subunit, putative	80.m02245
63	conserved hypothetical protein	113.m00009
64	zinc finger (CCCH type) protein, putative	59.m00013
65	hypothetical protein	44.m04681
66	tubulin beta chain	57.m00003
67	U1 small nuclear ribonucleoprotein, putative	541.m01233
68	malate:quinone oxidoreductase, putative	80.m00006
69	major sperm protein domain-containing protein	641.m01460
70	ribosomal protein L5, putative	641.m00186
71	40S ribosomal protein S28, putative	25.m02955
72	small nuclear ribonucleoprotein (snRNP), putative	583.m00614
73	ribosomal protein L34, putative	42.m03575
74	60s ribosomal protein L31, putative	57.m01771
75	adenylate kinase, putative	42.m00116
76	tubulin beta chain, putative	41.m00036
77	hypothetical protein	55.m10265
78	heat shock protein 70, putative	583.m00009
79	eukaryotic translation initiation factor 3, putative	83.m01278
80	histone H2A	145.m00002
81	calmodulin-related	50.m03141
82	40S ribosomal protein S24, putative	33.m01367
83	trypsin, putative	55.m05020
84	histone H4, putative	49.m03134
85	hypothetical protein	83.m01280
86	artoculins 4	41.m00021
87	RNA recognition motif domain-containing protein	645.m00319
88	hypothetical protein	80.m02161
89	dense granule protein 7	20.m00005
90	hypothetical protein	55.m05052
91	hypothetical protein	583.m05686
92	hypothetical protein	39.m00367
93	myosin light chain (TgMLC1)	55.m00021
94	60S ribosomal protein L10, putative	80.m00017
95	hypothetical protein	83.m00011
96	microneme protein 3 (MIC3)	641.m00002
97	phosphatidylinositol-4-phosphate 5-kinase, putative	583.m05359
98	60s ribosomal protein I6, putative	583.m05552
99	DnaJ domain-containing protein	55.m00016
100	hypothetical protein	49.m05663
101	40s ribosomal protein s19, putative	28.m00306
102	hypothetical protein	80.m03946
103	ribosomal protein S10, putative	64.m00338
104	inner membrane complex protein (IMC3)	35.m01595

105	calmodulin, putative	59.m03455
106	surface protein rhoptry, putative	583.m00003
107	conserved hypothetical protein	55.m00267
108	elongation factor 1-alpha, putative,coding) elongation factor 1-alpha, putative	3.m00013
109	hypothetical protein	44.m04616
110	40S ribosomal protein S4, putative	25.m00221
111	u1 small nuclear ribonucleoprotein 70 kda-related protein	20.m03892
112	hypothetical protein	67.m00007
113	pfs77-related	44.m00020
114	acid phosphatase, putative	38.m01061
115	delta-aminolevulinic acid dehydratase, putative	52.m01619
116	high mobility group protein	26.m00246
117	ribosomal protein L32, putative	57.m00005
118	ATP synthase gama chain, putative	44.m02684
119	40S ribosomal protein S5, putative	49.m00006
120	60S ribosomal protein L7a, putative	55.m00280
121	60S ribosomal protein L7a, putative	583.m05562
122	hypothetical protein	20.m03934
123	hypothetical protein	41.m03144
124	prohibitin-related	44.m00051
125	caltractin (centrin), putative	50.m03356
126	DnaJ domain-containing protein	44.m02699
127	porin, putative	55.m08198
128	hypothetical protein	583.m05356
129	ribosomal protein L37a	145.m00001
130	actin	25.m00007
131	small nuclear ribonucleoprotein, putative	57.m01848
132	nucleolar phosphoprotein (nucleolin), putative	80.m02340
133	EGF-like domain-containing protein	83.m01260
134	protein kinase domain-containing protein	55.m08191
135	fas apoptotic inhibitory molecule-related \ lifeguard protein-related	583.m00591
136	hypothetical protein	44.m02570
137	myosin light chain TgMLC1-related	583.m05420
138	40S ribosomal protein S26, putative	49.m03356
139	60s ribosomal protein L33-A, putative	50.m03286
140	40S ribosomal protein S8, putative	50.m00067
141	60S acidic ribosomal protein P0	38.m00002
142	hypothetical protein	44.m02826
143	protease-related	59.m03479
144	40S ribosomal protein S11, putative	42.m00125
145	lysophospholipase, putative	76.m01665
146	rhoptry antigen, putative,coding) rhoptry antigen ROP8	33.m01398
147	microneme protein 6 (MIC6)	38.m00003
148	ribosomal protein S23, putative	44.m02556
149	RNA recognition motif domain-containing protein	59.m03367
150	hypothetical protein	31.m00869
151	hypothetical protein	33.m01359
152	40S ribosomal protein S21, putative	55.m00116
153	RNA-binding protein, putative	20.m05985
154	hypothetical protein	28.m00303
155	DNA-binding protein HU, putative	42.m00103
156	ubiquinol-cytochrome C reductase complex 14 kDa protein-	80.m00018

	related	
157	60S acidic ribosomal protein P1, putative	55.m04872
158	endonuclease\exonuclease\phosphatase domain-containing protein	50.m03277
159	DnaJ domain-containing protein	583.m05418
160	nucleolar phosphoprotein p130, putative	25.m00202
161	hypothetical protein	583.m05525
162	conserved hypothetical protein	20.m08222
163	ras-GTPase-activating protein binding protein, putative	49.m05723
164	enhancer of rudimentary, putative	72.m00406
165	phosphoribosylpyrophosphate synthetase, putative	39.m00330
166	hypothetical protein	57.m03085
167	conserved hypothetical protein	583.m00625
168	ribosomal protein L23a, putative	49.m00008
169	eukaryotic translation initiation factor 3 subunit 7, putative	611.m00038
170	hypothetical protein	55.m05032
171	acetyl-coenzyme A synthetase, putative	57.m03124
172	conserved hypothetical protein	55.m00279
173	polypyrimidine track-binding protein, putative	80.m00057
174	centromere\microtubule binding protein, putative	33.m02213
175	hypothetical protein	46.m01721
176	eukaryotic translation initiation factor 4E, putative	42.m03276
177	RNA-binding protein, putative	27.m00080
178	40S ribosomal protein S15, putative	31.m00922
179	40S ribosomal protein S15a, putative	46.m00005
180	conserved hypothetical protein	38.m01082
181	calcium-dependent protein kinase, putative \ calmodulin-domain protein kinase, putative	20.m00372
182	splicing factor 3A subunit 2, putative	42.m00102
183	proliferation-associated protein 2G4, putative	69.m00140
184	60S ribosomal protein L24, putative	49.m00058
185	histone H3	55.m00013
186	thioredoxin, putative	50.m00069
187	hypothetical protein	42.m03543
188	ubiquitin \ ribosomal protein CEP52 fusion protein, putative	80.m02240
189	peroxiredoxin family protein\glutaredoxin, putative	76.m01670
190	hypothetical protein	44.m02644
191	60S ribosomal protein L18, putative	145.m00327
192	60S ribosomal protein L10a, putative	33.m01368
193	retinitis pigmentosa GTPase regulator-related	49.m03185
194	hypothetical protein	86.m00371
195	mitochondria-associated granulocyte macrophage CSF signaling molecule, putative	50.m03345
196	hypothetical protein	25.m02941
197	60S ribosomal protein L13, putative	55.m00189
198	conserved hypothetical protein	50.m03340
199	KH domain-containing protein	35.m00901
200	60S ribosomal protein L28, putative	44.m04692
201	hypothetical protein	72.m00399
202	OTU-like cysteine protease domain-containing protein	44.m02558
203	60S ribosomal protein L14, putative	57.m01833
204	ribosomal protein L27, putative	55.m05006
205	60S ribosomal protein L9, putative	76.m00009
206	hypothetical protein	583.m05370

207	hypothetical protein	33.m01384
208	eukaryotic translation initiation factor 2 alpha subunit, putative	583.m00670
209	40S ribosomal protein S2, putative	541.m00133

Appendix D: Monitoring parasite cell cycle stages via IFA

For IFA, use antibodies to TgIMC1, TgPCNA, and stain with DAPI

Description of parasite cell stages

Early G1: small parasites; small, single nuclei with dull DAPI staining

Mid to late G1: longer parasites; single, central nuclei with dull DAPI staining

S phase: large parasites; expanded nuclei with bright DAPI staining

Late S phase: large nuclei shifted to posterior; first indication of internal daughter cells development

Mitotic: posterior nuclei; 10 – 20 % daughter cell development

Late mitotic (telophase): U-shaped nuclei; 20 – 30% daughter cell development

Cytokinetic: divided nuclei; single parasite with two nuclei; 30 – 70% daughter cell development

Information from personal communication with Dr. Michael White (University South Florida)

References

1. Goldberg AD, Allis CD, Bernstein E (2007) Epigenetics: a landscape takes shape. *Cell* 128: 635-638.
2. Foth BJ, McFadden GI (2003) The apicoplast: a plastid in *Plasmodium falciparum* and other Apicomplexan parasites. *Int Rev Cytol* 224: 57-110.
3. Parsons M, Karnataki A, Derocher AE (2009) Evolving insights into protein trafficking to the multiple compartments of the apicomplexan plastid. *J Eukaryot Microbiol* 56: 214-220.
4. Black MW, Boothroyd JC (2000) Lytic cycle of *Toxoplasma gondii*. *Microbiol Mol Biol Rev* 64: 607-623.
5. Eisenberg JN, Seto EY, Colford JM, Jr., Olivieri A, Spear RC (1998) An analysis of the Milwaukee cryptosporidiosis outbreak based on a dynamic model of the infection process. *Epidemiology* 9: 255-263.
6. Hurtley S, Ash C, Roberts L (2010) Tuberculosis & malaria. Landscapes of infection. Introduction. *Science* 328: 841.
7. Hill DE, Chirukandoth S, Dubey JP (2005) Biology and epidemiology of *Toxoplasma gondii* in man and animals. *Anim Health Res Rev* 6: 41-61.
8. Dubey JP (2009) Toxoplasmosis in sheep--the last 20 years. *Vet Parasitol* 163: 1-14.
9. Roos DS, Crawford MJ, Donald RG, Fohl LM, Hager KM, et al. (1999) Transport and trafficking: *Toxoplasma* as a model for *Plasmodium*. *Novartis Found Symp* 226: 176-195; discussion 195-178.
10. Kim K, Weiss LM (2004) *Toxoplasma gondii*: the model apicomplexan. *Int J Parasitol* 34: 423-432.
11. Gajria B, Bahl A, Brestelli J, Dommer J, Fischer S, et al. (2008) ToxoDB: an integrated *Toxoplasma gondii* database resource. *Nucleic Acids Res* 36: D553-556.
12. Ferguson DJ (2009) Identification of faecal transmission of *Toxoplasma gondii*: Small science, large characters. *Int J Parasitol* 39: 871-875.
13. Dubey JP, Lindsay DS, Speer CA (1998) Structures of *Toxoplasma gondii* tachyzoites, bradyzoites, and sporozoites and biology and development of tissue cysts. *Clin Microbiol Rev* 11: 267-299.
14. Davis SW, Dubey JP (1995) Mediation of immunity to *Toxoplasma gondii* oocyst shedding in cats. *J Parasitol* 81: 882-886.

15. Dubey JP (2001) Oocyst shedding by cats fed isolated bradyzoites and comparison of infectivity of bradyzoites of the VEG strain *Toxoplasma gondii* to cats and mice. *J Parasitol* 87: 215-219.
16. Dubey JP, Jones JL (2008) *Toxoplasma gondii* infection in humans and animals in the United States. *Int J Parasitol* 38: 1257-1278.
17. Dubey JP (1996) Infectivity and pathogenicity of *Toxoplasma gondii* oocysts for cats. *J Parasitol* 82: 957-961.
18. Dubey JP (2009) History of the discovery of the life cycle of *Toxoplasma gondii*. *Int J Parasitol*.
19. Howe DK, Sibley LD (1995) *Toxoplasma gondii* comprises three clonal lineages: correlation of parasite genotype with human disease. *J Infect Dis* 172: 1561-1566.
20. Saeij JP, Boyle JP, Boothroyd JC (2005) Differences among the three major strains of *Toxoplasma gondii* and their specific interactions with the infected host. *Trends Parasitol* 21: 476-481.
21. Su C, Evans D, Cole RH, Kissinger JC, Ajioka JW, et al. (2003) Recent expansion of *Toxoplasma* through enhanced oral transmission. *Science* 299: 414-416.
22. Soete M, Fortier B, Camus D, Dubremetz JF (1993) *Toxoplasma gondii*: kinetics of bradyzoite-tachyzoite interconversion in vitro. *Exp Parasitol* 76: 259-264.
23. Grigg ME, Ganatra J, Boothroyd JC, Margolis TP (2001) Unusual abundance of atypical strains associated with human ocular toxoplasmosis. *J Infect Dis* 184: 633-639.
24. Luft BJ, Hafner R, Korzun AH, Leport C, Antoniskis D, et al. (1993) Toxoplasmic encephalitis in patients with the acquired immunodeficiency syndrome. Members of the ACTG 077p/ANRS 009 Study Team. *N Engl J Med* 329: 995-1000.
25. Renold C, Sugar A, Chave JP, Perrin L, Delavelle J, et al. (1992) *Toxoplasma* encephalitis in patients with the acquired immunodeficiency syndrome. *Medicine (Baltimore)* 71: 224-239.
26. Jones J, Lopez A, Wilson M (2003) Congenital toxoplasmosis. *Am Fam Physician* 67: 2131-2138.
27. Remington JS, McLeod R, Thullieux P, Desmonts G (2001) Toxoplasmosis. In: Remington JS, Klein J. O., editor. *Infectious diseases of the fetus and newborn infant*. 5th ed. Philadelphia: Saunders. pp. 205-346.
28. Liesenfeld O, Montoya JG, Kinney S, Press C, Remington JS (2001) Effect of testing for IgG avidity in the diagnosis of *Toxoplasma gondii* infection in pregnant women: experience in a US reference laboratory. *J Infect Dis* 183: 1248-1253.

29. Guerina NG (1994) Congenital infection with *Toxoplasma gondii*. *Pediatr Ann* 23: 138-142, 147-151.
30. McLeod R, Boyer K, Karrison T, Kasza K, Swisher C, et al. (2006) Outcome of treatment for congenital toxoplasmosis, 1981-2004: the National Collaborative Chicago-Based, Congenital Toxoplasmosis Study. *Clin Infect Dis* 42: 1383-1394.
31. Eckert GU, Melamed J, Menegaz B (2007) Optic nerve changes in ocular toxoplasmosis. *Eye (Lond)* 21: 746-751.
32. Commodaro AG, Belfort RN, Rizzo LV, Muccioli C, Silveira C, et al. (2009) Ocular toxoplasmosis: an update and review of the literature. *Mem Inst Oswaldo Cruz* 104: 345-350.
33. Dodds EM (2006) Toxoplasmosis. *Curr Opin Ophthalmol* 17: 557-561.
34. Holland GN (2004) Ocular toxoplasmosis: a global reassessment. Part II: disease manifestations and management. *Am J Ophthalmol* 137: 1-17.
35. Mets MB, Holfels E, Boyer KM, Swisher CN, Roizen N, et al. (1997) Eye manifestations of congenital toxoplasmosis. *Am J Ophthalmol* 123: 1-16.
36. Lum F, Jones JL, Holland GN, Liesegang TJ (2005) Survey of ophthalmologists about ocular toxoplasmosis. *Am J Ophthalmol* 140: 724-726.
37. Yolken RH, Dickerson FB, Fuller Torrey E (2009) *Toxoplasma* and schizophrenia. *Parasite Immunol* 31: 706-715.
38. Berdoy M, Webster JP, Macdonald DW (2000) Fatal attraction in rats infected with *Toxoplasma gondii*. *Proc Biol Sci* 267: 1591-1594.
39. Webster JP (2001) Rats, cats, people and parasites: the impact of latent toxoplasmosis on behaviour. *Microbes Infect* 3: 1037-1045.
40. Vyas A, Kim SK, Giacomini N, Boothroyd JC, Sapolsky RM (2007) Behavioral changes induced by *Toxoplasma* infection of rodents are highly specific to aversion of cat odors. *Proc Natl Acad Sci U S A* 104: 6442-6447.
41. Vyas A, Kim SK, Sapolsky RM (2007) The effects of *toxoplasma* infection on rodent behavior are dependent on dose of the stimulus. *Neuroscience* 148: 342-348.
42. Kozar Z (1953) [Studies on toxoplasmosis in mental diseases.]. *Biul Panstw Inst Med Morsk Trop J W Gdansk* 5: 134-137; English transl.
43. Torrey EF, Bartko JJ, Lun ZR, Yolken RH (2007) Antibodies to *Toxoplasma gondii* in patients with schizophrenia: a meta-analysis. *Schizophr Bull* 33: 729-736.
44. Hinze-Selch D, Daubener W, Eggert L, Erdag S, Stoltenberg R, et al. (2007) A controlled prospective study of *toxoplasma gondii* infection in individuals with schizophrenia: beyond seroprevalence. *Schizophr Bull* 33: 782-788.

45. Gaskell EA, Smith JE, Pinney JW, Westhead DR, McConkey GA (2009) A unique dual activity amino acid hydroxylase in *Toxoplasma gondii*. PLoS One 4: e4801.
46. Torrey EF, Yolken RH (2007) Schizophrenia and toxoplasmosis. Schizophr Bull 33: 727-728.
47. Yereli K, Balcioglu IC, Ozbilgin A (2006) Is *Toxoplasma gondii* a potential risk for traffic accidents in Turkey? Forensic Sci Int 163: 34-37.
48. Flegr J, Havlicek J, Kodym P, Maly M, Smahel Z (2002) Increased risk of traffic accidents in subjects with latent toxoplasmosis: a retrospective case-control study. BMC Infect Dis 2: 11.
49. Hill D, Dubey JP (2002) *Toxoplasma gondii*: transmission, diagnosis and prevention. Clin Microbiol Infect 8: 634-640.
50. Chio LC, Queener SF (1993) Identification of highly potent and selective inhibitors of *Toxoplasma gondii* dihydrofolate reductase. Antimicrob Agents Chemother 37: 1914-1923.
51. Donald RGK (2007) *Toxoplasma* as a Model System for Apicomplexan Drug Discovery. In: Kim LMWaK, editor. *Toxoplasma gondii*, The Model Apicomplexan: Perspective and Methods. 1 ed. London: Academic Press; Elsevier Ltd. pp. 507.
52. Petersen EaOL (2007) Clinical Disease and Diagnostics. In: Kim LMWaK, editor. *Toxoplasma gondii*, The model apicomplexan: Perspectives and Methods. 1 ed. London: Academic Press, Elsevier Ltd. pp. 93.
53. Rothova A, Bosch-Driessen LE, van Loon NH, Treffers WF (1998) Azithromycin for ocular toxoplasmosis. Br J Ophthalmol 82: 1306-1308.
54. Sukthana Y (2006) Toxoplasmosis: beyond animals to humans. Trends Parasitol 22: 137-142.
55. Fichera ME, Bhopale MK, Roos DS (1995) In vitro assays elucidate peculiar kinetics of clindamycin action against *Toxoplasma gondii*. Antimicrob Agents Chemother 39: 1530-1537.
56. Fichera ME, Roos DS (1997) A plastid organelle as a drug target in apicomplexan parasites. Nature 390: 407-409.
57. McAuley J, Boyer KM, Patel D, Mets M, Swisher C, et al. (1994) Early and longitudinal evaluations of treated infants and children and untreated historical patients with congenital toxoplasmosis: the Chicago Collaborative Treatment Trial. Clin Infect Dis 18: 38-72.
58. Goll MG, Bestor TH (2005) Eukaryotic cytosine methyltransferases. Annu Rev Biochem 74: 481-514.

59. Yang PK, Kuroda MI (2007) Noncoding RNAs and intranuclear positioning in monoallelic gene expression. *Cell* 128: 777-786.
60. Bernstein E, Allis CD (2005) RNA meets chromatin. *Genes Dev* 19: 1635-1655.
61. Stedman E (1950) Cell specificity of histones. *Nature* 166: 780-781.
62. Allfrey VG, Faulkner R, Mirsky AE (1964) Acetylation and Methylation of Histones and Their Possible Role in the Regulation of Rna Synthesis. *Proc Natl Acad Sci U S A* 51: 786-794.
63. Brownell JE, Zhou J, Ranalli T, Kobayashi R, Edmondson DG, et al. (1996) Tetrahymena histone acetyltransferase A: a homolog to yeast Gcn5p linking histone acetylation to gene activation. *Cell* 84: 843-851.
64. Sterner DE, Berger SL (2000) Acetylation of histones and transcription-related factors. *Microbiol Mol Biol Rev* 64: 435-459.
65. Chuikov S, Kurash JK, Wilson JR, Xiao B, Justin N, et al. (2004) Regulation of p53 activity through lysine methylation. *Nature* 432: 353-360.
66. de la Cruz X, Lois S, Sanchez-Molina S, Martinez-Balbas MA (2005) Do protein motifs read the histone code? *Bioessays* 27: 164-175.
67. Zeng L, Zhou MM (2002) Bromodomain: an acetyl-lysine binding domain. *FEBS Lett* 513: 124-128.
68. Wang Y, Fischle W, Cheung W, Jacobs S, Khorasanizadeh S, et al. (2004) Beyond the double helix: writing and reading the histone code. *Novartis Found Symp* 259: 3-17; discussion 17-21, 163-169.
69. Marmorstein R (2001) Structure of histone acetyltransferases. *J Mol Biol* 311: 433-444.
70. Nielsen PR, Callaghan J, Murzin AG, Murzina NV, Laue ED (2004) Expression, purification, and biophysical studies of chromodomain proteins. *Methods Enzymol* 376: 148-170.
71. Loidl P (1994) Histone acetylation: facts and questions. *Chromosoma* 103: 441-449.
72. Hong L, Schroth GP, Matthews HR, Yau P, Bradbury EM (1993) Studies of the DNA binding properties of histone H4 amino terminus. Thermal denaturation studies reveal that acetylation markedly reduces the binding constant of the H4 "tail" to DNA. *J Biol Chem* 268: 305-314.
73. Schwarz PM, Felthausen A, Fletcher TM, Hansen JC (1996) Reversible oligonucleosome self-association: dependence on divalent cations and core histone tail domains. *Biochemistry* 35: 4009-4015.
74. Norton VG, Imai BS, Yau P, Bradbury EM (1989) Histone acetylation reduces nucleosome core particle linking number change. *Cell* 57: 449-457.

75. Selvi BR, Batta K, Kishore AH, Mantelingu K, Varier RA, et al. (2010) Identification of a novel inhibitor of coactivator-associated arginine methyltransferase 1 (CARM1)-mediated methylation of histone H3 Arg-17. *J Biol Chem* 285: 7143-7152.
76. Balasubramanyam K, Varier RA, Altaf M, Swaminathan V, Siddappa NB, et al. (2004) Curcumin, a novel p300/CREB-binding protein-specific inhibitor of acetyltransferase, represses the acetylation of histone/nonhistone proteins and histone acetyltransferase-dependent chromatin transcription. *J Biol Chem* 279: 51163-51171.
77. Balasubramanyam K, Altaf M, Varier RA, Swaminathan V, Ravindran A, et al. (2004) Polyisoprenylated benzophenone, garcinol, a natural histone acetyltransferase inhibitor, represses chromatin transcription and alters global gene expression. *J Biol Chem* 279: 33716-33726.
78. Balasubramanyam K, Swaminathan V, Ranganathan A, Kundu TK (2003) Small molecule modulators of histone acetyltransferase p300. *J Biol Chem* 278: 19134-19140.
79. Poux AN, Cebrat M, Kim CM, Cole PA, Marmorstein R (2002) Structure of the GCN5 histone acetyltransferase bound to a bisubstrate inhibitor. *Proc Natl Acad Sci U S A* 99: 14065-14070.
80. Dekker FJ, Haisma HJ (2009) Histone acetyl transferases as emerging drug targets. *Drug Discov Today* 14: 942-948.
81. Dokmanovic M, Clarke C, Marks PA (2007) Histone deacetylase inhibitors: overview and perspectives. *Mol Cancer Res* 5: 981-989.
82. Yang XJ, Ogryzko VV, Nishikawa J, Howard BH, Nakatani Y (1996) A p300/CBP-associated factor that competes with the adenoviral oncoprotein E1A. *Nature* 382: 319-324.
83. Linares LK, Kiernan R, Triboulet R, Chable-Bessia C, Latreille D, et al. (2007) Intrinsic ubiquitination activity of PCAF controls the stability of the oncoprotein Hdm2. *Nat Cell Biol* 9: 331-338.
84. Kuo MH, Brownell JE, Sobel RE, Ranalli TA, Cook RG, et al. (1996) Transcription-linked acetylation by Gcn5p of histones H3 and H4 at specific lysines. *Nature* 383: 269-272.
85. Grant PA, Eberharter A, John S, Cook RG, Turner BM, et al. (1999) Expanded lysine acetylation specificity of Gcn5 in native complexes. *J Biol Chem* 274: 5895-5900.
86. Morris SA, Rao B, Garcia BA, Hake SB, Diaz RL, et al. (2007) Identification of histone H3 lysine 36 acetylation as a highly conserved histone modification. *J Biol Chem* 282: 7632-7640.

87. Tjeertes JV, Miller KM, Jackson SP (2009) Screen for DNA-damage-responsive histone modifications identifies H3K9Ac and H3K56Ac in human cells. *EMBO J* 28: 1878-1889.
88. Xie W, Song C, Young NL, Sperling AS, Xu F, et al. (2009) Histone h3 lysine 56 acetylation is linked to the core transcriptional network in human embryonic stem cells. *Mol Cell* 33: 417-427.
89. Tanner KG, Trievel RC, Kuo MH, Howard RM, Berger SL, et al. (1999) Catalytic mechanism and function of invariant glutamic acid 173 from the histone acetyltransferase GCN5 transcriptional coactivator. *J Biol Chem* 274: 18157-18160.
90. Berger SL, Pina B, Silverman N, Marcus GA, Agapite J, et al. (1992) Genetic isolation of ADA2: a potential transcriptional adaptor required for function of certain acidic activation domains. *Cell* 70: 251-265.
91. Marcus GA, Silverman N, Berger SL, Horiuchi J, Guarente L (1994) Functional similarity and physical association between GCN5 and ADA2: putative transcriptional adaptors. *EMBO J* 13: 4807-4815.
92. Balasubramanian R, Pray-Grant MG, Selleck W, Grant PA, Tan S (2002) Role of the Ada2 and Ada3 transcriptional coactivators in histone acetylation. *J Biol Chem* 277: 7989-7995.
93. Mohibullah N, Hahn S (2008) Site-specific cross-linking of TBP in vivo and in vitro reveals a direct functional interaction with the SAGA subunit Spt3. *Genes Dev* 22: 2994-3006.
94. Grant PA, Schieltz D, Pray-Grant MG, Yates JR, 3rd, Workman JL (1998) The ATM-related cofactor Tra1 is a component of the purified SAGA complex. *Mol Cell* 2: 863-867.
95. Grant PA, Sterner DE, Duggan LJ, Workman JL, Berger SL (1998) The SAGA unfolds: convergence of transcription regulators in chromatin-modifying complexes. *Trends Cell Biol* 8: 193-197.
96. Daniel JA, Torok MS, Sun ZW, Schieltz D, Allis CD, et al. (2004) Deubiquitination of histone H2B by a yeast acetyltransferase complex regulates transcription. *J Biol Chem* 279: 1867-1871.
97. Henry KW, Wyce A, Lo WS, Duggan LJ, Emre NC, et al. (2003) Transcriptional activation via sequential histone H2B ubiquitylation and deubiquitylation, mediated by SAGA-associated Ubp8. *Genes Dev* 17: 2648-2663.
98. Lee KK, Florens L, Swanson SK, Washburn MP, Workman JL (2005) The deubiquitylation activity of Ubp8 is dependent upon Sgf11 and its association with the SAGA complex. *Mol Cell Biol* 25: 1173-1182.

99. Huisinga KL, Pugh BF (2004) A genome-wide housekeeping role for TFIID and a highly regulated stress-related role for SAGA in *Saccharomyces cerevisiae*. *Mol Cell* 13: 573-585.
100. Govind CK, Zhang F, Qiu H, Hofmeyer K, Hinnebusch AG (2007) Gcn5 promotes acetylation, eviction, and methylation of nucleosomes in transcribed coding regions. *Mol Cell* 25: 31-42.
101. Kohler A, Pascual-Garcia P, Llopis A, Zapater M, Posas F, et al. (2006) The mRNA export factor Sus1 is involved in Spt/Ada/Gcn5 acetyltransferase-mediated H2B deubiquitinylation through its interaction with Ubp8 and Sgf11. *Mol Biol Cell* 17: 4228-4236.
102. Rodriguez-Navarro S, Fischer T, Luo MJ, Antunez O, Brettschneider S, et al. (2004) Sus1, a functional component of the SAGA histone acetylase complex and the nuclear pore-associated mRNA export machinery. *Cell* 116: 75-86.
103. Rodriguez-Navarro S (2009) Insights into SAGA function during gene expression. *EMBO Rep* 10: 843-850.
104. Nagy Z, Tora L (2007) Distinct GCN5/PCAF-containing complexes function as co-activators and are involved in transcription factor and global histone acetylation. *Oncogene* 26: 5341-5357.
105. Grant PA, Duggan L, Cote J, Roberts SM, Brownell JE, et al. (1997) Yeast Gcn5 functions in two multisubunit complexes to acetylate nucleosomal histones: characterization of an Ada complex and the SAGA (Spt/Ada) complex. *Genes Dev* 11: 1640-1650.
106. Eberharter A, Sterner DE, Schieltz D, Hassan A, Yates JR, 3rd, et al. (1999) The ADA complex is a distinct histone acetyltransferase complex in *Saccharomyces cerevisiae*. *Mol Cell Biol* 19: 6621-6631.
107. Wang YL, Faiola F, Xu M, Pan S, Martinez E (2008) Human ATAC Is a GCN5/PCAF-containing acetylase complex with a novel NC2-like histone fold module that interacts with the TATA-binding protein. *J Biol Chem* 283: 33808-33815.
108. Suganuma T, Gutierrez JL, Li B, Florens L, Swanson SK, et al. (2008) ATAC is a double histone acetyltransferase complex that stimulates nucleosome sliding. *Nat Struct Mol Biol* 15: 364-372.
109. Guelman S, Suganuma T, Florens L, Swanson SK, Kiesecker CL, et al. (2006) Host cell factor and an uncharacterized SANT domain protein are stable components of ATAC, a novel dAda2A/dGcn5-containing histone acetyltransferase complex in *Drosophila*. *Mol Cell Biol* 26: 871-882.
110. Lee TI, Causton HC, Holstege FC, Shen WC, Hannett N, et al. (2000) Redundant roles for the TFIID and SAGA complexes in global transcription. *Nature* 405: 701-704.

111. Helmlinger D, Marguerat S, Villen J, Gygi SP, Bahler J, et al. (2008) The *S. pombe* SAGA complex controls the switch from proliferation to sexual differentiation through the opposing roles of its subunits Gcn5 and Spt8. *Genes Dev* 22: 3184-3195.
112. Waterborg JH (2000) Steady-state levels of histone acetylation in *Saccharomyces cerevisiae*. *J Biol Chem* 275: 13007-13011.
113. Gunderson FQ, Johnson TL (2009) Acetylation by the transcriptional coactivator Gcn5 plays a novel role in co-transcriptional spliceosome assembly. *PLoS Genet* 5: e1000682.
114. Yu Y, Teng Y, Liu H, Reed SH, Waters R (2005) UV irradiation stimulates histone acetylation and chromatin remodeling at a repressed yeast locus. *Proc Natl Acad Sci U S A* 102: 8650-8655.
115. Ferreira JA, Powell NG, Karabetsou N, Mellor J, Waters R (2006) Roles for Gcn5p and Ada2p in transcription and nucleotide excision repair at the *Saccharomyces cerevisiae* MET16 gene. *Nucleic Acids Res* 34: 976-985.
116. Carre C, Szymczak D, Pidoux J, Antoniewski C (2005) The histone H3 acetylase dGcn5 is a key player in *Drosophila melanogaster* metamorphosis. *Mol Cell Biol* 25: 8228-8238.
117. Xu W, Edmondson DG, Evrard YA, Wakamiya M, Behringer RR, et al. (2000) Loss of Gcn5l2 leads to increased apoptosis and mesodermal defects during mouse development. *Nat Genet* 26: 229-232.
118. Bu P, Evrard YA, Lozano G, Dent SY (2007) Loss of Gcn5 acetyltransferase activity leads to neural tube closure defects and exencephaly in mouse embryos. *Mol Cell Biol* 27: 3405-3416.
119. Yamauchi T, Yamauchi J, Kuwata T, Tamura T, Yamashita T, et al. (2000) Distinct but overlapping roles of histone acetylase PCAF and of the closely related PCAF-B/GCN5 in mouse embryogenesis. *Proc Natl Acad Sci U S A* 97: 11303-11306.
120. Kikuchi H, Takami Y, Nakayama T (2005) GCN5: a supervisor in all-inclusive control of vertebrate cell cycle progression through transcription regulation of various cell cycle-related genes. *Gene* 347: 83-97.
121. Atanassov BS, Evrard YA, Multani AS, Zhang Z, Tora L, et al. (2009) Gcn5 and SAGA regulate shelterin protein turnover and telomere maintenance. *Mol Cell* 35: 352-364.
122. Patel JH, Du Y, Ard PG, Phillips C, Carella B, et al. (2004) The c-MYC oncoprotein is a substrate of the acetyltransferases hGCN5/PCAF and TIP60. *Mol Cell Biol* 24: 10826-10834.
123. Glozak MA, Sengupta N, Zhang X, Seto E (2005) Acetylation and deacetylation of non-histone proteins. *Gene* 363: 15-23.

124. Paolinelli R, Mendoza-Maldonado R, Cereseto A, Giacca M (2009) Acetylation by GCN5 regulates CDC6 phosphorylation in the S phase of the cell cycle. *Nat Struct Mol Biol* 16: 412-420.
125. Mateo F, Vidal-Laliena M, Pujol MJ, Bachs O (2010) Acetylation of cyclin A: a new cell cycle regulatory mechanism. *Biochem Soc Trans* 38: 83-86.
126. Gorlich D, Prehn S, Laskey RA, Hartmann E (1994) Isolation of a protein that is essential for the first step of nuclear protein import. *Cell* 79: 767-778.
127. Radu A, Blobel G, Moore MS (1995) Identification of a protein complex that is required for nuclear protein import and mediates docking of import substrate to distinct nucleoporins. *Proc Natl Acad Sci U S A* 92: 1769-1773.
128. Quimby BB, Dasso M (2003) The small GTPase Ran: interpreting the signs. *Curr Opin Cell Biol* 15: 338-344.
129. Yoneda Y, Hieda M, Nagoshi E, Miyamoto Y (1999) Nucleocytoplasmic protein transport and recycling of Ran. *Cell Struct Funct* 24: 425-433.
130. Jans DA, Xiao CY, Lam MH (2000) Nuclear targeting signal recognition: a key control point in nuclear transport? *Bioessays* 22: 532-544.
131. Frankel MB, Knoll LJ (2009) The ins and outs of nuclear trafficking: unusual aspects in apicomplexan parasites. *DNA Cell Biol* 28: 277-284.
132. Lange A, Mills RE, Lange CJ, Stewart M, Devine SE, et al. (2007) Classical nuclear localization signals: definition, function, and interaction with importin alpha. *J Biol Chem* 282: 5101-5105.
133. Lam MH, Briggs LJ, Hu W, Martin TJ, Gillespie MT, et al. (1999) Importin beta recognizes parathyroid hormone-related protein with high affinity and mediates its nuclear import in the absence of importin alpha. *J Biol Chem* 274: 7391-7398.
134. Truant R, Cullen BR (1999) The arginine-rich domains present in human immunodeficiency virus type 1 Tat and Rev function as direct importin beta-dependent nuclear localization signals. *Mol Cell Biol* 19: 1210-1217.
135. Palmeri D, Malim MH (1999) Importin beta can mediate the nuclear import of an arginine-rich nuclear localization signal in the absence of importin alpha. *Mol Cell Biol* 19: 1218-1225.
136. Harel A, Forbes DJ (2004) Importin beta: conducting a much larger cellular symphony. *Mol Cell* 16: 319-330.
137. Kalderon D, Richardson WD, Markham AF, Smith AE (1984) Sequence requirements for nuclear location of simian virus 40 large-T antigen. *Nature* 311: 33-38.

138. Dingwall C, Robbins J, Dilworth SM, Roberts B, Richardson WD (1988) The nucleoplasmin nuclear location sequence is larger and more complex than that of SV-40 large T antigen. *J Cell Biol* 107: 841-849.
139. Nair R, Carter P, Rost B (2003) NLSdb: database of nuclear localization signals. *Nucleic Acids Res* 31: 397-399.
140. Kosugi S, Hasebe M, Matsumura N, Takashima H, Miyamoto-Sato E, et al. (2009) Six classes of nuclear localization signals specific to different binding grooves of importin alpha. *J Biol Chem* 284: 478-485.
141. Santos-Rosa H, Valls E, Kouzarides T, Martinez-Balbas M (2003) Mechanisms of P/CAF auto-acetylation. *Nucleic Acids Res* 31: 4285-4292.
142. Blanco-Garcia N, Asensio-Juan E, de la Cruz X, Martinez-Balbas MA (2009) Autoacetylation regulates P/CAF nuclear localization. *J Biol Chem* 284: 1343-1352.
143. Thompson PR, Wang D, Wang L, Fulco M, Pediconi N, et al. (2004) Regulation of the p300 HAT domain via a novel activation loop. *Nat Struct Mol Biol* 11: 308-315.
144. Bhatti MM, Sullivan WJ, Jr. (2005) Histone acetylase GCN5 enters the nucleus via importin-alpha in protozoan parasite *Toxoplasma gondii*. *J Biol Chem* 280: 5902-5908.
145. Khan A, Taylor S, Su C, Mackey AJ, Boyle J, et al. (2005) Composite genome map and recombination parameters derived from three archetypal lineages of *Toxoplasma gondii*. *Nucleic Acids Res* 33: 2980-2992.
146. Cleary MD, Singh U, Blader IJ, Brewer JL, Boothroyd JC (2002) *Toxoplasma gondii* asexual development: identification of developmentally regulated genes and distinct patterns of gene expression. *Eukaryot Cell* 1: 329-340.
147. Radke JR, Behnke MS, Mackey AJ, Radke JB, Roos DS, et al. (2005) The transcriptome of *Toxoplasma gondii*. *BMC Biol* 3: 26.
148. Singh U, Brewer JL, Boothroyd JC (2002) Genetic analysis of tachyzoite to bradyzoite differentiation mutants in *Toxoplasma gondii* reveals a hierarchy of gene induction. *Mol Microbiol* 44: 721-733.
149. Meissner M, Soldati D (2005) The transcription machinery and the molecular toolbox to control gene expression in *Toxoplasma gondii* and other protozoan parasites. *Microbes Infect* 7: 1376-1384.
150. Iyer LM, Anantharaman V, Wolf MY, Aravind L (2008) Comparative genomics of transcription factors and chromatin proteins in parasitic protists and other eukaryotes. *Int J Parasitol* 38: 1-31.
151. Sullivan WJ, Jr., Hakimi MA (2006) Histone mediated gene activation in *Toxoplasma gondii*. *Mol Biochem Parasitol* 148: 109-116.

152. Bougdour A, Braun L, Cannella D, Hakimi MA (2010) Chromatin modifications: implications in the regulation of gene expression in *Toxoplasma gondii*. *Cell Microbiol*.
153. Dixon SE, Stilger KL, Elias EV, Naguleswaran A, Sullivan WJ, Jr. (2010) A decade of epigenetic research in *Toxoplasma gondii*. *Mol Biochem Parasitol*.
154. Saksouk N, Bhatti MM, Kieffer S, Smith AT, Musset K, et al. (2005) Histone-modifying complexes regulate gene expression pertinent to the differentiation of the protozoan parasite *Toxoplasma gondii*. *Mol Cell Biol* 25: 10301-10314.
155. Gissot M, Kelly KA, Ajioka JW, Greally JM, Kim K (2007) Epigenomic Modifications Predict Active Promoters and Gene Structure in *Toxoplasma gondii*. *PLoS Pathog* 3: e77.
156. Behnke MS, Radke JB, Smith AT, Sullivan WJ, Jr., White MW (2008) The transcription of bradyzoite genes in *Toxoplasma gondii* is controlled by autonomous promoter elements. *Mol Microbiol* 68: 1502-1518.
157. Balaji S, Babu MM, Iyer LM, Aravind L (2005) Discovery of the principal specific transcription factors of Apicomplexa and their implication for the evolution of the AP2-integrase DNA binding domains. *Nucleic Acids Res* 33: 3994-4006.
158. Jofuku KD, den Boer BG, Van Montagu M, Okamuro JK (1994) Control of *Arabidopsis* flower and seed development by the homeotic gene *APETALA2*. *Plant Cell* 6: 1211-1225.
159. De Silva EK, Gehrke AR, Olszewski K, Leon I, Chahal JS, et al. (2008) Specific DNA-binding by apicomplexan AP2 transcription factors. *Proc Natl Acad Sci U S A* 105: 8393-8398.
160. Yuda M, Iwanaga S, Shigenobu S, Kato T, Kaneko I (2010) Transcription factor AP2-Sp and its target genes in malarial sporozoites. *Mol Microbiol* 75: 854-863.
161. LaCount DJ, Vignali M, Chettier R, Phansalkar A, Bell R, et al. (2005) A protein interaction network of the malaria parasite *Plasmodium falciparum*. *Nature* 438: 103-107.
162. Dalmaso MC, Echeverria PC, Zappia MP, Hellman U, Dubremetz JF, et al. (2006) *Toxoplasma gondii* has two lineages of histones 2b (H2B) with different expression profiles. *Mol Biochem Parasitol* 148: 103-107.
163. Dalmaso MC, Onyango DO, Naguleswaran A, Sullivan WJ, Jr., Angel SO (2009) *Toxoplasma* H2A variants reveal novel insights into nucleosome composition and functions for this histone family. *J Mol Biol* 392: 33-47.
164. Sullivan WJ, Jr. (2003) Histone H3 and H3.3 variants in the protozoan pathogens *Plasmodium falciparum* and *Toxoplasma gondii*. *DNA Seq* 14: 227-231.
165. Sullivan WJ, Jr., Naguleswaran A, Angel SO (2006) Histones and histone modifications in protozoan parasites. *Cell Microbiol* 8: 1850-1861.

166. Sautel CF, Cannella D, Bastien O, Kieffer S, Aldebert D, et al. (2007) SET8-mediated methylations of histone H4 lysine 20 mark silent heterochromatic domains in apicomplexan genomes. *Mol Cell Biol* 27: 5711-5724.
167. Hettmann C, Soldati D (1999) Cloning and analysis of a *Toxoplasma gondii* histone acetyltransferase: a novel chromatin remodelling factor in Apicomplexan parasites. *Nucleic Acids Res* 27: 4344-4352.
168. Sullivan WJ, Jr., Smith CK, 2nd (2000) Cloning and characterization of a novel histone acetyltransferase homologue from the protozoan parasite *Toxoplasma gondii* reveals a distinct GCN5 family member. *Gene* 242: 193-200.
169. Bhatti MM, Livingston M, Mullapudi N, Sullivan Jr WJ (2006) Pair of unusual GCN5 histone acetyltransferases and ADA2 homologues in the protozoan parasite *Toxoplasma gondii*. *Eukaryot Cell* 5: 62-76.
170. Smith AT, Tucker-Samaras SD, Fairlamb AH, Sullivan WJ, Jr. (2005) MYST family histone acetyltransferases in the protozoan parasite *Toxoplasma gondii*. *Eukaryot Cell* 4: 2057-2065.
171. Sautel CF, Ortet P, Saksouk N, Kieffer S, Garin J, et al. (2009) The histone methylase KMTox interacts with the redox-sensor peroxiredoxin-1 and targets genes involved in *Toxoplasma gondii* antioxidant defences. *Mol Microbiol* 71: 212-226.
172. Sullivan WJ, Jr., Monroy MA, Bohne W, Nallani KC, Chrivia J, et al. (2003) Molecular cloning and characterization of an SRCAP chromatin remodeling homologue in *Toxoplasma gondii*. *Parasitol Res* 90: 1-8.
173. Vonlaufen N, Naguleswaran A, Coppens I, Sullivan WJ, Jr. (2010) MYST family lysine acetyltransferase facilitates ataxia telangiectasia mutated (ATM) kinase-mediated DNA damage response in *Toxoplasma gondii*. *J Biol Chem* 285: 11154-11161.
174. Bougdour A, Maubon D, Baldacci P, Ortet P, Bastien O, et al. (2009) Drug inhibition of HDAC3 and epigenetic control of differentiation in Apicomplexa parasites. *J Exp Med* 206: 953-966.
175. Darkin-Rattray SJ, Gurnett AM, Myers RW, Dulski PM, Crumley TM, et al. (1996) Apicidin: a novel antiprotozoal agent that inhibits parasite histone deacetylase. *Proc Natl Acad Sci U S A* 93: 13143-13147.
176. Singh SB, Zink DL, Liesch JM, Mosley RT, Dombrowski AW, et al. (2002) Structure and chemistry of apicidins, a class of novel cyclic tetrapeptides without a terminal alpha-keto epoxide as inhibitors of histone deacetylase with potent antiprotozoal activities. *J Org Chem* 67: 815-825.
177. Strobl JS, Cassell M, Mitchell SM, Reilly CM, Lindsay DS (2007) Scriptaid and suberoylanilide hydroxamic acid are histone deacetylase inhibitors with potent anti-*Toxoplasma gondii* activity in vitro. *J Parasitol* 93: 694-700.

178. Fan Q, An L, Cui L (2004) Plasmodium falciparum histone acetyltransferase, a yeast GCN5 homologue involved in chromatin remodeling. Eukaryot Cell 3: 264-276.
179. Fan Q, An L, Cui L (2004) PfADA2, a Plasmodium falciparum homologue of the transcriptional coactivator ADA2 and its in vivo association with the histone acetyltransferase PfGCN5. Gene 336: 251-261.
180. Cui L, Miao J (2007) Cytotoxic effect of curcumin on malaria parasite Plasmodium falciparum: inhibition of histone acetylation and generation of reactive oxygen species. Antimicrob Agents Chemother 51: 488-494.
181. Cui L, Miao J, Furuya T, Fan Q, Li X, et al. (2008) Histone acetyltransferase inhibitor anacardic acid causes changes in global gene expression during in vitro Plasmodium falciparum development. Eukaryot Cell 7: 1200-1210.
182. Roos DS, Donald RG, Morrissette NS, Moulton AL (1994) Molecular tools for genetic dissection of the protozoan parasite Toxoplasma gondii. Methods Cell Biol 45: 27-63.
183. Stripen BaDS (2007) Genetic Manipulation of *Toxoplasma gondii*. In: Kim LMWaK, editor. *Toxoplasma gondii*, The Model Apicomplexan: Perspectives and Methods. 1 ed. London: Academic Press; Elsevier Ltd. pp. 391-418.
184. Rottem S, Barile MF (1993) Beware of mycoplasmas. Trends Biotechnol 11: 143-151.
185. Rowe JA, Scragg IG, Kwiatkowski D, Ferguson DJ, Carucci DJ, et al. (1998) Implications of mycoplasma contamination in Plasmodium falciparum cultures and methods for its detection and eradication. Mol Biochem Parasitol 92: 177-180.
186. Fox BA, Ristuccia JG, Gigley JP, Bzik DJ (2009) Efficient gene replacements in Toxoplasma gondii strains deficient for nonhomologous end joining. Eukaryot Cell 8: 520-529.
187. Huynh MH, Carruthers VB (2009) Tagging of endogenous genes in a Toxoplasma gondii strain lacking Ku80. Eukaryot Cell 8: 530-539.
188. Donald RG, Carter D, Ullman B, Roos DS (1996) Insertional tagging, cloning, and expression of the Toxoplasma gondii hypoxanthine-xanthine-guanine phosphoribosyltransferase gene. Use as a selectable marker for stable transformation. J Biol Chem 271: 14010-14019.
189. Kim K, Soldati D, Boothroyd JC (1993) Gene replacement in Toxoplasma gondii with chloramphenicol acetyltransferase as selectable marker. Science 262: 911-914.

190. Donald RG, Roos DS (1993) Stable molecular transformation of *Toxoplasma gondii*: a selectable dihydrofolate reductase-thymidylate synthase marker based on drug-resistance mutations in malaria. *Proc Natl Acad Sci U S A* 90: 11703-11707.
191. van den Hoff MJ, Moorman AF, Lamers WH (1992) Electroporation in 'intracellular' buffer increases cell survival. *Nucleic Acids Res* 20: 2902.
192. Stols L, Gu M, Dieckman L, Raffin R, Collart FR, et al. (2002) A new vector for high-throughput, ligation-independent cloning encoding a tobacco etch virus protease cleavage site. *Protein Expr Purif* 25: 8-15.
193. (2003) NuPAGE Technical Guide. In: Technologies IL, editor.
194. Burg JL, Grover CM, Pouletty P, Boothroyd JC (1989) Direct and sensitive detection of a pathogenic protozoan, *Toxoplasma gondii*, by polymerase chain reaction. *J Clin Microbiol* 27: 1787-1792.
195. Costa JM, Pautas C, Ernault P, Foulet F, Cordonnier C, et al. (2000) Real-time PCR for diagnosis and follow-up of *Toxoplasma* reactivation after allogeneic stem cell transplantation using fluorescence resonance energy transfer hybridization probes. *J Clin Microbiol* 38: 2929-2932.
196. Makkerh JP, Dingwall C, Laskey RA (1996) Comparative mutagenesis of nuclear localization signals reveals the importance of neutral and acidic amino acids. *Curr Biol* 6: 1025-1027.
197. Aravind L, Landsman D (1998) AT-hook motifs identified in a wide variety of DNA-binding proteins. *Nucleic Acids Res* 26: 4413-4421.
198. Swan DG, Phillips K, Tait A, Shiels BR (1999) Evidence for localisation of a *Theileria* parasite AT hook DNA-binding protein to the nucleus of immortalised bovine host cells. *Mol Biochem Parasitol* 101: 117-129.
199. Swan DG, Stern R, McKellar S, Phillips K, Oura CA, et al. (2001) Characterisation of a cluster of genes encoding *Theileria annulata* AT hook DNA-binding proteins and evidence for localisation to the host cell nucleus. *J Cell Sci* 114: 2747-2754.
200. Schultz J, Milpetz F, Bork P, Ponting CP (1998) SMART, a simple modular architecture research tool: identification of signaling domains. *Proc Natl Acad Sci U S A* 95: 5857-5864.
201. Letunic I, Doerks T, Bork P (2009) SMART 6: recent updates and new developments. *Nucleic Acids Res* 37: D229-232.
202. Baker SP, Grant PA (2007) The SAGA continues: expanding the cellular role of a transcriptional co-activator complex. *Oncogene* 26: 5329-5340.
203. Dunker AK, Lawson JD, Brown CJ, Williams RM, Romero P, et al. (2001) Intrinsically disordered protein. *J Mol Graph Model* 19: 26-59.

204. Dunker AK, Cortese MS, Romero P, Iakoucheva LM, Uversky VN (2005) Flexible nets. The roles of intrinsic disorder in protein interaction networks. *FEBS J* 272: 5129-5148.
205. Uversky VN, Dunker AK (2010) Understanding protein non-folding. *Biochim Biophys Acta* 1804: 1231-1264.
206. Uversky VN, Oldfield CJ, Dunker AK (2005) Showing your ID: intrinsic disorder as an ID for recognition, regulation and cell signaling. *J Mol Recognit* 18: 343-384.
207. Mohan A, Oldfield CJ, Radivojac P, Vacic V, Cortese MS, et al. (2006) Analysis of molecular recognition features (MoRFs). *J Mol Biol* 362: 1043-1059.
208. Mohan A, Sullivan WJ, Jr., Radivojac P, Dunker AK, Uversky VN (2008) Intrinsic disorder in pathogenic and non-pathogenic microbes: discovering and analyzing the unfoldomes of early-branching eukaryotes. *Mol Biosyst* 4: 328-340.
209. Bhatti MM (2006) Functions of the unique N-terminus of a GCN5 histone acetylase in *Toxoplasma gondii*. Indianapolis: Indiana University.
210. Musselman CA, Kutateladze TG (2009) PHD fingers: epigenetic effectors and potential drug targets. *Mol Interv* 9: 314-323.
211. Banaszynski LA, Chen LC, Maynard-Smith LA, Ooi AG, Wandless TJ (2006) A rapid, reversible, and tunable method to regulate protein function in living cells using synthetic small molecules. *Cell* 126: 995-1004.
212. Banaszynski LA, Wandless TJ (2006) Conditional control of protein function. *Chem Biol* 13: 11-21.
213. Meissner M, Schluter D, Soldati D (2002) Role of *Toxoplasma gondii* myosin A in powering parasite gliding and host cell invasion. *Science* 298: 837-840.
214. Mital J, Meissner M, Soldati D, Ward GE (2005) Conditional expression of *Toxoplasma gondii* apical membrane antigen-1 (TgAMA1) demonstrates that TgAMA1 plays a critical role in host cell invasion. *Mol Biol Cell* 16: 4341-4349.
215. Huynh MH, Carruthers VB (2006) *Toxoplasma* MIC2 is a major determinant of invasion and virulence. *PLoS Pathog* 2: e84.
216. Mazumdar J, E HW, Masek K, C AH, Striepen B (2006) Apicoplast fatty acid synthesis is essential for organelle biogenesis and parasite survival in *Toxoplasma gondii*. *Proc Natl Acad Sci U S A* 103: 13192-13197.
217. Plattner F, Yarovsky F, Romero S, Didry D, Carlier MF, et al. (2008) *Toxoplasma* profilin is essential for host cell invasion and TLR11-dependent induction of an interleukin-12 response. *Cell Host Microbe* 3: 77-87.
218. Kessler H, Herm-Gotz A, Hegge S, Rauch M, Soldati-Favre D, et al. (2008) Microneme protein 8--a new essential invasion factor in *Toxoplasma gondii*. *J Cell Sci* 121: 947-956.

219. Herm-Gotz A, Agop-Nersesian C, Munter S, Grimley JS, Wandless TJ, et al. (2007) Rapid control of protein level in the apicomplexan *Toxoplasma gondii*. *Nat Methods* 4: 1003-1005.
220. Armstrong CM, Goldberg DE (2007) An FKBP destabilization domain modulates protein levels in *Plasmodium falciparum*. *Nat Methods* 4: 1007-1009.
221. Striepen B (2007) Switching parasite proteins on and off. *Nat Methods* 4: 999-1000.
222. Radke JR, Striepen B, Guerini MN, Jerome ME, Roos DS, et al. (2001) Defining the cell cycle for the tachyzoite stage of *Toxoplasma gondii*. *Mol Biochem Parasitol* 115: 165-175.
223. Gubbels MJ, White M, Szatanek T (2008) The cell cycle and *Toxoplasma gondii* cell division: tightly knit or loosely stitched? *Int J Parasitol* 38: 1343-1358.
224. Dang CV, Lee WM (1988) Identification of the human c-myc protein nuclear translocation signal. *Mol Cell Biol* 8: 4048-4054.
225. Zheng Y, Balasubramanyam K, Cebrat M, Buck D, Guidez F, et al. (2005) Synthesis and evaluation of a potent and selective cell-permeable p300 histone acetyltransferase inhibitor. *J Am Chem Soc* 127: 17182-17183.
226. Sturzu A, Heckl S (2009) The fluorinated and chlorinated nuclear localization sequence of the SV 40 T antigen. *Chem Biol Drug Des* 73: 127-131.
227. Kosugi S, Hasebe M, Entani T, Takayama S, Tomita M, et al. (2008) Design of peptide inhibitors for the importin alpha/beta nuclear import pathway by activity-based profiling. *Chem Biol* 15: 940-949.
228. Kaiser N, Lischka P, Wagenknecht N, Stamminger T (2009) Inhibition of human cytomegalovirus replication via peptide aptamers directed against the nonconventional nuclear localization signal of the essential viral replication factor pUL84. *J Virol* 83: 11902-11913.
229. Lopez-Ochoa L, Nash TE, Ramirez-Prado J, Hanley-Bowdoin L (2009) Isolation of Peptide aptamers to target protein function. *Methods Mol Biol* 535: 333-360.
230. Fu H, Subramanian RR, Masters SC (2000) 14-3-3 proteins: structure, function, and regulation. *Annu Rev Pharmacol Toxicol* 40: 617-647.
231. Morrison DK (2009) The 14-3-3 proteins: integrators of diverse signaling cues that impact cell fate and cancer development. *Trends Cell Biol* 19: 16-23.
232. Wang AH, Kruhlak MJ, Wu J, Bertos NR, Vezmar M, et al. (2000) Regulation of histone deacetylase 4 by binding of 14-3-3 proteins. *Mol Cell Biol* 20: 6904-6912.
233. Choudhary C, Kumar C, Gnad F, Nielsen ML, Rehman M, et al. (2009) Lysine acetylation targets protein complexes and co-regulates major cellular functions. *Science* 325: 834-840.

234. Stockinger EJ, Mao Y, Regier MK, Triezenberg SJ, Thomashow MF (2001) Transcriptional adaptor and histone acetyltransferase proteins in *Arabidopsis* and their interactions with CBF1, a transcriptional activator involved in cold-regulated gene expression. *Nucleic Acids Res* 29: 1524-1533.
235. Xu W, Edmondson DG, Roth SY (1998) Mammalian GCN5 and P/CAF acetyltransferases have homologous amino-terminal domains important for recognition of nucleosomal substrates. *Mol Cell Biol* 18: 5659-5669.
236. Sterner DE, Nathan D, Reindle A, Johnson ES, Berger SL (2006) Sumoylation of the yeast Gcn5 protein. *Biochemistry* 45: 1035-1042.
237. Braun L, Cannella D, Pinheiro AM, Kieffer S, Belrhali H, et al. (2009) The small ubiquitin-like modifier (SUMO)-conjugating system of *Toxoplasma gondii*. *Int J Parasitol* 39: 81-90.
238. Barlev NA, Poltoratsky V, Owen-Hughes T, Ying C, Liu L, et al. (1998) Repression of GCN5 histone acetyltransferase activity via bromodomain-mediated binding and phosphorylation by the Ku-DNA-dependent protein kinase complex. *Mol Cell Biol* 18: 1349-1358.
239. Zhang W, Bone JR, Edmondson DG, Turner BM, Roth SY (1998) Essential and redundant functions of histone acetylation revealed by mutation of target lysines and loss of the Gcn5p acetyltransferase. *Embo J* 17: 3155-3167.
240. Vernarecci S, Ornaghi P, Bagu A, Cundari E, Ballario P, et al. (2008) Gcn5p plays an important role in centromere kinetochore function in budding yeast. *Mol Cell Biol* 28: 988-996.
241. Orpinell M, Fournier M, Riss A, Nagy Z, Krebs AR, et al. (2010) The ATAC acetyltransferase complex controls mitotic progression by targeting non-histone substrates. *EMBO J*.
242. Morrisette NS, Sibley LD (2002) Disruption of microtubules uncouples budding and nuclear division in *Toxoplasma gondii*. *J Cell Sci* 115: 1017-1025.
243. Kvaal CA, Radke JR, Guerini MN, White MW (2002) Isolation of a *Toxoplasma gondii* cyclin by yeast two-hybrid interactive screen. *Mol Biochem Parasitol* 120: 187-194.

Curriculum Vitae

Stacy E. Dixon

Education

- 2004-present Medical Doctorate
Indiana University School of Medicine, Indianapolis, IN
Expected graduation date: May 2012
- 2006-2010 Doctor of Philosophy in Pharmacology
Indiana University, Indianapolis, IN
Dissertation topic: My research focuses on understanding gene expression in the protozoan parasite *Toxoplasma gondii*. In particular, I am investigating the role of the histone acetyltransferase GCN5 in parasite survival and differentiation using both biochemical techniques and molecular genetics.
Advisor: William J. Sullivan, Jr., PhD
- 2000-2004 Bachelor of Science in Microbiology, minor in Chemistry, with Distinction
Indiana University, Bloomington, IN

Grants/Fellowships

- 2008 PhRMA Foundation Pre-doctoral Fellowship, awarded December 2008
- 2008 K. K. Chen Fellowship (presented by the Department of Pharmacology and Toxicology IUSM), awarded December 2008

Honors and Awards

- 2007 Educational Enhancement Grant from IUPUI Graduate Student Organization (travel)
- 2007 Corporate Activities Program student travel grant award for the ASM General Meeting
- 2004 Jill Behrman Emerging Leader Scholarship, Indiana University

Publications

Dixon, Stacy E., Micah M. Bhatti, Vladimir N. Uversky, A. Keith Dunker, William J. Sullivan, Jr. "Regions of intrinsic disorder help identify a novel nuclear localization signal in *Toxoplasma gondii* histone acetyltransferase TgGCN5-B." accepted by *Molecular and Biochemical Parasitology*.

Dixon, Stacy E., Krista Stilger, Eliana V. Elias, Arunasalam Naguleswaran, William J. Sullivan, Jr. "A Decade of Epigenetics Research in *Toxoplasma gondii*." *Molecular and Biochemical Parasitology* 173(1), 2010: 1-9.

Narasimhan, Jana, Bradley R. Joyce, Arunasalam Naguleswaran, Aaron T. Smith, Meredith R. Livingston, **Stacy E. Dixon**, Isabelle Coppens, Ronald C. Wek, and William J. Sullivan, Jr. "Translation Regulation by eIF2 Kinases in the Development of Latent Cysts in *Toxoplasma gondii*." *Journal of Biological Chemistry* 283(24), 2008: 16591-16601.

Sullivan, Jr., William J., **Stacy E. Dixon**, Catherine Li, Boris Striepen, and Sherry F. Queener. "IMP Dehydrogenase from the Protozoan Parasite *Toxoplasma gondii*." *Antimicrobial Agents and Chemotherapy* 49 (2005): 2172-2179.

Presentations

Dixon SE, Naguleswaran A, Bhatti MM, and Sullivan Jr. WJ. "GCN5 histone acetyltransferases in *Toxoplasma*: two HATs are better than one." Molecular Parasitology Meeting, Woods Hole, MA (Sept. 2008), poster.

Dixon SE, Livingston MR, Bhatti MM, and Sullivan Jr. WJ. "Elucidating the Role of TgGCN5-B Histone Acetyltransferase in *Toxoplasma*." COST Action 857: 3rd Annual PhD Students' Retreat, Geneva, Switzerland (Nov. 2007), speaker.

Note: Attendance of this conference involved a competitive application process. I was the only American selected to attend with 14 European students.

Dixon SE, Livingston MR, Bhatti MM, and Sullivan Jr. WJ. "A Tale of Two GCN5s: An Unusual Pair of Histone Acetyltransferases in the Protozoan Parasite *Toxoplasma gondii*." Molecular Parasitology Meeting, Woods Hole, MA (Sept. 2007), poster.

Dixon SE, Livingston MR, Bhatti MM, and Sullivan Jr. WJ. "A Tale of Two GCN5s: An Unusual Pair of Histone Acetyltransferases in the Protozoan Parasite *Toxoplasma gondii*." 107th General Meeting of the American Society for Microbiology, Toronto, Ontario (May 2007), poster.

Dixon SE, Sullivan Jr. WJ, Li C, Striepen B, and Queener SF. "IMP Dehydrogenase from the Protozoan Parasite *Toxoplasma gondii*." Chicago Area Mycology and Parasitology meeting, Chicago, IL (Apr. 2005), poster.

Professional Development

July 2008 Kadner Institute, American Society for Microbiology, Boulder, CO
Note: Attendance of this institute involved a competitive application process.
June 2008 Preparing Future Faculty Summer Institute, IUPUI

Teaching Experience

The eight students were under my supervision during their rotations in the lab of Dr. William J. Sullivan, Jr. Each student was instructed on basic laboratory skills and molecular genetic techniques.

Clinical Experience

Aug. 2006 – July 2008 St. Vincent Hospital, Indianapolis, IN
Pre-doctoral Extern in the Emergency Department

Professional Affiliations

American Society for Microbiology
Phi Beta Kappa
Blue Key National Honor Fraternity
Golden Key International Honor Society

Service

2007-2009 President, Pharmacology and Toxicology Graduate Student Organization
and representative to IUSM Graduate Student Organization
2007-2009 Student Ambassador, Indiana University Biomedical Gateway program
2008-2009 Member, Universities Allied for Essential Medicines, Indiana University

Certifications

United States Medical Licensing Exam (USMLE) Step 1 completed 06/2006