FUNCTIONS OF THE UNIQUE N-TERMINUS OF A GCN5 HISTONE ACETYLASE IN *TOXOPLASMA GONDII*

Micah M. Bhatti

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

May 2007

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. - Chair

_____

Edward M. Chan, M.D.

_____

Sherry F. Queener, Ph.D.

Doctoral Committee

_____

Ahmad R. Safa, Ph.D.

_____

Anthony P. Sinai, Ph.D.

November 6, 2006

_____

Michael R. Vasko, Ph.D.

ACKNOWLEDGEMENTS

I cannot thank my mentor Dr. William Sullivan enough. He has been a wonderful teacher with an immeasurable passion for science and an equally immeasurable amount of patience. I doubt I will ever meet anyone else more devoted to his work, students, and lab personnel and yet he manages to find time to start a family. I have been truly blessed to have been able to work along side him.

I extend my deepest gratitude to my thesis committee members: Dr. Edward M. Chan, Dr. Sherry F. Queener, Dr. Ahmad R. Safa, Dr. Anthony P. Sinai and Dr. Michael R. Vasko. Your advice and suggestions were a tremendous help in completing my thesis project. I'd especially like to thank Dr. Sinai who made several commutes from University of Kentucky.

I also like to thank our collaborators Dr. Mohamed-Ali Hakimi and his laboratory at the National Centre for Scientific Research in Grenoble, France for their work and skills at biochemical purification of protein complexes.

Thanks to my lab mates in the Sullivan Lab (Kelli Bramlett, Stacy Dixon, Chunling Jiang, Meredith Livingston, and Aaron Smith) for their advice, humor, and 10X TBST. I look forward to more poker games; always, always remember to burn a card!!! I'd also like to thank Pam Torkelson in Dr. Queener's Lab for her help; a truly skilled technician.

I'd also like to thank all the students and faculty in the Department of Pharmacology and Toxicology.

Sincere thanks to my parents, Drs. Waqar and Martha Bhatti for their support and encouraging my love of science dating back to the Fisher Price microscope I received for my 8th birthday.

Lastly, I'd like to thank my wife, Emily. Without her love and support I would not have survived graduate school.

ABSTRACT

Micah M. Bhatti

FUNCTIONS OF THE UNIQUE N-TERMINUS OF A GCN5 HISTONE ACETYLASE IN *TOXOPLASMA GONDII*

GCN5 is a histone acetyltransferase (HAT) that remodels chromatin by acetylating lysine residues of histones. The GCN5 HAT identified in *Toxoplasma gondii* (TgGCN5) contains a unique N-terminal "extension" that bears no similarity to known proteins and is devoid of known protein motifs. The hypothesis of this thesis is the N-terminal extension is critical to the function of TgGCN5. Three possible roles of the N-terminus were investigated: nuclear localization, protein-protein interactions, and substrate recognition. Subcellular localization was determined via immunocytochemistry using parasites expressing recombinant forms of TgGCN5 fused to a FLAG tag. Initial studies performed with parasites expressing full length $_{FLAG}$TgGCN5 were positive for nuclear localization. Without the N-terminal extension ($_{FLAG}\Delta N_T$TgGCN5) the protein remains cytoplasmic. Additional studies mapped a six amino acid motif (RKRVKR) as the nuclear localization signal (NLS). When RKRVKR is fused to a cytoplasmic protein, it gains access to the nucleus. Furthermore, we have established the NLS interacts with *Toxoplasma* importin α, a protein involved in nuclear trafficking. Interaction with importin α provides evidence that the TgGCN5 N-terminal extension is involved in mediating protein-protein

interactions. In order to identify additional interacting proteins, FLAG affinity purification was performed on parasites expressing full length $_{FLAG}$TgGCN5 and $_{FLAG}\Delta N_T$TgGCN5. Upon comparing the results of the two purifications, proteins captured with only full length TgGCN5 may be interacting with the N-terminal extension. Full length TgGCN5 affinity purification indicates an interaction with histone proteins, two different homologues of Ada2 (adapter protein reported to interact with GCN5 homologues), and several heat shock proteins. With regard to substrate recognition, the N-terminal extension of TgGCN5 is dispensable for the acetylation of non-nucleosomal histones *in vitro*. However, the lysine acetylated by TgGCN5 is surprisingly unique. Other GCN5 homologues preferentially acetylate lysine 14 in histone H3, but TgGCN5 exclusively acetylates lysine 18 in histone H3 and has no activity on lysine 14. Taken together, these results argue that the N-terminal extension of TgGCN5 is critical for mediating protein-protein interactions, including those responsible for trafficking the HAT to the parasite nucleus but does not appear to be required for the acetylation of non-nucleosomal histones.

William J. Sullivan, Jr., Ph.D. - Chair

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

6-TX..... 6-thioxanthine
AAA ..... ATPase associated with a variety of activities
Ada ..... alteration/deficiency in activation protein or
ADA ..... Ada containing acetyltransferase complex
Ahc1 ..... ADA HAT complex component 1
AIDS ..... acquired immunodeficiency syndrome
ARF ..... auxin response factor
ARM ..... armadillo motifs
ATCC..... American Tissue Culture Company
β-Me ..... beta mercaptoethanol
βgal..... *E. coli* beta galactosidase
CARM1 ..... cofactor-associated arginine [R] methyltransferase 1
CAS ..... cellular apoptosis susceptibility gene
CBP ..... CREB binding protein
Cdc ..... cell division control
CDK..... cyclin-dependent kinase
ChIP ..... chromatin immunoprecipitation
CNS..... central nervous system
CREB ..... cAMP responsive element binding protein
CT..... computerized tomography
DAPI ..... 4',6'-diamino-2-phenylindole
DHFR ..... dihydrofolic acid reductase
DHFR-TS..... dihydrofolic acid reductase- thymidylate synthase
DHPS ..... dihydropteroate synthase
DMEM ..... Dulbecco's Modified Eagle Medium
DNA..... deoxyribonucleic acid
DTT ..... $_{DL}$-Dithiothreitol
ELISA ..... enzyme-linked immunosorbent assay
ENO..... enolase
EtOH..... ethyl alcohol

FAT..... factor acetyltransferase
FZ media ..... freezing media
GCN5 ..... general control non-derepressed 5
GDP..... guanosine diphosphate
GFP ..... green fluorescent protein
GTP ..... guanosine triphosphate
HA tag ..... hemagluttin tag
HAART ..... highly active anti-retroviral therapy
HAT ..... histone acetyltransferase
HC media ..... host cell media
HDAC ..... histone deacetylase
HIV ..... human immunodeficiency virus
HMG ..... high mobility group proteins
HRP ..... horseradish peroxidase
Hsp ..... heat shock protein
HXGPRT ..... hypoxanthine-xanthine-guanine phosphoribosyltransferase
IBB..... importin β binding domain
IFA..... immunofluorescent assay
IFN-γ..... interferon gamma
IPTG ..... isopropyl β-D-1-thiogalactopyranoside
iNOS..... inducible nitric oxide synthase
IMPα ..... importin alpha
IMPβ ..... importin beta
kDa ..... kiloDalton
LDH ..... lactate dehydrogenase
MALDI/TOF ..... matrix-assisted laser desorption/ionization time-of-flight
MAD ..... mitotic arrest deficient
MES..... 2-Morpholinoethanesulfonic acid
MIC..... microneme proteins of *Toxoplasma gondii*
MLL ..... mixed linage leukemia
MOF ..... male on first

MOPS ..... 3-(N-Morpholino) propanesulfonic acid
MOZ ..... monocytic leukemia zinc finger protein
MPA..... mycophenolic acid
MRI..... magnetic resonance imaging
mRNA..... messenger ribonucleic acid
MSL ..... male-specific lethal
MYST ..... MOZ, Ybf2/Sas3, Sas2, Tip60 histone acetylase family
NAD..... nicotinamide adenine dinucleotide
NCE..... nuclear cell extract
NLS ..... nuclear localization signal
NPC..... nuclear pore complex
NuA3 ..... nucleosome acetyltransferase of histone H3
NuA4 ..... nucleosome acetyltransferase of histone H4
PA media..... parasite media
PAGE ..... polyacrylamide gel electrophoresis
PCAF..... p300/CBP associating factor
PCR..... polymerase chain reaction
Ran..... Ras-related nuclear protein
RanGAP ..... Ran GTPase activating factor
RanGEF ..... Ran guanine nucleotide exchange factor
Rb..... retinoblastoma protein
RCC1..... regulator of chromosome condensation 1
RNA..... ribonucleic acid
RRM 1 ..... RNA recognition motif 1
RTS ..... Rubinstein-Taybi syndrome
SAG2A ..... surface antigen 2-A of *Toxoplasma gondii*
SAGA ..... Spt-Ada-GCN5 acetyltransferase complex
SALSA..... SAGA altered, Spt8 absent acetyltransferase complex
SLIK..... SAGA-like acetyltransferase complex
Sas2 ..... something about silencing 2
ScGCN5 ..... *Saccharomyces cerevisiae* homologue of GCN5

SDS ..... sodium dodecyl sulfate

SGF ..... SAGA associated factor

Spt ..... suppression of Ty insertions

TAF_{II} ..... TBP associated factors

Tat ..... trans-activator viral protein of HIV

TBP ..... TATA binding protein

TCA ..... trichloroacetic acid

TE ..... toxoplasmic encephalitis

TEN ..... toxic epidermal necrosis

TFTC ..... TATA-binding protein-free TAF-containing complex

TgGCN5 ..... *Toxoplasma gondii* homologue of GCN5

TgIMPα ..... *Toxoplasma gondii* homologue of importin alpha

TgCARM1 ..... *Toxoplasma gondii* homologue of CARM1

Tip60 ..... Tat-interactive protein, 60 kDa

Tra1 ..... transplantability associated gene 1

tRNA ..... transfer ribonucleic acid

TRRAP ..... transformation/transcription domain–associated protein

TSA ..... trichostatin A

Ty ..... translocatable yeast element

UV ..... ultraviolet

VCP ..... valosin-containing protein

VP16 ..... herpes simplex viral protein 16

WCE ..... whole cell extract

Ybf2/Sas3 ..... something about silencing 3

CHAPTER 1: INTRODUCTION

I. Phylum Apicomplexa and *Toxoplasma gondii*

There are over 4,000 members in the phylum Apicomplexa and all are parasitic organisms. Electron microscopy has revealed unique structural features which define this phylum (Figure 1). The defining characteristic of apicomplexa is a group of organelles found at the anterior end of the organism called the apical complex. The apical complex includes secretory organelles known as micronemes and rhoptries, polar rings composed of microtubules, and a conoid which lies within the polar rings (Blackman and Bannister, 2001). The apical organelles play a major role in parasite invasion of the host cell. Most apicomplexan parasites also possess an apicoplast, an essential organelle that appears to have been acquired by secondary endosymbiosis of a green alga (Waller and McFadden, 2005). Secondary endosymbiosis theorizes that certain organelles such as the mitochondrion and plastids were originally derived from an engulfed organism that was incorporated into the cellular structure rather than being digested (Marechal and Cesbron-Delauw, 2001). The 35 kilobase genome of the apicoplast encodes its own genes; however most of the proteome of the apicoplast is encoded in the parasites nucleus (Foth and McFadden, 2003).

Several medically important protozoan parasites belong to the phylum Apicomplexa. The most notorious is the *Plasmodium* spp., the causative agent of malaria. The various species of *Plasmodium* kill over 3 million people a year; one third of which are children (Snow et al., 1999). The current annual malaria death toll equates to approximately 340 deaths every hour. *Cryptosporidium parvum* is an apicomplexan that causes a self-limiting intestinal illness that can be life-threatening in immune compromised individuals. The largest outbreak of *C. parvum* in the US occurred in Milwaukee, Wisconsin in the spring of 1993. Approximately 400,000 people were infected and developed severe diarrhea (MacKenzie et al., 1994). The outbreak is believed to have been caused by ineffective filtration and treatment of the local water supply (MacKenzie et al., 1994). Of the 400,000 with disease, about 100 deaths occurred; all individuals

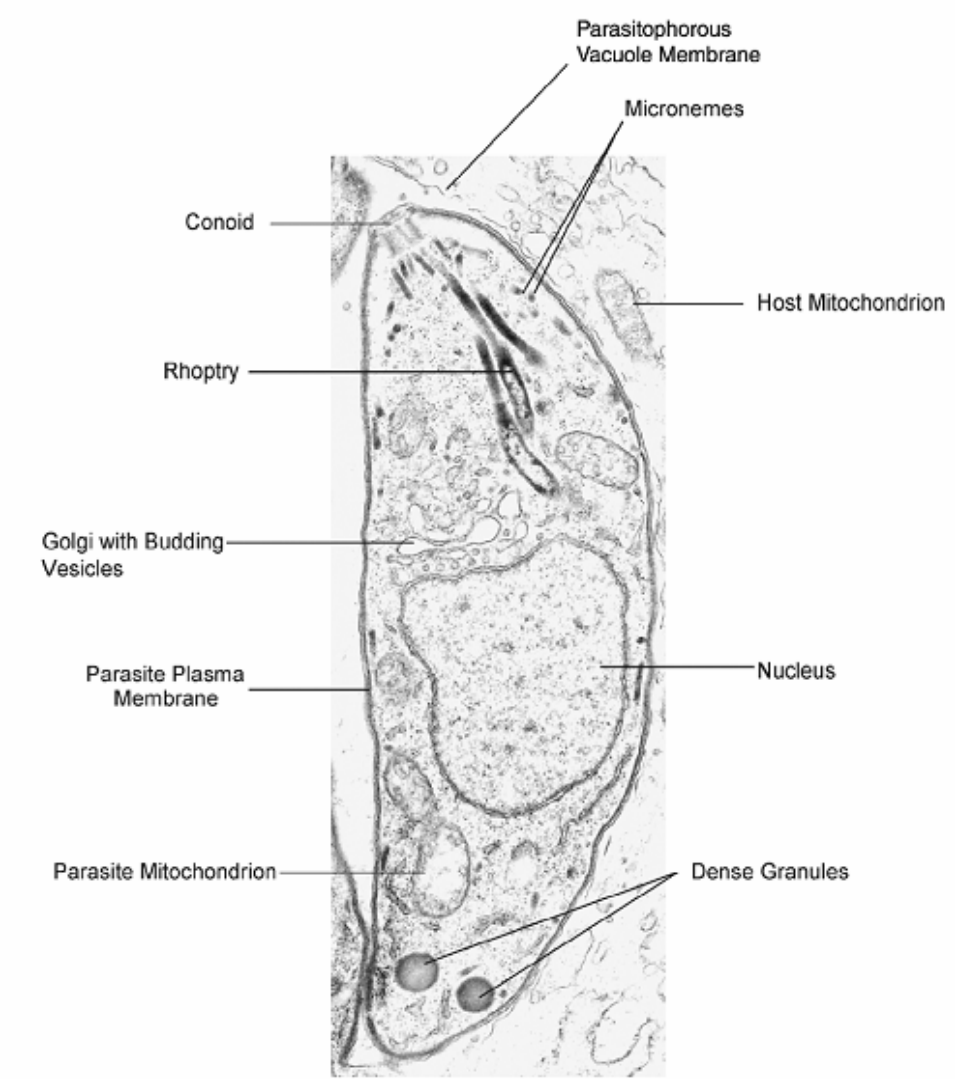


Figure 1: Transmission electron image of a tachyzoite within a host cell
Tachyzoites possess organelles present in all eukaryotes (mitochondria, nucleus, and Golgi complex). At the apical end, the micronemes, rhoptries and conoid form the apical complex, a group of organelles involved in host cell invasion. The micronemes and rhoptries are secretory organelles and the conoid is a cytoskeletal structure connected to the actin-myosin motility of the parasite. Although not a part of the apical complex, dense granules are secretory organelles that also function in the invasion process. The parasite is housed within a parasitophorous vacuole that is generated during host cell invasion. (Binder and Kim, 2004; reprinted with permission from Blackwell Publishing)

with immune deficiencies (Anderson, 1998). Currently, there are no effective antimicrobials capable of treating *C. parvum* infections (Lane et al., 1994).

Apicomplexans parasites also target non-human hosts causing significant economic burden for the livestock industry. Avian coccidiosis is an intestinal disease caused by parasites of the genus *Eimeria*, and is a constant threat to the poultry industry (Allen and Fetterer, 2002). The annual worldwide cost is estimated at about $800 million and about $450 million for the American poultry industry alone (Allen and Fetterer, 2002). These estimates include the costs of prophylactic medication and losses due to mortality by *Eimeria* infection (Allen and Fetterer, 2002).

Toxoplasma gondii, the most widely disseminated parasite in this phylum, infects virtually all warm-blooded vertebrates. It is estimated that between 16-40% of the United States population is infected and up to 80% of individuals in other countries (Hill and Dubey, 2002). While both the initial and resulting latent infection with *Toxoplasma* are asymptomatic, initial infections in pregnant women can result in serious obstetric complications and reactivation of the latent infection in immunocompromised patients can cause life-threatening illness (Guerina et al., 1994; Wong and Remington, 1993).

Every year between 400-4000 cases of congenital toxoplasmosis occurs in the US alone (Boyer, 1996). In nations with higher infection rates, the frequency of neonatal complications can be as high as 1 in every 1000 births (Guerina et al., 1994). *In utero Toxoplasma* infections can result in a wide range of severity from children born with asymptomatic disease to severe mental retardation and disability (Jones et al., 2003).

Growing numbers of immunocompromised individuals, largely due to the AIDS epidemic, has created an increase in toxoplasmosis seen in the health care setting. In the absence of drug therapy, 20-47% of HIV patients with latent *Toxoplasma* infections will develop toxoplasmic encephalitis (Wong and Remington, 1993). Between 1992 and 1998, over 6,000 HIV infected individuals died from toxoplasmosis (Jones et al., 2002).

Because immune suppression is necessary for organ transplantation, these individuals are also at risk for reactivation of latent *Toxoplasma*. In patients receiving heart transplants, 4-12% will suffer from complications stemming from reactivated toxoplasmosis (Hermanns et al., 2001; Wagner et al., 1994).

Unfortunately, for individuals threatened by *Toxoplasma* infections, the current therapies (page 14) are limited by serious side-effects (bone marrow suppression and hypersensitivity reactions). Furthermore, these drugs are unable to target the latent infection. Thus, novel therapeutics are seriously needed that possess fewer side effects and/or the ability to eradicate the latent infection.

In addition to its own medical significance, *Toxoplasma* also serves as an excellent model to study other apicomplexans (Roos et al., 1999). Clinically important apicomplexans such as the *Plasmodium* spp. and *C. parvum* are difficult to grow *in vitro* and are not amenable to genetic manipulation. *Toxoplasma* however, is relatively easy to cultivate *in vitro* and various molecular techniques and resources have been established to study this organism including transfection (Soldati and Boothroyd, 1993), gene knockout strategies (Donald and Roos, 1994), and an online genomic database (http://toxodb.org/; Kissinger et al., 2003). Thus, for many experimental questions, *Toxoplasma* remains the best model system to study the biology of the Apicomplexa. Our understanding of the mechanisms of drug resistance, the biology of the apicoplast, and the process of host cell invasion has been advanced by studies in *Toxoplasma* (Kim and Weiss, 2004). Heterologous expression of apicomplexan proteins in *Toxoplasma* has frequently facilitated further characterization of proteins that could not otherwise be easily studied (Kim and Weiss, 2004). Although findings in *Toxoplasma* will not always be applicable to other apicomplexan members, it is an important model system for understanding apicomplexan biology.

A. Life cycle of Toxoplasma gondii

Toxoplasma was first isolated by Nicolle and Manceaux in 1908, from the African rodent *Ctenodactylus gondii*, which shares the same species designation (Petersen and Dubey, 2001). Under the microscope, the parasite has a distinct crescent shape similar to a bow. The unusual shape of the organism was used to derived its genus name from the Greek word *toxon*, meaning "bow" and *plasma* meaning "the form" (Petersen and Dubey, 2001).

The life cycle consists of two major phases; an asexual and sexual cycle (Figure 2). While the parasites can grow asexually in almost any warm blooded vertebrate, the sexual cycle is restricted to cat intestinal tract (Dubey et al., 1970). In feline intestinal epithelium the parasite differentiates into either macro- or microgametes that fuse to form diploid oocysts which are subsequently excreted in the cat's feces (Ferguson et al., 1975). Oocysts have a very thick protective wall encasing the organism (Speer, 1998). Upon exposure to the external environment, the diploid organism inside the oocyst sporulate generating infectious sporozoites. The sporulation process (sporogony) consists of two meiotic divisions and a single mitotic division, producing eight haploid sporozoites. Sporulated oocysts present in the soil can remain infectious for more than a year (Wong and Remington, 1993).

When oocysts are ingested, the thick protective oocyst wall is degraded by host digestive enzymes, releasing the infectious sporozoites. If a cat consumes oocysts, the sexual cycle will take place, resulting in the formation of gametes and oocysts which are shed with the cat's feces. However, in all mammals (including cats) and birds that ingest oocysts, the released sporozoites will penetrate the intestinal epithelium and invade intestinal cells. During this initial invasion of host cells, the sporozoites differentiate into tachyzoites that disseminate throughout the host upon reaching the blood stream (Wong and Remington, 1993).

Once the parasite differentiates into a tachyzoite, it becomes an obligate intracellular parasite and cannot survive for long outside of a host cell. They are purine auxotrophs and require certain amino acids such as glutamine and

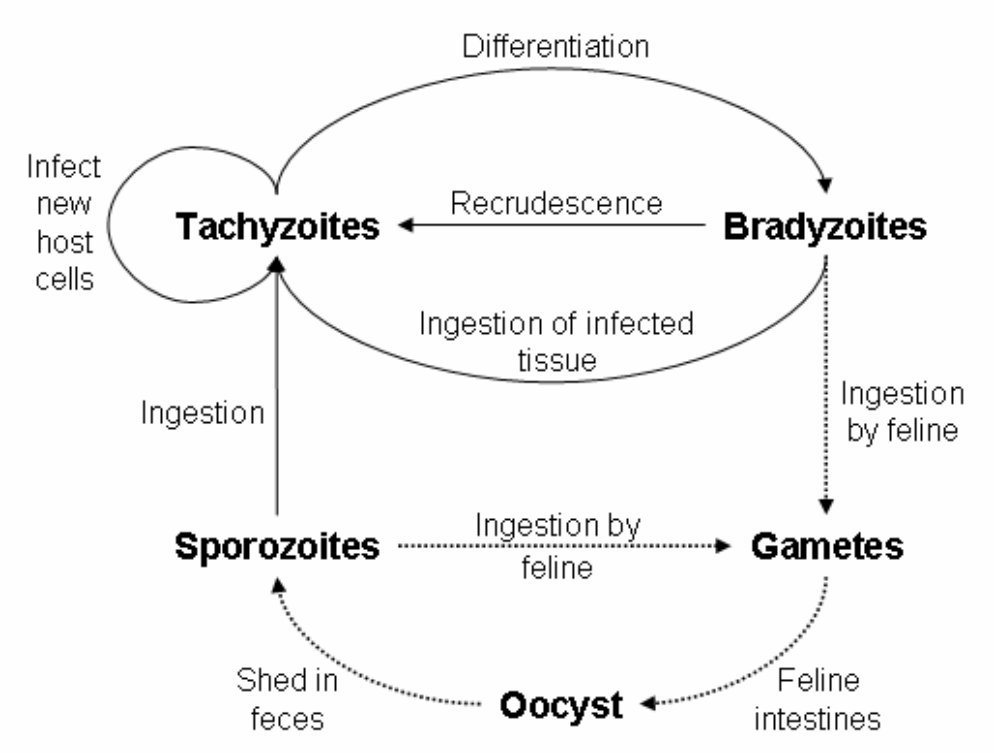


Figure 2: Life cycle of *Toxoplasma gondii*

The diagram depicts the sexual cycle (dotted line), which occurs only in the feline intestines, and the asexual cycle (solid line) which takes place in nucleated cells of all infected animals. During the asexual cycle, a small portion of the fast growing tachyzoites will differentiate into encysted slow growing bradyzoites. These bradyzoite cysts persist throughout the lifespan of the host resulting in a permanent latent infection. Furthermore, the bradyzoites can spontaneously recrudesce back to tachyzoites, potentially causing acute illness long after the initial infection has resolved.

arginine for survival (Sullivan et al., 1999; Fox and et al., 2004). These biomolecules are siphoned from the infected host cell. Because of its obligate intracellular lifestyle, *Toxoplasma* has developed a highly orchestrated invasion process that is quick and involves the use of the apical complex (Carruthers, 2002).

During the invasion process, the parasite surrounds itself within a vacuole derived largely from the host cell membrane (Leirião et al., 2004). However, many of the host membrane proteins appear to be absent and several parasite proteins released during the invasion process are incorporated into the parasite vacuolar membrane (Sibley and Krahenbuhl, 1988; Beckers et al., 1994). Through an unknown process, the parasite vacuole does not fuse with host cell lysosomes, protecting it from normal vacuole acidification (Sibley et al., 1985). Inside the vacuole, the parasite undergoes asexual replication by endodyogeny, a specialized type of division in which the two daughter cells form within the mother cell (Hu et al., 2002). The parasite continues to replicate inside the host cell with a generation time of six to eight hours until the vacuole becomes too large for the host cell and the cell ruptures (Black and Boothroyd, 2000). The tachyzoites emerge from the ruptured host cell and repeat the invasion and replication processes inside new host cells. The continuous cycle of invasion, replication, and cell lysis by tachyzoites leads to the tissue destruction associated with *Toxoplasma* infection.

About two weeks post infection, a portion of the invasive tachyzoites will differentiate into bradyzoites through a poorly understood process (Lyons et al., 2002). Bradyzoites replicate more slowly and are housed inside of a thick, rigid cyst-like membrane composed of polysaccharides and chitin (Weiss and Kim, 2000). The cysts vary in size (10-100μm) and form inside host cells, rendering detection by the immune system difficult and resulting in a permanent latent infection.

When an infected host animal is eaten, the bradyzoite cysts present in the host tissues are capable of causing infection in a manner similar to oocysts (Hill and Dubey, 2002). The cyst wall protects the bradyzoites from the digestive

enzymes present in the stomach. After passing through the stomach, the cyst wall is eventually degraded and the released bradyzoites infect the intestinal mucosa, differentiating back into tachyzoites (Dubey et al., 1970). The tachyzoites disseminate throughout the host animal. In the case of a feline, ingestion of bradyzoite cysts will also result in the formation of gametes producing oocysts which will be shed with the feces thus completing the life cycle (Dubey et al., 1970).

B. Infection and pathobiology

Humans can become infected through several routes. Young children can become infected through accidental ingestion by playing in soil contaminated with oocysts. People exposed to cats can acquire infection by handling cat liter containing oocysts. Consumption of improperly cooked meat containing bradyzoite cysts can result in infection. It is estimated that of the 750 deaths attributed to toxoplasmosis each year, 375 (50%) are believed to be caused by eating contaminated meat, making toxoplasmosis the third leading cause of food borne deaths in the United States (Lopez et al., 2000).

Once ingested, the oocysts or bradyzoite cysts release infective parasites that invade the intestinal epithelium and differentiate into tachyzoites. Tachyzoites disseminate throughout the body by infected macrophages and as extracellular parasites in the blood stream (Wong and Remington, 1993). The initial infection and parasitemia are largely asymptomatic in healthy individuals. However, some people may experience “flu-like” symptoms, the most common being fever, lymphadenopathy, and headache (Luft and Remington, 1988).

Due to the intracellular nature of the parasite, protective immunity to *Toxoplasma* is based on a cell-mediated immune response. A cell-mediated immune response relies on $CD4^+$ helper T-cells recognizing antigen presented by macrophages and dendritic cells (Germain and Stefanova, 1999). Infected macrophages and dendritic cells are unable to kill internal parasites, but a small amount of parasite specific protein is presented as antigen on the surface of phagocytic cells. Antigen presentation and IL-12 secretion by infected

macrophages and dendritic cells activates $CD4^+$ helper T-cells resulting in their production and release of interferon gamma (IFN-γ; Hunter and Reichmann, 2001). The released IFN-γ further activates macrophages and dendritic cells, leading to upregulation of inducible nitric oxide synthase (iNOS; Langermans et al., 1992). Macrophages require iNOS to produce reactive nitrogen intermediates which inhibit parasite replication (Sibley et al., 1991). Nonphagocytic cells are also affected by the release of IFN-γ which activates indolamine dioxygenase, starving intracellular parasites of the essential amino acid tryptophan (Pfefferkorn, 1984). The multiple actions of IFN-γ appear to be crucial in controlling acute *Toxoplasma* infections and in driving the infection into a quiescent latent stage.

During the initial infection, the cell-mediated response is accompanied by the production of antibodies. B-cells are stimulated by activated $CD4^+$ helper T-cells and begin producing low-affinity IgM antibodies. As the B-cell response matures and isotype switching occurs, high levels of high affinity anti-toxoplasma IgG antibodies can be detected (Hunter and Reichmann, 2001). However, because *Toxoplasma* is an intracellular pathogen, the utility of antibodies in controlling infection has been questioned. When the parasites lyse an infected cell, the extracellular parasites are susceptible to antibodies. During this intermittent extracellular phase, the antibodies may play a role in killing parasites by activating complement or by opsonizing the parasites for phagocytosis and killing by macrophages (Hammouda et al., 1995).

While the immune system is attempting to thwart the initial infection, a portion of the invading tachyzoites will differentiate into bradyzoites housed within a cyst inside of a host cell (Lyons et al., 2002). Tachyzoites that do not undergo differentiation will be eventually cleared by the immune system. On the other hand, bradyzoites are extremely slow growing and therefore do not provoke an inflammatory response (Derouin, 1992). Thus, an infected individual remains infected for life. Although the bradyzoite cysts can form in any cell in the body, they have a preference for skeletal and cardiac muscle and the central nervous system (Hill and Dubey, 2002). Currently, there is no adequate explanation for

this tissue preference. It has been a long standing presumption that the bradyzoite cysts present in these tissues does not alter their normal physiological function (Yolken et al., 2001). However, there is preliminary data showing a higher rate of *Toxoplasma* infection in the schizophrenic population suggesting that the bradyzoite cysts might alter CNS function and predispose individuals to mental illness (Yolken et al., 2001). At this time, these studies are incomplete as they do not take into account several important factors such as age of infection, severity of infection, and/or family history of mental illness.

Even after infected individuals have cleared the initial infection, antibody titers remain elevated (Ho-Yen, 1992). The relatively high antibody titer is best explained by a small portion of the bradyzoites cysts spontaneously reverting back to tachyzoites, termed recrudescence (Lyons et al., 2002). In response to the reemerging tachyzoites, the immune system is activated through similar mechanisms as the initial infection, only much quicker. Thus, the reemerging tachyzoites are efficiently eradicated before they can cause significant tissue destruction (Weiss and Kim, 2000). It is plausible that the constant priming of the immune system by recrudescing bradyzoites provides protection against reinfection, which is especially important in pregnant women as described in the next section.

C. Congenital infection

Virtually all births suffering complications due to *Toxoplasma* infection are the result of the mother being exposed for the first time during the pregnancy (Guerina et al., 1994). Otherwise healthy, pregnant women who have been previously infected with *Toxoplasma* are able to quickly control reexposure to ingested oocysts or bradyzoites before the parasites can spread to the fetus (Jones et al., 2003). The quick ability to control reinfection may be partly due to the primed immune system by recrudescing bradyzoites. Pregnant women who are infected for the first time during pregnancy need time to mount an effective immune response (Guerina et al., 1994). During this period, the parasite disseminates throughout the body, infecting various tissues including the

placenta. Once the placenta is infected, tachyzoites can cross over and infect the developing fetus (Boyer, 1996). Members of the cell-mediated immune system (macrophages and $CD4^+$ helper T-cells) are restricted by placental barriers and are unable to protect the fetus from infection. However, antibodies can cross placental barriers and may play a role in helping to control fetal infection.

At birth, the classic triad of signs suggestive of congenital toxoplasmosis includes chorioretinitis (infection of the choroid and retina of the eye), hydrocephalus, and intracranial calcifications (Hill and Dubey, 2002). The CNS appears to be the system suffering the most damage, resulting in potential cognitive and visual disabilities. Congenital disease is more severe when infection is acquired in the first trimester versus the second or third (Guerina et al., 1994). Later in gestation, the fetal immune system is better developed and more able to control in the infection. Interestingly, the risk of transmitting the infection to the fetus is higher during the later stages of pregnancy and may be related to placental development and increased blood supply during the later stages of pregnancy making transmission more likely (Guerina et al., 1994). In general, the overall risk of congenital infection from an initial acute *Toxoplasma* infection during pregnancy ranges from approximately 20 to 50 percent (Jones et al., 2003). When acute *Toxoplasma* infection is suspected, a diagnosis is made on the basis of antibody detection. Acute infection is confirmed by the presence of both IgG and IgM antibodies (Jones et al., 2003). If a pregnant woman is suffering from acute infection with *Toxoplasma*, the next step is to determine whether the fetus is infected, which is usually done via PCR testing of amniotic fluid (Jones et al., 2003).

D. Infection in HIV patients

While recrudescing bradyzoites may help prevent reinfection in healthy individuals, they can cause serious problems in immune suppressed individuals (Reiter-Owona et al., 2000). Without the protection of a fully functional immune system, reactivated tachyzoites can cause widespread tissue damage, blindness,

encephalitis, and even death (Black and Boothroyd, 2000). In recent decades, the population of immune suppressed individuals has become increasingly large due to the AIDS epidemic.

Patients suffering from end-stage AIDS clearly demonstrate the close association of T-cell function with resistance to *Toxoplasma*. Unfortunately for these individuals, the HIV virus attacks and destroys $CD4^+$ T-cells (Pattanapanyasat and Thakar, 2005). As mentioned earlier, $CD4^+$ T-cells are essential for controlling *Toxoplasma* infection and reactivation. There is strong correlation with the clinical presentation of *Toxoplasma* infection and $CD4^+$ cell counts. At the time they present with clinical toxoplasmosis, AIDS patients usually have $CD4^+$ T-cell counts that have fallen from 800-12000/mm^3 to less than 100/mm^3 (Hunter and Reichmann, 2001).

Toxoplasmic encephalitis (TE) without any other organ involvement is the most frequent manifestation of toxoplasmosis in patients with AIDS (Mariuz and Steigbigel, 2001). In one study of 86 HIV patients with TE, the most common general signs included fever and headache unresponsive to analgesics (Renold et al., 1992). The diagnosis of TE in HIV patients is usually made by the combination of neurological symptoms and findings on computerized tomographic (CT) scan or magnetic resonance imaging (MRI). The neurological symptoms can include cranial nerve palsies, confusion, lethargy, and ataxia. CT or MRI scans with contrast dye often reveal the presence of intracranial abscesses that can be as large as a tennis ball. If untreated, toxoplasmic encephalitis is uniformly fatal (Luft et al., 1993).

E. Infection in HIV-negative immune suppressed

Besides HIV infection, other individuals with suppressed immune systems are threatened by *Toxoplasma* infection or reactivation. Immune suppression in these individuals is the result of advances in the medical field and includes post operative organ transplant patients undergoing immunosuppressive therapy to prevent organ rejection and patients with neoplastic conditions receiving immunosuppressive chemotherapy (Wong and Remington, 1993).

The biggest concern in organ transplantation is the potential for the donor recipient to become infected via the donated organ. A report by Speirs et al. (1988) stated that 57% of heart and 20% of liver transplant recipients who were *Toxoplasma* negative before transplantation and who received organs from *Toxoplasma* positive donors acquired primary infection (Wreghitt and Joynson, 2001). Because the parasite has a tropism for cardiac tissue, heart and heart-lung recipients are at the greatest risk versus other transplanted tissues (Wreghitt et al., 1989). When a patient becomes infected from a donated organ, it is probably due to bradyzoite cysts present in the donated organ recrudescing back into tachyzoites in the recipient (Hermanns et al., 2001). When an uninfected recipient receives an infected organ, it is referred to as a "*Toxoplasma* mismatch" (Wreghitt and Joynson, 2001). "*Toxoplasma* mismatches" can be avoided by measuring antibody titers to determine if the recipient has had previous exposure to *Toxoplasma*. In cases where "*Toxoplasma* mismatches" cannot be avoided, immediate treatment for toxoplasmosis (page 14) may reduce the severity of acute *Toxoplasma* infection (Wreghitt and Joynson, 2001).

Another concern in organ transplantation is the post operative immune suppression necessary to prevent organ rejection. In the early years of organ transplantation, large amounts of immunosuppressive drugs were used resulting in high risk *Toxoplasma* reactivation (Wreghitt and Joynson, 2001). Newer, less aggressive therapies including low dose cyclosporine A, steroids, and azathioprine have significantly reduced toxoplasmosis seen in organ transplant patients (Wreghitt and Joynson, 2001). However, patients are often placed on prophylaxis therapy with trimethoprim/sulfamethoxazole to prevent toxoplasmosis and other opportunistic diseases (Wreghitt and Joynson, 2001).

Toxoplasmosis can occur in patients with immune suppression caused by certain malignancies, particularly Hodgkin's lymphoma, with extended doses of corticosteroids, or certain chemotherapies. While such infections are usually rare, patients must undergo treatment for toxoplasmosis (page 14) and alternations made in the dosage of chemotherapy that can limit the treatment of the malignancy (Wreghitt and Joynson, 2001).

F. Treatment of toxoplasmosis

The standard treatment for toxoplasmosis is with pyrimethamine (Daraprim) and a sulphonamide, usually sulphadiazine. Both drugs target two different enzymes involved in the folate pathway and therefore work synergistically to inhibit replicating tachyzoites. *Toxoplasma* cannot uptake preformed folic acid and therefore must synthesize its own (Derouin, 2001). The first step in the synthesis is carried out by dihydropteroate synthase (DHPS) which converts *para*-aminobenzoic acid to dihydrofolic acid, which then is converted to tetrahydrofolic acid by dihydrofolic acid reductase (DHFR). Sulphonamides are competitive inhibitors of DHPS and pyrimethamine inhibits DHFR (Tracy and Webster, 1996). However, both of these drugs have potentially serious side effects.

DHPS is not expressed in mammalian cells and thus there are few serious adverse reaction associated with administering sulphonamides (Derouin, 2001). Hypersensitivity reactions are the most common and can vary in severity from mild dermatitis to sever dermatological reactions and anaphylaxis (McCabe, 2001). The development of severe mucocutaneous reactions such as Steven-Johnson syndrome and toxic epidermal necrosis (TEN) is also a concern when administering sulphonamides (Schmidt-Westhausen et al., 1998). TEN is characterized by widespread blisters that involve over 90% of human body. The blisters usually burst and the skin sloughs off causing increased risk of dehydration, hypothermia, and infection secondary to substantial skin loss (Fritsch and Sidoroff, 2000). Furthermore, there is evidence that adverse skin reactions are higher among patients with AIDS than those without AIDS (Kimura et al., 1991). Because TEN is a severe, potentially fatal skin reaction, sulfonamide-containing drugs should be given cautiously to patients with AIDS (Kimura et al., 1991)

Unlike DHPS, DHFR is expressed in mammalian cells. While pyrimethamine possesses a much higher affinity for the *Toxoplasma* DHFR there is still an effect on mammalian cells at therapeutic doses (Derouin, 2001). In particular, the rapidly dividing progenitor blood cells in the bone marrow are

suppressed resulting in decreased platelets (thrombocytopenia), macrophages (granulocytopenia), and red blood cells (megaloblastic anemia; Winstanley et al., 1995). These types of anemia can lead to bleeding problems, further infections, and difficulty oxygenating tissues (cyanosis). The likelihood of bone marrow suppression can be reduced with folic acid (Leucovorin) supplements in conjunction with pyrimethamine, but it does not prevent suppression (Tracy and Webster, 1996). Folic acid supplementation only benefits mammalian cells because parasites cannot import it (Tracy and Webster, 1996).

AIDS patients and other immune suppressed individuals suffering from reactivated toxoplasmosis are usually treated with pyrimethamine and sulphadiazine. Once the acute infection has resolved, these individuals are usually placed on prophylaxis treatment to prevent reactivated tachyzoites from causing further tissue destruction (Nath and Sinai, 2003). Because there is no drug capable of eradicating the encysted bradyzoites the prophylaxis treatment is usually life long (Lane et al., 1994). Unfortunately, prolonged treatment with sulphonamides increases the risk of developing a hypersensitive reaction. If an adverse reaction with sulphadiazine occurs, prophylaxis treatment can be maintained with pyrimethamine and a macrolide antibiotic such as clindamycin (Lane et al., 1994). These compounds block protein synthesis in bacteria but their target in *Toxoplasma* is unknown. It is theorized that the bacterial-like ribosomes present in the apicoplast may be inhibited by these antibiotics (Fichera et al., 1995).

Highly active anti-retroviral therapy (HAART) has significantly reduced the number of AIDS-related cases of toxoplasmosis (Samuel et al., 2002). These drugs slow the destruction of $CD4^+$ cells by the HIV virus. If $CD4^+$ cell counts are maintained above $100/mm^3$, AIDS patients can usually be taken off prophylaxis treatment (Hunter and Reichmann, 2001). However, the HIV virus mutates rapidly and if the virus develops resistance to current antiretroviral drugs, the incidence of AIDS-associated toxoplasmosis will likely rise again (Sacktor, 2002).

If a fetus is suspected of being infected with toxoplasmosis, treatment must begin immediately. Pyrimethamine is not used in pregnant women because

at high doses it has been shown to be teratogenic in animals (McCabe, 2001). The preferred treatment is with the macrolide antibiotic spiramycin (Jones et al., 2003). The exact mechanism of action of this antibiotic in *Toxoplasma* is unknown, but it is preferred because high concentrations in the placenta are achieved and it is extremely safe (Nath and Sinai, 2003). Spiramycin has no known adverse effects on the fetus (McCabe, 2001). Despite the relative safety of spiramycin, it is not readily available in the United States. The drug can only be obtained through direct purchase from the manufacturer which generates logistical barriers to fast and effective treatment of pregnant women infected with *Toxoplasma* (Many and Koren, 2006). If the *Toxoplasma* infection is thought to have occurred before the 16th week of pregnancy, or if the fetus shows evidence of hydrocephalus on ultrasound, abortion should be considered because of the high likelihood of severe fetal central nervous system damage (McCabe, 2001).

G. Bradyzoite differentiation and transcriptional control

The bradyzoite cyst is critical to the pathogenesis of *Toxoplasma*, responsible for the permanent latent stage of infection. Currently, there are no therapies capable of targeting and eradicating this form of the parasites. The differentiation process from tachyzoite to bradyzoite and back to tachyzoite is poorly understood. In order for better therapeutics to be designed to target bradyzoites conversion, a better understanding of the differentiation process must be established.

Studying bradyzoite differentiation was initially limited by the ability to generate *in vitro* bradyzoites. In 1986, Jones et al. discovered that addition of IFN-γ to *in vitro* cultures of parasites results in the formation of bradyzoite cysts. As previously mentioned, it is theorized that IFN-γ starves the parasites of tryptophan. Other mediators of *in vitro* differentiation that have since been discovered include alkaline treatment (pH 8.1), heat shock (42°C), and chemical stress (2μM sodium arsenite) (Soête, 1994). All of these treatments share a common feature: they are all stressors slowing down the growth of the host cell

and parasite (Soête, 1994). Thus, it would seem that bradyzoite differentiation is a stress mediated event.

In addition to the development of *in vitro* differentiation protocols, bradyzoite-specific antiserums have been developed. Western blot and ELISA studies using these antiserums have confirmed that tachyzoites and bradyzoites have unique as well as shared antigens (Weiss and Kim, 2000). The first bradyzoite specific antigen discovered was BAG1 (bradyzoite antigen; Bohne et al., 1995). Since then there have been 6 other bradyzoite-specific antigens and two stage-specific enzymatic paralogues (Singh et al., 2002). The two enzymatic paralogues are lactate dehydrogenase 2 (LDH2) and enolase 1 (ENO1). The respective counterparts to these paralogues, LDH1 and ENO2, are expressed only in tachyzoites. The bradyzoite-specific LDH2 and ENO1 have different enzyme kinetics than the tachyzoite-specific counterparts, and it is believed that these differences help facilitate the slow metabolic rate of bradyzoites (Yang and Parmley, 1997; Dzierszinski et al., 2001, respectively). Additional studies are underway attempting to characterize the promoters of these and other bradyzoite-specific genes. Bradyzoite-specific promoters will aid in the identification of additional bradyzoite-specific genes and transcription factors that regulate these promoters.

There are currently no known transcription factors in *Toxoplasma* that control or are involved in the bradyzoite differentiation process. Additionally, the signaling pathways involved in triggering the differentiation process are unknown. As bradyzoite differentiation is unique to parasites, there is potential that transcription factors and/or signaling proteins involved in mediating differentiation are parasite specific and would make excellent drug targets. Drugs capable of targeting proteins involved in controlling differentiation could halt the differentiation process and thus prevent the latent stage of infection. However, searches of apicomplexan genome databases suggests there are a relatively small number of transcription factors present in *Toxoplasma* and other apicomplexans (Saksouk et al., 2005; Templeton et al., 2004).

In addition to transcription factors, other classes of proteins known to be involved in regulating gene expression should be investigated in *Toxoplasma* as they may participate in the differentiation process. Chromatin remodeling proteins in other eukaryotic organisms play a profound in role in gene expression, some of which are instrumental in mediating stress responses (Huisinga and Pugh, 2004). Analogous chromatin remodeling homologues exist in *Toxoplasma* (Sullivan and Hakimi, 2006) and may also be involved in stress responses which, as mentioned above, initiate bradyzoite conversion. Currently, very little is known about chromatin remodeling proteins in apicomplexans. Identification and characterization of chromatin remodeling proteins in *Toxoplasma* will facilitate our understanding of gene regulation in this parasite and other apicomplexans, and may shed new light on the bradyzoite differentiation process.

II. Transcription Control and Chromatin Remodeling Proteins

A. Chromatin structure and nucleosomes

Chromatin is defined as a complex of DNA, histones, and non-histone proteins from which eukaryotic chromosomes are formed (Alberts et al., 2002). The fundamental unit of chromatin is the nucleosome which is composed of a core of eight histones (two molecules of H2A, H2B, H3 and H4; Figure 3). Around this protein core, 146 base pairs of DNA is wrapped much like thread around a spool (Loid, 2001). The core histones are small proteins with a basic charge that share the same basic structure: a globular C-terminal domain critical to nucleosome formation and a flexible N-terminal tail that protrudes from the nucleosome core (Santos-Rosa and Caldas, 2005). The flexible N-terminal tails are enriched with basic residues which are theorized to interact with the negatively charged backbone of DNA (Allfrey et al., 1963).

In-between each nucleosome is a short stretch of DNA termed linker DNA. The nucleosomes and their linker DNA form chains like a long string of beads,

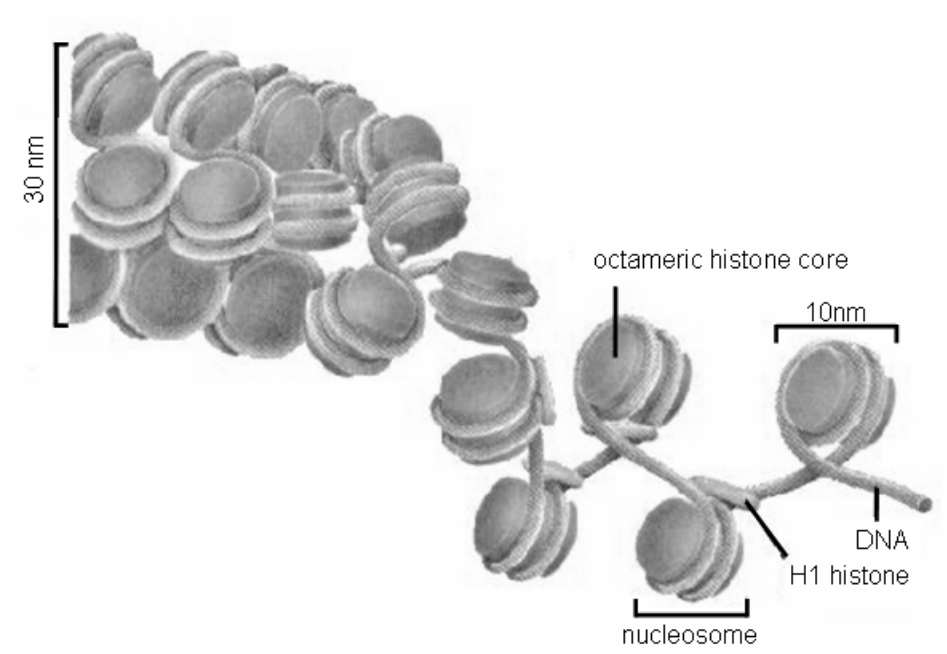


Figure 3: Nucleosome structure

DNA is wrapped around the octameric histone core forming the nucleosome. Each nucleosome is tethered together by linker DNA which associates with one H1 protein forming the 10nm “bead on string” fiber. Through interactions between the H1 histones the fiber coils into a solenoid structure with a diameter of 30nm. (adapted from Grunstein, 1992; reprinted with permission from author and *Scientific American*)

known as the 10-nm chromatin fiber (Loid, 2001). These fibers are then packed into a higher order structure, facilitated by another histone protein, H1. Histone H1 binds to the linker DNA and interacts with neighboring H1 proteins to form a spiral, with 6-8 nucleosomes per turn, generating a solenoid or 30nm chromatin fiber (Figure 3; Grunstein, 1992). There are higher orders of structure that eventually lead to the formation of the metaphase chromosomes (Grunstein, 1992). However, these higher order structures have not been completely resolved.

B. Histone acetylation and transcriptional control

In the early 1960s, two papers by Allfrey and colleagues were the first to demonstrate the impact histones had on gene expression. In 1963, Allfrey and colleagues confirmed that DNA-histone complexes (nucleosomes) fail to act as primers for RNA synthesis (Allfrey et al., 1963). The high basic charge of histones was theorized to interact strongly with DNA preventing DNA from being transcribed by a polymerase. Allfrey also developed a trypsin treatment that selectively depleted nuclear histones from cells. Treating calf thymus cells with his trypsin protocol resulted in a large increase in RNA synthesis. The increase in transcription was caused by increases in existing message *and* by the transcription of new mRNA not present before histone depletion (Allfrey et al., 1963). Allfrey deduced that the new transcripts were being generated because the digested histones were no longer able to repress transcription.

Based on observations made by Phillips in 1963 of the substantial presence of acetylated histones in the nucleus, Allfrey examined the effect of acetylated histones on RNA synthesis (Allfrey et al., 1964). Unlike unmodified histones, DNA-acetylated histone complexes were not inhibitory to RNA synthesis (Allfrey et al., 1964). It was argued that acetylation attenuated the highly basic charge of histone proteins preventing them from strongly interacting with DNA. However, the failure to inhibit RNA synthesis was not caused by acetylated histones completely dissociating from DNA. Allfrey demonstrated that

acetylated histones were still able to protect DNA from thermal denaturation, suggesting an interaction is still present (Allfrey et al., 1964).

The studies conducted by Allfrey and colleagues clearly demonstrated that histones possess the ability to control gene expression. Furthermore, unmodified histones are repressive to RNA synthesis and acetylated histones are correlated with active transcription. Although acetylated histones still interact with DNA, they generate a more favorable environment for gene transcription. Despite this interesting discovery, the proteins responsible for acetylation of nuclear histones would not be elucidated for several decades.

C. Histone acetyltransferases and GCN5

Proteins capable of acetylating histones are termed histone acetyl transferases or HATs. HATs covalently modify histones by transferring an acetyl group from acetyl-coenzyme A to the ε-amino group of specific lysine side chains within the N-terminal tails of histones (Sterner and Berger, 2000).

GCN5 (General Control of Non-derepressed 5) was the first discovered protein capable of acetylating nuclear histones. GCN5 was initially identified by Penn et al. in a yeast mutation screen identifying genes involved in amino acid biosynthesis (Penn et al., 1983). Mutations were induced by UV-light irradiation and clones unable to grow in media deficient in preformed amino acids were considered to have disruptions in the regulation of their amino acid biosynthesis pathways. The screen lead to the discovery of 5 genes involved in the derepression (induction) and expression of proteins involved in the synthesis of histidine, arginine, tryptophan, and methionine (Penn et al., 1983). The genes were named *GCN1* through *GCN5*.

Additional work by Hope and Struhl lead to the discovery that GCN4 was the main transcription factor of the five *GCN* genes responsible for upregulation of amino acid biosynthesis genes under starvation conditions (Hope and Struhl, 1985; Hope and Struhl, 1987). Increasing the amount of GCN4 protein rescues cells suffering from mutations in *GCN1*, *GCN2*, and *GCN3*, suggesting that these genes are involved in the upregulation and expression of GCN4 (Hinnebusch,

1985). However, mutations in *GCN5* are not rescued by alterations in levels of GCN4 (Georgakopoulos and Thireos, 1992). Using a *GCN4* promoter driven reporter system, it was determined that GCN5 is required by GCN4 to fully activate its target genes (Georgakopoulos and Thireos, 1992). Thus, GCN5 is involved in potentiating GCN4 mediated transcription. However, the mechanism by which GCN5 functioned remained unknown. Georgakopoulos and Thireos cloned the *GCN5* gene, which encoded a protein of 439 amino acids. They noted the presence of a putative domain in the C-terminus, between residues 348 and 422, which shared similarity with a class of transcriptional factors related to DNA helicases. Although GCN5 does not share any similarity to DNA helicases, this observation lead Georgakopoulos and Thireos to postulate that GCN5 may be involved in regulating chromatin structure (Georgakopoulos and Thireos, 1992).

While the true function of GCN5 was attempting to be ascertained, others were working on trying to identify proteins responsible for acetylating histones. Although many attempts had been made trying to purify HATs from eukaryotic cells, all were unsuccessful due to low level of endogenous expression. In 1995, success in identifying a HAT was finally achieved by Brownell and Allis using an "in gel" assay system (Brownell and Allis, 1995). Core histones were incorporated into standard polyacrylamide gels during acrylamide polymerization. Nuclear extracts from *Tetrahymena* (a ciliated freshwater protozoan) were resolved on the histone-SDS gels under native conditions. Following electrophoresis, the gels were incubated in a buffer containing tritiated acetyl-CoA. Upon exposure to film, a single band at 55 kilodaltons was detected suggesting the presence of a single HAT of this size termed p55 (Brownell and Allis, 1995). Peptide sequencing of p55 was performed following several fractionation techniques on large amounts of *Tetrahymena* nuclear extracts. Using the peptide sequence data, they were able to clone the p55 gene (Brownell et al., 1996). The full-length p55 gene encoded a protein of 421 amino acids and possessed unequivocal homology to yeast GCN5 (Brownell et al., 1996). This observation established a clear biochemical function for yeast GCN5

as a HAT and provided a mechanistic link between histone acetylation and gene activation (Brownell et al., 1996).

Since its initial discovery, GCN5 proteins have been identified and characterized in many eukaryotic organisms including humans, mouse, *Drosophila*, and *Arabidopsis* (Candau et al., 1996; Driessen et al., 1997; Smith et al., 1998a; Stockinger et al., 2001; respectively). Figure 4 contains diagrams depicting the different GCN5 homologues in these organisms and the conserved domains.

All the GCN5 homologues known to date possess three conserved domains. The catalytic domain is responsible for acetyltransferase activity. Initial enzymatic studies demonstrated that yeast GCN5 selectively acetylates histone H3 in nucleosomal substrates *in vivo* (Ruiz-Garcia et al., 1997). Additional studies with recombinant yeast GCN5 indicated that acetylation occurred on lysines in the N-termini of histones H3 *and* H4 (Kuo et al., 1996). Although the N-termini of H3 and H4 contain several conserved lysine residues that can be acetylated, recombinant yeast GCN5 preferentially acetylates lysine 14 of H3 and to a lesser degree, lysine 8 and lysine 16 of H4 *in vitro* (Kuo et al., 1996).

The two other domains, the Ada2 binding domain and the bromodomain, are downstream of the catalytic domain. The Ada2 binding domain is the region where the Ada2 protein binds to GCN5. Ada2 belongs to the Ada (alteration/deficiency in activation) group of proteins that complex with GCN5 *in vivo* (Berger et al., 1992). Proteins that interact with GCN5 including Ada proteins are discussed further in a Section II-E (page 29).

The bromodomain is a conserved motif present in many transcription and chromatin regulators and is the same domain first noted in the C-terminus of yeast GCN5 by Georgakopoulos and Thireos in 1992. Further analysis of the bromodomain has revealed it binds to specific acetylated lysines in histones H4 and H3 (Dhalluin et al., 1999). Thus, the bromodomain appears to be involved in recognizing modified histones and may serve to direct the function of GCN5 to areas of chromatin containing modified histones (Marmorstein, 2001). The

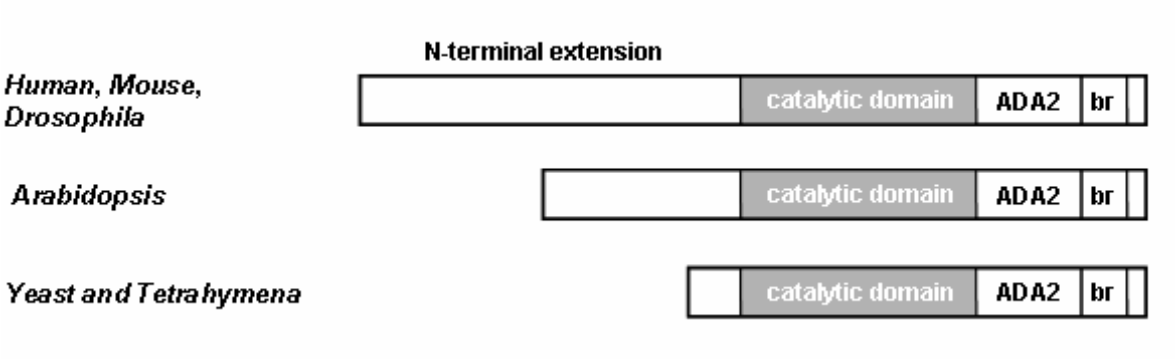


Figure 4: GCN5 homologues and domains

Schematic diagrams comparing homologues of GCN5. All GCN5 proteins contain three conserved domains: the catalytic domain responsible for acetylating histones, the Ada2 binding domain (ADA2) where the Ada2 protein interacts with GCN5, and the bromodomain (br) which has been shown to interact with acetylated histones (Dhalluin et al., 1999). Upstream of the catalytic domain is referred to as the N-terminal extension. The N-terminus is highly conserved between metazoans but not between metazoans, plants, yeast or protozoa.

potential link between the bromodomain and the enzymatic activity of GCN5 will be addressed in Section II-G (page 36).

The sequence upstream of the HAT catalytic domain varies in size and is sometimes referred to as an N-terminal extension. Early branching eukaryotes generally possess a short sequence of less than 100 residues upstream of the HAT catalytic domain (Brownell et al., 1996). The N-terminal extension in plant homologues are of moderate length, between 150 and 250 amino acids (Bhat et al., 2003). The N-terminal extensions in plants and lower eukaryotes have little homology to each other or other HAT proteins. The GCN5 homologues present in metazoans contain rather long N-terminal extensions, approximately 500 residues in length. The relatively long N-terminal extensions present in mouse and human GCN5 have been shown to be important in substrate recognition. When the N-terminal extension is removed from human and mouse GCN5, they are no longer able to recognize and acetylate nucleosomal histones *in vitro* (Xu et al., 1998). The ability to acetylate free histones in solution is unaffected by removal of the N-terminal extension. Another role for the N-terminal extension might be nuclear localization. The maize homologue of GCN5 (ZmGCN5) has an N-terminal extension of about 160 residues and appears to be involved in targeting the HAT to the nucleus (Bhat et al., 2003). Without the N-terminal extension, ZmGCN5 remains cytoplasmic. Furthermore, when the first 175 amino acids of ZmGCN5 are fused to GFP, the fusion protein accumulates in the nucleus (Bhat et al., 2003). The involvement of the N-terminal extension in the nuclear localization of other GCN5 homologues has not been investigated.

In addition to amino acid starvation, GCN5 is also important in the activating response elements during the stress of phosphate starvation. The GCN5 knock-out (Δgcn5) generated in yeast has shown GCN5 to be important for remodeling chromatin at the inducible *PHO5* and *PHO8* promoters under phosphate starvation (Gregory et al., 1998; Gregory et al., 1999). Unlike wild-type yeast, the *PHO5* and *PHO8* promoters in Δgcn5 are resistant to DNase digestion, suggesting the promoters are condensed and not properly opened during starvation conditions (Gregory et al., 1998; Gregory et al., 1999).

In addition to stress remediation pathways, several studies on GCN5 in various organisms have implicated this HAT in more essential and global processes. In yeast, Δgcn5 cells accumulate in G2/M cell cycle checkpoint indicating that GCN5 is important for normal cell-cycle progression (Zhang et al., 1998). GCN5-null mouse embryos die 10.5 days post conception. The embryos failed to form mesodermal-derived structures (precursors to muscles and bones) and rarely formed a recognizable neural tube implicating a critical role for GCN5 in mammalian development (Xu et al., 2000). In *Arabidopsis*, GCN5-null seedlings fail to form a proper root structure and have shorter stature, suggesting that GCN5 is critical in plant development and growth (Vlachonasios et al., 2003).

Another HAT with striking homology to GCN5 termed PCAF (p300/CBP Associating Factor) has been identified and characterized. PCAF was originally identified from a human cDNA database on the basis of its homology to GCN5 (Sterner and Berger, 2000). There is an 86% homology between the C-terminal portion of PCAF (containing the HAT domain, Ada2 binding domain, and bromodomain) and the C-terminal end of GCN5 (Yang et al.,1996). To date, PCAF has only been found in vertebrates and has overlapping and separate roles from GCN5 (Sterner and Berger, 2000).

PCAF also possess an N-terminal extension that is 66% homologous to the N-terminal extension found in human and mouse GCN5 proteins (Xu et al., 1998). Just as in GCN5, the N-terminal extension in mouse and human PCAF is required for the recognition of nucleosomal substrates *in vitro* (Xu et al., 1998). PCAF possess the ability to acetylate its own N-terminus which augments its enzymatic activity (Herrera et al., 1997). The acetylation of the N-terminus occurs by transacetylation; acetylation by second PCAF protein (Santos-Rosa et al., 2003). There are five lysines in the PCAF N-terminus that are acetylated; lysines 416, 428, 430, 441, and 442. Human GCN5 does not appear to acetylate its N-terminus (Herrera et al., 1997). Only two of the five lysines that are acetylated in human PCAF are conserved in human GCN5 (lysines 428 and 441). Lysines 428, 430, 441, and 442 in human PCAF also appear to be involved in nuclear localization (Santos-Rosa et al., 2003). When these four

lysines are mutated to arginine, PCAF is no longer targeted to the nucleus (Santos-Rosa et al., 2003). These results suggest that transacetylation of PCAF may play a role in targeting it to the nucleus.

As its name implies, PCAF interacts *in vivo* with another transcriptional regulator p300/CBP (CREB binding protein), which is another HAT. The histone acetylation properties of p300/CBP will be addressed in the next section (page 28). Interestingly, PCAF binds to the same site on p300/CBP as does the oncogenic viral protein E1A (Yang et al., 1996). In fact, the mitogenic actions of E1A occur by disrupting the interaction between PCAF and p300/CBP, suggesting a role for PCAF in cell cycle regulation through its interaction with p300/CBP (Yang et al., 1996).

Recombinant PCAF, like GCN5, acetylates histones, preferring lysine 14 of histone H3, and more weakly lysine 8 of histone H4 (Schiltz et al., 1999). Multiple studies have shown PCAF to be important in several cellular processes including myogenesis and nuclear receptor-mediated activation (Sterner and Berger, 2000).

In addition to acetylating histones, PCAF also acetylates several non-histone proteins involved in transcriptional regulation, a processes referred to factor acetyltransferase (FAT) activity. The non-histone factors targeted by PCAF acetylation include the cell cycle regulator p53 (Liu et al., 1999), the general transcription factors $TF_{II}E$ and $TF_{II}F$ (Imhof et al., 1997), and the chromatin associating protein HMG17 (Herrera et al., 1999). The acetylation of p53 by PCAF increases the ability of p53 to bind DNA, and the acetylation of p53 increases in response to DNA-damaging agents (Liu et al., 1999). Thus, PCAF appears to be involved in stress remediation pathways through its FAT activity.

Unlike GCN5, mice lacking PCAF are developmentally normal without a distinct phenotype (Yamauchi et al., 2000). In PCAF null homozygous mice, protein levels of GCN5 are drastically elevated, suggesting that GCN5 can functionally compensate for the loss of PCAF (Yamauchi et al., 2000). However, PCAF is unable to compensate for the loss GCN5, as mice lacking GCN5 are not

viable. Therefore, PCAF and GCN5 may have overlapping functions but are not functionally equivalent, especially in embryogenesis.

D. Other HATs and HAT families

In addition to GCN5 and PCAF, there are two other groups of HATs: the p300/CBP family and the MYST family. The p300/CBP family is named after the two proteins p300 and CBP [CREB (cAMP responsive element binding protein) binding protein]. Currently, these are the only two proteins in this family. Due to the considerable structural and functional homology between the two proteins, they are often referred to as a single entity, p300/CBP (Roth et al., 2000). Several studies have also determined that the two proteins have many overlapping functions (Goodman and Smolik, 2000). The two proteins are found only in multi-cellular animals (from worms up to humans) and are highly conserved across these organisms (Marmorstein, 2001). CBP and p300 have been implicated in many cellular processes including cell cycle control, differentiation, and apoptosis. Mutations in CBP are associated with various types of cancer, further demonstrating their regulatory importance in cell cycle (Timmermann et al., 2001). As mentioned in the previous section, p300/CBP binds to the HAT PCAF, an interaction that is interrupted by the adenoviral oncoprotein E1A (page 27).

p300/CBP is a large protein of about 300kD and is capable of acetylating all four core histones. *In vitro* enzymatic assays revealed it acetylates all lysines on the four core histones known to be acetylated *in vivo.* However, it shows the strongest activity on lysines 14 and 18 of histone H3 and lysines 5 and 8 of histone H4 (Schiltz et al., 1999). The actions of p300/CBP are not always linked to histone acetylation. p300/CBP can also function as a FAT acetylating non-histone proteins such as p53. Acetylation of p53 by p300 increases the binding affinity of p53 to DNA (Gu and Roeder, 1997). The global acetylation capabilities of p300/CBP versus other HATs is congruent with its multiple regulatory functions.

The MYST family of HATs is named for its founding members: MOZ (monocytic leukemia zinc finger protein; Borrow et al., 1996), Ybf2/Sas3 (something about silencing 3; Reifsnyder et al., 1996), Sas2 (something about silencing 2; Reifsnyder et al., 1996), Tip60 (Tat-interactive protein, 60kD; Kamine et al., 1996). The MYST members share a highly conserved catalytic domain and although similar to the GCN5 catalytic domain, it is structurally divergent and may explain the differing acetylation patterns between GCN5 and MYST HATs (Utley and Cote, 2003). MYST members do not share the exact same acetylation specificities, but in general, they preferentially acetylate histone H4. Some MYST HATs can also acetylate H3 and H2A to a lesser degree (Utley and Cote, 2003). The divergence in acetylation patterns may explain the differing regulatory functions in which the various MYST members are involved.

The first MYST family member to be linked to histone acetylation was the *Drosophila* protein MOF (male absent on first; Hilfiker et al., 1997). MOF is involved in dosage compensation of the "X" sex chromosome in male flies. Dosage compensation refers to a regulatory mechanism that ensures the equalization of X-linked gene products in males and females (Lucchesi et al., 2005). In addition to MOF, other MYST HATs have been implicated in other regulatory processes. *Saccharomyces cerevisiae* Esa1 (essential SAS2-related acetyltransferase) is a MYST member shown to be important for cell cycle progression. Deletion of Esa1 resulted in strains that grew 40 times slower and eventually died (Smith et al., 1998b). MOZ was identified to cause acute myeloid leukemia when fused to CBP due to a chromosomal translocation (Timmerman et al., 2001). It is believed that the resulting fusion protein is unable to be properly regulated and disrupts the cell cycle through aberrant histone acetylation (Borrow et al., 1996).

E. HAT complexes

Although histone acetylases have been studied extensively *in vitro*, these proteins do not function as single entities *in vivo* but as members of large multi-subunit complexes (Roth et al., 2001). The two best characterized HAT

complexes are SAGA (Spt-Ada-GCN5 acetyltransferase) and ADA (Ada containing complex). Both complexes were initially discovered in yeast but homologous complexes have also been isolated in other eukaryotic organisms such as humans, mouse, and *Drosophila* (Sterner and Berger, 2000). The catalytic core of both SAGA and ADA is GCN5 (Grant et al., 1997). In metazoans, it is theorized that PCAF can also serve as the catalytic core in GCN5 containing complexes (Roth et al., 2001).

The ADA complex was discovered first and is much smaller than SAGA, approximately 800kD (Eberharter et al., 1999). The complex is comprised of several Ada (alteration/deficiency in activation) proteins (Berger et al., 1992). The Ada proteins were discovered using a mutational screen looking for yeast resistant to the growth suppressive effects of VP16 (herpes simplex viral protein 16). VP16 causes growth arrest by sequestering multiple general transcription factors, preventing them from properly regulating gene expression (Berger et al., 1992). The ability of VP16 to bind multiple transcriptional regulators is facilitated by a group of endogenous adapter proteins. Using a chemical mutagenesis screen, 5 adapter proteins were isolated and determined to be involved in the sequestering effects of VP16 (Berger et al., 1992). When any one of these five proteins is mutated, yeast cells become resistant to the growth suppressive effect of the VP16 (Berger et al., 1992). These proteins were termed Ada1-5. Upon further analysis, it was discovered that the Ada4 isolated in the mutagenesis screen was identical to GCN5, providing a link between GCN5 and the Ada proteins (Berger et al., 1992).

Only Ada2 and Ada3 are found in the ADA complex (Pina et al., 1993). In fact, there is a conserved domain in the C-terminus of all known GCN5 proteins where Ada2 interacts with GCN5, termed the Ada2 domain. Ada3 interacts only with Ada2 within the complex and not with GCN5 (Pina et al., 1993). In addition to the two Ada proteins, the complex also contains the protein Ahc1 (Ada HAT complex component 1; Eberharter et al., 1999). Although the Ada proteins are present in other HAT complexes, the Ahc1 protein is not, indicating that the ADA complex is a unique GCN5-containing acetylation complex (Eberharter et al.,

1999). *In vitro* acetylation assays using purified yeast ADA complex revealed it can acetylate histone H3 at lysines 14 and 18 (Grant et al., 1999). When the Ahc1 protein is knocked out, the ADA complex can no longer be purified from yeast, indicating the complex is not present. The Ahc1 knockout does not have a readily observable phenotype indicating the ADA complex is not essential in yeast (Eberharter et al., 1999). A physiological function of the ADA complex remains to be elucidated. There is difficulty teasing apart its potential function from the much larger SAGA complex.

SAGA is a 1.8 MDa HAT complex and 15 different subunits have been characterized, but it is expected that at least several more remain to be identified (Figure 5; Sterner and Berger, 2000). In addition to GCN5, the complex contains The four remaining Ada proteins (Ada1-3 and Ada5), several Spt proteins (suppression of Ty insertions; Spt 1, Spt 7, Spt 8, and Spt 20), a subset of TAF_{II}s [TBP (TATA binding protein) associated factors], and Tra1 (transplantability associated gene 1) (Sterner and Berger, 2001).

The Spt proteins were discovered in a mutagenesis screen isolating proteins capable of suppressing the effects of Ty (translocatable yeast element) insertions (Silverman and Fink, 1984). Knockout studies targeting specific Spt family members have revealed these proteins function as important general transcription regulators. Strains harboring *Spt3*, *Spt7*, *Spt8*, or *Spt15* mutations showed reduced viability and had multiple transcriptional defects (Marcus et al., 1996). Upon further analysis, it was discovered that many of the Spt proteins were TBP associating factors (Dudley et al., 1999). In fact, Spt15 is identical to TBP, TATA binding protein (Schroeder and Weil, 1998). Interestingly, Ada5 is identical to Spt20 and is the only Ada protein to date to display a Spt phenotype providing a link between Spt and Ada proteins (Marcus et al., 1996).

To date, several TAF_{II} proteins have been isolated from SAGA including TAF_{II}20/17, TAF_{II}25/23, TAF_{II}60, TAF_{II}68/61, and TAF_{II}90 (Eberharter et al., 1999). Not only are those proteins integral components of the SAGA complex, but the TAF_{II}s found in SAGA resemble histones H3 (TAF_{II}20/17), H4

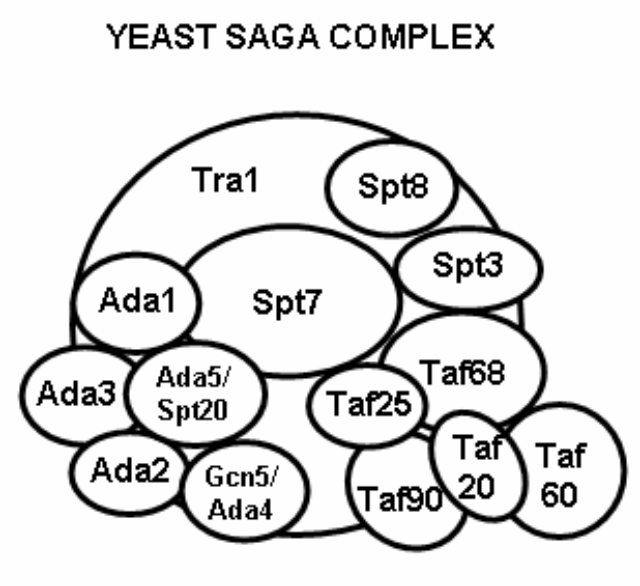


Figure 5: Yeast SAGA Complex

Above is a schematic diagram of known subunits of the yeast SAGA complex. The yeast SAGA complex contains Gcn5 as its HAT catalytic subunit. SAGA has been shown to interact with acidic activation domains mediated by its adaptor components (Ada2, Ada3, Ada5, and Gcn5/Ada4). Another subset of SAGA proteins (Ada1, Spt7, and Spt20/Ada5) are required for structural integrity. The Spt3 and Spt8 subunits have been implicated in interaction with TBP. The TAF_{II} (Taf) group of subunits share structural homology to the four core histones. (adapted from Sterner and Berger, 2000; reprinted with permission from *Microbiology and Molecular Biology Reviews*)

($TAF_{II}25/23$), and H2B ($TAF_{II}68/61$) (Roth et al., 2001). The significance of the homology to histones is unknown. Further analysis of $TAF_{II}68$ revealed it was required for both SAGA-dependent nucleosomal HAT activity and transcriptional activation from chromatin templates *in vitro* (Grant et al., 1998). These results illustrate a role for certain TAF_{II} proteins in the regulation of gene expression through their association with the SAGA complex (Grant et al., 1998).

TRA1 (transplantability associated gene 1) is an essential yeast gene which encodes a 3744-amino acid protein that belongs to a group of proteins possessing carboxy-terminal regions related to phosphatidylinositol 3-kinases (Saleh et al., 1998). Despite possessing a kinase-like domain, no kinase activity has ever been elicited from Tra1 or its homologues (Saleh et al., 1998). The human homologue of Tra1, TRRAP (transformation/transcription domain–associated protein), has been shown to be essential for c-Myc and E2F-mediated oncogenic transformation in human cells (McMahon et al., 1998). The identification of TRRAP as an essential cofactor for oncogenic transcription factor pathways implies that the recruitment of TRRAP by DNA-binding activators directly regulates gene expression (Grant et al., 1998). Given the relatively large size of TRRAP, it is likely a central molecule in the formation and stability of the SAGA complex (Saleh et al., 1998).

A relatively new collection of proteins that interact with the SAGA complex are SGF (SAGA associated factor) proteins (Powell et al., 2004). They do not possess a common element except that are believed to function as part of the SAGA complex. SGF11 (11-kD SAGA associated factor) was the first SGF protein to be identified and it does not appear to play a role in SAGA-mediated histone acetylation (Powell et al., 2004). SAGA purified from an SGF11 deletion strain has reduced amounts of the deubiquitylation protein Ubp8p (Powell et al., 2004). Histone H2B ubiquitylation and Ubp8-mediated deubiquitylation are both required for transcriptional activation (Ingvarsdottir et al., 2006). Therefore, it has been suggested that Ubp8 and SGF11 represent a new function for the SAGA complex that involves gene regulation through H2B deubiquitylation (Ingvarsdottir et al., 2006). Another SGF protein recently identified is SGF73, a protein with

unknown molecular function (Titz et al., 2006). The human homologue of SGF73 is ataxin-7 which is defective in the neurodegenerative disorder spinocerebellar ataxia type 7 (Helmlinger et al., 2004). Ataxin-7 is a subunit of the human SAGA-like complex, TFTC (TATA-binding protein-free TAF-containing complex; Helmlinger et al., 2004). However, the role of SGF73 or ataxin-7 within their respective complexes is unknown.

Unlike the ADA complex, loss of the SAGA complex results in severe phenotypic changes demonstrating that the SAGA complex is an essential gene regulator (Roth et al., 2001). The SAGA complex has a more expanded acetylation pattern than the ADA complex. *In vitro* HAT assays using purified complexes, revealed both SAGA and ADA can acetylate at lysines 14 and 18 on histone H3 (Grant et al., 1999). However, SAGA can also acetylate histone H3 at lysines 9 and 23, which ADA cannot (Grant et al., 1999).

Given the diversity of the subunits comprising the SAGA complex, it is highly likely that this complex is designed to act as a co-activator. It bridges basal transcription factors such TBP and TAF_{II}s with transcriptional activators such as GCN5 and the Ada proteins (Roth et al., 2001). The transcriptional activators are designed to augment basal transcription factors and help facilitate regulation of gene expression. With regard to GCN5, it is theorized that HAT complexes guide the HAT activity to certain promoters during cellular events (Sterner and Berger, 2000). HAT complexes provide a connection to histone acetylation from signal transduction cascades initiated by a cellular stimulus. For example, the SAGA complex has been shown to be required for the proper remodeling and activation of the inducible yeast phosphate gene *PHO8* (Gregory et al., 1999). Without the SAGA complex, the *PHO8* promoter is not properly opened and there is minimal transcription even at maximal induction. Members of the SAGA complex interface with stress remediation pathways activated during phosphate starvation resulting in the activation of the *PHO8* gene (Gregory et al., 1999). How SAGA is regulated and how it interacts with various signaling events is still largely unknown.

There are complexes containing other HATs besides GCN5. PCAF also functions in a complex that is very similar to SAGA. In fact, many of the same proteins are found in both complexes including Ada2, Ada3, Spt3, and Tra1 and also some of the same TAF_{II} proteins (Kotani et al., 1998). As the deletion of PCAF does not cause a different phenotype from wild-type, it is plausible that SAGA can compensate for the loss of PCAF-containing complexes (Yamauchi et al., 2000). However, studies to fully answer this question have not been conducted.

There are three known MYST containing complexes: NuA3 (nucleosome acetyltransferase of histone H3; Utley and Cote, 2003), NuA4 (nucleosome acetyltransferase of histone H4; Utley and Cote, 2003), and MSL (male-specific lethal; Hilfiker at el., 1994). The NuA3 and NuA4 are yeast complexes and MSL was isolated from *Drosophila* (Utley and Cote, 2003). Currently, the MYST complexes are not as well defined as SAGA or ADA. However, there are some similar proteins found between the NuA3, NuA4 and SAGA complexes. All three complexes contain Spt and TAF_{II} proteins and NuA4 contains Tra1 (Sterner and Berger, 2000). NuA4 acetylates histones H4 and H2A, and NuA3 acetylates predominantly histone H3 (Eberharter et al., 1998). The MSL complex contains MOF as its catalytic core and is integral to the dosage compensation of the X-chromosome in male fruit flies (Taipale and Akhtar, 2005). The MSL complex does not appear to possess subunits found in other HAT complexes, but rather contains proteins specific to interacting with the X-chromosome (Taipale and Akhtar, 2005).

F. Histone deacetylases

Histone acetylation is a reversible modification. In order to reestablish the nucleosome structure, histone deacetylases (HDACs) are recruited to deacetylate the histones (Kurdistani and Grunstein, 2003). HDACs are divided into three main classes. The enzymatic domains between class I and II HDACs are highly conserved (Ekwall, 2005). The class III enzymes differ in that they require nicotinamide adenine dinucleotide (NAD; Blander and Guarente, 2004).

Just like HATs, HDACs also function as subunits of large complexes, often referred to as repressor complexes (de Ruijter et al., 2003).

In addition to deacetylating histones, HDACs can also deacetylate transcription factors. The mammalian histone deacetylase, HDAC-1 is capable of deacetylating p53 acetylation by p300/CBP and thus down regulating its activity (Juan et al., 2000).

Ultimately, in order for a cell to achieve a proper functioning state, there must be an appropriate balance between HDACs and HATs (Kurdistani and Grunstein, 2003). When this balance is lost, it often leads to cancer, developmental disorders, or other detrimental cellular states resulting in cell death (Timmermann et al., 2001). As mention above, dysregulation of CBP through a translocation has been implicated in leukemia and the disruption of the interaction between PCAF and p300/CBP by E1A results in oncogenesis. Dysfunction of all three classes of HDACs has been implicated in loss of cell cycle control and tumorgenesis (Ekwall, 2005). In fact, treatment of colon cancer cells with HDAC inhibitors results in an upregulation of p21 expression and subsequent growth arrest (Timmerman et al., 2001).

G. The histone code

The discovery that histone acetylation facilitated gene transcription led to the immediate hypothesis that acetylation of lysine residues neutralized the positive charge of histones attenuating their interaction of DNA. However, Allfrey and colleagues demonstrated that acetylated histones are still able to protect DNA from thermal denaturation (Allfrey et al., 1964). Therefore, it would appear that acetylated histones maintain an interaction with DNA.

In addition to acetylation, histones can also be methylated, phosphorylated, ubiquitinated, ADP-ribosylated, and sumolated (Strahl and Allis, 2000). The large number of post-translation modifications that can be made to histones led to the development of the histone code hypothesis. The histone code hypothesis suggests that these modifications serve as a code which is "read" by various transcription factors and chromatin associating proteins (Strahl

and Allis, 2000). For example, a particular pattern of acetylation, methylation, and/or phosphorylation may lead to the recruitment of specific transcription factors that regulate gene expression.

There are two major lines of support for this hypothesis. First is the discovery of protein domains capable of binding to post-translationally modified histones. For examples, bromodomains bind acetylated lysine residues and chromodomains recognize methylated histone residues (Bottomley, 2004). The bromodomain was the first domain discovered to bind modified histones. It is found in several classes of transcription factors including the HATs GCN5, PCAF, and p300/CBP. The SAGA complex is able to bind and anchor to nucleosomal arrays by interacting with acetylated H3 via the bromodomain of GCN5. If the GCN5 bromodomain is deleted, the SAGA complex is still enzymatically active but cannot anchor itself to the nucleosomal array (Hassan et al., 2002). In addition, a crystal structure for the GCN5 bromodomain binding histone H4 acetylated at lysine 16 has been resolved indicating a stable interaction (Owen et al., 2000). Acetylation of H4 at lysine 16 is one of the major targets of the MYST family of HATs. The ability of GCN5 to recognize MYST modifications suggests that histone modifications made by MYST HATs may direct the enzymatic activity of GCN5-containing complexes through the GCN5 bromodomain. The bromodomain has been shown to interact with acetylated non-histone proteins as well. Acetylated p53 has been shown to be involved in recruiting p300/CBP and PCAF to promoters through interactions with their bromodomains (Barlev et al., 2001).

The second major support for the histone code hypothesis is the discovery that certain modifications take place in predictable and reproducible patterns. However, deciphering the histone code is still in its beginning stages. Phosphorylation of histone H3 at serine-10 (H3S10) occurs at high frequency resulting in chromatin condensation as cells enter and pass through mitosis (Bottomley, 2004). However, H3S10 is also phosphorylated in response to growth factors causing chromatin to unfold and allowing for increased transcription (Bottomley, 2004). The same modification may thus cause two

completly different effects. Upon further analysis, it was discovered that after H3S10 is phosphorylated in response to growth factors, H3 is heavily acetylated (Bottomley, 2004). The presence of hyperacetylation allows for the upregulation of gene transcription.

It has also been recognized that newly synthesized DNA incorporates histone H4 acetylated at lysines 5 and 12 (Turner, 2000). However, the importance of such modifications is unknown. Mutagenesis of these two residues does not adversely affect DNA synthesis. Another example is the activation of the pS2 promoter by estrogen. Within 15 minutes of estrogen stimulation, CBP is recruited to the promoter and acetylates histone H3 at lysine 18 (H3K18) followed by acetylation of lysine 23 (H3K23) and recruitment of the arginine histone methyltransferase CARM1 (cofactor-associated arginine [R] methyltransferase 1) that methylates histone 3 at arginine 17 (Daujat et al., 2002). The combination of these modifications leads to the upregulation of estrogen sensitive promoters (Daujat et al., 2002).

H. Histone variants

Histone variants are expressed in eukaryotic cells, adding another level of complexity to the histone code hypothesis (Henikoff et al., 2004). The four canonical core histones that comprise the nucleosome can be replaced by slightly different histones termed histone variants (Henikoff et al., 2004). Histone variants exist for histones H2A, H2B, and H3; to date no variant exists for histone H4 (Kamakaka and Biggins, 2005). Histone variants differ from core histones in several aspects. There are usually multiple copies of genes encoding the core histone, but only a single copy of each gene encoding a histone variant (Kamakaka and Biggins, 2005). The mRNA for variant histones contains introns, is polyadenylated, and is constitutively expressed (Pusarla and Bhargava, 2005). Canonical core histone mRNA is intronless, is not polyadenylated, and is only expressed during S-phase for incorporation into newly synthesized DNA (Kamakaka and Biggins, 2005). How histone variants are regulated and the

mechanisms involved in directing their incorporation into the genome are poorly understood (Pusarla and Bhargava, 2005).

CenH3 and H3.3 are two well known variants of H3 (Pusarla and Bhargava, 2005). CenH3 is a conserved essential protein that binds to centromeres, the DNA locus that directs formation of the kinetochore protein structure that mediates chromosome segregation in eukaryotes (Kamakaka and Biggins, 2005). H3.3 is the least divergent variant, containing only four amino acid differences compared to core histone H3 in *Drosophila*. However, unlike the core H3 histone, H3.3 is expressed throughout the cell cycle and often localizes to transcriptionally active regions of the chromosome (Ahmad and Henikoff, 2002).

Of the core histones, H2A has the largest number of variants in higher eukaryotes. The two major variants are H2A.Z and H2A.X. H2A.X and H2A.Z are constitutively expressed and localize throughout the genome (Pusarla and Bhargava, 2005). Interestingly, the core H2A protein in *Saccharomyces cerevisiae* is more similar to the mammalian H2A.X variant than to the mammalian core H2A (Malik and Henikoff, 2003).

Histone H2B has only a few variants and they appear to have very specialized functions in chromatin compaction and transcription repression during gametogenesis (Lewis et al., 2003). Despite extensive searching, there are no known variants for histone H4 (Henikoff et al., 2004).

III. Histone Modification and Apicomplexan Parasites

Although there are drugs capable of treating acute *Toxoplasma* infection and reactivation, the current therapies are limited by adverse toxicities. Furthermore, no currently approved therapy is capable of eradicating the latent bradyzoite cysts. Therefore, research continues looking for novel therapeutics capable of targeting bradyzoites and/or with less severe side effects. In the search for better therapeutics to treat apicomplexan diseases, Darkin-Rattray et al. (1996) discovered a novel antiprotozoal agent, termed apicidin. Apicidin is a cyclic tetrapeptide isolated from a Costa Rican fungus (*Fusarium* spp.; Darkin-

Rattray et al., 1996). *In vitro* studies showed that apicidin was a broad spectrum antiprotozoal inhibiting the growth of drug-resistant human malaria, *C. parvum*, and *Toxoplasma* (Darkin-Rattray et al., 1996). The mechanism of action of apicidin is the non-competitive inhibition of a histone deacetylase (Singh et al., 2002). Unfortunately, apicidin also inhibits human HDACs, causing an anti-proliferative effect which is detrimental to human cells, and thus limits its clinical potential. However, the potent broad spectrum antiprotozoal effect of apicidin illustrates that chromatin remodeling enzymes may provide urgently needed targets for drug research (Darkin-Rattray et al., 1996).

A. GCN5 homologue in Toxoplasma

Before the discovery of apicidin, chromatin remodeling proteins had not been investigated in *Toxoplasma*. In 1999, Hettman and Soldati described a homologue of GCN5 in *Toxoplasma* termed TgGCN5. They reported a cDNA of 2.3kb, which encoded a 474 amino acid protein. TgGCN5 possesses a high degree of homology to the GCN5 present in yeast and fellow Alveolate *Tetrahymena*, with a short N-terminal extension of 141 residues.

However, in 2000 Smith and Sullivan published their cloning results of GCN5 in *Toxoplasma,* which indicated that the actual transcript was considerably larger. The new sequence contained a longer, unique N-terminal extension encoded by a single exon in the genomic locus (Sullivan and Smith, 2000). A Northern blot containing mRNA purified from *Toxoplasma* tachyzoites was probed with 2.6kb of sequence capable of hybridizing to both the long form and the previously reported short form of TgGCN5. The Northern results showed hybridization to only one band consistent with the larger version of TgGCN5 (Bhatti and Sullivan, 2005). Furthermore, Hettman and Soldati reported that the truncated form of TgGCN5 does not localize to the nucleus when over-expressed in *Toxoplasma* tachyzoites (Hettman and Soldati, 1999). In contrast, the full-length protein with the longer N-terminal extension does localize to the parasite nucleus when fused to green fluorescent protein (Sullivan, unpublished data).

Collectively, the above data argues that *Toxoplasma* expresses only the form of TgGCN5 consistent with the size reported by Sullivan and Smith.

B. The N-terminal extension of TgGCN5

Figure 6 shows a comparison of TgGCN5 with other GCN5 homologues present in other organisms. The conserved Ada2 and bromodomains are present in TgGCN5. However, there is a striking difference in the N-terminal sequence between TgGCN5 and other GCN5 homologues. As depicted in Figure 6, the large N-terminal extension of TgGCN5 goes against the general trend that GCN5 proteins from lower eukaryotes have shorter N-termini. The entire TgGCN5 N-terminal extension is comprised of 820 amino acids which is larger than the extensions found in metazoan GCN5 homologues (approximately 500 residues). The closely related *Tetrahymena* possesses a small N-terminal extension. Furthermore, unlike the large N-terminal extensions present in metazoan GCN5 and PCAF, the lengthy N-terminal extension present in TgGCN5 bears no similarity to any known protein and is devoid of known protein motifs.

The presence of a parasite-specific domain in a highly conserved protein generates many questions regarding its function(s). Given what is known about other GCN5 proteins, the N-terminus may interact with other proteins to form HAT complexes not present in other eukaryotes. The unique N-terminal extension may play a role in regulating the acetylation activity of TgGCN5. Perhaps TgGCN5 has a unique histone acetylation pattern directed by the N-terminal extension. The N-terminus may also be involved in substrate recognition and facilitate the acetylation of non-histone proteins in the parasite. There are many theoretical functions of the unique N-terminus of TgGCN5. However, from the two initial reports of TgGCN5, it can be deduced that at least one known function of the N-terminus is involvement in nuclear localization. As mentioned, the N-terminal extension present in the maize homologue of GCN5 is also required for nuclear targeting, but the exact sequence and mechanism

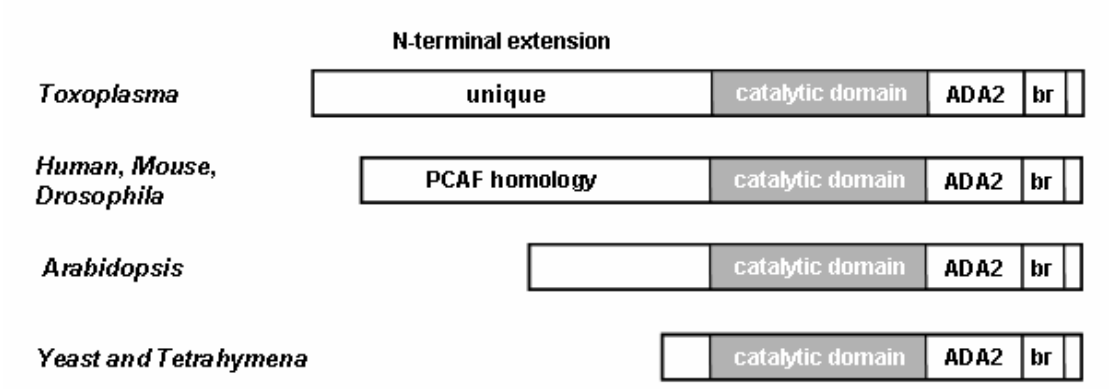


Figure 6: TgGCN5 and other GCN5 homologues

Above are schematic diagrams comparing known homologues of GCN5 to TgGCN5. Compared to other species, TgGCN5 possess a lengthy unique N-terminal extension. The N-terminal extensions present in human, mouse, and *Drosophila* are highly homologous to the N-terminal extension found in PCAF. The N-terminal extension in TgGCN5 shares no homology to any known protein. TgGCN5 also possesses the conserved catalytic domain, responsible for acetylating histones, an Ada2 binding domain (Ada2), and a bromodomain (br).

responsible for transporting the protein to the nucleus was not determined (Bhat et al., 2003).

IV. Nuclear Import

The histone acetylase and gene regulatory functions of GCN5 have been rigorously studied. However, little attention has been directed at determining how the HAT gets into the cell nucleus. Clearly the unique N-terminal extension of TgGCN5 is involved in nuclear targeting making it an excellent tool to resolve how this protein is target to the parasite nucleus. Understanding how TgGCN5 and other transcription factors enter the parasite nucleus may lead to the discovery of novel elements present in nuclear trafficking pathways in *Toxoplasma* and other apicomplexans. Novel elements not present in higher eukaryotes may be exploited in the design of more selective therapeutic agents. Blockade of parasite nuclear trafficking could subvert parasite differentiation and other processes essential for parasite survival. Additionally, the study of these pathways in protozoa provides a unique perspective on how these systems evolved in early eukaryotic cells.

A. Nuclear localization signals

In all eukaryotic organisms, the nucleus is separated from the cytoplasm by the nuclear envelope. The nuclear envelope is perforated by many pores through which water-soluble molecules enter and leave the nucleus. Each pore is an elaborate multi-protein structure termed the nuclear pore complex (NPC), which regulates the movement of macromolecules across the nuclear envelope (Pante and Kann, 2002). The nuclear pore complex is immense, approximately 12.5 million Daltons (Weis, 2003). Ions, small metabolites and globular proteins less than 40kD can diffuse through the nuclear pore complex (Pante and Kann, 2002). However, large proteins and complexes cannot diffuse in and out of the nucleus. Therefore, they must be actively transported through the nuclear pore complex. Many proteins actively imported through the nuclear pore contain a nuclear localization signal (NLS; Quimby and Corbett, 2001).

The first NLS was discovered during mutational analysis of the large T-antigen, an early viral protein of the simian virus SV40 (Kalderon et al., 1984a). In infected cells, large T-antigen normally accumulates in the nucleus. Mutations within a short sequence of basic residues KKKRK (K=lysine, R=arginine) prevented large T-antigen from entering the nucleus (Kalderon et al., 1984a). Therefore the basic pentapeptide defined a region of the protein involved in nuclear localization. When the basic penta-peptide was fused to two very large cytosolic proteins, beta-galactosidase and pyruvate kinase, both were targeted to the cell nucleus instead of remaining in the cytosol (Kalderon et al., 1984b).

Since the initial discovery in large T-antigen, NLSs have been isolated in numerous other proteins imported into the nucleus (Quimby and Corbett, 2001). They are similar to the SV40 large T-antigen NLS in that they contain consecutive basic amino acids. A large study looking at different NLSs by Boulikas reported nine distinct forms of strongly basic hexapeptides characteristic of nuclear proteins among all eukaryotic species (Boulikas, 1994). The hexapeptides contained at least four arginines and/or lysines and one or two non-polar residues. The four most common NLS motifs were θθθθ, θθxθθ, θθθxθ, and θθxθxθ, where θ equals a basic residue (arginine or lysine) and "x" represents a non-polar residue. Boulikas suggested that these streches of basic residues are different versions of a core NLS (Boulikas, 1994).

Targeting proteins to the nucleus utilizing clusters of basic residues is a highly conserved eukaryotic phenomenon that occurred early in eukaryote evolution. Using phylogenetic analysis with 16s rRNA sequence, *Giardia lamblia* represents the earliest diverging lineage in the eukaryotic line of descent (Sogin et al., 1989). When GFP is expressed in *Giardia lamblia*, is localizes throughout the cytoplasm, but if the SV40 large T-antigen NLS is fused to GFP, it is targeted to the *Giardia lamblia* nuclei (Elmendorf et al., 2000). Therefore, this early branching eukaryote contains machinery capable of recognizing and transporting proteins to the nucleus through recognition of an NLS composed of basic residues.

B. Importin alpha

Additional experiments studying nuclear localization revealed two proteins involved in nuclear import through interaction with NLSs composed of basic residues: importin α and importin β (Gorlich et al., 1995). Importin α (IMPα) and importin β (IMPβ) are also referred to as karyopherin α and karyopherin β (Chook and Blobel, 2001).

The function of importin α is to recognize and bind to NLSs comprised of basic residues (Gorlich et al., 1994). Importin α is a highly conserved protein present in virtually all eukaryotic organisms (Chook and Blobel, 2001). A text search of the annotated proteins in the *Giardia* genomic database (www.mbl.edu/Giardia) revealed the presence of an importin α homologue present in this primitive eukaryote. From its early origins, the importin α gene family has undergone considerable expansion during the course of eukaryotic evolution. Whereas the yeast *Saccharomyces cerevisiae* genome encodes a single importin α, the human genome encodes six different genes (Chook and Blobel, 2001). Phylogenetic analyses indicate that a single ancestral animal importin α gave rise to the additional paralogues in higher eukaryotes (Goldfarb et al., 2004). The presence of multiple importin α genes in metazoans suggests they appeared during the evolution of multicellular animals and presumably perform cell and tissue specific roles (Goldfarb et al., 2004).

Crystallographic analysis revealed that importin α has a rather unique structure that facilitates its ability to bind to an NLS (Kobe, 1999). The structure of importin α can be divided into two parts. The first part is a short N-terminal domain of about 60 residues in length called the importin β binding (IBB) domain (Goldfarb et al., 2004). As the name implies, the IBB domain is the region of the protein that interacts with importin β. The remainder of the ~60kD protein is comprised of repeating armadillo (ARM) motifs (Teh et al., 1999). Each ARM motif is composed of three α-helices and the repeating ARM motifs associate together to form a superhelical tunnel-like structure. The superhelical structure exposes highly conserved acidic residues on its surface which interact with the basic residues present in the NLS of nuclear proteins (Teh et al., 1999).

Although importin-α is capable of recognizing and binding an NLS, it is unable to translocate a protein into the nucleus. In order to transport a protein into the nucleus importin α must complex with importin β through its IBB domain (Gorlich et al., 1995). Importin α must also be bound by importin β before it can effectively bind an NLS (Goldfarb et al., 2004). Further analysis of the crystal structure of importin α indicates that the IBB domain is autoinhibitory (Kobe, 1999). When not bound to an NLS, the IBB domain folds over and interacts with the ARM repeats (Kobe, 1999). The IBB contains an NLS-like motif, KKR. Thus, it appears that importin α is autoinhibited by an internal NLS. Once importin α is bound by importin β, the autoinhibition is negated allowing an NLS to be bound (Kobe, 1999). Once the NLS is bound by importin α, importin β facilitates translocation into the nucleus through interaction with components of the nuclear pore complex (Chook and Blobel, 2001).

Several other proteins that interact with importin α and importin β have been discovered, eventually leading to the development of a model for the import of cytosolic proteins containing a basic NLS (Figure 7). Once the trimeric complex consisting of importin β, importin α, and the nuclear protein are translocated through the NPC, importin β is bound by Ran-GTP causing the complex to dissociate (Quimby and Corbett, 2001). Ran (ras-related nuclear protein) is a small GTPase that is involved in recycling importin β and importin α back to the cytoplasm (Bischoff and Ponsting, 1991; Chook and Blobel, 2001). When importin β is bound by Ran-GTP, it is transported out of the nucleus back into the cytoplasm where Ran-GTP interacts with RanGAP (Weis, 2003). RanGAP (Ran GTPase activating factor) facilitates the GTPase activity of Ran, causing it to hydrolyze GTP to GDP (Stewart and Rhodes, 1999). The hydrolysis of GTP causes Ran to release importin β. Ran-GDP is then translocated back into the nucleus where it is acted upon by RCC1 (regulator of chromosome condensation; Bischoff and Ponsting, 1991). RCC1 (or RanGEF; Ran guanine nucleotide exchange factor), catalyzes the release of GDP from Ran and the regeneration of Ran-GTP (Quimby and Corbett, 2001).

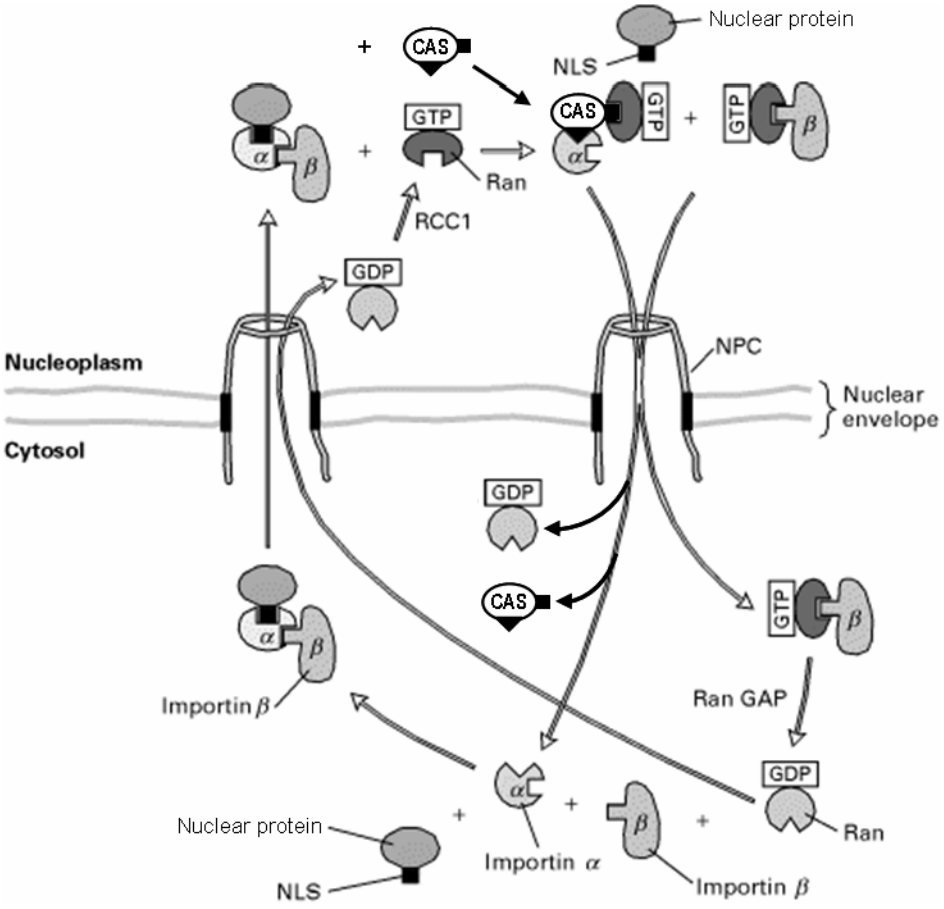


Figure 7: Importin α and β nuclear import pathway

A diagram of the proposed mechanism for the transport of proteins containing a basic NLS into the nucleus. Importin α, Importin β, and the nuclear protein form a trimeric complex, that is transported through the NPC into the nucleoplasm. Ran-GTP and importin β are transported back to the cytoplasm. Importin α is also recycled back to the cytoplasm through interactions with CAS and Ran-GTP. In the cytoplasm, RanGAP stimulates conversion of Ran-GTP to Ran-GDP resulting in Ran to dissociate from Import β and CAS. The released Importin β and α can now interact with another nuclear protein bearing a basic NLS. Ran-GDP is transported back to the nucleus, where RCC1 causes it to release GDP and bind GTP. (modified from Lodish et al., 2000; reprinted with permission from H. Freeman & Company)

Importin α is recycled in a similar manner, but involves a nuclear export protein, CAS (cellular apoptosis susceptibility gene; Stewart and Rhodes, 1999). CAS binds to importin α, and then CAS is bound by Ran-GTP. The trimeric complex is exported out the nucleus where Ran-GTP interacts with RanGAP causing GTP hydrolysis and dissociation of the complex. CAS and Ran-GDP are subsequently transported back into the nucleus. The majority of Ran-GTP in the nucleus and Ran-GDP in the cytoplasm establishes a gradient that drives the recycling of importin α and β facilitating continuous nuclear import (Weis, 2003).

C. Nuclear import in Apicomplexans

Nuclear trafficking is virtually an untapped area of research in apicomplexans. Homologues of importin β and importin α have been identified and characterized in *Plasmodium* (Mohmmed et al., 2003). *Plasmodium* also possesses Ran and RCC1 (Ji et al., 1998). However, no functional studies have been performed on any of these nuclear trafficking proteins. There have been no reports on the identification of any proteins involved in nuclear transport in *Toxoplasma*. The N-terminal extension of TgGCN5 may provide an excellent tool to begin studying the nuclear trafficking pathways in apicomplexans.

D. Nuclear trafficking of GCN5

To date, very little is known about how GCN5 is targeted to the nucleus in any organism. The maize GCN5 homologue (ZmGCN5) is the only GCN5 protein for which an NLS has been proposed. When the entire 175 amino acid N-terminal extension of ZmGCN5 is fused to GFP, GFP is targeted to the maize nucleus (Bhat et al., 2003). Near the beginning of the N-terminal extension of ZmGCN5 there is a motif comprised of four basic residues, RKRK, that has been proposed to be the NLS for ZmGCN5 (Bhat et al., 2003). However, no additional studies have been performed to confirm that RKRK is indeed the NLS for ZmGCN5. The 118 amino acid yeast GCN5 N-terminal extension reveals no stretches of 3 or more consecutive basic residues. How yeast GCN5 is targeted to the nucleus is unknown, but it is doubtful to involve the N-terminal extension.

A study investigating the ability of a human importin α paralogue to target nuclear proteins to the nucleus used the GCN5 related HAT, PCAF as one of many bait proteins (Kohler et al., 1999). The study demonstrated that importin α is required for PCAF to be transported to the nucleus. However, the binding site (NLS) of PCAF was not identified in the study. TgGCN5 contains several clusters of basic residues in its N-terminal extension that may be functioning as an NLS capable of associating with a putative *Toxoplasma* homologue of importin α and targeting it the parasite nucleus.

V. Thesis Goals

Bradyzoite differentiation can be replicated *in vitro* using physiological and chemical stressors such as heat, sodium arsenite, basic pH, and IFNγ. GCN5 in yeast is involved in mediating stress responses, particularly amino acid and phosphate starvation. Thus, it is quite possible that TgGCN5 may be involved in mediating stress responses in *Toxoplasma* and therefore involved in the bradyzoite differentiation process. Characterization of TgGCN5 will facilitate our understanding of chromatin remodeling in apicomplexans and may bring us closer to understanding the differentiation process in *Toxoplasma*.

TgGCN5 contains a unique N-terminal extension upstream of the catalytic domain that shares no homology to analogous domains found in other GCN5 proteins, nor is it similar to any sequence in protein databases. The goal of this thesis is to better understand the purpose of the unique N-terminus. **Our hypothesis is that this unique N-terminal extension is critical to the function of TgGCN5.** Three possible roles of the N-terminal domain will be investigated in the following aims:

A. AIM 1: Determine how the N-terminal domain mediates nuclear localization

It is clear that the N-terminal extension is involved in mediating the nuclear localization of TgGCN5. Therefore, TgGCN5 makes an excellent model to study how a GCN5 protein is translocated into the nucleus. Previous studies performed in the laboratory have narrowed the region involved in nuclear

localization to be between amino acids 58 and 260. Between residues 58 and 260 there is a single stretch of basic residues that may be function as an NLS. Truncations will be made in the N-terminal extension to isolate this stretch of basic residues to determine if it is functioning as the NLS. Localization of the truncation mutants will be determined using immunocytochemistry via a fused FLAG tag. If this stretch of basic residues is indeed the NLS, it will be determined if it is sufficient to target a cytoplasmic protein to the parasite nucleus. To date, an NLS has never been mapped for any *Toxoplasma* protein. Therefore, the NLS elucidated in TgGCN5 will be the first NLS described in *Toxoplasma*. In addition, it will be the first NLS described in a GCN5 homologue.

B. AIM 2: Assess the involvement of the N-terminal extension in protein-protein interactions

There are two different groups of proteins that may interact with the N-terminus of TgGCN5. First group of proteins would be those involved in mediating the translocation of GCN5 into the parasite nucleus. Once the NLS has been elucidated in AIM 1, it will be used to guide our search for potential nuclear trafficking proteins. If the NLS is indeed comprised of basic residues, we will search the *Toxoplasma* genomic database for a homologue of importin α. If present, the homologue will be cloned and protein expressed. Using an *in vitro* pull down assay, it will determined if the importin α homologue interacts with TgGCN5 and if the interaction occurs at the NLS elucidated in AIM 1. To date, no one has determined if nuclear trafficking proteins exist in *Toxoplasma*.

The second group proteins that may interact with TgGCN5 are those involved in forming HAT complexes analogous to the SAGA and ADA complexes found in yeast. To identify proteins interacting with N-terminus, co-immunoprecipitation experiments using anti-FLAG affinity resin will be performed on transgenic parasites that over-express recombinant full length TgGCN5 fused to FLAG tag ($_{FLAG}$TgGCN5) and TgGCN5 lacking the N-terminal extension also fused to a FLAG tag ($_{FLAG}\Delta N_T$TgGCN5). Proteins isolated in pull down assays from $_{FLAG}$TgGCN5 and that are not present in pull down assays from

$_{FLAG}\Delta N_T$TgGCN5 will be deemed to be interacting with the N-terminal extension of TgGCN5.

C. AIM 3: Determine if the N-terminal extension modulates enzymatic (HAT) function

The long extensions present in metazoan homologues of GCN5 are involved in substrate recognition. Even though the N-terminal extensions are not homologous, the unusual N-terminal domain of TgGCN5 may have a role in regulating substrate specificity and/or acetylase function. *In vitro* HAT assays will be performed using full length recombinant TgGCN5 and TgGCN5 lacking the N-terminal extension on free histone substrates. Antibodies to specific acetylated lysine residues will be used to determine the substrate specificity of TgGCN5 with and without the N-terminal extension.

CHAPTER 2: MATERIALS AND METHODS

I. Culture and Parasite Techniques

Below is an explanation of these well established techniques, which are described in greater detail in Roos et al., 1994.

A. Cell and parasite culture

The RH *Toxoplasma* strain obtained from the AIDS Research and Reference Reagent Program (Division of AIDS, NIAID, National Institutes of Health; http://www.aidsreagent.org) was used for all experiments in this thesis. Tachyzoites were cultivated using confluent monolayers of human foreskin fibroblast (HFF) cells in T-25cm^2 flasks (Fisher #10-126-30). HFF cells (American Tissue Culture Company; #CRL-2522) were grown to confluency in host cell (HC) media. HC media consists of Dulbecco's Modified Eagle Medium (DMEM; Invitrogen #11965-126) supplemented with 10% heat-inactivated fetal bovine serum (Invitrogen #16000-044) and 25µg/L gentamicin (Invitrogen #15710-064). After combining the components of HC media, it is filter sterilized using a 0.2µm filtering device (Fisher #09-740-25A). HFF cells were grown at 37°C in a humidified 5% CO_2 environment. All tissue culture work was performed under sterile conditions inside a laminar flood hood.

Additional HFF flasks were generated using trypsin digestion. The HC media from a T-150cm^2 flask (Fisher #10-126-32) containing confluent HFF cells was aspirated. The monolayer was washed with 3.0ml of 37°C sterile PBS which was subsequently aspirated. Following the wash, 3.0ml of trypsin (Invitrogen #25200-056) was added to the flask. The flask was rocked several times to allow the trypsin to adequately coat the monolayer and then approximately two-thirds of the trypsin was aspirated. The flask was placed at 37°C for one minute. The monolayer was disrupted by banging the corner of the flask against the palm of the hand several times. The trypsinization was squelched by pipetting 25ml of fresh 37°C HC media into the flask. The wall of the flask containing the monolayer was pipetted against several times to completely suspend all of the

trypsinized cells. The trypsinized cells were diluted up to 400ml using fresh 37°C HC media. For making additional T-25cm^2 flasks, 10ml of the trypsinized cells were added to each new T-25cm^2 flasks and for making additional T-150cm^2 flasks, 50ml of the resuspended trypsinized cells were added to new T-150cm^2 flasks. New flasks containing resuspended trypsinized cells were placed at 37°C in a humidified 5% CO_2 environment. The caps on the flasks were left slightly loose to allow for gas exchange. Cell growth of the new monolayers was monitored daily using a Leica DM IL inverted microscope at 250X and 400X magnification with a phase filter. Once the monolayers were confluent, the flask lids were closed tightly.

Because HFF cells are not an immortalized cell line, they cannot be passed an infinite number of times. Therefore, the number of times HFF cells were trypsinized was recorded on each flask. HFF cells usually senesced and would no longer grow to confluency after 20 rounds of trypsin digestion. When this occurred, lower passage HFF cells were thawed from liquid nitrogen stocks as outline below (page 56). HFF cells are routinely chosen for parasite cultivation because they exhibit strong contact inhibition and can be stored in incubators for up to 4 weeks without losing their ability to support parasites.

For immunofluorescence assays, 12 well plates containing coverslips were required. Tissue culture grade 12 well plates (Fisher #07-200-81) and coverslips (Fisher #12-545-83) were used. The coverslips were sterilized by autoclaving prior to use in cell culture. Using sterile flat tweezers (Fisher #09-753-50), a single cover slip was placed in each well of the 12 well plate. Host cells from a T-150cm^2 flask were trypsinized as outline on the previous page. After squelching the trypsinization with 25ml of fresh 37°C HC media, 1.0ml of the trypsinized cells were diluted into 25ml of 37°C HC media in a sterile 50ml conical tube (Fisher #14-959-49A). The diluted cells were mixed gently by inversion. Two milliliters of the diluted cells were added to each well and the plate was placed at 37°C in a humidified 5% CO_2 environment. Cell growth within each well was monitored daily using a Leica DM IL inverted microscope at 250X and 400X magnification with a phase filter.

For cloning by limiting dilution assays, 96 well plates were required. Tissue culture grade 96 well plates (Fisher #07-200-89) were used. Host cells from a T-150cm^2 flask were trypsinized as outline above (page 52). After squelching the trypsinization with 25ml of fresh 37°C HC media, 5.0ml of the trypsinzed cells were diluted into 100ml of 37°C HC media in a sterile T-75cm^2 (Fisher #10-126-31). The diluted cells were mixed gently by inversion. 20ml of the diluted cells were poured into a sterile trough (Fisher #13-681-101) and using a Brinkmann EasyPet automatic pipetter (Fisher #13-688-177), 200µl of diluted cells were added to each well generating a total of 5, 96 well plates. The plates were placed at 37°C in a humidified 5% CO_2 environment. Cell growth of the monolayers within each well was monitored daily using a Leica DM IL inverted microscope at 250X magnification without a phase filter.

Due to the obligate intracellular nature of the tachyzoite, they must constantly be provided with new host cells in order to sustain themselves. Once a monolayer in a T-25cm^2 flask has been infected with tachyzoites, the monolayer will be destroyed or lysed in a matter of days depending on the size of the inoculum. If the parasites are not provided with a new monolayer, the parasites will die. Prior to parasite inoculation, the old HC media in the flasks was aspirated and replaced with 9ml of parasite media (PA) which consists of DMEM, 1% heat-inactivated fetal bovine serum, and 25µg/L gentamicin, filter sterilized using a 0.2µm filtering device (Fisher #09-740-25A). The reduction in fetal bovine serum is thought to minimize the exposure of parasites to serum antibodies and does not have detrimental effect on confluent host cells.

The flask containing the infected monolayer is scrapped with a sterile spatula (Fisher #08-773-2) to completely rupture any remaining parasite vacuoles and to suspend the parasites into the media. For routine parasite upkeep, 1.0ml of media containing the parasites from an infected flask was transferred to a T-25cm^2 flask containing a new confluent HFF monolayer. Infection and subsequent destruction of the new monolayer would typically occur within 2-3 days yielding approximately 1X10^7 to 5X10^7 parasites. Once the HFF monolayer is destroyed parasites must be transferred to a new flask containing a

confluent HFF monolayer. Parasites which remain extracellular for longer than ten to twelve hours begin to die. When large numbers of parasites were required (e.g. for immunoprecipitation) a T-150cm^2 flask of HFF cells was inoculated with 3-4ml of parasites yielding at least 1X10^8 parasites. All flasks containing infected monolayers were grown at 37°C in a humidified 5% CO_2 incubator. The lids on the flasks were left loose overnight (approximately twelve to eighteen hours) to allow media to equilibrate with the 5% CO_2 environment inside the incubator. The following day, the lids were tightened to minimize the risk of spillage and culture contamination. Parasite infection of new host cells was monitored daily using a Leica DM IL inverted microscope at 400X magnification with a phase filter.

B. Purification and quantification of parasites

Before genetic manipulation or biochemical experiments were performed using parasites, tachyzoites had to be purified away from host cell debris. Flasks containing infected monolayers were scrapped with a sterile spatula to completely rupture any remaining parasite vacuoles and to suspend the parasites into the media. The suspended parasites were transferred into the barrel of a 60ml plastic syringe (Fisher #14-841-36). The syringe was attached to a filter cartridge (Fisher #NC9671597) containing a 3µm pore-size polycarbonate filter disk (Fisher #NC9655197). The parasites were passed through the filter cartridge removing most of the host cell debris. If the filtered parasites were to be maintained in tissue culture an unopened sterile syringe and autoclaved sterile filter cartridge were used. The filtrate containing parasites was centrifuged at 800 x g for ten minutes. The supernatant was aspirated off and the parasite pellet was resuspended in PBS (pH 7.0). When necessary, tachyzoites were quantitated using an improved Neubauer phase hemocytometer (Fisher #02-671-54) viewed with a Leica DM IL inverted microscope at 400X magnification with a phase filter. Following quantitation, the parasite were centrifuged again 800 x g for ten minutes. The PBS was aspirated off and the parasite pellet was

resuspended in an appropriate buffer dictated by the intended experimental design.

C. Freezing and thawing cells and parasites

For long-term storage, tachyzoites were frozen in liquid nitrogen using DMSO (Sigma #D4540) as a cryoprotectant. Because tachyzoites do not survive very long out side of a host cell, parasites were frozen before completely destroying the host cell monolayer. An infected monolayer in T-150cm^2 flask was scraped at about 70-80% destruction with a sterile spatula (Fisher #08-773-2). The suspended infected cells were transferred to a 15ml conical (Fisher #14-959-70C) and centrifuged for ten minutes at 800 x g at 4°C and the pellet was resuspended in 1.5ml of ice cold PA media (page 54). An equal volume of freezing media (FZ; DMEM with 20% heat-inactivated fetal bovine serum and 20% DMSO) was added. The culture was aliquoted into round-bottom, cryovials (Fisher #09-761-72) and placed in a pre-chilled styrofoam freezer box at -80°C. Frozen aliquots were moved to liquid nitrogen storage the following day.

HFF cells were frozen in a similar manner from a confluent T-150cm^2 flask. Trypsinized cells were resuspended in 1.5ml of ice cold HC media and 1.5ml of FZ media. The culture was divided into 0.5ml aliquots in round-bottom, cryotubes (Fisher #09-761-72) and placed in a pre-chilled styrofoam freezer box at -80°C. Frozen aliquots were moved to liquid nitrogen storage the following day.

When necessary, parasites and HFF cells were thawed quickly to ensure adequate viability. Frozen aliquots were removed from liquid nitrogen and immediately placed into 37°C water. Close to complete thawing, the tube was doused with 70% ethanol and dried (to minimized contamination). Thawed HFF cells were quickly transferred to a T-75cm^2 flask (Fisher #10-126-31) containing 30ml of host cell media. Thawed parasites were inoculated into a T-25cm^2 of confluent HFF cells with 9.5ml PA media. Freshly thawed HFF cells took up to a week to become confluent. Thawed parasites could also take up to a week to destroy a monolayer. After the thawed parasites had lysed the initial monolayer

1.0ml of parasites was transferred to a fresh HFF monolayer and the parsites were maintained in culture as outlined above in Section-A (page 54).

D. Parasite transfection

For the purposes of generating transgenic parasites, the RH parasite in which the HXGPRT (hypoxanthine-xanthine-guanine phosphoribosyltransferase) enzyme had been deleted was used (RHΔHX; Donald et al., 1996). The RHΔHX parasite line can be selected for by the addition of the prodrug 6-thioxanthine (6-TX; 320µg/ml Sigma #S448524 in 0.5M KOH) to PA media (Pfefferkorn et al., 2001). Parasites are transfected with a plasmid containing a copy of the HXGPRT gene and are selected for using mycophenolic acid [MPA; 25mg/ml Sigma #M3536 in 100% EtOH (ethyl alcohol)] supplemented with xanthine (50mg/ml Sigma #X4002 in 0.5M KOH). Under MPA selection, RHΔHX parasites that did not incorporate the plasmid, no longer possess a mechanism to generate guanine nucleotides and die (Donald et al., 1996). Transformed parasites can utilize their restored HXGPRT salvage pathway to generate guanine nucleotides from the supplemented xanthine.

RHΔHX parasites for transfection were harvested from a freshly lysed T-25cm^2 flask and filter purified from host cells as described above. Parasites were washed in 5.0ml of filter sterilized cytomix electroporation buffer (Soldati et al., 1993). After centrifugation, the pellet was resuspended in 300µl of cytomix supplemented with fresh 2mM ATP (Sigma #A6419) and 5mM glutathione (Sigma #G6013). Between 30-50µg of sterile DNA was resuspended in 100µl of cytomix supplemented with 2mM ATP and 5mM glutathione and was added to the 300µl of parasites. The parasites and DNA were mixed thoroughly and transferred to a 2.0mm gap cuvette (Fisher #BTX620) and electroporated using a single 1500V pulse with a resistance setting of 25Ω on a BTX model 630 Electro Cell Manipulator. Electroporated parasites were allowed to recover for fifteen minutes at room temperature. The transfected parasites were split between two T-25cm^2 flasks containing a confluent monolayer of HFF cells. After complete destruction of the monolayer, 2.0ml of parasites were passed onto a fresh

monolayer. Following the destruction of the second monolayer, 2.0ml of parasites were passed onto a fresh monolayer adding 10µl of MPA (25mg/ml Sigma #M3536 in 100% EtOH) supplemented with 25µl xanthine (50mg/ml Sigma #X4002 in 0.5M KOH) for selection of parasites possessing the transfected plasmid.

E. Clonal dilution

Despite the use of MPA to select for parasites possessing the transfected plasmid, the population of transfected parasites is never homogenous. In order to obtain a homogenous population of transfected parasites a process termed clonal dilution was employed to generate a parasite population originally derived from a single parasite.

Following at least two rounds of MPA selection, parasite were harvested and filtered under sterile conditions to remove host cell debris. The parasites were resuspended in PM media containing MPA and xanthine. The concentration of the parasites was determined using an improved Neubauer phase hemocytometer (Fisher #02-671-54) viewed with a Leica DM IL inverted microscope at 400X magnification with a phase filter. Ten microliters of the suspended parasites was added to the hemocytometer. The hemocytometer contains a 5 X 5 grid generating 25 boxes when viewed on the microscope. The number of parasites within 5 boxes is counted and averaged. The counting is repeated with a total of 4 ten-microliter samples. The average number of parasite from each round of counting is then averaged. The purpose of multiple rounds of counting is to reduce the amount of error. The final average from four rounds of counting is multiplied by 25 because there are 25 boxes on the grid. The resulting product is multiplied by 1000 because the volume of the grid in the hemocytometer is approximately 1 microliter. The final number is concentration of parasite per milliliter in the flask. The concentration is usually around $1X10^6$ to $5X10^6$ parasites per milliliter.

A 1:1000 dilution must be performed to bring the concentration of the parasites down to a useable concentration. Ten microliters of the parasites is

added to 10ml of PM buffer containing MPA supplemented with xanthine. The parasite concentration in the flask is divided by 1000 giving the concentration of the diluted parasites (usually about 1-5X10^3 parasites per milliliter). The concentration of diluted parasites can be rewritten as parasites per microliter (usually about 1 to 5 parasites per microliter). In order to separate out the parasite individually, 96 well plates are generated with a confluent monolayer of HFF cells at the bottom of each well as outlined above (page 54). The goal is to get one parasite into each well. For simplicity, the 96 wells in a single plate are rounded up to 100 wells which means 100 parasites are needed. Therefore, the number of parasites (100) is divided by the concentration of the parasites (100 wells / 1 to 5 parasites/μL = 20 to 100μL) giving the volume of the 1:1000 dilution that contains 100 parasites. Because of the high amount of error associated with parasite counting, the calculated volume containing 100 parasites is usually multiply by 2 to ensure enough parasites are used to seed the well plate (~40 to 200μl). The final volume, which contains about approximately 100 parasites is added to 20ml of PA media containing MPA supplemented with xanthine (20ml is the volume of media necessary to fill one 96 well plate). The 20ml containing approximately 100 parasites is then divided evenly over one 96 well plate by placing 200μl into each well. The plates are allowed to grow for about seven days without being disturbed. The goal will be to look for individual wells that contain a single plaque of destroyed host cells caused be a single parasite. Multiple plaques in a well indicate more than one parasite was placed in the well and does not represent a population derived from a single parasite. If the plates are moved excessively it will suspend the parasites in the wells causing the appearance of multiple plaques making it difficulty to tell which wells contain populations derived from a single parasite and which one do not.

Several wells, usually between 4 to 7, were picked and the parasites present in the well were transferred to a T-25cm^2 flask containing a fresh monolayer of host cells with PA media containing MPA supplemented with xanthine. Because of the small inoculum, it took between 7 to 14 days for the parasites to destroy the monolayer. The parasites are passed through at least

two more T-25cm^2 flasks to ensure the parasites numbers were sufficient and then the clonal parasites were assessed for the presence of the transfected plasmid. For the purposes of this thesis, immunofluorescent assays were performed to assess the presence and expression of the transfected plasmid. Clonal dilution was used to obtain clonal parasite lines expressing full length TgGCN5 and TgGCN5 lacking the N-terminal extension (ΔN_TTgGCN5).

II. Vector Construction

A. General PCR protocol

All primers were ordered from Invitrogen and were resuspended at ~100pm/µl in double-distilled deionized 18 milliohm water. All of the reaction components except for the template DNA were assembled on ice in a designated PCR only area using barrier tips to minimize contamination. The 50µl reactions were assembled in thin walled 500µL PCR tubes (Fisher #E0030 124 600). For all reactions approximately ~100ng of primer was used. All PCRs were conducted with 1.0µL of the proofreading DNA polymerase *Pfu*Ultra™ (Stratagene #600380), 5.0µl of the accompanying 10X reaction buffer, 0.5µl of 100µM dNTP mixture, and 1.5µl of DMSO. The ice bucket was moved out of the designated clean area and ~50ng of template was added to the reaction. The PCR was mixed by gently flicking the tube and was spun down in a microfuge and placed in the thermalcycler.

The general thermal-cycling protocol for all PCR is as follows:

1. 95°C for two minutes
2. 95°C for thirty seconds
3. T_m for thirty seconds [T_m calculated using Cybergene Primer design utility software (http://www.cybergene.se/primer.html)]
4. 72°C for one minute per kilobase of amplicon
5. Repeat steps 2-4 for a total of 25 cycles
6. 72°C for ten minutes
7. 4°C indefinitely

PCR reactions were separated by agarose gel electrophoresis using a 0.8% agarose gel impregnated with 0.5 µg/ml ethidium bromide run at 110V for approximately twenty-five mintues. Agarose gels were place on plate glass and visualzed by transluminating UV light at 312nm wavelength through the glass. Appropriate bands were excised using a clean razor blade and purified using a gel extraction kit (Marligen Bioscience #11456-019). Purified PCR products were subsequently TOPO® ligated into the pCR®-Blunt II-TOPO® (Invitrogen #K2800-20) vector.

B. Tranformation into E. coli

After the TOPO® ligation, One Shot® TOP10 chemically competent *E. coli* (Invitrogen #C4040-03) were transformed with the pCR®-Blunt II-TOPO® vector containing the PCR product using the heat shock method outline in the product literature. Following transformation, the bacteria were plated out on LB agar plates containing 50 µg/ml kanamycin and incubated inverted overnight in a dry 37°C incubator. Eight colonies were picked using sterile loops to inoculate 2ml LB culture media with 50 µg/ml kanamycin. Liquid cultures were incubated for eight to twelve hours at 37°C in a shaking incubator. 1.5ml of each culture was pipetted into a 2.0ml microfuge tube and spun down at 13,000g for ten minutes. The remaining 0.5ml of each culture was placed at 4°C. After centrifugation, the liquid media was aspirated off and plasmids from the bacterial pellet were isolated using the Qiagen Miniprep Kit (Qiagen #27144) according to the manufacture's instructions. Ten microliters of each purified plasmid was analyzed using restriction digest mapping based on restriction sites incorporated into the primers. The restriction digests were incubated at overnight at 37°C. Digest reactions were separated by agarose gel electrophoresis using a 0.8% agarose gel impregnated with 0.5 µg/ml ethidium bromide run at 110V for approximately twenty-five mintues. Agarose gels were place on plate glass and visualized by transilluminating UV light at 312nm wavelength through the glass. One of the plasmids that produced the predicted fragments was sequenced verified at the Biochemistry Biotechnology Facility (Indiana University School of

Medicine, IN) using the M13 primer sites present in the pCR®-Blunt II-TOPO® plasmid. In addition, the remaining 0.5ml of culture at 4°C containing the plasmid sent for sequencing was used to generate a glycerol stock by transferring 100µl to a 1.5ml microfuge tube and adding 10µl of sterile glycerol. The tube was mixed by through shaking and inversion and stored at -80°C.

Once sequencing of the PCR product was verified as correct, the remaining pCR®-Blunt II-TOPO® vector containing the PCR product was digested with restriction enzymes based on the restriction sites incorporated into the primers to cleave the PCR product from the pCR®-Blunt II-TOPO® vector. In parallel, the destination vector was also digested with the same restriction enzymes. Both digests were incubated overnight at 37°C. Digest reactions were separated by agarose gel electrophoresis using a 0.8% agarose gel impregnated with 0.5 µg/ml ethidium bromide run at 110V for approximately twenty-five minutes. Agarose gels were place on plate glass and visualized by transilluminating UV light at 312nm wavelength through the glass. Bands corresponding to the liberated PCR product and the cut destination vector were excised using a clean razor blade and purified using a gel extraction kit (Marligen Bioscience #11456-019). One microliter of the purified PCR product and destination vector were analyzed by agarose gel electrophoresis using a 0.8% agarose gel impregnated with 0.5 µg/ml ethidium bromide run at 110V for approximately twenty-five mintues to establish proper ratios for the ligation reaction. Agarose gels were place on plate glass and visualzed by transluminating UV light at 312nm wavelength through the glass. Typically a 1:3 or 1:5 ratio of vector to PCR product was used. Ligation reactions were completed using using T4 DNA Ligase (Promega #M179A) as outline in the product literature and were incubated at 15°C overnight. Two microliters of each ligation reaction were transformed into One Shot® TOP10 chemically competent *E. coli* (Invitrogen #C4040-03) using the heat shock method outline in the product literature (Invitrogen). Following transformation the transformed cells were plated out on LB agar plates containing the appropriate antibiotic selection and incubated at 37°C overnight. Liquid cultures from 8 different colonies present on

the agar plates were grown the following day and plasmids were isolated using the Qiagen Miniprep Kit (#27144). Each purified plasmid was analyzed using restriction digest mapping as outlined above. A plasmid displaying the appropriate predicted bands was sent for sequencing and a glycerol stock was made as outline above.

For large amounts of plasmid (e.g. transfection), the glycerol stock was thawed on ice and used to inoculate a 100 to 250ml liquid culture. The liquid culture was incubated for twelve to sixteen hours at 37°C in a shaking incubator at 250rpm. Plasmid was purified using the Qiagen Hi-Speed MIDI-Prep Kit (Qiagen #12643) as outlined in the product literature.

For the purposes of bacterial expression the expression vector was made as outline above. Once sequencing analysis confirmed the correct sequence had been ligated into the destination vector, BL21-CodonPlus® (DE3) cells were transformed with the expression vector obtained from a Qiagen Hi-Speed MIDI-Prep plasmid purification. Depending on the base pair composition of the gene to be expressed, either BL21-CodonPlus® (DE3)-RIL (Novagen #230245) or BL21-CodonPlus® (DE3)-RP (Novagen #230255) competent cells were used for the transformation. The transformation was performed as outlined in the product literature (Novagen).

C. Constructs mapping the TgGCN5 NLS

Constructs for immunolocalization of TgGCN5 and NLS mutants were based on a *Toxoplasma* expression vector built into *p*miniHXGPRT (AIDS Research and Reference Reagent Program, Division of AIDS, NIAID, National Institutes of Health; http://www.aidsreagent.org). The *Toxoplasma TUB1* promoter was amplified from *p*tubP30-GFP/*sag*CAT (Striepen et al., 1998) using primers 1 and 2 (listed in Table I). These primers incorporate a BamHI site on the 5' end of the PCR product and a polylinker region at the 3' end containing sites for BglII, NdeI, EcoRV, and AvrII. The 3'UTR of *Toxoplasma* DHFR was amplified using primers 3 and 4. The resulting PCR product contained AvrII and NotI sites that were digested for a 3-piece ligation with the *TUB1* amplicon above

Table I: Primers used in PCRs detailed in Vector Construction

No.	primer sequence 5'-3'
1.	BamHI ggatccACACACTTCGTGCAGCATGTGCCCC
2.	AvrII EcoRV NdeI BglII cctaggatatcatatgagatctAAAAGGGAATTCAAGAAAAAATGCC
3.	AvrII cctagg(F)GAAGCTGCCCGTCTCTCGTTTTCC
4.	NotI TGgcggccgcTCTAGACCTAGTGGATCG
5.	NdeI catatgAAAATG(F)GAGACTGTCGAAGTGCCTGCATTCCTCG
6.	NdeI catatgAAAATG(F)AAAGGCGCTCCAACAGGTCTGGGG
7.	NdeI catatgAAAATG(F)AGGAAGCGTGTGAAGCGCGAGTGGGC
8.	NdeI catatgAAAATG(F)GAGTGGGCAAGTGGAGGTCGACTGGAGC
9.	AvrII cctaggTCAGAAACTCCCGAGAGCCTCGACC
10.	BglII agatctAAAATGGTCGTTTTACAACGTCGTGACTGG
11.	BglII agatctAAAATGAGGAAGCGTGTGAAGCGCGTCGTTTTACAACGTCGTGACTGG
12.	AvrII cctaggTTTTTGACACCAGACCAACTGGTAATGG
13.	NdeI catatgGAGACTGTCGAAGTGCCTGCATTCCTCG
14.	NdeI catatgGAGTGGGCAAGTGGAGGTCGACTGG
15.	XmaI cccgggTCAGAAACTCCCGAGAGCCTCGACCTTGG
16.	NheI gctagcGCTCTGTCCGCTGTCTCCATCCCTTGACG
17.	NheI gctagcGAGTGGGCAAGTGGAGGTCGACTGGAGC
18.	NdeI catatgGAGCGCAAGTTGGCCGATCGTCGATCG
19.	XmaI cccgggCTACTGGCCGAAGTTGAAGCCTCCCTGAGG

(F) denotes FLAG epitope tag sequence (5'GACTACAAGGACGACGACGACAAG)

[cut with BamHI (New England Biolabs #R0136S) and AvrII (New England Biolabs #R0174S)] and the *p*miniHXGPRT vector [cut with BamHI (New England Biolabs #R0136S) and NotI (New England Biolabs #R0189S)]. We have designated this expression construct as *p*tubX$_{FLAG}$::HX, the first "X" represents where the gene of interest is to be installed (Figure 8). The "HX" refers to the hypoxanthine-xanthine-guanine phosphoribosyltransferase (HXGPRT) gene expressed by the vector, which is used as a selection marker following transfection. It is important to note that while a C-terminal FLAG tag option is available in this vector, it was used only for the βgal constructs below. An N-terminal FLAG tag was incorporated into the primers generating TgGCN5 sequences to be tested for localization (primers 5-8 in Table I)

The sense primers 5 through 8 with the antisense primer 9 from Table I were used to amplify the different form of TgGCN5 using the pHAT vector (Sullivan and Smith, 2000) as template DNA. PCRs were performed as outlined above (page 60). All of the TgGCN5 derived inserts were cut from the pCR®-Blunt II-TOPO® vector by digestion with NdeI (New England Biolabs #R0111S) and AvrII (New England Biolabs #R0174S) and were ligated into the *p*tubX$_{FLAG}$::HX vector at the same restriction sites. Prior to electroporation, ~50µg of plasmid DNA was linearized by digestion with NotI (New England Biolabs #R0189S) overnight at 37°C. Following digestion, the cut DNA was ethanol precipitated using 2.5 volumes of 100% ice cold ethanol and 0.1 volumes of sodium acetate (3M, pH 5.2). The microfuge tube was mixed by inverting 15-20 times resulting in a visible clump of DNA precipitating out of solution. The microfuge tube was placed at -20°C for at least thirty minutes. The precipitated DNA was spun down using a microfuge in a 4°C cold room at maximum speed (~13,000 X g) for ten minutes. The supernatant was carefully poured off and the DNA pellet was washed with 500µl of 70% ethanol. The DNA was spun down again using a microfuge in a 4°C cold room at maximum speed (~13,000 X g) for five minutes. In the tissue culture hood, supernatant was carefully poured off onto paper towels. A kemwipe twised to a point is used to remove large droplets

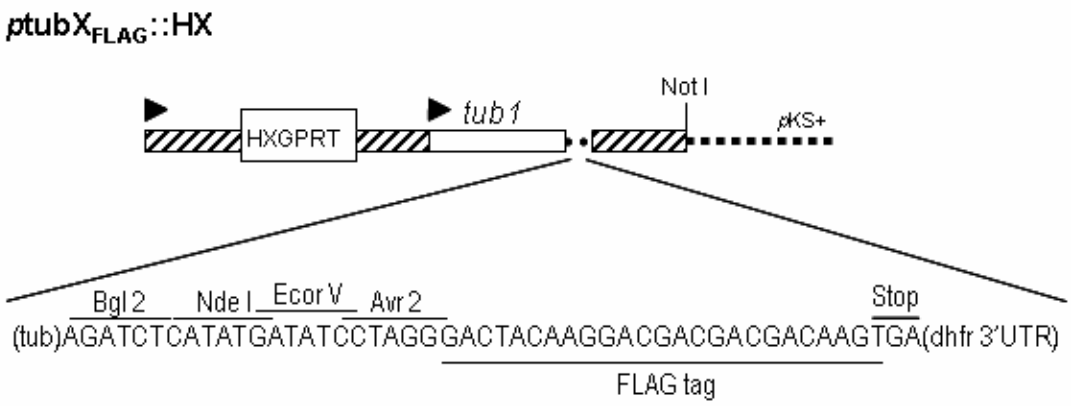


Figure 8: Schematic diagram of the *Toxoplasma* expression vector *p*tubX$_{FLAG}$::HX

Vector was built into the pBluescript KS$^+$ (pKS$^+$; Stratagene #212205) containing an ampicillin selection marker for plasmid propagation in *E. coli*. Striped boxes refer to the untranslated regions of the *Toxoplasma* DHFR-TS gene (dihydrofolic acid reductase-thymidylate synthase; Donald and Roos, 1993). *TUB1* = *Toxoplasma* tubulin promoter (Striepen et al., 1998); HXGPRT = hypoxanthine-xanthine-guanine phosphoribosyltransferase

of ethanol present on the sides of the microfuge tube. The pellet is dried in tissue culture hood for several hours or overnight. The microfuge tube was capped while in hood and transferred to -20 for temporary storage until parasites were ready for electroporation.

To create the beta galactosidase (βgal) constructs verifying the NLS, two PCRs were performed with the sense primers 10 and 11 in conjunction with antisense primer 12 from Table 1 utilizing the pTUB-βgal vector (Seeber and Boothroyd, 1996) as template DNA. The βgal derived inserts were cut from the pCR®-Blunt II-TOPO® vector by digestion with BglII (New England Biolabs #R0144S) and AvrII (New England Biolabs #R0174S) enzymes. The cut PCR products were ligated into the *p*tubX$_{FLAG}$::HX vector at the same two restriction sites. Unlike the TgGCN5 derived inserts, the βgal inserts do not contain a 3' stop codon and therefore utilized the C-terminal FLAG tag incorporated into the *p*tubX$_{FLAG}$::HX vector. Prior to electroporation, ~50µg of plasmid DNA was linearized by digestion with NotI (New England Biolabs #R0189S), ethanol precipitated, and dried under sterile conditions as outlined in the previous paragraph.

D. pGBK-TgGCN5 and pGAD-TgIMPα vectors

For the production of *in vitro* translated protein, the pGADT7 (Clontech #630442) and pGBKT7 (Clontech #630443) vectors were used. PCR was performed with sense primers 13 and 14 and the antisense primer 15 from Table I (page 64) using tub$_F$TgGCN5:HX as template to amplify inserts encoding full-length TgGCN5 and TgGCN5 lacking the first 99 amino acids (Δ99TgGCN5), respectively. The inserts were cut from the pCR®-Blunt II-TOPO® vector using NdeI and XmaI sites incorporated in the primers. Full-length TgGCN5 and Δ99TgGCN5 were ligated into the pGBKT7 vector using the NdeI and XmaI sites.

ΔNLS-TgGCN5 was generated by a ligation of three DNA fragments: two PCR products and the destination vector, pGBKT7. The first PCR product encodes TgGCN5 amino acids 1-93 and was amplified using the primer 13 (contains a NdeI site) and 16 (contains a NheI site) from Table I (page 64) with tub$_F$TgGCN5:HX as template. The

first PCR product was cut from the pCR®-Blunt II-TOPO® vector by digestion with NdeI (New England Biolabs #R0111S) and NheI (New England Biolabs #R0131S). The second PCR fragment encodes amino acids 100-1169 and was generated with primer 17 (contains a NheI site) and primer 15 (contains an XmaI site) from Table I (page 64). pGBKT7 was digested with NdeI (New England Biolabs #R0111S) and XmaI (New England Biolabs #R0180S). The ligation of the two PCR products and the pGBKT7 vector resulted in a plasmid encoding a form of TgGCN5 in which RKRVKR was replaced by an alanine-serine dipeptide (encoded by the NheI site).

TgIMPα insert was generated by single tube RT-PCR using primers 18 and 19 from Table I (page 64) with *Toxoplasma* mRNA purified from RH tachyzoites. The mRNA was purified from RH tachyzoites using the Ambion Poly(A)Pure™ Kit (Fisher #NC9343859). The reverse transcription step was carried out using the Omniscript Reverse Transcriptase (Qiagen #205110) following the single tube PCR protocol as outlined in the product literature (Qiagen). The TgIMPα insert was cut from pCR®-Blunt II-TOPO® by digesion with NdeI (New England Biolabs #R0111S) and XmaI (New England Biolabs #R0180S). The insert was ligated into the pGADT7 vector in frame with an N-terminal hemagglutinin (HA) tag using the same two restriction sites.

E. pET28-ScGCN5 Vector

Sacchromyces cerevisiae GCN5 (ScGCN5) was amplified from a previously made yeast cDNA library by PCR using the sense primer: 5'-ATCGAgctagcGTCACAAAACATCAGATTGAAGAGG-3' and the antisense primer: 5'-ATCGActcgagTAATCAATAAGGTGAGAATATTCAGG-3'. The sense primer encodes a NheI restriction site (lower case) and the antisense primer encodes an XhoI site (lower case). The two restriction sites were used to digest the pCR®-Blunt II-TOPO® vector to liberate the PCR product. The same two restriction sites (NdeI and XhoI) were also used as the insertion sites for ligation into the pET28 bacterial expression vector (Novagen #69865-3) in frame with an N-terminal poly-histidine tag.

F. pET28-TgGCN5 vector

Toxoplasma gondii (TgGCN5) was amplified from the *p*tubX$_{FLAG}$::HX vector containing full length TgGCN5 (*p*tub$_{FLAG}$TgGCN5::HX) by PCR using the sense primer: 5'-ATCGAcatatgGAGACTGTCAAGGTGCCTGCATTCC-3' and the antisense primer: 5'-ATCGAgcggccgcTCAGAAACTCCCGAGAGCCTCGACC-3'. The sense primer encodes an NdeI restriction site (lower case) and the antisense primer encodes a NotI site (lower case). The two restriction sites were used to cut the PCR product out of the pCR®-Blunt II-TOPO® vector and were also use as the insertion sites for ligation into the pET28 bacterial expression vector (Novagen #69865-3) in frame with an N-terminal poly-histidine tag.

III. Bioinformatics

Sequencing alignments and phylogeny trees of importin α homologues were generated using AlignX, a component of the Vector NTI Advance 9.0 (Informax) program. Other protein sequencing alignments were performed using the ClustalW program (http://www.ch.embnet.org/software/ClustalW.html) where indicated. Protein motif searches were performed using the Pfam database (ver. 14.0; http://pfam.wustl.edu/) and/or PROSITE database (ver. 18.35; http://us.expasy.org/prosite/).

IV. Biochemical Techniques

A. Immunofluorescence assay

For immunofluorescence assays (IFA), 12 well plates containing coverslips were made as outlined above (Section I-A, page 53). Prior to infection, the lid of the well plate(s) were labeled with the parasite line to be infected into each well. The old HC (host cell) media in the wells was aspirated and replaced with 1.0ml of PA media (parasite media) which consists of DMEM, 1% heat-inactivated fetal bovine serum, and 25µg/L gentamicin). The T-25cm^2 flask containing the infected monolayer was scrapped with a sterile spatula (Fisher #08-773-2) to completely rupture any remaining parasite vacuoles and to suspend the parasites into the media. Each well was inoculated with 200µl of

parasites from the T-25cm^2 flask. To maintain the parasites in culture, 1.0ml of parasites from the scraped flask was also is transferred to a T-25cm^2 flask containing a new confluent HFF monolayer as mentioned above (Section I-A, page 54).

The infected 12 well plates were incubated at 37°C in a humidified 5% CO_2 environment for eighteen to twenty-four hours. Parasite infection of the well plate was monitored using a Leica DM IL inverted microscope at 400X magnification with a phase filter. Once the infection had progressed to the point where each parasite vacuole contained an average of 4-8 parasites per vacuole, the well plate(s) were removed from the tissue culture incubator and processed for IFA.

The PA media from each well was aspirated and each well was washed with 1.0ml of PBS three times. Cultures were fixed by adding 1.0ml of ice-cold methanol to each well and incubating the well plate at -20°C for ten minutes. The methanol was aspirated and each well was washed with 1.0ml PBS three times. Cells were blocked with 0.5ml of PBS containing 3% Fraction-V BSA (Sigma #A3059), 5% goat serum (Invitrogen #10000C), and 1% fish gelatin (Sigma #G7765) for one hour at 25°C or overnight at 4°C. After aspirating the blocking buffer, 0.5ml of polyclonal anti-FLAG (Sigma #F7425) at 1:1000 in 3% BSA (Sigma #A3059) in PBS was added to each well and incubated for one hour at 25°C. After 3 washes with 1.0ml PBS, 0.5ml of goat anti-rabbit Alexa 488 (Molecular Probes #A11034) was added to each well at 1:3000 in 3% BSA (Sigma #A3059) in PBS. Immediately after adding the Alexa conjugated antibody, the well plate was wrapped in aluminum foil and incubated for one hour in the dark at 25°C. Each well plate was washed with 1.0ml PBS and then 0.5ml of 0.3µM DAPI (Invitrogen #D1306) was added to each well and incubated, wrapped in foil for five minutes at 25°C. Each well was washed three times with 1.0ml of PBS. The last 1.0ml of PBS was left in each well to aid in removing the coverslips from the wells.

A single drop of mounting reagent (~10µl) containing 50% glycerol with Mowiol 4-88 (Calbiochem #81381) and DABCO (Sigma #10981) to retard

photobleaching was placed on a microscope slide. The next paragraph contains the protocol for making the mounting reagent. The coverslips were removed from each well using non-sterile flat tweezers (Fisher #09-753-50) and placed inverted on top of the drop of mounting agent on the microscope slide. Slides were air dried in the dark for approximately thirty minutes. Slides were viewed with a Leica DMLB scope with a 100X HCX Plan Apo oil immersion objective. Images were captured using a monochrome SPOT-RTSE (Model 12) Camera and Spot Diagnostic Software 4.0.9 and pseudocolored using Adobe Photoshop 7.0.

The mounting reagent was made by dissolving 2.4 grams of Mowiol 4-88 (Calbiochem #81381) in to 12ml of 50% glycerol in a 50ml conical (Fisher #14-959-49A). The solution was mixed well by vigorous vortexing. After the Mowiol 4-88 has dissolved, 0.5 grams of DABCO (Sigma #10981) was added. The conical was taped down on the plate form of a shaking incubator. The solution was shaken for two hours at room temperature to dissolve the Mowiol 4-88 and DABCO. The conical was removed from the shaker and 12ml of 200mM Tris-HCl (pH 8.5) was added to the conical. The solution was incubated at 50°C with accasional mixing to until the Mowiol 4-88 was completely dissolved (approximately three hours). The solution was transferred to a Beckman tube and spun at 5000 x g for fifteen minutes at room temperature. The mounting reagent was aliquoted out into 1.5ml microfuge tubes and stored at -20°C. Prior to use, the solution was thawed by hand warming an aliquot.

B. Production of TgGCN5 antiserum

The pET19-TgGCN3000 vector had been previously generated in Sullivan lab and had been transformed into BL21(DE3) *E. coli* cells. The pET19-TgGCN3000 vector contains a fragment encoding 30kD of the C-terminal end of TgGCN5 (amino acids 976-1169) in-frame with an N-terminal polyhistidine tag.

A glycerol stock was used to inoculate 100ml of LB media which was grown at 37°C shaking at 250rpm until the OD_{600} reading was apprimately 0.5 (usually around two to three hours). Bacterial protein expression in the 100ml

culture was induced by adding 100mM IPTG stock solution to a final concentration of 0.5mM. The culture was placed back at 37°C for three hours shaking at 250rpm. At the end of induction, the 100ml induction was spun down in large Beckman centrifuge at 4°C at 10,000rpm for ten minutes to pellet the bacteria. After centrifugation, the media was poured off and the pellet was resuspended into 10ml of cold PBS and transferred to a 25ml Beckman tube. The reususpended pellet was spun down at 4°C at 10,000Xg for ten minutes. Following spin, the PBS was poured off and the pellets were stored at -20°C. The histidine fustion protein was purified as outlined below (Section D, page 75). The purified protein was lyophilized and sent to Pocono Rabbit Farms & Laboratories (Canadensis, PA). The purified protein was used as antigen to produce polyclonal antisera in rabbits.

Upon receiving the antiserum from Pocono Rabbit Farms & Laboratories, it was tested against recombinant antigen in a western blot (Figure 9). Approxmiately 0.02µg and 0.2µg of antigen were resolved on a NuPAGE gel (Invitrogen #NP0335BOX) and transferred to a PVDF membrane (Invitrogen #LC2005). The membrane was blocked overnight at 4°C in TBST (20mM Tris, 150mM NaCl, 0.1% v/v Tween-20) containing 5% milk. The membrane was cut in half and one half was incubated with TgGCN5 antiserum at 1:10,000 in TBST containing 5% milk for one hour and the other half was incubated for one hour in pre-immune rabbit serum obtained prior to antibody production also diluted 1:10,000 in TBST containing 5% milk. Both blots were washed 3 times, 5 minutes each with TBST. The secondary antibody, goat derived anti-rabbit conjugated to HRP (horseradish peroxidase; Amersham #NA934), was incubated with both membranes at 1:2500 dilution in 5% Milk-TBST for one hour, washed 3 times, five minutes each with TBST.

The blots were treated with 1.5ml of Amersham's ECL™ detection reagent (#RPN2209). Excess reagent was removed by gently shaking the blot over paper towels. The blot was then wrapped in cellophane and taped down in a metal film tray. In a photography dark room, the blot was exposed to High Senstive Blue photographic film (RPS Imaging #33-0810). Following exposure,

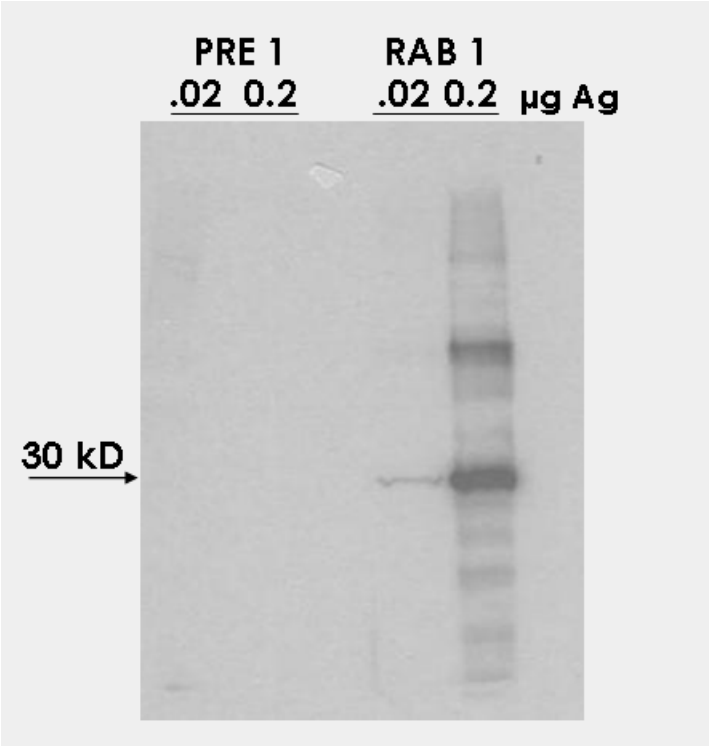


Figure 9: Evalulation of TgGCN5 antisera

Western blot shows 1:10,000 dilution of antiserum (RAB 1) reacts specifically to 0.02 and 0.2µg of TgGCN5 antigen (Ag; ~30kD). Pre-immune (PRE) serum was tested as a control. Secondary antibody (goat derived anti-rabbit, Amersham #NA934) was used at 1:2500. Detection was performed using Amersham ECL reagent (#RPN2209) with High Sensitive Blue photographic film (RPS Imaging #33-0810). kD = kilodaltons

the film was removed and run through an automated developer. As shown in Figure 9, the TgGCN5 antiserum recognizes <0.1µg of recombinant antigen at a dilution of 1:10,000, while pre-bleed sera is non-reactive.

C. In vitro translation and co-immunoprecipitation

In vitro translated proteins representing full-length TgGCN5, Δ99TgGCN5, ΔNLS-TgGCN5, and *Toxoplasma* importin α (TgIMPα) were generated using the T_NT® T7 Coupled Reticulocyte Lysate System (Promega #L4610) in the presence of Redivue™ L-[^{35}S]methionine (Amersham Pharmacia #AG1094) and RNase Out (Invitrogen #10777-019). Translation reaction mixtures with a total volume of 50µl were assembled and incubated as described by manufacturer (Promega). Co-immunoprecipitations (CoIP) were performed by combining 20µl of the TgGCN5 *in vitro* translation reaction with 20µl of the TgIMPα *in vitro* translation reaction in a fresh microfuge tube. The mixture was mixed by gently tapping the side of the tube and subsequently incubated at 25°C for one hour. Initially, CoIPs were performed by adding 10µl of a polyclonal HA antibody (Santa Cruz #SC-7392) and 10µl of monoclonal cMyc antibody (Santa Cruz #SC-40). When these antibodies failed to provide positive results, subsequent CoIPs were performed by adding 1.0µl of TgGCN5 polyclonal antiserum (page 71) to each mixtures, mixing by tapping the tube, and incubating at 25°C for one hour. EZ-View Protein A agarose (Sigma #P6486) was washed two times with PBS and 10µl washed agarose slurry was added to all CoIPs, mixed by tapping tube, and incubated at 25°C for one hour.

A similar protocol was followed for immunoprecipitation of single proteins using 10µl of completed translation reaction and 1µl of TgGCN5 polyclonal antisera for TgGCN5 proteins and 10µl of a polyclonal HA antibody (Santa Cruz #SC-7392) for TgIMPα. EZ-View Protein A agarose (Sigma #P6486) was washed twice with PBS and 10µl of washed agarose was added to immunoprecipitations, mixed by tapping tube, and incubated at 25°C for one hour

For both co-immunoprecipitations and immunoprecipitations, the protein-A agarose was washed 5 times with 200µl of CoIP Buffer 1 [10mM HEPES (pH

7.7), 500mM KCl, 0.1mM CaCl2, 1mM MgCl2, 50mM Sucrose, 0.5% Nonidet P-40 (nonionic detergent)] and twice with 300µl of CoIP Buffer 2 (10mM HEPES (pH 7.7), 100mM KCl, 0.1mM CaCl2 , 1mM MgCl2, and 50mM Sucrose). Bound proteins were eluted by boiling the protiein A agarose in NuPAGE loading dye (Invitrogen #NP0007) containing 5.0% β-mercaptoethanol (Sigma #M7154) and resolved on 4-12% NuPAGE Bis-Tris SDS gels (Invitrogen #NP0335BOX). Ten microliters of SeeBlue® Plus2 Protein Standard (Invitrogen #LC5925) was run on the gel and was used to estimate protein molecular weight after gel electrophoresis. Gels were fixed in an aqueous solution of 5% isopropanol and 5% glacial acetic acid overnight and rinsed in running distilled water for one hour. Gels were incubated in Autoflour (National Diagnostics #LS-315) for two hours, wrapped in gel wrap(The Gel Company #EJA331-050), and air dried overnight. The excess gel wrap was trimmed away and the gel was taped down in a metal film tray. In a photography dark room, a piece of High Sensitive Blue photographic film (RPS Imaging #33-0810) was placed over the gel and the film tray was closed and locked. The film tray was placed at -80°C for twenty-four hours. Following exposure, the film tray was removed and allowed to warm to room temperature. In a photography dark room, the film was removed and run through an automated developer.

D. Bacterial expression and purification of ScGCN5 and TgGCN5

The glycerol stocks of the BL21-CodonPlus® (DE3)-RIL *E. coli* cells containing the pET28-ScGCN5 vector and BL21-CodonPlus® (DE3)-RP *E. coli* cells containing the pET28-TgGCN5 generated following transformation were thawed on ice and 10µl of each was used in inoculate 4ml of LB with 50µg/ml chloramphenicol (selection for proprietary plasmid coding for extra tRNAs) and 35µg/ml kanamycin (pET28 vector selection).

For both ScGCN5 and TgGCN5, The liquid culture was incubated overnight at 37°C shaking at 250rpm. A 250ml flask containing 100ml of LB media with 35µg/ml kanamycin was inoculated with 1.0ml of the overnight culture. The 100ml culture was grown at 37°C shaking at 250rpm until the OD_{600}

reading was approximately 0.5 (usually around two to three hours). Before inducing protein expression, a 1.0ml sample of the 100ml culture was obtained and placed into 1.5ml microtube. The 1.0ml sample was centrifuged at maximum speed for a few minutes to pellet the bacteria. The LB media was aspirated off, and the record OD_{600} reading of culture was written on the side of tube, and the bacterial pellet was stored at -20°C.

Bacterial protein expression in the 100ml culture was induced by adding 100mM IPTG stock solution to a final concentration of 0.5mM. The bacteria culture of cells expressing ScGCN5 was placed back at 37°C for three hours shaking at 250rpm. The bacteria culture of cells expressing TgGCN5 was incubated at 15°C for sixteen hours shaking at 250rpm. At end of induction, two 1.0 samples of the induced culture were obtained into 2 separate 1.5ml microtubes. One of the samples was centrifuged at maximum speed for a few minutes to pellet the bacteria. The media was aspirated off and stored at -20°C for subsequent analysis. The second tube was used to determine the OD_{600} of the induced culture. The sample was too turbid to perform a direct reading. Therefore, it was diluted 1:10 by pipeting 100µl of the culture into the cuvette and adding 900µl of H_2O. The dilution was mixed by inversion and OD_{600} reading was obtained and recorded.

The remaining 100ml induction was spun down in large Beckman centrifuge at 4°C at 10,000rpm for ten minutes to pellet the bacteria. After centrifugation, the media was poured off and the pellet was resuspended into 10ml of cold PBS and transferred to a 25ml Beckman tube. The resuspended pellet was spun down at 4°C at 10,000Xg for ten minutes. Following spin, the PBS was poured off and the pellets were stored at -20°C.

The two 1.0ml pellets were used to verify successful induction before attempting to purify the protein from the bacteria. The two 1.0ml samples were thawed one ice for five minutes and were resuspended each in 100µl of COLD PBS. The bacterial pellets were sonicated using a microprobe 3 times for 15 seconds with a 30 second recovery on ice between each pulse. The sonicated samples were spun down using a microfuge in the 4°C cold room at maximum

speed (~13,000 X g) for ten minutes. The supernatant (~100µl) was transferred into pre-chilled microfuge tubes and labeled as the soluble fraction.

For SDS-PAGE analysis, 21.0µl of the soluble fraction sample was mixed with 7.5µl of NuPAGE Loading Dye (Invitrogen #NP0007) and 1.5µl of beta-mercaptoethanol (β-Me; Sigma #M7154). The samples were heated at 70°C for ten minutes and the entire sample was loaded onto a NuPAGE gel (Invitrogen #NP0335BOX). The insoluble fraction (pellet) was resuspended in 25.0µl of NuPAGE Loading Dye (Invitrogen #NP0007), 5.0µl of β-Me (Sigma #M7154) and 70.0µl of ddH_2O. If the sample was very viscous, it was sonicated 3 times for 15 seconds with a 30 second recovery on ice between each pulse. Each sample was incubated at 70°C for ten minuntes and 30µl was loaded onto the NuPAGE gel (Invitrogen #NP0335BOX). The NuPAGE gel was run using MOPS buffer (Invitrogen #NP0001) according to manufacture's (Invitrogen) directions. Ten microliters of SeeBlue® Plus2 Protein Standard (Invitrogen #LC5925) was run on the gel and was used to estimate protein molecular weight after gel electrophoresis.

Following gel electrophoresis, the gel was removed from the gel mold and was washed in deionized water three times, for ten minutes each. To visualize protein in the gel, the gel was stained overnight with SimplyBlue™ (Invitrogen #LC6060). The gel was destained by washing in deionized water three times, for ten minutes each. If analysis of the 1.0ml samples indicated that induction was successful in producing protein, the large bacterial pellet was processed for protein purification by virtue of the polyhistidine fusion tag using nickel resin.

The frozen bacterial pellet from the 100ml culture was thawed on ice for fifteen minutes. The pellet resuspended in 4ml of lysis buffer (50mM NaH2PO4 (pH 8.0), 300mM NaCl, 10mM of imidazole) with 4mg lysozyme (Sigma #L6876) and 80µl of His-Tag Protease Inhibitor Cocktail (Sigma #P8849). The lystate was incubated on ice for thirty minutes with occasional mixing every five to seven minutes. The lysate was transferred to a 15ml conical tube (Fisher #14-959-70C) and sonicated 6 times at 15 second bursts with a 30 second recovery on ice between each burst. Following sonication, lysate was transferred into centrifuge

tube and spun at 4°C and 10,000 X g for thirty minutes. The supernatant was poured into a 15ml conical tube and placed on ice. An additional 40µl of His-Tag Protease Inhibitor Cocktail (Sigma #P8849) was added to the lysate in the conical and mix thoroughly. Approxmiately 1.0ml of Ni-NTA agarose slurry (Qiagen #30210) was added to the conical and mixed by inversion. The conical was taped to the platform of shaking incubator in a 4°C cold room and was incubated at 15°C one hour, shaking at 200rpm.

During the shaking incubation, a column was prepared by packing a small wad of glass wool into the bottom of a 5ml syringe. The syringe was clamped to a ring stand in the cold room. After incubation, the contents of the conical was poured slowly into the syringe column and the column was pack by gravity. Once the supernatant flowed through the column, the column was washed twice with 4.0ml of wash buffer (50mM NaH2PO4 (pH 8.0), 300mM NaCl, 50mM of imidazole). After the washes, protein was eluted off the column using the following buffers:

Buffer A: 50mM NaH_2PO_4 (pH 8.0), 300mM NaCl, 100mM imidazole

Buffer B: 50mM NaH_2PO_4 (pH 8.0), 300mM NaCl, 200mM imidazole

Buffer C: 50mM NaH_2PO_4 (pH 8.0), 300mM NaCl, 300mM imidazole

Protein was eluted with two treatments of 500µl of buffer A and B and three treatments of buffer C. Each elution was captured in 1.5ml microfuge tubes. An additional 5µl of protease inhibitors (Sigma #P8849) was added to the all elutions. Each elutions was analyzed by SDS-PAGE using NuPAGE gels (Invitrogen #NP0335BOX). For PAGE analysis, 21.0µl of each elution was mixed with 7.5µl of NuPAGE Loading Dye (Invitrogen #NP0007) and 1.5µl of beta-mercaptoethanol (β-Me) and incubated at 70°C for ten minutes. The NuPAGE gel was run using MOPS buffer (Invitrogen #NP0001) according to manufacturer's (Invitrogen) directions. Ten microliters of SeeBlue® Plus2 Protein Standard (Invitrogen #LC5925) was run on the gel and was used to estimate protein molecular weight after gel electrophoresis.

Following gel electrophoresis, the gel was removed from mold and was washed in deionized water three time for ten minutes each. To visualize protein

in the gel, the gel was stained with SimplyBlue™ (Invitrogen #LC6060) overnight. The gel was destained by washing in deionized water three times for ten minutes each. The elutions containing recombinant protein were pooled and concentrated using an Amicon Ultra-4 Centrifugal Filter Device (Fisher #UFC801008) as outlined in the product literature. The concentrated elutions were transferred to a Slide-A-Lyzer® dialysis cassette (Pierce #PI66380) for dialysis to remove the imidazole. The dialysis cassettes was placed in 1.0L of 50mM Tris·HCl (pH 8.0) at 4°C for fifteen minutes with gentle stirring using a magnetic stir plate. The cassette was transferred to another 1.0L of 50mM Tris·HCl (pH 8.0) at 4°C for forty-five minutes with gentle stirring using a magnetic stir plate. The sample was removed from the dialysis cassette and quantitated using a standard Bradford assay (Fisher #PI23225).

After protein concentration has been determined, approximately two samples of 1.0µg of ScGCN5 were resolved on a NuPAGe gel. The gel was cut and one sample was stained with SimplyBlue™ (Invitrogen #LC6060) and the other gel was transferred to a PVDF membrane (Invitrogen #LC2005) for western blot detection. After the transfer was complete, the blot was blocked overnight at 4°C in TBST (20mM Tris, 150mM NaCl, 0.1% v/v Tween-20) containing 5% milk. The following day, the membrane was incubated with monoclonal anti-histidine antibody (Clontech #PT3359-2) at a dilution of 1:5000 in TBST containing 5% milk for one hour then washed 3 times, five minutes each with TBST. The secondary antibody, goat derived anti-mouse conjugated to HRP (horseradish peroxidase; Amersham #NA931), was incubated with membrane at a 1:5000 dilution in 5% Milk-TBST for one hour, washed 3 times, five minutes each with TBST.

In the case of TgGCN5, two samples of 500ng of protein were resolved on a NuPAGE gel. The gel was cut and one sample was stained with SimplyBlue™ (Invitrogen #LC6060) and the other gel was transferred to a PVDF membrane (Invitrogen #LC2005) for western blot detection. After the transfer was complete, the blot was blocked overnight at 4°C in TBST containing 5% milk. The following day, the membrane was incubated with TgGCN5 antiserum at 1:10,000 in TBST

containing 5% milk for one hour then washed 3 times, five minutes each with TBST. The secondary antibody, goat derived anti-rabbit conjugated to HRP (Amersham #NA934), was incubated with membrane at 1:2500 dilution in 5% Milk-TBST for one hour, washed 3 times, five minutes each with TBST.

The ScGCN5 and TgGCN5 blots were treated with 1.5ml of Amersham's ECL™ detection reagent (#RPN2209). Excess reagent was removed by gently shaking the blot over paper towels. The blot was then wrapped in cellophane and taped down in a metal film tray. In a photography dark room, the blot was exposed to a High Sensitive Blue photographic film (RPS Imaging #33-0810). Following exposure, the film was removed and run through an automated developer. SeeBlue® Plus2 Protein Standard (Invitrogen #LC5925) transferred to the membrane were used to estimate protein molecular weight after gel electrophoresis.

E. FLAG affinity purification of $_{FLAG}$TgGCN5 and $_{FLAG}\Delta N_T$TgGCN5 from parasites

Parasites overexpressing FLAG tagged TgGCN5 ($_{FLAG}$TgGCN5) and/or TgGCN5 lacking the N-terminal extension ($_{FLAG}\Delta N_T$TgGCN5) were harvested from infected T-150cm^2 flasks by scraping the monolayer with a sterile spatula (Fisher #08-773-2) to completely rupture any remaining parasite vacuoles and to suspend the parasites into the media. The suspended parasites were filtered under non-sterile conditions and washed with PBS as previously outlined (Section I-B, page 55). Following PBS washing, the parasite pellet was resuspended in 1.0ml of lysis buffer [50mM Tris-HCl (pH 7.4) 150mM NaCl, 1mM EDTA, 1% Triton X-100, and 20µl of mammalian protease inhibitors (Sigma #P8340)]. Lysate was sonicated 3 times for 15 seconds with 30 second recovery on ice between each pulse. The lysate was placed on rocker at 4°C for thirty minutes. Lysates were spun at maximum speed (~13,000 X g) at 4°C for ten minutes. The supernatant was transferred to a new microtube and 60µl of EZview™ Red Protein A Affinity Gel (Sigma #P6486) was added to the supernatant. The mixture was placed at 4°C on a rocker for thirty minutes. The purposes of this step was to pre-clear the lysate by removing "sticky" proteins

that non-specifically interact with affinity resins. The lysate was spun at 8200 X g for one minute and the supernatant was transferred to a new tube with 60µl of EZview™ Red ANTI-FLAG® M2 Affinity Gel (Sigma #F2426) and placed on rocker overnight at 4°C.

The next day, affinity gel was washed 3 times with 750µl of wash buffer [50mM Tris-HCl (pH 7.4), 300mM NaCl, and 1mM EDTA) plus 10µl/ml of fresh mammalian protease inhibitors (Sigma #P8340), and then washed twice with 750µl of HAT assay buffer (50mM Tris-HCl (pH 8.0), 50mM KCl, 0.1mM EDTA, 1.0mM DTT, and 10% glycerol). During the second wash with 750µl of HAT assay buffer, 250µl (~10µl packed resin) was transferred into a separate tube for western blot analysis to insure that immunoprecipitation was successful. The remaining 500µl (~20µl packed resin) was used in a HAT assay (see next section). Samples were spun down at 8200 X g for one minute and the HAT buffer was aspirated.

Twenty microliters of 2X NuPAGE Loading Dye (Invitrogen #NP0007) containing 2.5µl β-mercaptoethanol (β-Me; Sigma #M7154) was added to the 10µl of packed resin for western blot analysis. Samples were incubated at 70°C for ten minutes and then loaded on a 4-12% NuPAGE Bis-Tris (Invitrogen #NP0335BOX) gel and run in MOPS buffer according to the manufacture's (Invitrogen) directions. SeeBlue® Plus2 Protein Standard (Invitrogen #LC5925) was used to estimate protein molecular weight after gel electrophoresis.

Samples were then transferred to a PVDF membrane (Invitrogen #LC2005) and blocked overnight at 4°C in TBST (20mM Tris, 150mM NaCl, 0.1% v/v Tween-20) containing 5% milk. The following day, the membrane was incubated with polyclonal anti-FLAG (Sigma #F7425) at 1:1000 in TBST containing 5% milk for one hour then washed 3 times, five minutes each with TBST. The secondary antibody, goat derived anti-rabbit conjugated to HRP (horseradish peroxidase; Amersham #NA934) was incubated with membrane at 1:2500 dilution in 5% Milk-TBST for one hour, washed 3 times, five minutes each with TBST.

The blot was treated with 1.5ml of Amersham's ECL™ detection reagent (#RPN2209). Excess reagent was removed by gently shaking the blot over paper towels. The blot was then wrapped in cellophane and taped down in a metal film tray. In a photography dark room, the blot was exposed to a High Sensitive Blue photographic film (RPS Imaging #33-0810). Following exposure, the film was removed and run through an automated developer. SeeBlue® Plus2 Protein Standard (Invitrogen #LC5925) transferred to the membrane were used to estimate protein molecular weight after gel electrophoresis.

F. Radioactive histone acetylase assays

The *in vitro* HAT assays used in the thesis were based on those originally reported by Browell et al. (1996). All radioactive HAT assays were 50µl reactions containing 2µl of ^{3}H-Acetyl-CoA (Amersham #TRK688) and 4µl of chicken erythrocyte core histones (1µg/µl in H_2O; Upstate #13-107). The amount of 5X HAT assay buffer (250mM Tris-HCl (pH 8.0), 250mM KCl, 0.5mM EDTA, 5.0mM DTT, and 50% glycerol; modified from Fan et al., 2004b), and deionized water depended on the protein source.

HAT assays performed with recombinant ScGCN5 contained ~1.0µg of protein, 10µl of 5X HAT assay buffer, and water sufficient to 50µl. HAT assays performed with recombinant TgGCN5 generated in bacteria contained ~500ng of protein, 10µl of 5X HAT assay buffer, and water sufficient to 50µl. Assays to assess the activity of FLAG tagged proteins immunoprecipitated from parasites contained ~20µl of packed resin from a FLAG affinity purification mixed with 6µl of 5X HAT assay buffer, and 18µl of deionized water. A negative control reaction was made with 10µl of 5X HAT assay buffer and 34µl of water.

All of the reactions were mixed thoroughly and incubated at 30°C for sixty minutes. Reactions were remixed by gently tapping the sides of the tubes every ten minutes. Reactions were stopped by adding 15µl of 2X NuPAGE loading dye (Invitrogen #NP0007) containing 2.5µl of β-Me (Sigma #M7154) and incubated at 70°C for ten minutes before resolving on a NuPAGE 10% Bis-Tris gel (Invitrogen #NP0315BOX) using MES buffer according to the manufacture's (Invitrogen)

directions. Ten microliters of SeeBlue® Plus2 Protein Standard (Invitrogen #LC5925) was run on the gel and was used to estimate protein molecular weight after gel electrophoresis.

The gel was removed from the mold and was washed in deionized water 3 times, five minutes per wash. Gels were fixed in an aqueous solution of 5% isopropanol and 5% glacial acetic acid overnight and rinsed in running distilled water for one hour. Gels were incubated in Autoflour (National Diagnostics #LS-315) for two hours. A piece of Whatman paper, slighter larger than the gel, was added to tray with Autoflour and gently agitated for an additional five minutes. The gel was placed on top of the soaked Whatman paper and sandwiched between two pieces of gel wrap (The Gel Company #EJA331-050) and air dried overnight. The excess gel wrap was trimmed away and the gel was taped down in a metal film tray. In a photography dark room, a piece of High Sensitive Blue photographic film (RPS Imaging #33-0810) was placed over the gel and the film tray was closed and locked. The film tray was placed at -80°C for twenty-four hours to one week. Following exposure, the film tray was removed and allowed to warm to room temperature. In a photography dark room, the film was removed and run through an automated developer.

G. Non-Radioactive histone acetylase assays

Non-radioactive HAT assays were performed the same as the radioactive assays outlined in the previous section, with the following differences: 1µl of 1mM acetyl CoA and 1µl of recombinant histone H3 (1µg/µl in water) were used in place of tritiated acetyl CoA and core histones. The 1mM acetyl CoA was generated by dissolving acetyl CoA sodium salt (Sigma #A2056) in 10mM sodium acetate (pH 4.5) to a final concentration of 1mM. The recombinant histone H3 (Upstate #14-411) is *Xenopus laevis* Histone H3 generated in *E. coli*. The total volume of each reaction remained 50µl and the amount of water added to each reaction was adjusted accordingly.

The same positive and negative controls were used in the non-radioactive assay but with histone H3 substrate instead of core histones. Reactions were

resolved via NuPAGE 10% Bis-Tris gel (Invitrogen #NP0315BOX) using MES buffer according to the manufacture's (Invitrogen) directions. Ten microliters of SeeBlue® Plus2 Protein Standard (Invitrogen #LC5925) was run on the gel and was used to estimate protein molecular weight after gel electrophoresis.

Following electrophoresis, the proteins were transferred to a PVDF membrane (Invitrogen #LC2005) using an XCell II™ Blot Module (Invitrogen #EI9051) with 1X NuPAGE® Transfer Buffer (Invitrogen #NP0006-1) at 25V for two hours. After transfer, the membrane was blocked overnight at 4°C in TBST (20mM Tris, 150mM NaCl, 0.1% v/v Tween-20) containing 5% non-fat dry milk.

The following day, the membrane was incubated with a polyclonal antibody against acetylation at lysine 14 of histone H3 Polyclonal [α-AcH3(K14), Upstate #07-353] at 1:1000 in TBST containing 5% milk for one hour then washed 3 times, five minutes each with TBST. The secondary antibody, goat derived anti-rabbit conjugated to HRP (horseradish peroxidase; Amersham #NA934) was incubated with membrane at 1:2500 dilution in 5% Milk-TBST for one hour, washed 3 times, five minutes each with TBST. The blot was treated with 1.5ml of Amersham's ECL™ detection reagent (#RPN2209). Excess reagent was removed by gently shaking the blot over paper towels. The blot was then wrapped in cellophane and taped down in a metal film tray. In a photography dark room, the blot was exposed to a High Sensitive Blue photographic film (RPS Imaging #33-0810). Following exposure, the film was removed and run through an automated developer. SeeBlue® Plus2 Protein Standard (Invitrogen #LC5925) transferred to the membrane were used to estimate protein molecular weight after gel electrophoresis.

Following western blot detection, the blot was removed from the cellophane and placed in tray with ~25ml of TBST (enough to cover the blot well). The blot was washed twice, five minutes each, with TBST for at room temperature. The blot was covered with stripping solution and incubated at 50 to 55°C for thirty minutes with gentle rocking. The stripping buffer was made immediately prior to using it by combining stripping buffer pre-mix (67.5mM Tris (pH 6.7) and 2% SDS) with β-Me to a final concentration of 100mM.

Commercially available β-Me is 14.3M (Sigma #M7154). Therefore, to make 50ml of strip solution, 350µl of stock β-Me was added to 50ml of the strip solution pre-mix. After incubation, the strip solution was poured down the drain with lots of water. The blot was washed twice with TBST for five minutes each at room temperature. The blot was blocked with 5% non-fat dry milk dissolved in TBST for one hour at room temp or overnight at 4°C. Blots were stripped no more than two times. Stripping a blot more than twice removed too much protein from the blot to provide adequate detection by subsequent staining with additional antibodies.

After blocking the membrane, it was incubated with either a polyclonal antibody against acetylation at lysine 9 of histone H3 [α-AcH3(K9), Upstate #07-352] or polyclonal antibody recognizing acetylation at lysine 9 and 18 of histone H3 [α-AcH3(K9/K18), Upstate #07-593]. Both antibodies were used at a dilution of 1:1000 in TBST containing 5% milk for one hour then washed 3 times, five minutes each with TBST. The secondary antibody, goat derived anti-rabbit conjugated to HRP (horseradish peroxidase; Amersham #NA934) was incubated with membrane at 1:2500 dilution in 5% Milk-TBST for one hour, washed 3 times, five minutes each with TBST. The blot was treated with 1.5ml of Amersham's ECL™ detection reagent (#RPN2209). Excess reagent was removed by gently shaking the blot over paper towels. The blot was then wrapped in cellophane and taped down in a metal film tray. In a photography dark room, the blot was exposed to a High Sensitive Blue photographic film (RPS Imaging #33-0810). Following exposure, the film was removed and run through an automated developer. SeeBlue® Plus2 Protein Standard (Invitrogen #LC5925) transferred to the membrane were used to estimate protein molecular weight after gel electrophoresis.

H. FLAG co-immunoprecipitations on $_{FLAG}$TgGCN5 and $_{FLAG}\Delta N_T$TgGCN5 transgenic parasites

Figure 10 shows the general scheme used by Dr. Mohamed-Ali Hakimi's laboratory at the French National Centre for Scientific Research in Grenoble,

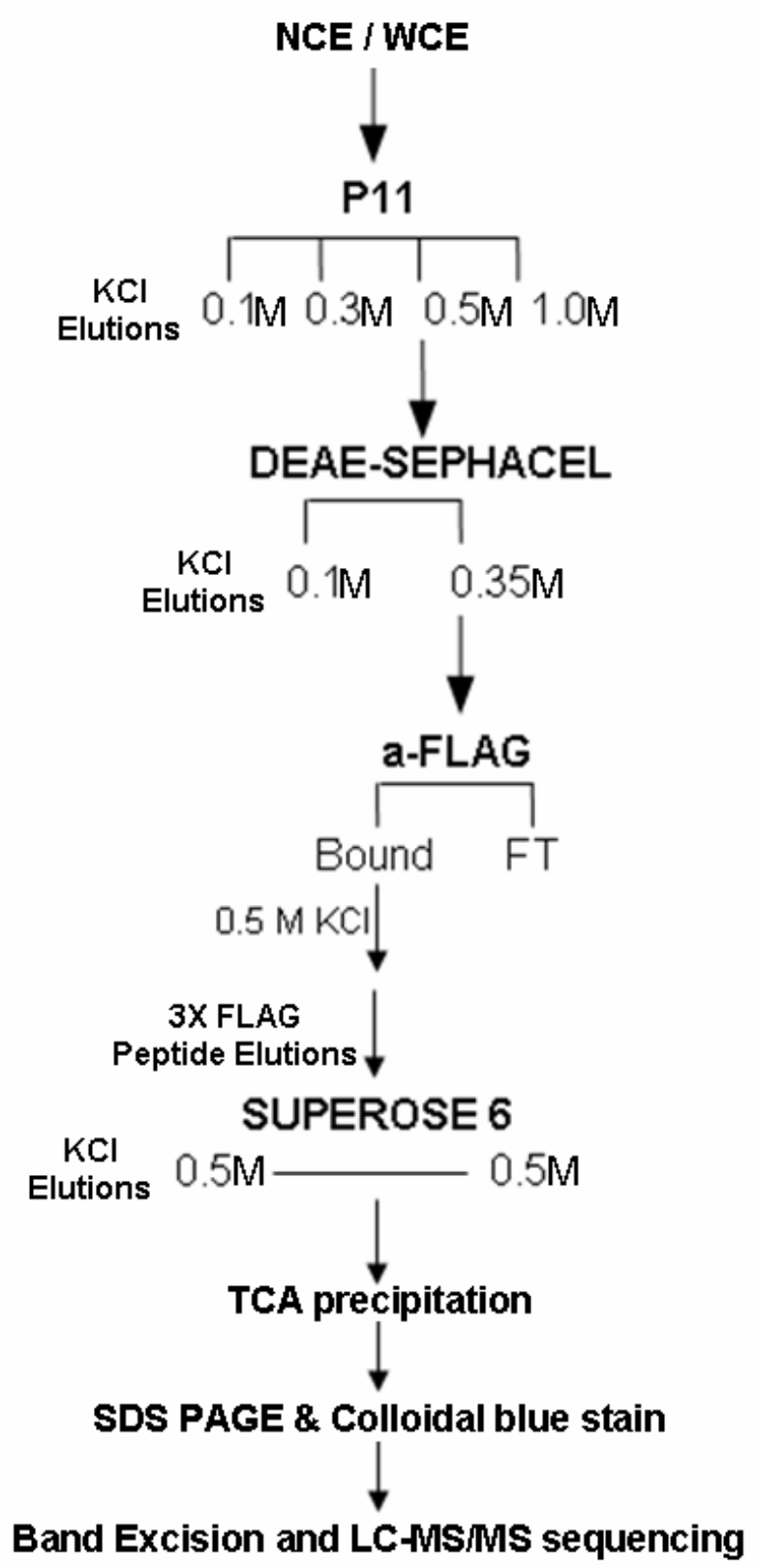

NCE / WCE
P11
KCl Elutions
0.1M
0.3M
0.5M
1.0M
DEAE-SEPHACEL
KCl Elutions
0.1M
0.35M
a-FLAG
Bound
FT
0.5 M KCl
3X FLAG Peptide Elutions
SUPEROSE 6
KCl Elutions
0.5M
0.5M
TCA precipitation
SDS PAGE & Colloidal blue stain
Band Excision and LC-MS/MS sequencing

Figure 10: Purification scheme of proteins interacting with $_{FLAG}$TgGCN5 or $_{FLAG}$ΔN$_T$TgGCN5.

Scheme was developed by Dr. Mohamed-Ali Hakimi's laboratory at the French National Centre for Scientific Research in Grenoble, France. Nuclear cell extract (NCE) or whole cell extract (WCE) from filter-purified transgenic tachyzoites expressing $_{FLAG}$TgGCN5 or $_{FLAG}$ΔN$_T$TgGCN5 was fractionated by chromatography as outlined above. The 0.35M potassium chloride (KCl) elution of DEAE-Sephacel column was purified using an anti-FLAG M2 affinity resin. The bound proteins were further separated by fractionation on a Superose 6 gel filtration column. The horizontal lines indicate stepwise elution. Concentrations are given in molars (M). P11 = Whatman phosphocellulose; FT = flow through; TCA = trichloroacetic acid; LC-MS/MS = nanocapillary liquid chromatography coupled with tandem mass spectrometry

France for purifying protein complexes (Saksouk et al., 2005). The same method was employed to identify proteins that may be interacting with full length FLAG tagged TgGCN5 ($_{FLAG}$TgGCN5) and/or TgGCN5 lacking the N-terminal extension ($_{FLAG}\Delta N_T$TgGCN5) stably expressed in *Toxoplasma*. Co-immunoprecipitations were performed using nuclear cell extract (NCE) from parasites expressing $_{FLAG}$TgGCN5 and using whole cell (WCE) from parasites expressing $_{FLAG}\Delta N_T$TgGCN5.

The appropriate lysate from freshly harvested and filtered transgenic extracellular tachyzoites was loaded onto a 250ml column of phosphocellulose (P11, Whatman) and fractionated stepwise by the indicated potassium chloride (KCl) concentrations (Figure 10) in buffer A [20mM Tris.HCl (pH 7.9), 0.2mM EDTA, 10mM β-Me, 10% glycerol, 0.2mM PMSF (phenylmethanesulfonyl fluoride)]. The P11 0.5M KCl fraction was loaded on a 45ml DEAE-Sephacel column (diethylaminoethyl; Pharmacia) and eluted with 0.35M KCl. The 0.35M KCl elution was incubated with 1ml of anti-FLAG M2 affinity gel (Sigma #A2220) for two hours at 4°C. Beads were washed with 50ml of BC500 buffer (20mM Tris (pH 8.0), 0.5M KCl, 10% glycerol, 1mM EDTA, 1mM DTT, 0.1% NP40, 0.5mM PMSF, and 1µg/ml aprotinin, 1µg/ml leupeptide, and 1µg/ml pepstatin). Beads were subsequently washed once with 10ml BC100 buffer (20mM Tris (pH 8.0), 0.1M KCl, 10% glycerol, 1mM EDTA, 1mM DTT, 0.1% NP40, and 1µg/ml aprotinin, 1µg/ml leupeptide, and 1µg/ml pepstatin). Protein bound to the anti-FLAG M2 affinity gel proteins were eluted stepwise with 350µg/ml of 3X FLAG peptide (Sigma #F4799) diluted in BC100 buffer. Elutions enriched in $_{FLAG}$TgGCN5 or $_{FLAG}\Delta N_T$TgGCN5 were fractionated on a Superose 6 HR 10/30 (Pharmacia) and equilibrated in 0.5M KCl in buffer A containing 0.1% NP-40, 1µg/ml aprotinin, 1µg/ml leupeptin, and 1µg/ml pepstatin. Proteins were TCA precipitated and resolved using SDS-PAGE and stained with SimplyBlue™ (Invitrogen #LC6060).

I. Mass spectrometry peptide sequencing

Dr. Mohamed-Ali Hakimi's laboratory at the French National Centre for Scientific Research in Grenoble, France excised protein bands from SimplyBlue™ (Invitrogen #LC6060) stained gels. The exised gel pieces were treated with 7% H_2O_2 and subjected to in-gel tryptic digestion as described in the literature (Ferro et al., 2003). Extracted proteins were analyzed by nanocapillary liquid chromatography coupled with tandem mass spectrometry (LC-MS/MS). Tandem mass spectra were searched against a compiled *Toxoplasma* database with their in-house software, Mascot (Matrix Sciences, London).

CHAPTER 3: RESULTS

I. Elucidating How TgGCN5 is Targeted to the Parasite Nucleus

This section reports the mapping of the first NLS to be identified and validated for any apicomplexan protein and in the process defined a function of the long N-terminal extension of TgGCN5. Furthermore, the NLS elucidated in TgGCN5 is the first completely mapped NLS to be defined for any GCN5 protein. The mechanism involved in targeting GCN5 proteins to the nucleus is poorly understood. Understanding how TgGCN5 is targeted to the nucleus may be applicable to other GCN5 proteins.

In addition, we searched the *Toxoplasma* and related apicomplexan databases for proteins containing similar NLS motifs. The goal of the search was two-fold. If results of the search found other known or predicted nuclear proteins that contained a similar NLS-like sequence it would provide further evidence that this is a true NLS. Secondly, the NLS may assist in the annotation of predicted proteins in apicomplexan databases with no homology or known function as well as the continued characterization of known proteins.

A. Mapping the NLS of TgGCN5

Figure 11 is the result of immunocytochemistry on transgenic parasites expressing the full length TgGCN5 and TgGCN5 lacking the N-terminal extension (residues 1-697; ΔN_TTgGCN5) both fused to a FLAG tag. To determine the localization of the recombinant proteins, immunocytochemistry was performed using a polyclonal antibody against the FLAG tag (Sigma #F-7425) followed by an anti-rabbit antibody conjugated to the fluorochrome Alexa-488 (Molecular Probes #A-11034). As a reference, parasite nuclei were visualized with the fluorescent nuclear stain, 4',6'-diamino-2-phenylindole (DAPI; Molecular Probes #D-1306). Recombinant protein that co-localizes with the DAPI stain is in the parasite nucleus. As shown in Figure 11, full length TgGCN5 co-localizes with the DAPI stain, indicating it is targeted to the nucleus. ΔN_TTgGCN5 does not co-localize

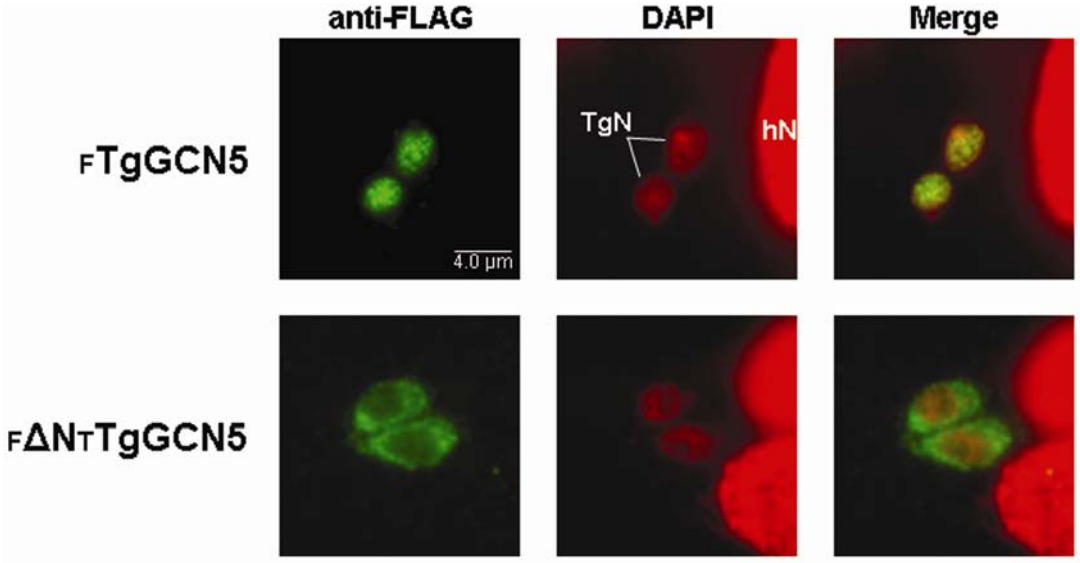


Figure 11: The N-terminal extension of TgGCN5 is required for nuclear localization

Recombinant protein in transgenic parasites expressing either full-length $_{F}$TgGCN5 or TgGCN5 lacking the N-terminal extension ($_{F}\Delta N_{T}$TgGCN5) was detected by immunofluorescence assay using anti-FLAG antibodies followed by staining with anti-rabbit Alexa 488 (green). 4',6'-Diamino-2-phenylindole (DAPI) was used as a nuclear co-stain (red). Images were obtained at 1000X magnification using a Leica DMLB scope with a 100X HCX Plan Apo oil immersion objective. Images were captured using a monochrome SPOT-RTSE (Model 12) Camera and Spot Diagnostic Software 4.0.9 and pseudocolored using Adobe Photoshop 7.0. hN = host cell nucleus = TgN, parasite nucleus, F = FLAG

with DAPI and is restricted to the parasite cytoplasm. These results confirm the extension's involvement in nuclear localization as indicated in previous studies.

The 697 amino acid sequence for the N-terminal extension of TgGCN5 known to be involved in nuclear localization is shown in Figure 12. Motif searches of the N-terminal extension suggested an NLS between residues 488 and 504 (Sullivan and Smith, 2000). However, removal of this 18-amino acid motif from full-length TgGCN5 fused to a C-terminal green fluorescent protein tag did not subvert nuclear localization (Sullivan and Striepen, unpublished observations). As shown in Figure 12, the N-terminal extension contains several clusters of basic residues that may be functioning as an NLS. Previous studies performed in our laboratory have narrowed the location of the NLS to be between amino acids 58 and 260. Within this region, there is only one stretch of basic residues, the sequence RKRVKR. The RKRVKR sequence fits one of the nine groups of NLS hexapeptide motifs devised by Boulikas, θθθxθθ where θ equals a basic residue (arginine or lysine) and x represents a non-basic residue (Boulikas, 1994). However, the NLS motif, θθθxθθ, was found to be a rare motif not commonly used by nuclear proteins identified in yeast or multicellular organisms (Boulikas, 1994).

In order to determine if the RKRVKR sequence contributed to the nuclear localization of TgGCN5, two truncated mutants of TgGCN5 were expressed in tachyzoites. An N-terminal FLAG tag was fused to both recombinant proteins. The first truncation was at residue 99, just downstream of the RKRVKR sequence ($_{FLAG}$Δ99TgGCN5). The second truncation removed the first 93 residues which includes all amino acids upstream of the RKRVKR sequence ($_{FLAG}$Δ93TgGCN5). If RKRVKR is the NLS of TgGCN5, $_{FLAG}$Δ93TgGCN5 should localize to the nucleus and $_{FLAG}$Δ99TgGCN5 should not. To determine the localization of the truncated mutants, immunocytochemistry was performed using the same antibodies as in Figure 11 (page 91). As shown in Figure 13, $_{FLAG}$Δ93TgGCN5 co-localizes with the DAPI stain, indicating it is in the nucleus, but $_{FLAG}$Δ99TgGCN5 does not co-localize with DAPI and cannot enter the parasite nucleus.

METVEVPAFLEAANSAEAPRPASPARELDVQLTVPCPRDGADAGEAVGTRKRSRGPADRAEEASG 65

RTHSAAEEPRAKEGGLSPASRDGDSGQSRKRVKREWASGGRLEQRGPLGVAAHAVEWRRASAEAR 130

GDARQTPDAARGSSSRVDRFAGEDGVLDEGDLFSPSPFWPFFSACDPVALAPAPTVCPDTPSVSV 195

VSACSRLLRAPPRDLASEATLDSLAGNVHVEESLWRGCESPNILEILRDILVVFAKDPSTQLSKM 260

FHTASLLAPYQAASCHAASDAERDKKGDEADSARRRDHHEKERNHGGGPLAVGESADSEGRHPGA 325

TEHAAGEAAGEAADGVRTADGDACQERRAEGAGAANGVAPLGMPEPAHAVGDRERENGSLEVKKA 390

EREDAERLENGGGGQTAGEDRDDVNDDPPDTEREPGPQQAGRRCLEANGVRGVKQEQENAGTDEE 455

RVLYLQLKEVVLSVAAALEIQQLEPPRKRKDPKKKPHSSSASSACRRGRGNAVSSSVSCFSASAN 520

CTSEDGERSGGDAAGARANGEIKAASDETYSDRRVPGGLHEESEKNEAAREKTREESGQTRTPAS 585

ESGSNAFVGFAQQSESRANSGVSSRPAEERGNEETFPDKEEEGAPAQREKHVAWAETTGAETARN 650

SDESEARRDPRLETCPLCDCQEDDAAWTRHGILPYDILYEKWKAFLM 697

Figure 12: Putative NLSs in the N-terminal extension of TgGCN5

The 697 amino acids comprising the N-terminal extension of TgGCN5 required for nuclear localization. Consecutive stretches of three or more basic residues are highlighted in gray. Previous work performed in the laboratory narrowed the region involved in mediating nuclear localization to be between residues 58 and 260 (denoted by the underlined sequence). The box outlines the NLS predicted by motif search analysis reported by Smith and Sullivan in 2000, which was later demonstrated to be uninvolved in nuclear localization (Sullivan and Striepen, unpublished observations).

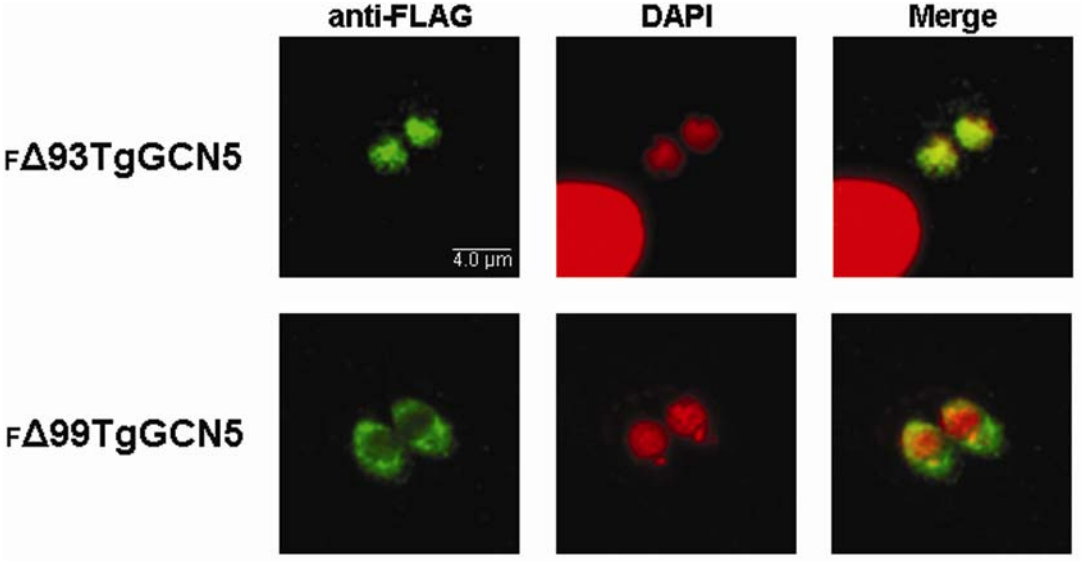


Figure 13: The hexapeptide RKRVKR (amino acids 94-99) is necessary for nuclear localization of TgGCN5

Parasites were transfected with FLAG-tagged forms of TgGCN5 lacking the first 93 ($_F$Δ93TgGCN5) or 99 ($_F$Δ99TgGCN5) amino acid residues. Immunofluorescence assay was carried out as described in the legend to Figure 11 (page 91). F = FLAG

These results demonstrate that the hexapeptide RKRVKR is required for the nuclear localization of TgGCN5. To verify whether the hexapeptide is a complete and sufficient NLS, RKRVKR was fused to a large cytoplasmic protein. The confirmation studies of the NLS from the SV40 large T antigen were performed fusing the basic NLS pentapeptide to beta-galactosidase, a large cytoplasmic protein (Kalderon et al., 1984b). Therefore, the TgGCN5 NLS was fused to the N-terminus of *E. coli* beta-galactosidase (βgal) and expressed in *Toxoplasma*. As shown in Figure 14, when βgal is expressed in parasites (βgal_{FLAG}), it does not co-localize with DAPI. However, when βgal is fused to the RKRVKR sequence (NLS-βgal_{FLAG}), it co-localizes with DAPI indicating it can now access the parasite nucleus. Interestingly, NLS-βgal_{FLAG} is not found only in the nucleus, but in both the nucleus and the cytoplasm.

There are several explanations for why NLS-βgal_{FLAG} is present in both the nucleus and cytoplasm. Beta-galactosidase might be interacting with other proteins in *Toxoplasma*. During these interactions, the NLS may not be properly exposed for binding by nuclear trafficking proteins, but when β-galactosidase is not involved in any interactions it is transported into the nucleus. Another explanation may be that beta-galactosidase contains a nuclear export signal and the NLS fusion protein is being shuttled back and forth between the nucleus and the cytoplasm. A different possibility may be due to the high overexpression of NLS-βgal_{FLAG} which is driven by the *Toxoplasma* beta-tubulin promoter, an extremely potent promoter (Striepen et al., 1998). Overexpression of NLS-βgal_{FLAG} may be saturating the nuclear trafficking pathways causing the protein to accumulate in the cytoplasm. However, saturation of nuclear trafficking pathways by a heterologous protein would mean that endogenous nuclear proteins would be unable to be effectively transported into the parasite nucleus. In other eukaryotic cell, a global disruption in nuclear trafficking is usually detrimental to cell survival (Zaidi et al., 2004). There did not appear to be a gross defect in the growth of parasites expressing NLS-βgal_{FLAG} indicating no such global defects in nuclear localization were occurring. Nonetheless, forcing

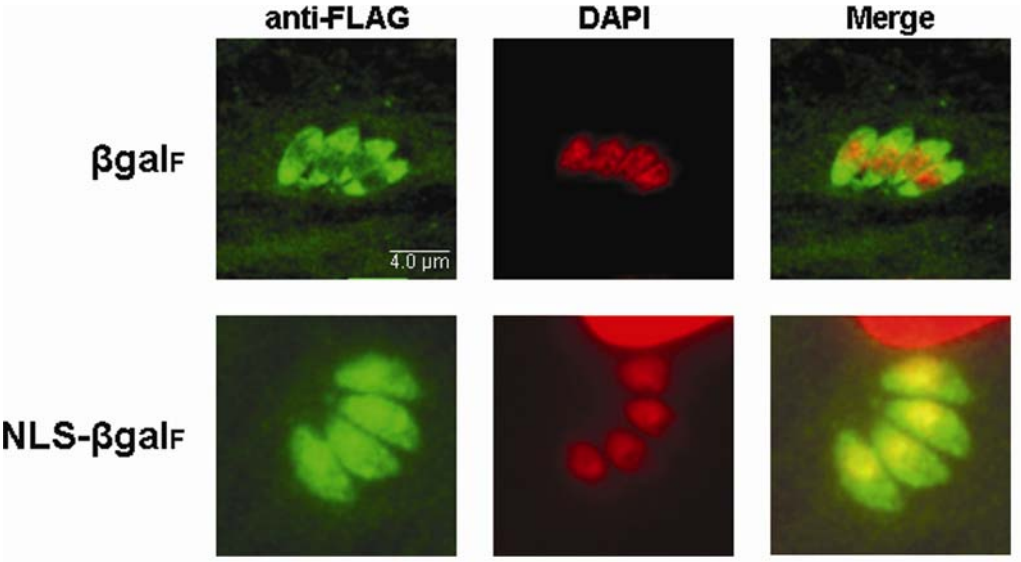


Figure 14: The TgGCN5 NLS is sufficient to translocate a heterologous cytoplasmic protein to the parasite nucleus

Parasites transfected with *E. coli* βgal fused to FLAG with or without the TgGCN5 NLS (βgal_F and NLS-βgal_F, respectively) were examined by immunofluorescence assay as described in the legend to Figure 11 (page 91). F = FLAG

the nuclear localization of a highly expressed heterologous protein could create perturbations in parasite nuclear trafficking resulting in the homogenous staining seen in Figure 14. Regardless of the explanation, there is clearly a difference in the nuclear localization between βgal_{FLAG} and NLS-βgal_{FLAG}; βgal_{FLAG} is restricted from the parasite nucleus while NLS-βgal_{FLAG} is capable of being targeted to the parasite nucleus.

It should be noted that the FLAG epitope tag in the βgal constructs was moved to the C-terminal end of the protein. The FLAG epitope tag was moved for two reasons. In the truncated mutant $_{FLAG}\Delta$93TgGCN5, the FLAG tag was placed directly upstream of the RKRVKR sequence. The FLAG tag sequence, DYKDDDDK, (D = aspartate, Y = tyrosine, K = lysine) contains several basic and acidic resides that might generate artifactual localization. Therefore, in the NLS-βgal constructs, the FLAG tag was moved to the opposite end of the protein to show that the FLAG tag is not causing artifactual localization. Secondly, placing the FLAG tag and the NLS sequence at opposite ends prevents the misinterpretation of localization data due to the cleavage of the NLS-βgal fusion protein. It is possible that protein degradation of NLS-βgal in *Toxoplasma* could result in the cleaving off of the RKRVKR sequence. If the FLAG tag was present on the N-terminus upstream of the NLS, the small cleaved product could freely diffuse into the nucleus. As mentioned in the Chapter 1 Section IV, small globular proteins less than 40kD can freely diffuse into the nucleus (page 43). The antibodies utilized in immunocytochemistry would bind to the FLAG tag present in the cleavage products located in the parasite nucleus, and it would erroneously appear as if the NLS-βgal fusion protein was nuclear. With the FLAG tag and NLS motifs on opposite ends, it insures that the difference in localization patterns between βgal_{FLAG} and NLS-βgal_{FLAG} is being generated by intact protein, not cleavage by-products.

B. Searches of apicomplexan databases for proteins containing the TgGCN5 NLS

To facilitate apicomplexan research, online genomic databases have been established for *Toxoplasma* (ToxoDB, http://toxodb.org; Kissinger et al., 2003), *Plasmodium* spp. (PlasmoDB, http://plasmodb.org/; Bahl et al., 2003), and *Cryptosporidium parvum* (CryptoDB, http://cryptodb.org; Puiu et al., 2004). These databases contain the results of extensive genomic sequencing efforts in these parasites and searchable EST libraries. Each database also contains information about known genes and predicted and known proteins. Predicted proteins are generated using gene prediction algorithms to generate proteins theorized to exist in the parasite. These proteins are only theorized to exist in the parasite; they have not been confirmed *in vivo*.

The NLS mapped in TgGCN5 is the first NLS to be elucidated for any apicomplexan protein. As this is a new motif identified in apicomplexans, it would be important to search the databases looking for proteins that contain an identical motif. Proteins that contain the sequence RKRVKR present in *Toxoplasma* and other apicomplexans should also contain other domains indicating a nuclear function. Predicted nuclear parasite proteins containing the sequence RKRVKR would provide additional evidence that RKRVKR is indeed a nuclear localization signal.

In additional to verifying RKRVKR as an NLS, searches of apicomplexan databases for identical or similar motifs may assist in the characterization of novel parasite-specific proteins. Despite the active research into apicomplexans, the protein databases contain large numbers of proteins with no known function. Of the estimated 10,585 genes in *Toxoplasma*, 74% are considered to be unique parasite-specific proteins with no homology to known proteins and are devoid of any discernable protein motifs (Li et al., 2003). In *Plasmodium* and *Eimeria*, it is estimated that over 50% of the genes present in both of these apicomplexan organisms have no clear predictable function (Li et al., 2003). These proteins may be the most worthwhile to study as they might be involved in parasite-specific functions such as virulence and differentiation. As new functional motifs

are elucidated, the databases should be searched for other predicted proteins that contain similar domains. Parasite-specific proteins without any recognizable function that contain the sequence RKRVKR would indicate a function within the parasite nucleus, thus aiding in the characterization of novel proteins.

Unfortunately, searches of the ToxoDB using the TgGCN5 NLS against known and predicted proteins revealed no exact matches, save TgGCN5 itself. No known or predicted proteins present in the PlasmoDB or in CryptoDB possessed the motif RKRVKR, either.

The fact that no proteins contain the exact same motif is not all together surprising. The particular NLS motif θθθxθθ ("θ" representing arginine or lysine and the "x" representing any amino acid) is not a commonly utilized NLS motif by nuclear proteins (Boulikas, 1994). Furthermore, searching for an exact match may not be appropriate. Within any NLS motif arginine can be substituted for lysine and vice versa and the non-basic residue can be substituted for any amino acid (Boulikas, 1994; Hodel et al., 2001). Therefore, our search parameters were expanded to include these permutations.

When the motif RKRxKR was used ("x" defined as any amino acid), a total of 19 potential hits emerged from the *Toxoplasma* set of predicted proteins based on the TIGR Draft3 prediction algorithm. Eight of the proteins had no homology to known proteins nor contained any known protein motifs. Of the remaining 11, nine possessed domains or homology indicating a nuclear function and only two had domains or homology indicating a cytoplasmic function. Thus, over 80% of the proteins containing the motif RKRxKR with additional motifs or homology appear to have a nuclear function providing further evidence that the motif RKRxKR is a nuclear localization signal. The nine putative nuclear proteins are listed on Table II.

As shown in Table II, 20.m03745 contains RKRKKR and possesses several RRM 1 (RNA recognition motif 1) domains indicating it may function as a nuclear RNA binding protein mostly likely involved in mRNA processing (Maris et al., 2005). Protein 38.m01112 contains a region with homology to the catalytic

Table II: Predictive value of the TgGCN5 NLS in Apicomplexa

Identifier	Putative NLS	Predicted Homology / Motifs
Toxoplasma		
20.m03745	RKRKKR	RRM_1 repeats (RNA recognition)
38.m01112	RKRGKR	tRNA synthetase class II core domain
42.m03605	RKRIKR	auxin response element
44.m02721	RKRMKR	WD40 repeats
49.m03309	RKRDKR	MAD (mitotic checkpoint protein)
50.m05631	RKRWKR	thioredoxin
55.m05013	RKRRKR	peptidase_C50
80.m02168	RKRTKR	DNA polymerase X
162.m00318	RKRQKR	protein kinase
Plasmodium		
PFC0425w	RKRNKR	PHD (zinc finger)
PY07179	KRKNKK	histone deacetylase domain
PF10_0175	RKKRKK	tRNA pseudouridylate synthase
PF10_0362	KKKMKK	DNA polymerase B family
Cryptosporidium		
CpEST_AA224688	RKRNKR	histone H2B
CpEST_AA253596	KKKRRK	putative nucleolar protein

core domain of a tRNA synthetase. While originally thought to be cytoplasmic, there is mounting evidence that tRNA synthetases are also present in the nucleus (Schimmel and Wang, 1999). Interestingly, several studies have implicated an NLS composed of basic residues (arginine and lysine) in targeting tRNA synthetases to the nucleus (Mucha, 2002). To date, none of the proposed NLS-like motifs present in other tRNA synthetases are exact matches to the motif in 38.m01112 (Mucha, 2002). Nonetheless, the fact that the search yielded another protein that appears to enter the nucleus via a NLS composed of basic residues in other organisms provides further support that RKRxKR is an NLS motif. The exact function of tRNA synthetases within the nucleus is still unknown (Mucha, 2002).

Protein 42.m03605 not only contains an NLS like motif but also contains an auxin response factor (ARF) which modulates gene expression in response to the plant hormone auxin (Liscum and Reed, 2002). No studies have been conducted to determine if *Toxoplasma* generates or responds to auxin. The apicoplast present in the parasites is of plant ancestry and ARF proteins may be apart of its function.

Proteins containing WD40 repeats like 44.m02721 form a large family of proteins found in all eukaryotes, and are named because the repeats usually end in a tryptophan-aspartic acid (W-D) dipeptide (Smith et al., 1999). This family of nuclear proteins is implicated in a variety of functions including transcription regulation and cell cycle control (Smith et al., 1999).

The MAD (mitotic arrest deficient) domain present in 49.m03309 suggests a function within the nucleus and the motif, RKRDKR may be its NLS. Proteins with MAD domains function as mitotic checkpoint proteins in other eukaryotic organisms (Glotzer, 1996).

The predicted protein 50.m05631 contains a thioredoxin domain. Proteins possessing thioredoxin domains are important in protection against reactive oxygen species (Byrne and Welsh, 2005). Although thioredoxin proteins are typically found in the cytoplasm, stress conditions have been shown to cause the proteins to translocate into the nucleus (Karimpour et al., 2002). The mechanism

behind how these proteins are targeted to the nucleus is still unknown, but the motif KRDKR in 50.m05631 may be involved.

Protein 55.m05013 contains a peptidase family C50 domain which is a type of cysteine peptidase of the separase family (Uhlmann, 2003). The separases are caspase-like proteases present in the nucleus, and they play a central role in chromosome segregation (Uhlmann, 2003). Interestingly, the NLS in protein 55m05013 is located within the peptidase motif (Figure 15). Usually, NLSs are not found within other domains as demonstrated by other proteins in Figure 15. However, an exception is DNA and RNA binding proteins in which there is often overlap between the binding domain and the NLS (LaCasse and Lefebvre, 1995). Although the mechanism for nuclear localization in other peptidase C50 family members has not been deduced, it is plausible given their function that the NLS is imbedded within the peptidase domain as is with other DNA binding proteins. However, the hexapeptide RKRRKR in 55m05013 may simply be a functional part of the peptidase domain and does not function as an NLS.

The protein 80.m02168 is undoubtedly a DNA polymerase and contains the motif, RKRTKR, which may contribute to its nuclear compartmentalization. The protein kinase 162.m00318 contains an NLS-like motif suggesting that the target of its phosphorylation may be within the parasite nucleus.

Figure 15 shows schematic representations of TgGCN5 and the nine *Toxoplasma* proteins listed in Table II. The NLS has been depicted as a black vertical line and the domains indicating nuclear function are in gray boxes. The image was generated to determine if there was a trend in the position of NLSs within *Toxoplasma* proteins. Figure 15 demonstrates that the putative NLS-like domains can be found anywhere within a protein. Six of the proteins have the NLS present in the first half of the protein with three of them being with in the first 5% of the protein. The other three proteins have the NLS in the C-terminal half with 2 of them located within the last 10% of the protein. Thus, there does not seem to be a clear consensus on the location of this particular NLS. These nine proteins are no doubt a very small fraction of all the nuclear proteins present in

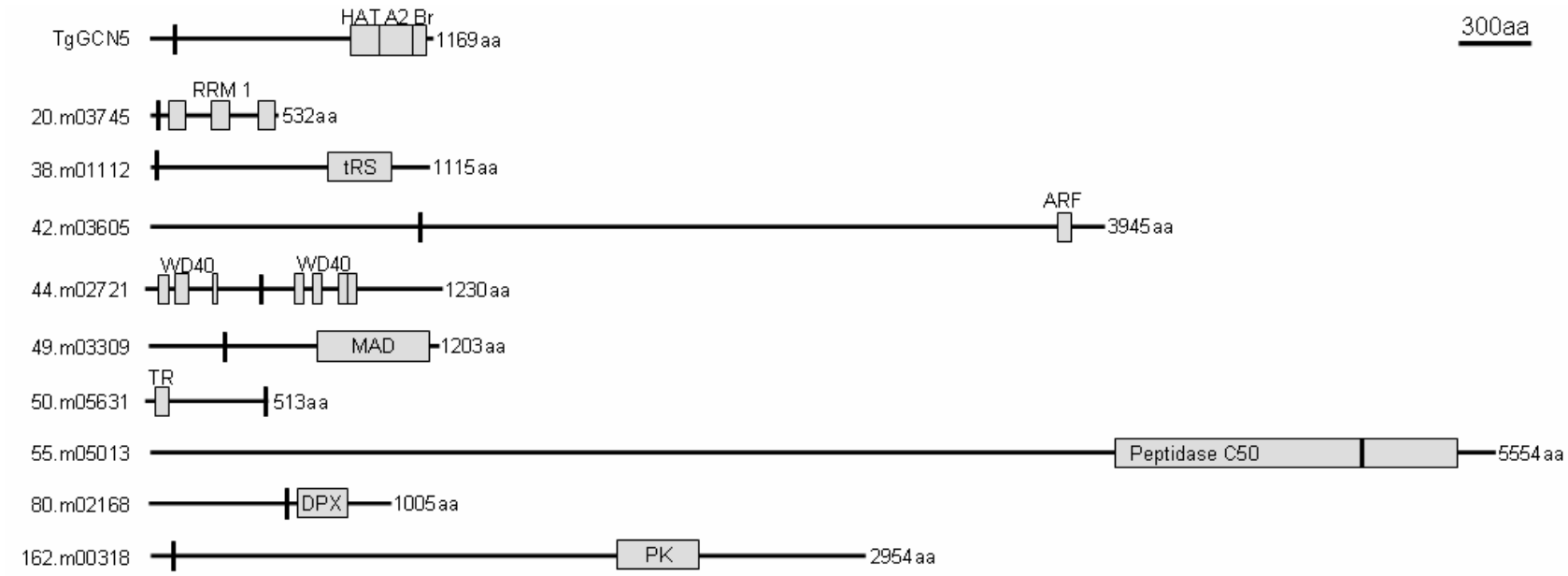

TgGCN5
HAT A2 Br
1169aa
300aa
20.m03745
RRM 1
532aa
38.m01112
tRS
1115aa
42.m03605
ARF
3945aa
44.m02721
WD40
WD40
1230aa
49.m03309
MAD
1203aa
50.m05631
TR
513aa
55.m05013
Peptidase C50
5554aa
80.m02168
DPX
1005aa
162.m00318
PK
2954aa

Figure 15: TgGCN5 and nine putative *Toxoplasma* nuclear proteins
Schematic diagrams of TgGCN5 and the nine putative nuclear proteins in *Toxoplasma* from Table II comparing the location of the NLS within each protein. Each protein is depicted by a horizontal black line. As indicated in the top right corner, 1cm is equal to 300 amino acids. The NLS-like motif (RKRxKR) is depicted as a single vertical black line. Other motifs indicating nuclear function present in the proteins are represented as gray boxes. The motifs within each protein are labeled according to their function. As is seen in other eukaryotes, the NLS in these *Toxoplasma* proteins can be located anywhere within the protein. aa = amino acid, HAT = histone acetylase, A2 = Ada2 domain, Br = bromodomain, RRM1 = RNA recognition motif 1 repeats, tRS = tRNA synthetase, ARF = auxin response factor, WD40 = tryptophan-aspartic acid domain repeats, MAD = mitotic arrest deficient domain, TR = thioredoxin domain, DPX = DNA polymerase X, PK = protein kinase

Toxoplasma. As more NLSs are identified, a trend may emerge indicating that NLSs with *Toxoplasma* proteins tend to be located in one particular area within a protein. It should be noted that in other organisms there does not appear to be a general trend in the position of an NLS within other nuclear proteins (Dingwall and Laskey, 1991).

The *Toxoplasma* proteins listed in Table II (page 100) contain domains implicating function within the nucleus which supports the hypothesis that the TgGCN5 NLS has value in predicting the localization of proteins. Therefore, the eight *Toxoplasma* proteins with unknown function containing RKRxKR motifs should be targeted to the parasite nucleus. However, two of the 19 proteins from the database search contained ribosomal motifs in addition to an NLS-like motif. Despite the presence of an NLS-like motif, the ribosomal domains suggest a cytoplasmic function, which demonstrates the limitation of database searches. The presence of an NLS-like domain within a protein does not guarantee it is a nuclear protein. The presence of an NLS-like domain must be taken into context with other domains present in the protein. Ultimately, localization studies must be performed to confirm if any of these hits are truly nuclear proteins.

Table II also displays results from similar searches of the *Plasmodium* database with permutations of the TgGCN5 NLS. For example, searching the *P. falciparum* annotated protein data set with RKRxKR revealed 14 hits, including the predicted nuclear protein PFC0425w which contains three PHD fingers. PHD fingers are zinc finger-like motifs present in nuclear proteins that are involved in transcriptional regulation (Aasland et al., 1995).

Searches with the less stringent criteria allowing for the basic residues to be either lysine or arginine (R/K)(R/K)(R/K)X(R/K)(R/K) revealed 1500 hits in the annotated protein data sets from *P. falciparum*. Cursory analysis of some of these candidate nuclear proteins revealed a putative histone deacetylase (PY07179; KRKNKK), a nuclear protein tRNA pseudouridine synthase (PF10 0175; RKKRKK), and a putative DNA B family polymerase with a zeta catalytic subunit (PF10_0362; KKKMKK).

Searches of the *Cryptosporidium* expressed sequence tag database in CrytptoDB for instances of (R/K)(R/K)(R/K)X(R/K)(R/K) were also successful in identifying nuclear proteins; among the nine hits were histone H2B (CpEST_AA224688; RKRRKR) and a putative nucleolar protein (CpEST_AA253596; KKKLRK).

The fact that a *Toxoplasma* NLS is able to identify nuclear proteins in other apicomplexans indicates that the *Plasmodium* and *Cryptosporidium* may also utilize a similar NLS to target nuclear proteins to the nucleus.

II. Identifying Proteins Interacting with the N-terminus of TgGCN5

Two groups of proteins proposed to be interacting with the N-terminus of TgGCN5 are those involved in nuclear trafficking and those forming HAT complexes analogous to the yeast SAGA and ADA complexes.

The previous section presented data that the hexapeptide RKRVKR in the N-terminus of TgGCN5 is required for nuclear localization. In other organisms, proteins containing NLSs composed of basic residues are transported into the nucleus through an interaction with the nuclear trafficking protein, importin α (Gorlich et al., 1994; Figure 7, page 47). Therefore, it is likely that the NLS within the N-terminus of TgGCN5 interacts with importin α. However, an importin α homologue in *Toxoplasma* has not been identified.

A. Cloning and characterization of an importin α homologue in Toxoplasma

The *Toxoplasma* database (Release 2.1; http://toxodb.org) was screened for sequences encoding possible importin α homologues using the BLAST algorithm and text queries. The genomic database entry Tgg_994532 in Release 3.0 (52.m00001 in version 4.0) exhibited a high degree of similarity to importin α proteins from other species. Primers designed to the genomic sequence data were employed to obtain the full length cDNA for *Toxoplasma* importin α (TgIMPα, AY267540). Based on the results of the cDNA sequence, we have deduced that the ~9.5-kb TgIMPα genomic locus is comprised of 11 exons and 10 introns.

TgIMPα possesses an open reading frame of 1638 nucleotides with a consensus start ATG (Kozak, 1991) that is preceded 156 nucleotides upstream by an in frame stop codon. The deduced 545 amino acid sequence has a calculated molecular mass of ~60kD and shows unequivocal similarity to importin α orthologues present in other eukaryotic organisms, especially the fellow apicomplexan parasite *P. falciparum* (Mohammed et al., 2003; Figure 16A). TgIMPα contains the well conserved importin β binding (IBB) domain and armadillo (ARM) repeats found in all other importin α proteins known to date (Herold et al., 1998; Figure 16A). The autoinhibitory motif (KRR) observed in other importin α proteins is present in the importin β binding (IBB) domain, but the corresponding sequence in TgIMPα is more akin to the one in plants (KKR). The *Toxoplasma* IMPα also possesses a CAS (cellular apoptosis susceptibility) binding site within its C terminus required to recycle importin α back to the cytoplasm (Kobe, 1999).

Phylogenic analysis of apicomplexan importin α proteins (Figure 16B) illustrates that they are more similar to those found in plants, a characteristic that has been noted in other apicomplexan proteins (Huang et al., 2004).

B. Interaction of TgIMPα with TgGCN5 though the NLS

Co-immunoprecipitation studies were performed to assess whether TgGCN5 physically interacts with TgIMPα and, furthermore, whether such an interaction occurred by virtue of the NLS, RKRVKR.

Recombinant HA-tagged TgIMPα and full-length cMyc tagged TgGCN5 were generated using a commercially available *in vitro* translation system (Promega; Chapter 2, Section IV-C, page 74). The *in vitro* translation system incorporates radiolabeled methionine into the synthesized proteins. The two proteins were purified from the translation reaction using commercially available anti-HA and anti-cMyc affinity resins (Clontech). The immunoprecipitated proteins were resolved by SDS-PAGE gel and visualized by autoradiography. Figure 17 shows the purification of each protein generated by the *in vitro* translation system in detectable quantities and effectively purified from the

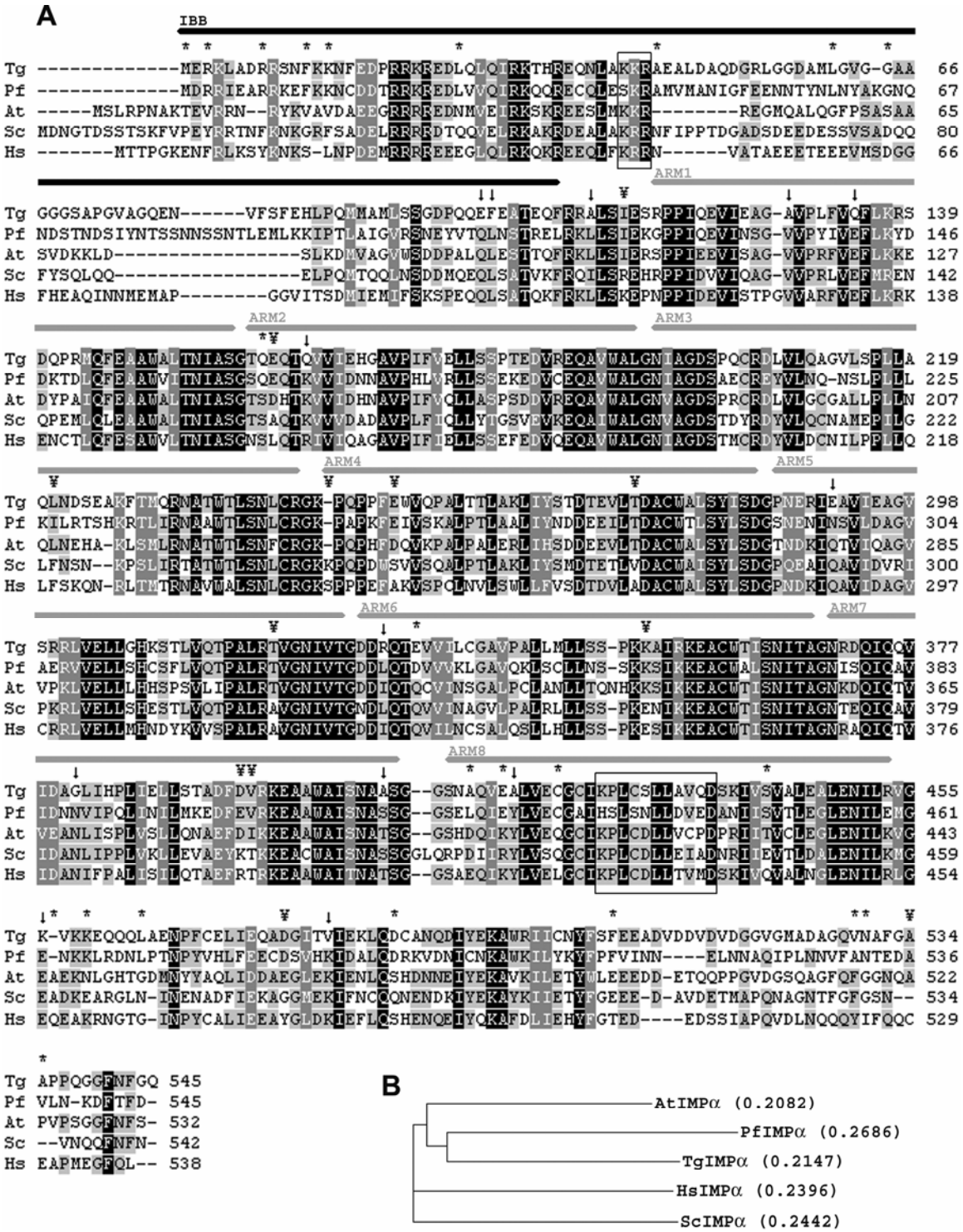

A
IBB
ARM1
ARM2
ARM3
ARM4
ARM5
ARM6
ARM7
ARM8
Tg
Pf
At
Sc
Hs
B
AtIMPα (0.2082)
PfIMPα (0.2686)
TgIMPα (0.2147)
HsIMPα (0.2396)
ScIMPα (0.2442)

Figure 16: *Toxoplasma* importin α (TgIMPα)

A – protein sequences of importin α proteins from various species were aligned using the AlignX module in Vector NTI Advance 9.0. Tg, *Toxoplasma* (AY267540); Pf, *P. falciparum* (AAO85774); At, *Arabidopsis thaliana* (Q96321); Sc, *Saccharomyces cerevisiae* (Q02821); Hs, *Homo sapiens* (P52294). Black inversed letters indicate identical residues in all species, gray inversed letters are functionally conserved residues, and gray highlighted letters are identical or functionally conserved residues in more than 50% of the species examined; ↓ denotes residues not conserved in *Toxoplasma*; * denotes conservation in apicomplexans only, and ¥ denotes elements conserved between apicomplexans and *A. thaliana*. IBB, importin β binding domain; ARM, armadillo repeats 1–8. The boxed region within the importin β binding domain is the autoinhibitory sequence, and the boxed region within armadillo repeat 8 is the CAS binding domain involved in nuclear export. **B** – phylogenic analysis obtain from the AlignX module in Vector NTI Advance 9.0 of importin α orthologues from representative species with the calculated distance values in parentheses (Bhatti and Sullivan, 2005).

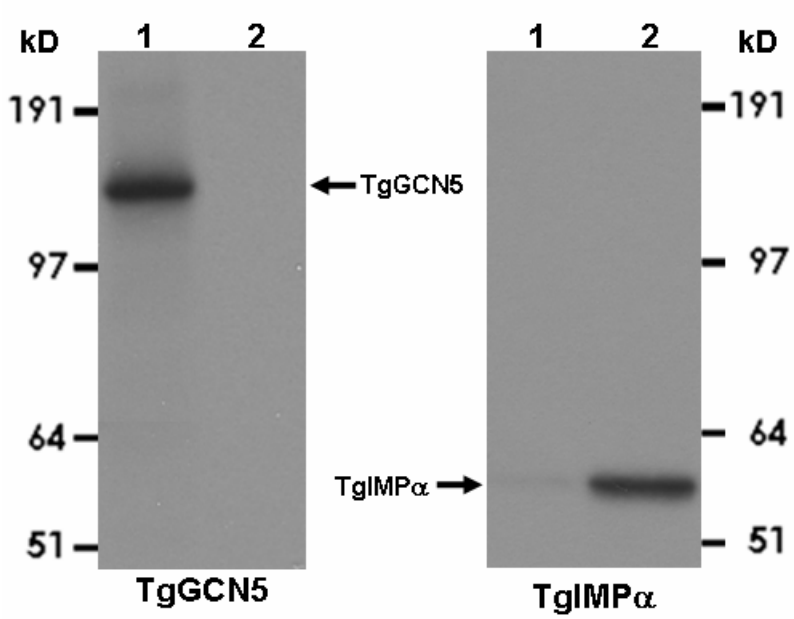


Figure 17: Successful *in vitro* translation and purification of TgGCN5 and TgIMPα

Autoradiography results from the purification of TgGCN5 (129kD) and TgIMPα (60kD) using both anti-cMyc (lane 1) and anti-HA (lane 2) affinity resins. TgGCN5 and TgIMPα were generated using *in vitro* translation (Promega #L4610) and gel was exposed to High Sensitive Blue photographic film (RPS Imaging #33-0810). Molecular weights are displayed in kilodaltons (kD).

reaction without significant degradation. As a negative control, Figure 17 shows that TgGCN5 cannot be isolated with anti-HA affinity resin and TgIMPα cannot be purified with anti-cMyc affinity resin. This negative control was necessary to rule out potential cross reactivity between the recombinant tags and affinity resins. Therefore, if TgIMPα and TgGCN5 co-immunoprecipitate, it cannot be argued that it is due to non-specific interaction between the proteins and the affinity resin.

Once protein production was verified, a small amount of the translation reaction containing TgIMPα was mixed with a small amount of the reaction containing TgGCN5. The mixture was then split between the anti-HA and anti-cMyc resin. The anti-cMyc affinity resin will pull down TgGCN5 by virtue of its cMyc epitope tag. If TgIMPα interacts with TgGCN5 then it will be pulled down with the TgGCN5 resulting in two bands when resolved via SDS-PAGE and visualized by autoradiography. The reverse would happen with the anti-HA resin. TgIMPα would be immunoprecipitated by its HA-tag and if the two proteins interacted, TgGCN5 would be co-precipitated. Again, two protein bands would be visualized by autoradiography. However, if the two proteins do not interact, only single bands would be present corresponding to the immunoprecipitated protein by virtue of the recombinant tag (i.e. HA or cMyc).

Figure 18 contains the results of TgGCN5 and TgIMPα co-immunoprecipitation experiments using cMyc and HA affinity resins. Unfortunately, the co-immunoprecipitation efficiency was very low. The band intensity in lane 1 is very strong for the protein containing the recombinant tag. There is a faint band in lane 1 in both co-immunoprecipitations corresponding to a potential co-precipitated protein. However, it is not nearly intense enough to state these two proteins interact with one another. Either something is interfering with the interaction or TgGCN5 and TgIMPα do not strongly interact with each other. Lane 2 contains proteins not bound by the affinity resin and demonstrates that both proteins were produced in sufficient quantities in the *in vitro* translation system to allow for good purification. Therefore, it cannot be argued that protein production was insufficient.

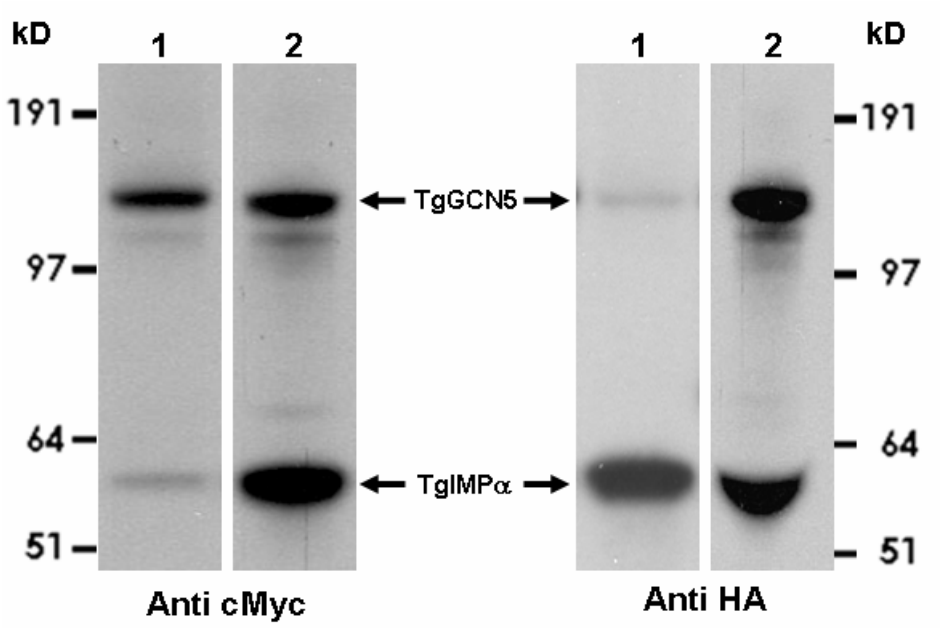


Figure 18: TgGCN5 and TgIMPα co-immunoprecipitations with cMyc and HA affinity resins

The above figure is a representative of three replicate experiments. TgGCN5 (129kD) and TgIMPα (60kD) were expressed using *in vitro* translation and subsequently mixed together. Co-immunoprecipitations were performed with anti-cMyc and anti-HA affinity resin. Following resolution using SDS-PAGE, gel was exposed to High Sensitive Blue photographic film (RPS Imaging #33-0810). Lane 1 contains the proteins eluted off the affinity resin and lane 2 contains proteins from the *in vitro* translation reaction that did not bind to the affinity resin. Molecular weights are displayed in kilodaltons (kD).

The possibility that TgGCN5 and TgIMPα do not strongly interact was not initially surprising. Biophsyical experiments have shown that the interaction between importin α and an NLS is extremely weak (Catimel et al., 2001). The weak interaction is due to the importin β binding domain, which possesses an autoinhibitory function (Kobe, 1999). When importin α is bound by importin β through the importin β binding domain, the autoinhibition is attenuated allowing importin α to strongly interact with an NLS (Goldfarb et al., 2004). While this model is the general excepted method of NLS binding by importin α, there are exceptions. In plants, it has been shown that importin α is capable of binding an NLS and transporting a protein to the nucleus without importin β (Hubner et al., 1999). Because apicomplexans have a propensity to possess plant like functions (Huang et al., 2004), it still seemed plausible that TgIMPα could bind to TgGCN5 without requiring an importin β protein. As noted in the previous section, phylogenic analysis of TgIMPα illustrated that it was more closely related to plant importin α proteins (Figure 16B, page 108). Therefore, the possibility that something was interfering with the interaction or with the co-immunoprecipitation needed to be ruled out.

The cMyc and HA epitope tags fused to TgGCN5 and TgIMPα that are utilized for co-immunoprecipitation are located at the N-terminus of both proteins. Thus, the epitiope tags are proximal to where the two proteins may be interacting and therefore masked during an interaction. If the epitopes were masked during the interaction, it would prevent them from being purified by the affinity resin. In order to determine if this was occurring, the epitope tags would have to be moved to the C-terminal end, or a different method of purification needed to be employed. Others working in the Sullivan Lab had generated a polyclonal TgGCN5 antiserum against the C-terminus of TgGCN5 (Chapter 2, Section IV-B, page 71). Since the TgGCN5 NLS lies within the N-terminal extension, a C-terminal antibody should bind to an area of the protein not involved in a potential TgGCN5-TgIMPα interaction. Therefore, it would be an excellent tool to determine if TgGCN5 and TgIMPα actually interact. As shown in Figure 19, Lane 1 shows the antiserum effectively purifies TgGCN5 from the *in vitro* translation

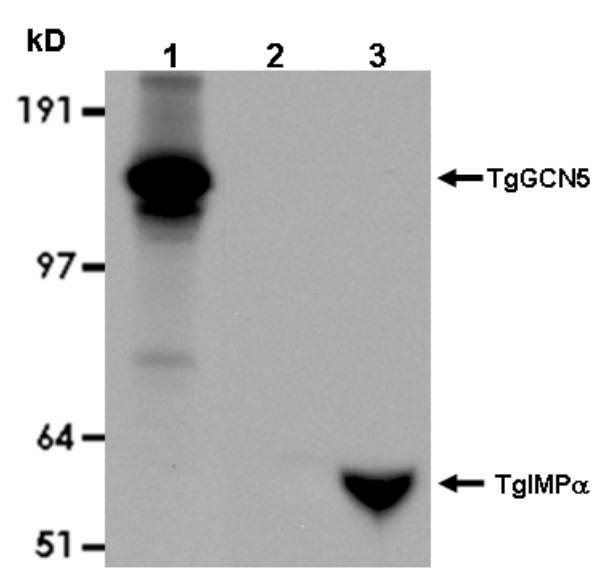


Figure 19: Immunoprecipitation with TgGCN5 antiserum

An autoradiogram of immunoprecipitation reactions preformed with the TgGCN5 antiserum. Following resolution using SDS-PAGE, gel was exposed to High Sensitive Blue photographic film (RPS Imaging #33-0810). Lane 1 is immunoprecipitation of full-length TgGCN5 and Lane 2 is of TgIMPα. Lane 3 is the flow through or unbound proteins from the TgIMPα immunoprecipitation with the TgGCN5 antiserum. Molecular masses are displayed in kilodaltons (kD).

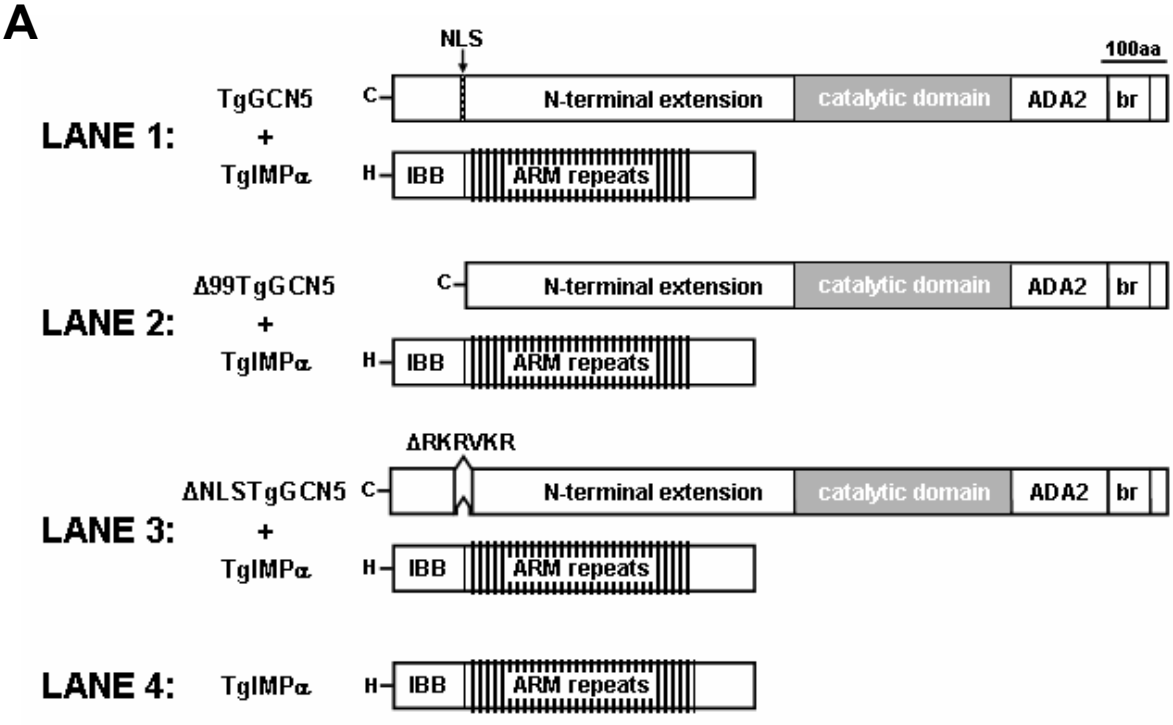
reaction mixture but does not cross react with TgIMPα (Lane 2). Lane 3 in Figure 19 is a small sample of the unbound fraction from Lane 2 indicating that TgIMPα was synthesized by *in vitro* translation in sufficient quantities.

After evaluating the TgGCN5 antiserum, TgGCN5 and TgIMPα proteins were generated and mixed together. Co-immunoprecipitation was performed using the TgGCN5 antiserum. The immunoprecipitated proteins were resolved by SDS-PAGE gel and visualized by autoradiography. If TgGCN5 interacts with TgIMPα, the antiserum will pull down TgIMPα through its interaction with TgGCN5. When resolved on a gel, two bands will be present, one from TgGCN5 and the other corresponding to TgIMPα. If the two proteins do not interact, then the antiserum will only pull down TgGCN5 resulting in a single band on the gel. Lane 1 in Figure 20-B is the result of mixing full-length TgGCN5 with TgIMPα. Two bands are present indicating the two proteins interact *in vitro*.

Once it had been established that TgGCN5 and TgIMPα interacted with each other, a truncated version of TgGCN5 starting just downstream of the RKRVKR NLS (Δ99TgGCN5; Figure 20-A) was expressed in the *in vitro* translation system and mixed with TgIMPα. Lane 2 in Figure 20-B shows only one band corresponding to the size of this truncated TgGCN5 (118kD). The single band indicates that without the first 99 residues, TgGCN5 no longer interacts with TgIMPα. The single band runs slightly lower than full-length TgGCN5 indicating a lower molecular weight congruent with it being the truncated form.

To definitively show that the NLS is the interaction point between TgGCN5 and TgIMPα, a deletion mutant was generated removing the RKRVKR hexapeptide (ΔNLSTgGCN5). It was replaced with a restriction enzyme site to facilitate construction of the expression construct which encodes the two amino acids alanine and serine (Chapter 2, Section II-D page 67). Thus, the hexapeptide RKRVKR was replaced with an alanine and serine dipeptide. Following *in vitro* translation, ΔNLSTgGCN5 was mixed with TgIMPα. Lane 3 of Figure 20-B shows the results of co-immunoprecipitation with only one band

A


NLS
100aa
TgGCN5
C
N-terminal extension
catalytic domain
ADA2
br
LANE 1:
+
TgIMPα
H
IBB
ARM repeats
Δ99TgGCN5
LANE 2:
ΔRKRVKR
ΔNLSTgGCN5
LANE 3:
LANE 4:

B

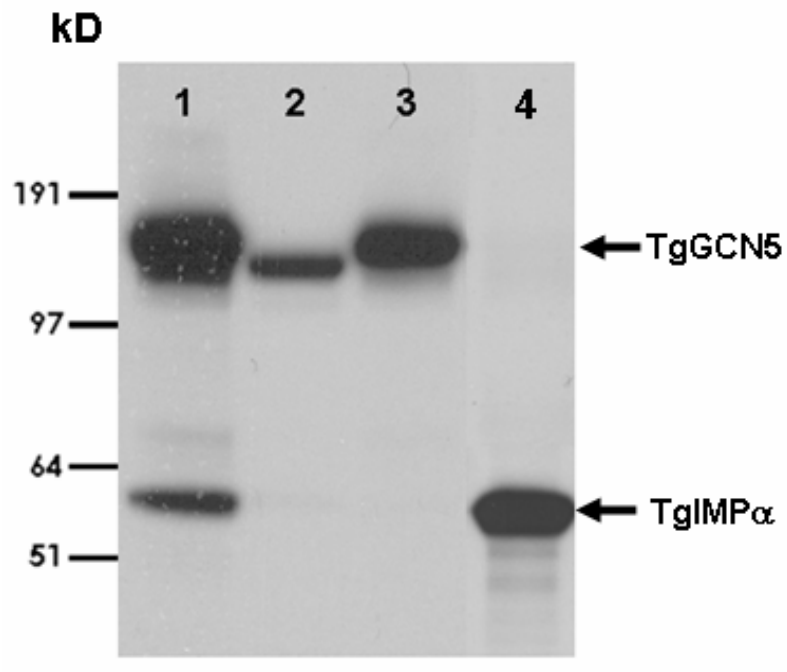

kD
1
2
3
4
191
97
64
51
TgGCN5
TgIMPα

Figure 20: TgGCN5 interacts with TgIMPα via the NLS RKRVKR

A – Schematic diagrams drawn to scale depicting the different co-immunoprecipitations that were performed to demonstrate that TgGCN5 interacts with TgIMIPα via the NLS RKRVKR. Each diagram corresponds to a lane in the gel below (C = cMyc epitope tag, NLS = nuclear localization signal, ADA2 = Ada2 binding domain, br = bromodomain, H = HA epitope tag, IBB = importin β binding domain, ARM = armadillo superhelical repeats). **B** – Autoradiogram of immunoprecipitation reactions using High Sensitive Blue photographic film (RPS Imaging #33-0810). Lane 1, full-length TgGCN5 + HA-TgIMPα; lane 2, Δ99TgGCN5 + HA-TgIMPα; lane 3, ΔNLSTgGCN5 + HA-TgIMPα; lane 4, HA-TgIMPα. Lanes 1-3 were immunoprecipitated with TgGCN5 antiserum and lane 4 was precipitated with anti-HA affinity resin. The top arrow points to the forms of TgGCN5 used (~118 - 130kD), and the bottom arrow points to HA-TgIMPα (~60kD). Molecular masses are displayed in kilodaltons (kD).

present, corresponding to the size of ΔNLSTgGCN5 (128kD). The lack of TgIMPα in this lane confirms that TgIMPα interacts with TgGCN5 at the NLS elucidated in the previous section. Lane 4 in Figure 20-B is an anti-HA immunoprecipitation of the TgIMPα translation reaction used in the co-immunoprecipitations in the previous lanes. The immunoprecipitation provides evidence of adequate TgIMPα protein production.

C. Identification of proteins interacting with TgGCN5 to form HAT complexes

It is plausible that the N-terminus of TgGCN5 also interacts with parasite proteins to form complexes analogous to yeast SAGA and ADA. Given the divergent nature of the N-terminal extension of TgGCN5, it may also interact with novel, parasite-specific components. Several methods were considered in trying to identify proteins interacting with the N-terminus of TgGCN5.

Yeast two-hybrid could be employed using the N-terminus as a bait protein to screen a cDNA library. However, yeast two-hybrid has several limitations. First, yeast two-hybrid screens notoriously generate false positives. Second, not all *Toxoplasma* proteins are going to be expressed properly in yeast because of codon bias, improper folding, and post-translational modification. The expression of certain *Toxoplasma* proteins may also be lethal to yeast.

Another option would be to co-immunoprecipitate protein complexes from wild-type parasite lysate. Our lab has developed an antiserum directed at the C-terminal end of TgGCN5 with the potential to immunoprecipitate TgGCN5 containing complexes. However, attempts to purify native HAT complexes using endogenously expressed protein in other organisms have been largely unsuccessful because HAT proteins are not expressed at high levels.

One of the more promising methods for purifying complexes is to over-express one of the complex members fused to an epitope tag. The N-terminal extension portion of TgGCN5 could be over-expressed in the parasites fused to a FLAG tag. Using affinity resin, anti-FLAG co-immunoprecipitations could be performed on parasite nuclear extract. The proteins would be eluted from the affinity resin and resolved via SDS-PAGE. Bands present in the gel would be

excised and sent for sequencing analysis. Unfortunately, multiple attempts at trying to stably express just the N-terminal extension of TgGCN5 in *Toxoplasma* were unsuccessful. It appears that expression of only the N-terminus of TgGCN5 is deleterious to the parasite. If the N-terminus interacts with proteins in the parasite, expression of the N-terminus could have acted as a dominant negative. The recombinant N-terminal extension may have sequestered proteins away from fully functional complexes causing a depletion of HAT complexes resulting in dysregulation of histone acetylation and parasite death.

An alternate strategy was developed using the viable transgenic parasites expressing recombinant full length TgGCN5 and TgGCN5 lacking the N-terminal extension (ΔN_TTgGCN5) fused to a FLAG tag. As shown in a previous section, both of these proteins can be stably expressed in the parasites. FLAG co-immunoprecipitations could be performed on these two parasites lines. Proteins that are present in purifications using full length TgGCN5 and absent in purifications with ΔN_TTgGCN5 would be considered to be interacting with the N-terminus of TgGCN5.

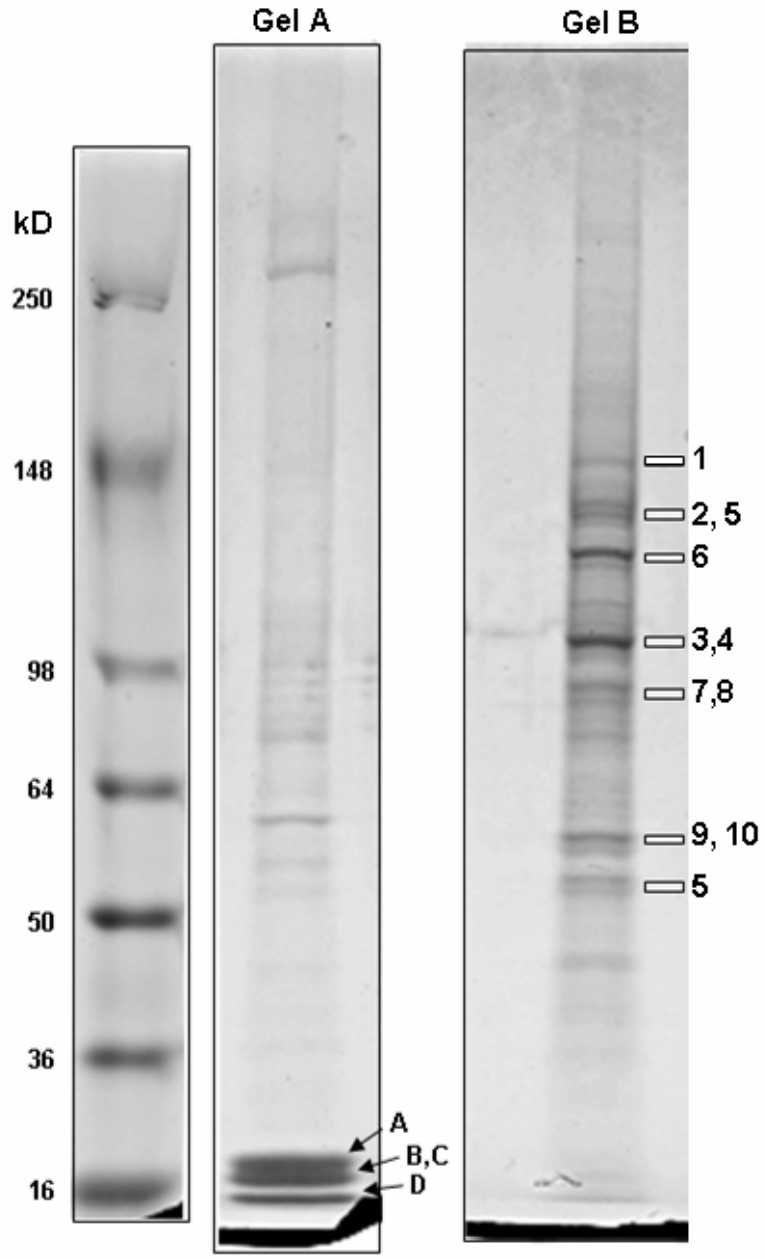
To isolate TgGCN5 containing complexes from these two transgenic parasite lines, a collaboration was established with Dr. Mohamed-Ali Hakimi's laboratory at the French National Centre for Scientific Research in Grenoble, France. His laboratory has developed a successful method to purify histone deacetylase complex members from the parasites using anti-FLAG co-immunoprecipitation. However, Dr. Hakimi's protocol requires enormous amounts of parasite lysate; approximately 100 T-150cm^2 cell culture flasks. The research facility at Grenoble has a large centralized tissue culture facility allowing for the growth and cultivation of large numbers of parasites. In addition, laboratories at the research center have free access to a protein sequencing facility. Therefore, the two transgenic parasite lines over-expressing $_{FLAG}$TgGCN5 and $_{FLAG}\Delta N_T$TgGCN5 were sent to Hakimi's laboratory for a preliminary attempt to co-immunoprecipitate interacting proteins. Unfortunately, the protein sequencing facility in Grenoble does not release the results of its

protein sequencing data and thus we are unable to confirm or make our own interpretations of the protein sequencing data.

Large quantities of both of our parasite lines were cultivated by Dr. Hakimi's laboratory. Co-immunoprecipitations with anti FLAG affinity resin was performed on parasite lysate from the transgenic parasites using the same protocol that had been successful in purifying members of histone deacetylation complexes (Saksouk et al., 2005; Chapter 2, Section IV-H, page 85). Once the purification was complete, elutions containing TgGCN5 were pooled, TCA precipitated, and resolved by SDS-PAGE.

Figure 21 contains two gels (A and B) loaded with two different sets of pooled elutions from the affinity purification of nuclear cell extract from parasites over-expressing full-length $_{FLAG}$TgGCN5. Gel A contains proteins present in the first few elutions off the purification column. The gel does not contain significant amounts of TgGCN5, but has very intense bands of low molecular weight that correspond to the size of histone proteins. The four bands labeled with arrows A, B, C, and D were excised and sent for sequencing analysis. The results of the sequencing confirmed that they are histone proteins. Band D is the core histone H4. The remaining bands are not the typical core histones, but rather variant histones H3.3, H2A.1, and H2AF/Z. No known histones variants exist for histone H4. The existence of histone variants has been previously established in *Toxoplasma* (Sullivan et al., 2006). As mentioned in the introduction, histone variants differ from core histones in several ways. The main difference is that variant histones are associated with restructuring transcriptionally active chromatin (Kamakaka and Biggins, 2005). Finding variant histones interacting with TgGCN5 agrees with the predicted role for TgGCN5 being involved in remodeling chromatin.

Gel B in Figure 21 represents a second pool of fractions that contained large amounts of TgGCN5. The numbered rectangles correspond to locations in the gel that were excised and sent for sequencying analysis by nanocapillary liquid chromatography coupled with tandem mass spectrometry. The sequencing



Gel A Sequencing Results:
A) Histone H3.3 (TgGlmHmm_1329)
B) Histone H2A.F/Z (TgGlmHMM_0487)
C) Histone H2A.1(TgTigrScan_0798)
D) Histone H4 (TgGlmHMM_3485)

Gel B Sequencing Results:
1. FLAGTgGCN5
2. FLAGTgGCN5
3. ADA2-A
4. ADA2-B
5. Unknown function (TgTwinScan_2624)
6. Unknown function (TgTigrScan_8186)
7. Hsp 90-like (TgGlmHMM_3720)
8. Cdc48 (TgGlmHMM_4686)
9. Hsp70 (TgGlmHMM_4715)
10. Hsp60 (TgGlmHMM_1548)

Figure 21: Chromatographic purification and identification of proteins interacting with $_{FLAG}$TgGCN5

Gel A and Gel B contain different sets of pooled elutions resolved by SDS-PAGE (4% - 12%) and visualized by colloidal blue staining. Molecular weight markers in kilodaltons (kD) are indicated on the left. In Gel A, the arrows point to four low molecular weight bands that were excised and sent for sequencing analysis. The results of the sequencing are listed below the gel. In Gel B, the white rectangles represent the location and size of the excised bands sent for sequencing. Each rectangle has a number that corresponds to the sequencing results listed below in the gel. Two numbers next to a rectangle indicates both proteins were identified in the same excised gel slice. Data preseted in this figure was generated by Dr. Mohamed-Ali Hakimi's laboratory at the French National Centre for Scientific Research in Grenoble, France.

data was used to determine the identity of the proteins present within the gel slice. A total of 7 gel slices were sequenced. The numbers refer to the protein identity listed below the gel. In some of the gel slices, more than one protein was present. In cases where sequencing analysis detected two proteins, two numbers were placed next to the rectangle to denote the identity of two distinct proteins.

The band labeled "2" corresponds to the size of $_{FLAG}$TgGCN5 (~130kD). Sequencing data for band 1 suggests that this is also $_{FLAG}$TgGCN5, but it is running about 30kD too large. It is likely that this is a post-translationally modified version of TgGCN5 causing it to migrate slower. There is evidence that GCN5 can be phosphorylated (Barlev, 1998) and the related HAT PCAF can be transacetylated (Herrera, 1997). These same modifications may occur to TgGCN5.

The gel slice of the band observed at about 100kD yielded sequence data consistent with two different Ada2 proteins (proteins 3 and 4 in Figure 21). It has been shown that humans and plants possess more than one homologue of Ada2 (Stockinger et al., 2001; Barlev et al., 2003), but to date all unicellular organisms such as yeast only express one Ada2 protein. Searches of ToxoDB have revealed the presence of two Ada2 proteins (TgAda2-A and TgAda2-B), which our lab has subsequently cloned (Bhatti et al., 2006). Consistent with our preliminary data, protein 3 agrees with the predicted size for TgAda2-A (~102kD), but protein 4 is much smaller than the predicted size of 238kD for TgAda2-B. It is possible that protein 4 is an alternative splice variant of TgAda2-B or a proteolytically processed form resulting in a smaller protein. Alternatively, it could be a degradation product. Splice variants are known to exist for Ada2 proteins but their functional importance is still unknown (Kusch et al., 2003).

Protein 9 possesses homology to heat shock protein 70 (hsp70), which has been shown to be involved in bradyzoite to tachyzoite conversion in *Toxoplasma* (Weiss et al., 1998; Silva et al., 1998). The Hsp70 family contains a number of highly-related protein isoforms ranging in size from 66kD to 78kD (Tavaria et al., 1996). Protein 9 is migrating too low to fit within this size range

and therefore is either a unique shorter hsp70 member or breakdown from a larger hsp70 member. Nonetheless, the presence of hsp70 interacting with TgGCN5 is a strong possibility despite the lack of full intact protein. Transcriptional studies have shown that histone acetylation can significantly enhance both the basal and the inducible expression of hsp70 gene in *Drosophila* (Chen et al., 2002). It is plausible that a TgGCN5 containing complex is involved in remodeling and regulating hsp70 expression. A similar phenomenon may occur in *Toxoplasma*. A TgGCN5 containing complex may be activated under stress conditions resulting in the upregulation of hsp70 expression and ultimately leading to bradyzoite differentiation. Following the increase in hsp70 expression, the hsp70 protein may interact with TgGCN5 containing complexes to further direct their activity to other stress responsive promoters or it may act as a negative feed back, repressing the activity of the TgGCN5 containing complex once TgGCN5 differentiation is complete.

Proteins 7 and 10 have homology to previously described heat shock proteins 90 and 60 in *Toxoplasma* (Weiss and Kim, 2000). These two heat shock proteins have also been linked to bradyzoite differentiation (Echeverria et al, 2005; Toursel et al., 2000). If this is a true interaction, it links TgGCN5 to stress-induced pathways and possibly bradyzoite differentiation. It should be noted that heat shock proteins are among the most abundant proteins in eukaryotic cells (Schlesinger, 1990) and are often contaminants in co-immunoprecipitations. Co-immunoprecipitations with *in vitro* translated heat shock proteins and full length TgGCN5 can be utilized to verify if these interactions are real.

Protein 8 was identified as a homologue of yeast cdc48 (cell division control) which is a member of the type II AAA proteins (ATPase associated with a variety of activities; Neuwald et al., 1999). As the class name suggests, cdc48 has been implicated in many cellular functions including cell cycle regulation, stress response, transcriptional regulation, and protein degradation. All of these activities seem to revolve around its involvement in the ubiquitin-proteosome system (Hershko and Ciechanover, 1998). A link between acetylation and

ubiquitination pathways was established when cdc48 was shown to interact with a histone deacetylase (HDAC) complex in mice (Seigneurin-berny et al., 2001). Additionally, the yeast deubiquitinating enzyme, Ubp8 is a component of SAGA and mediates the removal of ubiquitin from histone H2B (Daniel et al., 2004). The deubiquitination of H2B by SAGA coincides with its acetylation of histone H3 at lysines 9 and 14 (Daniel et al., 2004). Therefore, ubiquitination may direct the function of GCN5 containing complexes. A similar process might exist in *Toxoplasma* between TgGCN5 containing complexes and ubiquitination. Thus, cdc48 may be involved in ubiquitinating histones while a TgGCN5 containing complex is responsible for removing the ubiquitin modification and acetylating histones. However, cdc48 is highly abundant in cells (great than 1% of total protein) making it a potential contaminant in co-immunoprecipitations (Wang et al., 2004). Therefore, this potential interaction must be confirmed in additional experiments such as co-immunoprecipitation of purified recombinant proteins.

The two remaining proteins, 5 and 6, do not have homology to any known proteins nor do they possess any recognizable motifs. These two proteins may be parasite-specific components of a TgGCN5 acetylase complex. However, these two proteins are migrating between 100 to 150kD, which is much higher than their predicted molecular weights according to the *Toxoplasma* database (ToxoDB; ~61kD for protein 5 and ~71kD for protein 6). Another band corresponding to protein 5 migrated to approximately 55kD which is close to its predicted molecular weight. The larger version of proteins 5 and 6 may be alternative splice variants. Further characterization of the ORF (open reading frame) encoding these two proteins is needed to clarify the discrepancy in sizes of these parasite proteins.

D. Identification of proteins interacting with ΔN_TTgGCN5 to form HAT complexes

Similar purifications of whole cell lysate from parasites over-expressing $_{FLAG}\Delta N_T$TgGCN5 were analyzed (Figure 22). Despite using the same purification protocol, the final elutions following purification of $_{FLAG}\Delta N_T$TgGCN5 whole cell

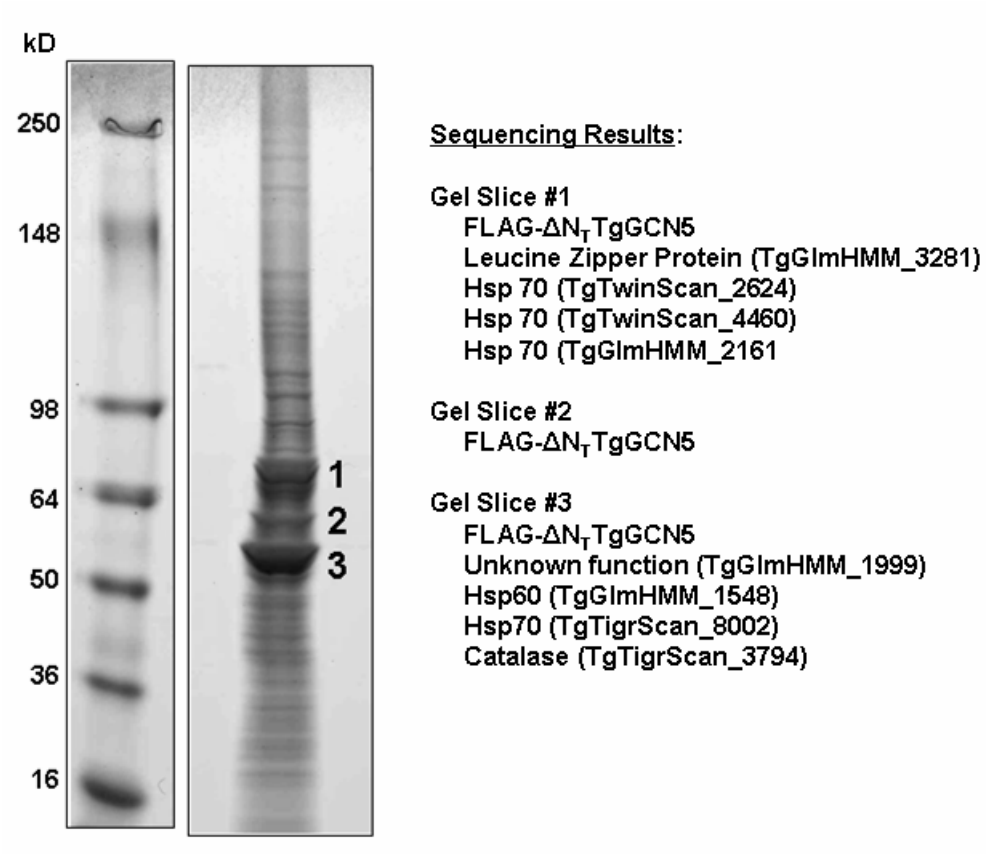


Figure 22: Chromatographic purification of proteins that may interact with $_{FLAG}\Delta N_T$TgGCN5

Elutions were pooled, TCA precipitated, resolved by SDS-PAGE (4% - 12%), and visualized by colloidal blue staining. Molecular weight markers in kilodaltons (kD) are indicated on the left. Gel slices were made corresponding to the location of the numbers on the gel and were sent for protein sequencing. Proteins identified by mass spectrometry sequencing are indicated on the right. Data presented in this image was generated by Dr. Mohamed-Ali Hakimi's laboratory at the French National Centre for Scientific Research in Grenoble, France.

extract were not as pure as the elutions following purification of full length TgGCN5 nuclear cell extract. The three most intense bands were excised from the gel and sent for sequencing. All three bands appear to contain $_{FLAG}\Delta N_T$TgGCN5. The estimated size of $_{FLAG}\Delta N_T$TgGCN5 is 54kD, which corresponds to the size of band 3. The slower migrating TgGCN5 found in bands 1 and 2 may be posttranslationally modified.

Also found in large abundance in the first gel slice is an unknown predicted protein (TgGlmHMM_3281) containing a leucine zipper domain. Leucine zipper domains are commonly found in gene regulatory proteins (Robinson and Lopes, 2000). However, it is running smaller than the 140 kilodaltons predicted in the ToxoDB. The protein present as this size may be degradation product or it could be a protein prediction error.

Heat shock protein 70 is also present in band one. As mentioned earlier, this protein is associated with stress pathways potentially linking TgGCN5 to bradyzoite conversion. All three predicted hsp70 proteins appear to be different and are not overlapping entries. Thus, the possibility exists for multiple hsp70 proteins in *Toxoplasma*. TgGlmHMM_2161 is predicted to be approximately 133kD, but this may be an error in the prediction algorithm used by the *Toxoplasma* database.

The third gel slice contains several bands in addition to TgGCN5. However, the only protein migrating at its predicted size is TgGlmHMM_1548, an hsp60 homologue. The same homologue was found in the co-immunoprecipitation with full length TgGCN5. TgGlmHMMM_1999 is a protein of unknown function and possesses no known motifs but is predicted to be 43kD. It is possible that this protein is a parasite-specific component of an acetylase complex. Predicted heat shock protein 70 and catalase are present but both of these proteins are running smaller than predicted size (72kD for hsp70 and 88kD for catalase), indicating these may be breakdown products or potential splice variants. The hsp70 protein overlaps with the TgTwinScan_7546 entry found in band one and therefore maybe a splice variant from the same genomic loci. Splice variants for heat shock proteins are known to exist in mammals, but their

function in stress remediation is still unknown (Yamada et al., 1999). To date, the possibility of splice variants of the heat shock proteins 60, 70 and 90 that are involved in bradyzoite conversion has not been addressed in *Toxoplasma*.

Catalase is an enzyme involved in oxidative stress remediation responsible for catalyzing the conversion of hydrogen peroxide into water and oxygen (Chelikani et al., 2004). A homologue has been characterized in *Toxoplasma* and it appears to serve a protective function against macrophage reactive oxygen specie formation (Murray and Cohn, 1979; Murray et al., 1980) and maybe involved in parasite virulence (Ding et al., 2004). A potential interaction between GCN5 and catalase provides another link between HATs and stress remediation in the parasite. However, catalyse is an abundant cytoplasmic protein in *Toxoplasma* (Ding et al., 2004) and therefore it could be a contaminant in the purification of $_{FLAG}\Delta N_T$TgGCN5.

III. Characterization of the enzymatic activity of TgGCN5

All GCN5 homologues and the closely related HAT PCAF preferentially acetylate histone H3 at lysine 14. To a lesser degree, GCN5 and PCAF can also acetylate lysines 9 and 18 on histone H3 *in vitro* (Grant, 2001). Given the high homology of the catalytic domain of TgGCN5 to other GCN5 proteins (Sullivan and Smith, 2000), it was expected to also preferentially acetylate histone H3 at lysine 14. Therefore, recombinant TgGCN5 was obtained and *in vitro* HAT assays were performed to determine its acetylation pattern on core histones.

The N-terminal extension in mammalian GCN5 proteins is not required for the acetylation of free histones (Xu et al., 1998). As the N-terminal extension of TgGCN5 is divergent from mammalian homologues, it is important to investigate this function for the N-terminal extension in TgGCN5. If the N-terminus of TgGCN5 is required for the acetylation of free histones, it would provide evidence that the N-terminus has a parasite-specific function not present in its mammalian counterparts.

A. Development of a positive control for in vitro HAT assays

Before the enzymatic activity of TgGCN5 could be assessed, a positive control needed to be created for the *in vitro* HAT assay. The catalytic activity of the *Saccharomyces cerevisiae* GCN5 homologue (ScGCN5) has been rigorously studied and well characterized using published *in vitro* assays (Sternglanz and Schindelin, 1999). Therefore it seemed excellently suited to be used as a positive control for our purposes. Recombinant protein is commercially available but it is extremely expensive. Thus, it was decided to generate our own recombinant yeast GCN5 protein using a bacterial cell line.

As outlined in Chapter 2 (Section IV-D, page 75), ScGCN5 was expressed using a BL21-CodonPlus® *E. coli* strain (Stratagene) in frame with an N-terminal 6X histidine fusion tag to allow for purification using nickel-NTA resin. The CodonPlus bacterial strains have been designed to solve the problem of codon bias which can limit efficient expression of heterologous proteins in bacteria.

Codon bias refers to the observation that there is a bias in the genome of an organism for one or two codons for almost all degenerate codon families (Makrides, 1996). The codon bias is different between organisms, reflected in the abundance of their cognate tRNAs. Therefore, heterologous genes enriched with codons that are rarely used by *E. coli* may not be efficiently expressed (Gustafsson et al., 2004). Forced high-level expression of heterologous proteins can deplete the pool of rare tRNAs and stall translation. BL21-CodonPlus® strains are engineered to contain extra copies of genes that encode the tRNAs that most frequently limit translation of heterologous proteins in *E. coli* (Stratagene, 2002). Increased availability of these rare tRNAs allows for high-level expression of recombinant genes in BL21-CodonPlus® cells that are poorly expressed in conventional BL21 strains (Stratagene, 2002).

There are two different BL21-CodonPlus® stains, RIL and RP. CodonPlus-RIL® cells contain extra copies of the *argU*, *ileY*, and *lueW* tRNA genes (Stratagene, 2002). These genes encode tRNAs that recognize the arginine codons AGA and AGG, the isoleucine codon AUA, and the leucine codon CUA, respectively. The CodonPlus-RIL® strains express extra tRNAs that

most frequently restrict translation of heterologous proteins from genes that are AT-rich (Stratagene, 2002). BL21-CodonPlus-RP® cells contain extra copies of the *argU* and *proL* genes (Stratagene, 2002). These genes encode tRNAs that recognize the arginine codons AGA and AGG and the proline codon CCC, respectively. The BL21-CodonPlus-RP® strains are enriched with tRNAs that most frequently restrict translation of heterologous proteins of organisms that have GC-rich genomes (Stratagene, 2002). For the expression of yeast GCN5, the BL21-CodonPlus-RIL® strain was used because the yeast GCN5 gene has an AT ratio of 62% suggesting that the BL21-CodonPlus-RIL® strain would be best suited for optimal protein expression.

The CodonPlus® strains have been engineered using the *lac operon* for IPTG (isopropyl β-D-1-thiogalactopyranoside) induced protein expression. Protein induction was carried out as outlined in Chapter 2 (Section IV-D, page 75). Figure 23-A shows SDS-PAGE analysis of a small scale pilot induction of a BL21-CodonPlus-RIL bacterial clone expressing ScGCN5. The first two lanes are from the uninduced (UN) bacteria to serve as a negative control. The first lane contains the soluble (S) fraction and the second lane contains the pellet (P) or insoluble fraction from the bacterial cells just prior to protein induction. The last two lanes are after three hours of protein induction (IN) at 37°C. Again, the first lane contains the soluble (S) fraction and the second lane is the pellet (P) or insoluble fraction. When compared to the uninduced sample, there is a band at approximately 55 kilodaltons in the induced sample that is ScGCN5. The calculated size of ScGCN5 is 48kD, but ScGCN5 has been shown to migrate slightly higher than its calculated size during SDS-PAGE analysis (Candau et al., 1997).

The induced protein appears to be in sufficient quantities in the soluble fraction to be processed for purification. Recombinant protein within the soluble fraction is desirable because it is more likely to possess enzymatic function (Makrides, 1996). Protein in the pellet or insoluble fraction is contained within inclusion bodies and is less likely to be enzymatically active due to improper folding (Makrides, 1996).

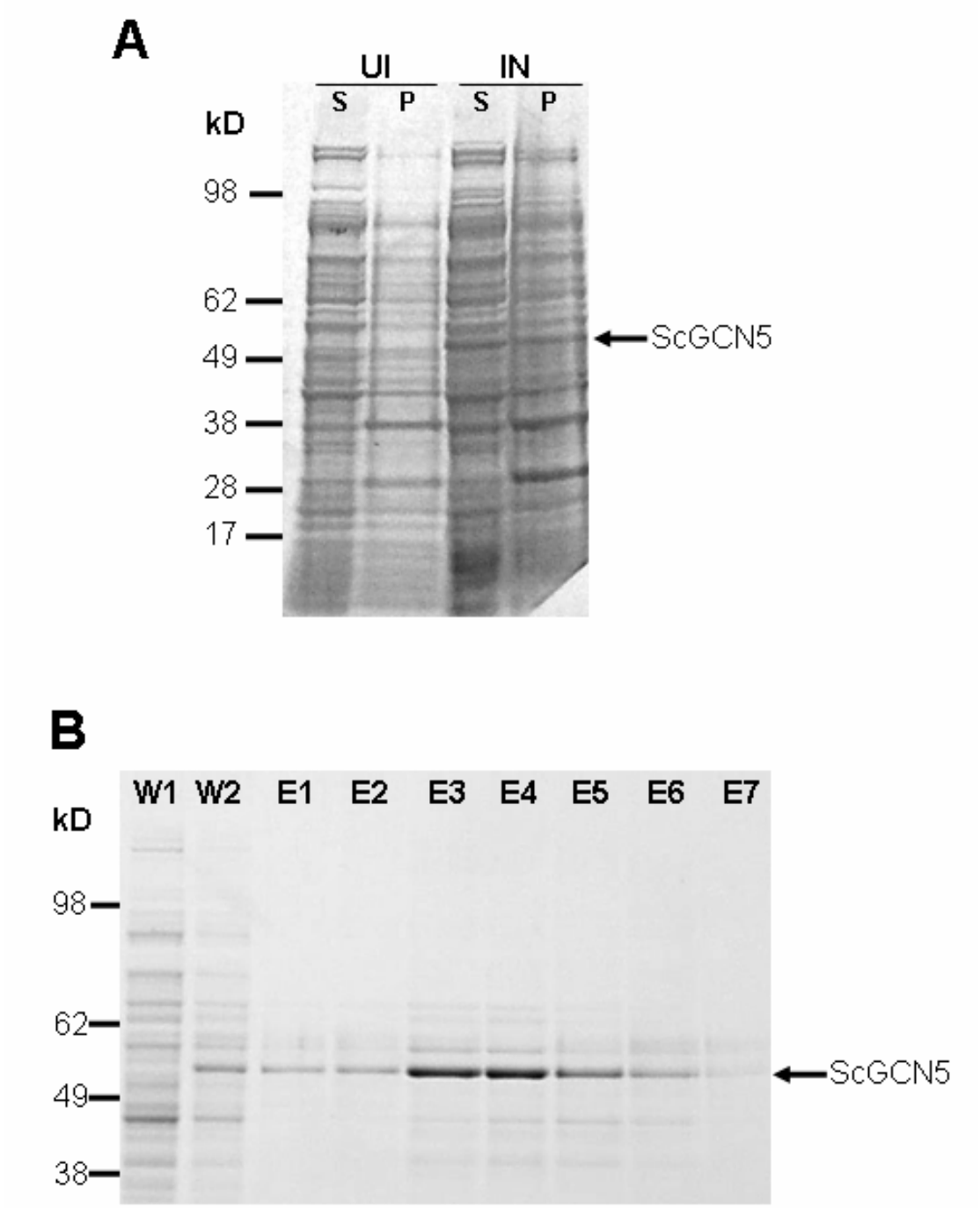

A
UI
IN
S
P
S
P
kD
98
62
49
38
28
17
ScGCN5
B
W1
W2
E1
E2
E3
E4
E5
E6
E7
kD
98
62
49
38
ScGCN5

Figure 23: Protein induction and purification of ScGCN5

A – SDS-PAGE of a pilot induction of ScGCN5 in BL21-CodonPlus-RIL stained with Simply Blue (Invitrogen) colloidal stain. The first two lanes are from uninduced (UN) bacteria cells just prior to protein induction. The last two lanes are after three hours of protein induction (IN) at 37°C. (S = soluble fraction; P = pellet or insoluble fraction). The band present at 55kD presumed to be ScGCN5 is denoted by an arrow. **B** – SDS-PAGE of washes and subsequent elutions from the purification of a large scale induction of ScGCN5 stained with Simply Blue. Following incubation with bacterial lysate, the nickel-NTA resin was washed twice (W1 and W2) with 20mM imidazole. Recombinant protein was eluted from the resin using increasing concentrations of imidazole: E1 and E2 = 100mM imidazole, E3 and E4 = 200mM imidazole, and E5, E6, and E7 = 300mM.

Once expression of soluble protein had been confirmed, a large scale induction was performed and the ScGCN5 protein was purified by virtue of the 6X histidine tag using nickel-NTA resin (Qiagen) as outlined in Chapter 2 (Section IV-D, page 75). Figure 23-B shows the results of the nickel purification. Elutions E2 through E6 contain significant portions of ScGCN5 in reasonable purity. Therefore, these five elutions were pooled and concentrated using an Amicon Ultra-4 Centrifugal Filter Device (Fisher #UFC801008). The protein concentration of the pooled elutions was determined using a standard Bradford assay.

Figure 24-A shows an SDS-PAGE analysis of ~1.0µg of ScGCN5 protein and a Western blot stained with an anti-histidine antibody confirming the production of a histidine tagged protein. The SDS-PAGE and Western blot show several additional bands that are not full-length ScGCN5. The bands are either contaminating bacterial proteins that co-purified with the ScGCN5 during nickel resin purification or possibly ScGCN5 degradation products that were generated during purification. A commercially available His-Tag Protease Inhibitor Cocktail (Sigma #P8849) was added to all solutions used during purification to minimize degradation. Alternatively, the lower weight bands could also be truncation products from incomplete translation of the recombinant protein. As the 6X histidine tag is located at the N-terminus, truncation products would co-purify with full length protein.

The presence of contaminating bacterial proteins does not pose a serious concern for the use of ScGCN5 in the *in vitro* HAT assay for our purposes. Bacteria do not possess histones, so it is highly unlikely that the contaminating bacterial proteins will adversely affect the enzymatic activity of ScGCN5 in the HAT assay. To date, no bacterial proteins have been discovered that are capable of acetylating histones (Roth et al., 2001). The lack of histone modifying enzymes in bacteria is one of the advantages of expressing HAT proteins in bacteria. Any contaminating bacterial proteins that co-purify with the recombinant protein should not cause aberrant acetylation in the HAT assay.

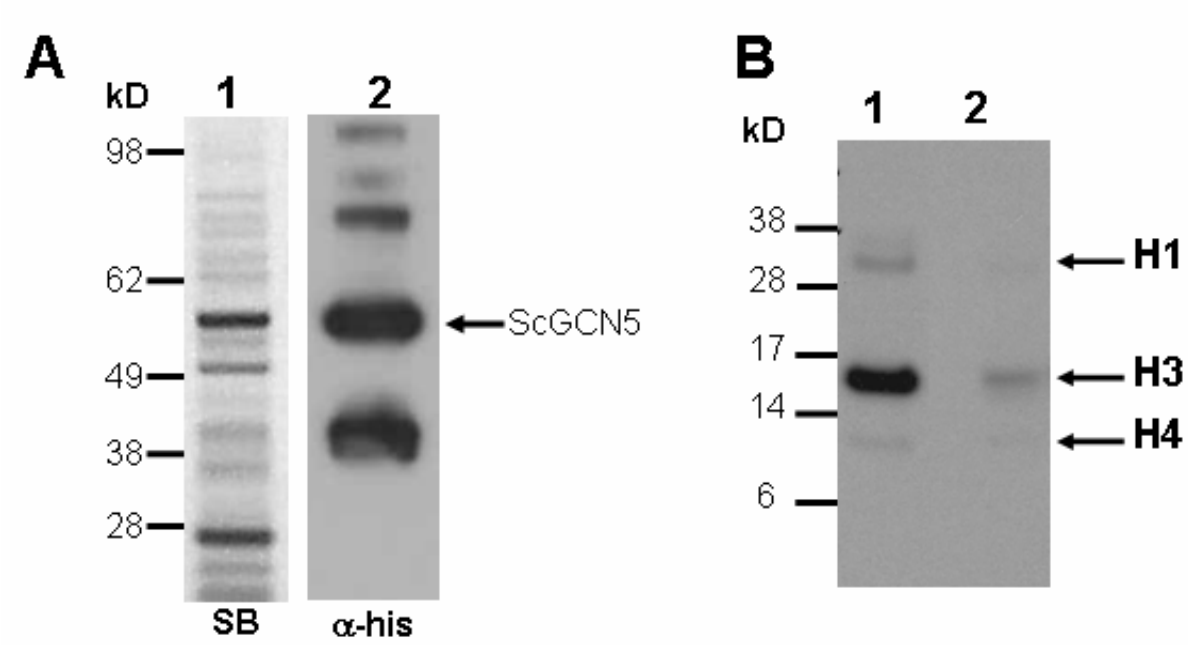


Figure 24: Recombinant ScGCN5 possesses HAT activity

A – Lane 1 is an SDS-PAGE of ~1.0µg of ScGCN5 stained with Simply Blue (SB, Invitrogen) colloidal stain. Lane 2 is a Western of 500ng of TgGCN5 stained with a monoclonal anti-histidine antibody (α-his, Clontech #PT3359-2) at a dilution of 1:5000 and the secondary antibody was goat derived anti-mouse (Amersham #NA931) used at 1:5000. Detection was performed using Amersham ECL reagent (#RPN2209) with High Sensitive Blue photographic film (RPS Imaging #33-0810). **B** – Autoradiograph of *in vitro* HAT assay with ~1.0µg of ScGCN5 and core histones exposed to High Sensitive Blue photographic film (RPS Imaging #33-0810). Markers are displayed on left in kilodaltons (kD).

It is doubtful that the truncated proteins and/or degradation products will interfere with full length enzymatically active ScGCN5 in the HAT assay. The truncated proteins and degradation products probably do not possess enzymatic function. However, the presence of truncated protein, degradation products, and contaminating bacterial proteins prevents accurate determination of the concentration of full length ScGCN5. The amount of full length ScGCN5 protein present in the pooled elutions is an unknown percentage of the total protein concentration determined by Bradford assay. Therefore, the actual amount of full length ScGCN5 protein used in our HAT assays is unknown. Fortunately, our experiments do not require the exact amount of active HAT protein used in the HAT assay to be known.

After obtaining recombinant ScGCN5 in sufficient quantity, the functionality was assessed in an *in vitro* HAT assay. The mostly commonly used *in vitro* HAT assay reported in the literature measures HAT activity using tritiated acetyl CoA and core histones (Brownell et al., 1996). A protein that possesses HAT activity will transfer the tritiated acetyl group from the acetyl CoA to the core histones present in solution. The reaction is then resolved by SDS-PAGE and processed for autoradiography. Histones modified with the radioactive tritiated acetyl group will form bands when the gel is exposed to film.

Figure 24-B contains the autoradiograph of a HAT assay with ~1.0µg of ScGCN5 (lane 1) demonstrating that the recombinant ScGCN5 possesses HAT activity in our *in vitro* HAT assay based the assay published by Brownell et al. (1996). As a negative control (lane 2), a HAT assay was performed containing only histones and tritiated acetyl CoA. Clearly, the HAT activity of the recombinant ScGCN5 far exceeds the minor activity within the negative control. The small activity present in the negative control suggests that these histones are being acetylated. Background activity has been a constant problem in HAT assays in other labs (Cote, J; unpublished observations). The histones used in the HAT assay are commercially available histones (Upstate) purified from chicken erythrocytes. It is plausible that they are contaminated with minor amounts of chicken HATs which generate minor background activity in the assay.

Regardless of the cause of the background activity, ScGCN5 possesses activity much higher than the activity in the negative control. Thus, we have successfully reproduced the published *in vitro* HAT assay with a functionally relevant positive control. As predicted from the literature, ScGCN5 shows a strong preference for acetylating histone H3 and has minor transferase activity on histone H4 (Kuo et al., 1996). ScGCN5 also appears to acetylate histone H1. Other GCN5 homologues have been shown to also acetylate histone H1 *in vitro* but not *in vivo* (Herrera et al., 1997).

B. Impact of the TgGCN5 N-terminus on histone acetylation

Once the *in vitro* HAT assay had been successfully tested and a working positive control had been generated, the next objective was to obtain pure recombinant TgGCN5 protein. As a result of the success in expressing ScGCN5 in BL21-CodonPlus cells, an attempt was made to also express TgGCN5 in bacteria fused to an N-terminal 6X histidine tag. Unlike ScGCN5, TgGCN5 has a GC ratio of 60% making the RP strain of the BL21-CodonPlus cell line better suited for expression.

TgGCN5 is over twice the size of ScGCN5, therefore the protein induction protocol was altered to maximize expression of soluble protein in the bacteria. Lowering the temperature of bacterial cultures during induction has been shown to facilitate increased expression of soluble protein (Shirano and Shibata, 1990). Therefore, the protein induction of TgGCN5 was performed at 15°C instead of 37°C. Lowering the temperature of induction slows the bacterial translation machinery allowing for more efficient translation (Makrides, 1996). The drawback of using a lower temperature is a reduction in the total amount of protein expressed. However, a greater fraction of the protein that is expressed will be in the soluble fraction and be more likely to possess enzymatic activity (Shirano and Shibata, 1990). To compensate for the reduction in the rate of protein synthesis the duration of induction was extended from three hours to sixteen hours. The induction protocol used for generating TgGCN5 is outlined in greater detail in Chapter 2 Section IV-D (page 75).

Figure 25-A shows an SDS-PAGE analysis of a pilot induction of a BL21-CodonPlus-RP bacterial clone expressing TgGCN5. As a negative control, the first two lanes are from uninduced (UN) bacteria. The first lane contains the soluble (S) fraction and the second lane contains the pellet (P) or insoluble fraction from the bacterial cells just prior to induction. The last two lanes are after sixteen hours of protein induction (IN) at 15°C. The first lane contains the soluble (S) fraction and the second lane is the pellet (P) or insoluble fraction. When compared to the uninduced sample, there is a very abundant band between 105kD and 160kD which is congruent with the 130kD size of TgGCN5. A sufficient amount of recombinant protein is present in the soluble fraction to be process for purification.

After confirming adequate expression of soluble protein, a large scale induction was performed and TgGC5 was purified via the 6X histidine tag using nickel-NTA resin (Qiagen) as outlined in Chapter 2 Section IV-D (page 75). Figure 25-B shows the results of the nickel resin purification. Elutions E1 through E7 contain significant portions of TgGCN5 in reasonable purity. Therefore, all seven elutions were pooled and concentrated using an Amicon centrifugal filter concentrator (Millipore). The protein concentration of the pooled elutions was determined using a standard Bradford assay.

Figure 26-A shows SDS-PAGE analysis of 500ng of TgGCN5 protein and a Western blot stained with an anti-histidine antibody (Clontech) confirming the production of a histidine tagged protein. The SDS-PAGE and Western blot show several additional bands that are not full-length ScGCN5. The bands are either contaminating bacterial proteins that co-purified with the TgGCN5 during nickel resin purification or are truncated products from incomplete translation of the recombinant protein. Despite the presence of additional proteins, the major protein present is full length TgGCN5.

The enzymatic activity of the purified recombinant TgGCN5 was assessed in an *in vitro* HAT assay (Figure 26-B). Lane 1 is the positive control performed with 1.0µg of ScGCN5 and Lane 2 is the HAT assay performed with 500ng of TgGCN5. As a negative control (lane 3), a HAT assay was performed containing

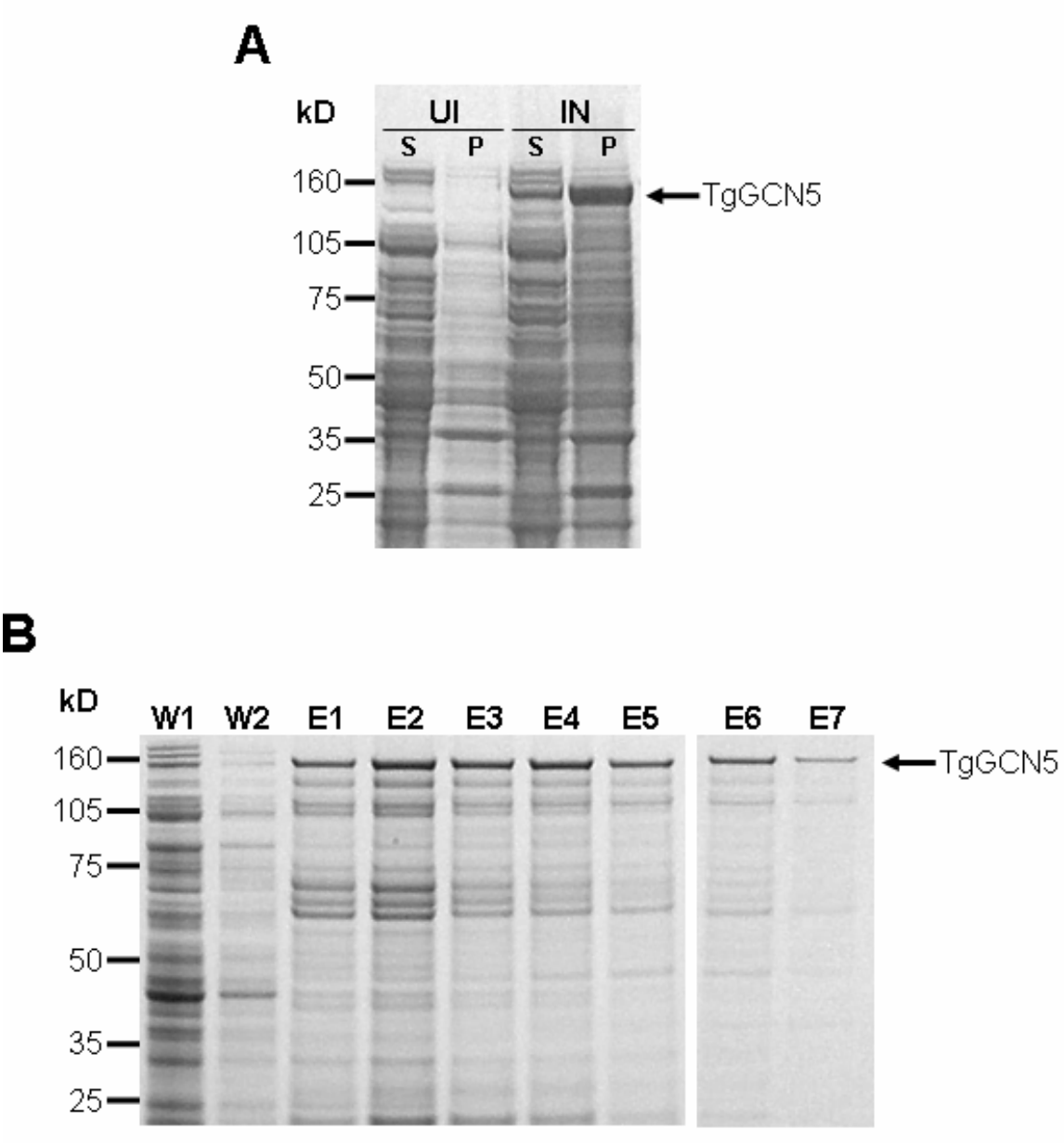

A
kD
UI
IN
S
P
S
P
160
105
75
50
35
25
TgGCN5
B
kD
W1
W2
E1
E2
E3
E4
E5
E6
E7
160
105
75
50
35
25
TgGCN5

Figure 25: Bacterial protein induction and purification of TgGCN5

A – SDS-PAGE of a pilot induction of TgGCN5 in BL21-CodonPlus-RP stained with Simply Blue colloidal stain (Invitrogen). The first two lanes are from uninduced (UN) bacteria cells just prior to protein induction. The last two lanes are after an overnight protein induction (IN) at 15°C. (S = soluble fraction; P = pellet or insoluble fraction). The band present at 130kD presumed to be TgGCN5 is denoted by an arrow. **B** – SDS-PAGE of washes and subsequent elutions from the purification of a large scale induction of TgGCN5 stained with Simply Blue. Following incubation with bacterial lysate, the nickel-NTA resin was washed twice (W1 and W2) with 20mM imidazole. Recombinant protein was eluted from the resin using increasing concentrations of imidazole: E1 and E2 = 100mM imidazole, E3 and E4 = 200mM imidazole, and E5, E6, and E7 = 300mM. Markers are on left displayed in kilodaltons (kD).

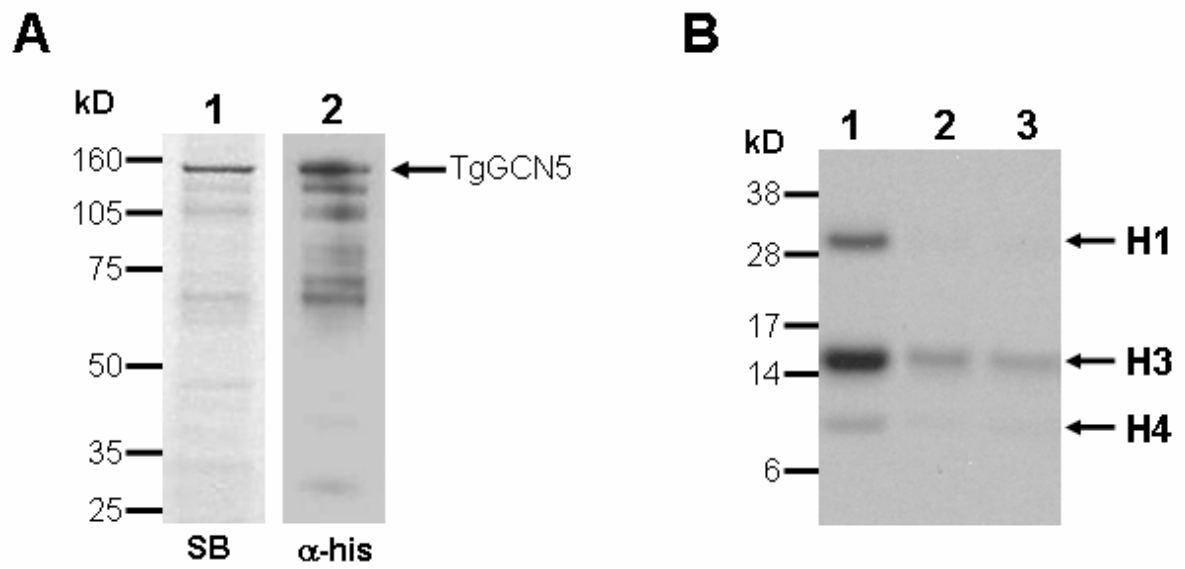


Figure 26: Recombinant TgGCN5 produced in bacteria lacks HAT activity
A – Lane 1 is an SDS-PAGE of 500ng of TgGCN5 stained with Simply Blue colloidal stain (SB, Invitrogen). Lane 2 is a Western of 500ng of TgGCN5 stained with a monoclonal anti-histidine antibody (α-his, Clontech #PT3359-2) at a dilution of 1:5000 and the secondary antibody was goat derived anti-mouse (Amersham #NA931) used at 1:5000. Detection was performed using Amersham ECL reagent (#RPN2209) with High Sensitive Blue photographic film (RPS Imaging #33-0810). **B** – *In vitro* HAT assay with 1µg of ScGCN5 (Lane 1), 500ng of TgGCN5 (Lane 2), and core histones alone as a negative control (Lane 3) exposed to High Sensitive Blue photographic film (RPS Imaging #33-0810). Markers are on left displayed in kilodaltons (kD).

only histones and tritiated acetyl CoA. Unfortunately, the recombinant TgGCN5 produced in BL21-CodonPlus-RP strain does not appear to possess significant HAT activity above background. The positive control (ScGCN5, Lane 1) shows intense HAT activity on core histones suggesting that no error was made in performing the HAT assay. The HAT assay was repeated two more times but no HAT activity was detected. Recombinant TgGCN5 was purified from two additional BL21-CodonPlus-RP bacterial clones but the recombinant protein from these different clones did not possess HAT activity in the *in vitro* assay, either.

Several explanations are possible for the lack of activity of TgGCN5. The most plausible is that the BL21-CodonPlus-RP strain is unable to produce enzymatically active protein. Unlike eukaryotic organisms, bacteria lack the ability to perform many of the posttranslational modifications found in eukaryotic proteins and they have limited ability to facilitate extensive disulfide bond formation (Makrides, 1996). Bacteria do not possess chaperone proteins that may be required to assist in folding proteins into their proper conformation. Protocols have been established to solve the problem of disulfide bond formation and proper protein folding increasing the likelihood of obtaining functionally active recombinant protein from bacteria. Many of these protocols involve the use of different bacterial strains, additional fusion tags or co-expression of chaperone proteins (Derman et al., 1993; Cole, 1996; Terpe, 2003). While it is possible that these methods may have been successful in producing enzymatically active TgGCN5 in bacteria, a large amount of time would have been spent optimizing protein expression conditions without any guarantee of success. Therefore it seemed best to try alternative methods for obtaining recombinant protein.

To obtain active TgGCN5 protein, several options were possible. One method would be to purify TgGCN5 directly from parasite lysate. A TgGCN5 antiserum has been developed in our lab and it was used in previous experiments exploring the interaction between TgGCN5 with importin α. It is theoretically possible to purify endogenous TgGCN5 from wild type parasite using the same TgGCN5 antiserum. However, low level expression of endogenous TgGCN5 would limit our ability to obtain sufficient amounts of the

protein to use in an *in vitro* assay. An alternative would be to express recombinant protein in a eukaryotic expression system such as yeast or baculovirus-insect cell systems (Schmidt and Hoffman, 2002). Eukaryotic cells possess chaperon proteins and are capable of readily generating disulfide bonds offering a greater likelihood that the recombinant protein would possess enzymatic activity (Schmidt and Hoffman, 2002). However, these expression systems require laborious setup and optimization for efficient expression of large quantities of protein. A third possibility, that had the potential to offer abundant pure protein without significant setup time, was to express the TgGCN5 protein in *Toxoplasma*. Others have had success in expressing and purifying recombinant proteins directly from transgenic parasites (Donald and Liberator, 2002). The advantage of expressing recombinant protein in *Toxoplasma* is that it guaranteed the proper folding and posttranslational modifications required for full HAT activity. Furthermore, the two parasite lines necessary for our experiments had already been generated as part of the TgGCN5 NLS mapping experiments: (1) full length TgGCN5 fused to a FLAG tag ($_{FLAG}$TgGCN5) and (2) the truncated from lacking the N-terminal extension also fused to a FLAG tag ($_{FLAG}\Delta N_T$TgGCN5). Recombinant protein from the parasites could be purified from parasite lysate via the N-terminal FLAG tag using commercially available anti-FLAG affinity resin.

Commercially available anti-FLAG affinity resin (Sigma) was used to purify $_{FLAG}$TgGCN5 and $_{FLAG}\Delta N_T$TgGCN5 from transgenic parasites (Chapter 2, Section IV-E, page 80). The purified recombinant protein was eluted from the affinity resin using competitive elution via a concentrated solution of 3XFLAG peptide (Sigma). The elution peptide was removed by dialysis and a standard Bradford assay was used to determine the concentration of the recombinant protein. Unfortunately, no protein was detectable by Bradford assay in either of the dialyzed elutions of $_{FLAG}$TgGCN5 or $_{FLAG}\Delta N_T$TgGCN5. The purification was repeated several more times with similar results. The most likely causes of low protein yield were either inefficient affinity purification and/or sample loss during

dialysis. To determine the efficiency of the affinity purification, a Western blot stained with anti-FLAG was performed on transgenic parasite lysate, the elution before dialysis and the affinity resin from which protein had been eluted (Figure 27). To analyze the resin by Western blot, the resin was boiled in SDS-PAGE loading buffer following the elution step. The loading buffer was run on a gel and transferred to a Western blot. As shown in Figure 27, Lane 2, there appears to be no $_{FLAG}$TgGCN5 or $_{FLAG}\Delta N_T$TgGCN5 present in the elutions that is detectable by Western blotting. However, in Figure 27, Lane 3 there is a strong band at the size expected for $_{FLAG}$TgGCN5 (130kD) and a faint band at the size expected for $_{FLAG}\Delta N_T$TgGCN5 (52kD) protein bound to the resin. Therefore, it appears that the affinity purified protein is not being efficiently eluted from the affinity resin. An alternative method of elution using acidic glycine was attempted, but the elutions also failed to have detectable protein following dialysis. Harsher conditions could have been employed to elute the protein such as high salt concentration, adding detergents, or using denaturants such as urea or guanidine to elute $_{FLAG}$TgGCN5 and $_{FLAG}\Delta N_T$TgGCN5. However, these elution conditions would have partially or completely denatured the HAT protein making it unusable for enzymatic analysis (Rutkowska and Skowron, 1999).

The Western blot results in Figure 27 indicate that the anti-FLAG affinity resin appears to be purifying sufficient protein from the parasite lysate for enzymatic assays. Therefore, an *in vitro* HAT assay was attempted without eluting the protein from the resin. The resin with $_{FLAG}$TgGCN5 or $_{FLAG}\Delta N_T$TgGCN5 still bound was added directly to the HAT assay. Following purification, resin containing the bound $_{FLAG}$TgGCN5 or $_{FLAG}\Delta N_T$TgGCN5 was split between Western blot analysis and the HAT assay. One third of the resin was boiled in SDS-PAGE loading buffer and processed for Western blot analysis. The remaining two-thirds of resin was used in an *in vitro* HAT assay. By dividing up the resin in this manner, a qualitative comparison can be made between the amount of protein obtained from affinity purification and the overall activity elicited in the HAT assay.

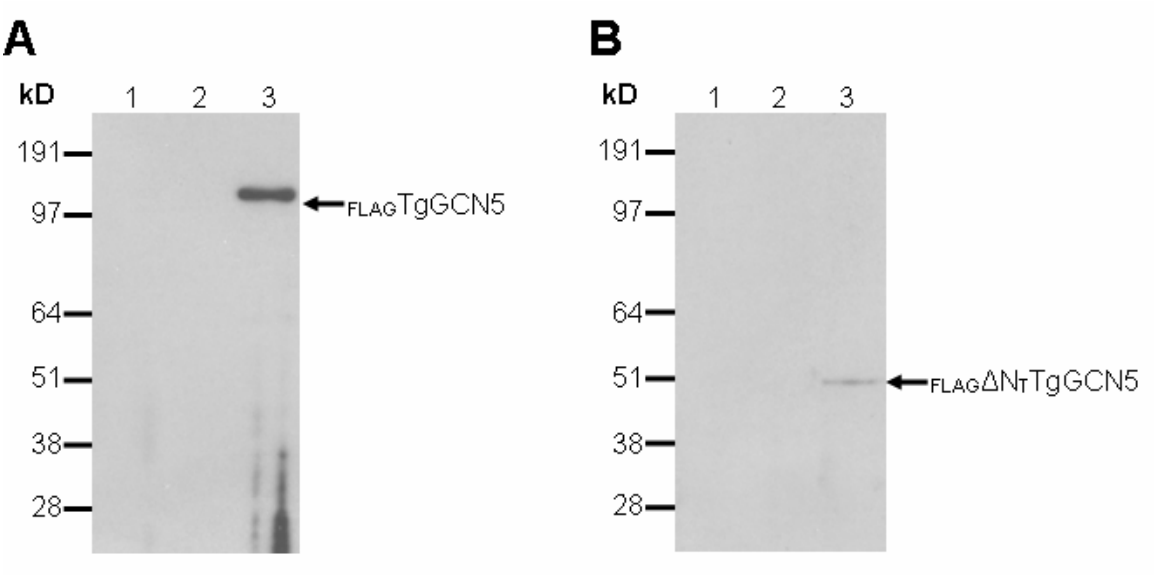


Figure 27: Elution of TgGCN5 and ΔN_TTgGCN5 from FLAG affinity resin
A – Western blot assessing the efficiency of the anti-FLAG affinity purification and elution of $_{FLAG}$TgGCN5. Lane 1 = lysate from parasites overexpressing $_{FLAG}$TgGCN5, Lane 2 = elution from anti-FLAG resin using 3X FLAG peptide, Lane 3 = Resin boiled in SDS loading buffer following elution with 3X peptide.
B – Western blot assessing the efficiency of the anti-FLAG affinity purification and elution of $_{FLAG}\Delta N_T$TgGCN5. Lane 1 = lysate from parasites overexpressing $_{FLAG}\Delta N_T$TgGCN5, Lane 2 = elution from anti-FLAG resin using 3X FLAG peptide, Lane 3 = Resin boiled in SDS loading buffer following elution with 3X peptide.
For both Western blots, the primary antibody was polyclonal anti-FLAG (Sigma #F7425) used at 1:1000 and the secondary antibody was goat derived anti-rabbit (Amersham #NA934) used at 1:2500. Detection was performed using Amersham ECL reagent (#RPN2209) with High Sensitive Blue photographic film (RPS Imaging #33-0810). Markers are on left displayed in kilodaltons (kD)

As shown in Figure 28A, the affinity purification pulled down adequate amounts of $_{FLAG}$TgGCN5 (Lane 1) and $_{FLAG}\Delta N_T$TgGCN5 (Lane 2) protein. Lane 3 in Figure 28A is of wild type lysate carried through anti-FLAG affinity purification to serve as a negative control in the HAT assy. As expected, it contains no FLAG tagged proteins.

Both $_{FLAG}$TgGCN5 and $_{FLAG}\Delta N_T$TgGCN5 while still bound to the affinity resin are enzymatically active and acetylate core histones (Figure 28B). In figure 28-B, Lane 1 is a HAT assay performed with ~1.0µg ScGCN5 to serve as a positive control. Lane 4 is a negative control HAT assay performed with anti-FLAG affinity resin from the purification of wild type lysate. In addition to the recombinant HATs expressed in the parasites, *Toxoplasma* also expresses other HAT proteins (Sullivan and Hakimi, 2006). The negative control demonstrates that the anti-FLAG affinity resin is not capable of purifying these endogenous *Toxoplasma* HATs through non-specific interactions. The results shown in Figure 28-A and -B are representative of three duplicate experiments which all yielded similar results.

As shown in Figure 28-B, both TgGCN5 and ΔN_TTgGCN5 are able to acetylate free non-nucleosomal histones in solution. As expected for a GCN5 HAT, the preferred substrate of both proteins is histone H3 and to a lesser extent H4. Full length TgGCN5 appears to be able to acetylate H1. The truncated form of TgGCN5 also shows detectable activity on H1 as well. The acetylation of H1 and H4 by GCN5 are believed to occur *in vitro* but not *in vivo* (Herrera et al., 1997).

The HAT assay results also suggest that full length TgGCN5 has higher activity than the truncated form. Although there appears to be more ΔN_TTgGCN5 present in the Western blot versus full length TgGCN5, ΔN_TTgGCN5 appears to have much weaker activity than TgGCN5 in the HAT assay. It is possible that the N-terminal extension stabilizes full length protein and provides stronger catalytic activity. However, the FLAG tag in both recombinant proteins is fused to the amino-terminal end. In $_{FLAG}\Delta N_T$TgGCN5,

this places the catalytic domain in close proximity to the affinity resin, possibly impeding its activity. Finding a

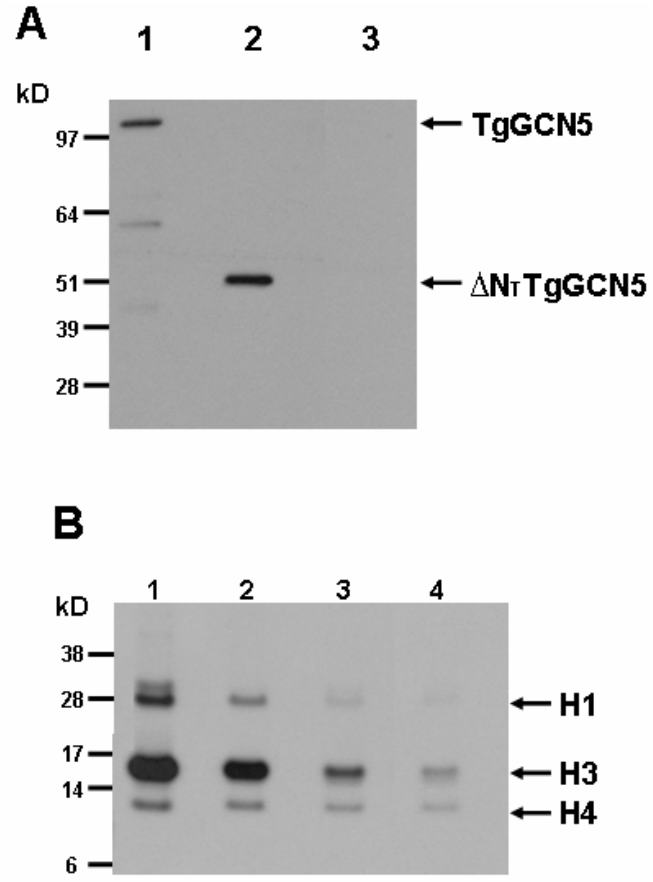


Figure 28: HAT activity of TgGCN5 and ΔN_TTgGCN5

A – Western blot of the anti-FLAG affinity purification. Primary antibody (polyclonal anti-FLAG, Sigma #F7425) used at 1:1000 and secondary antibody (goat derived anti-rabbit, Amersham #NA934) used at 1:2500. Detection was performed using Amersham ECL reagent (#RPN2209) with High Sensitive Blue photographic film (RPS Imaging #33-0810). **B** – Autoradiogram with High Sensitive Blue photographic film (RPS Imaging #33-0810) of an *in vitro* HAT assay using ^{3}H-Acetyl CoA (Amersham #TRK688) with chicken erythrocyte histones H1 (~30kD), H2A (~14kD) H2B (~15kD), H3 (~17kD), and H4 (~10kD). LANES: 1 = ~1.0g of ScGCN5, 2 = over-expressing $_{FLAG}$TgGCN5, 3 = over-expressing $_{FLAG}\Delta N_T$TgGCN5, and 4 = wild-type lysate as a negative control. Markers are on left displayed in kilodaltons (kD).

method to effectively elute the proteins off the affinity resin would give a better comparison. Alternatively, the FLAG tag could be moved the C-terminus in both proteins and the activity reassessed. Nevertheless, it is clear that both forms of TgGCN5 preferentially acetylate histone H3 *in vitro*.

C. Determining the lysine(s) acetylated by TgGCN5

The development of antibodies capable of recognizing specific modifications to histone proteins has made it possible to easily determine which lysine residues are targeted by HAT proteins using standard Western blot techniques (White et al., 1999). Once it had been determined that TgGCN5 preferentially acetylated histone H3, additional HAT assays were performed to identify which lysine(s) were being acetylated.

Since Western analysis was employed to determine the residues that are acetylated by TgGCN5, radioactively labeled acetyl CoA was not required in the HAT assay. Detection was obtained using chemiluminescence via a secondary antibody conjugated to horseradish peroxidase and exposing the blot to film. Therefore, the HAT assay protocol was altered, switching the radioactively labeled acetyl CoA with a similar solution of non-radioactive acetyl CoA (Sigma) dissolved in acidic sodium acetate. In addition, the substrate was changed from core histones to commercially available recombinant histone H3, the preferred substrate for GCN5 proteins. A detailed protocol of the non-radioactive HAT assay is in Chapter 2, Section IV-G (page 83).

As in the previous HAT assays, anti-FLAG affinity purification was performed on parasites overexpressing $_{FLAG}$TgGCN5 or $_{FLAG}\Delta N_T$TgGCN5. Affinity purification with anti-FLAG resin was also performed on wild type parasites to serve as a negative control for the HAT assay. Following purification, the affinity resin was split between Western analysis (Figure 29-A) and HAT assay (Figure 29-B). ScGCN5 continued to serve as a positive control in these HAT assays (Figure 29-B, Lane 1). Following the HAT assay, the reactions were run on a gel and transferred to a PVDF membrane. Acetylated lysines were detected by Western analysis using antibodies capable of detecting

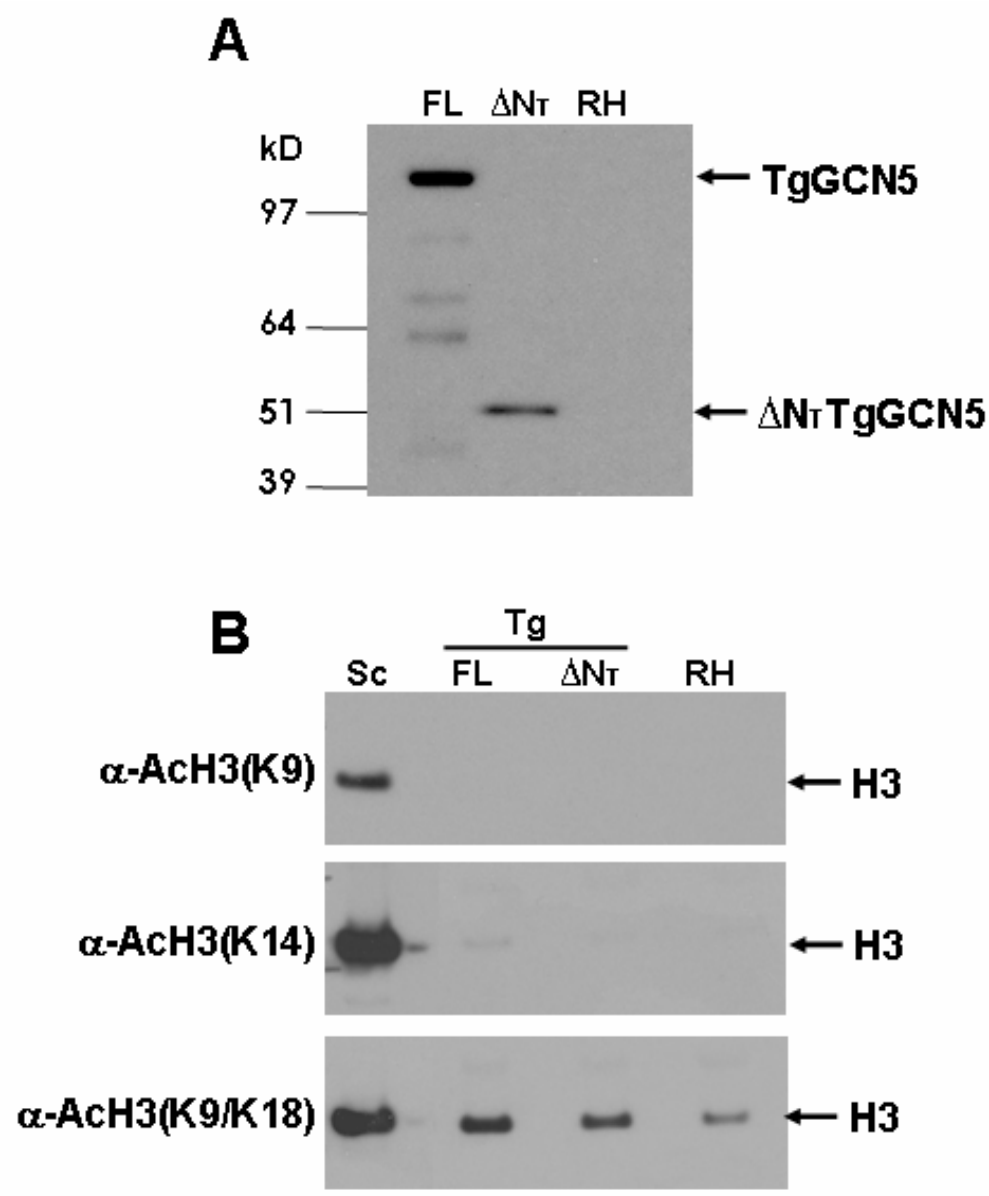

A
FL
ΔNT
RH
kD
97
64
51
39
TgGCN5
ΔNTTgGCN5
B
Tg
Sc
FL
ΔNT
RH
α-AcH3(K9)
α-AcH3(K14)
α-AcH3(K9/K18)
H3
H3
H3

Figure 29: Specific lysine acetylation of TgGCN5 and ΔN_TTgGCN5

A – Western blot of the anti-FLAG affinity purification. Primary antibody (polyclonal anti-FLAG, Sigma #F7425) used at 1:1000 and secondary antibody (goat derived anti-rabbit, Amersham #NA934) used at 1:2500. Detection was performed using Amersham ECL reagent (#RPN2209) with High Sensitive Blue photographic film (RPS Imaging #33-0810). Markers are on left in kilodaltons (kD). **B** – Western blots of *in vitro* HAT assay stained with different antibodies to specific acetylated (Ac) lysine residues. Primary antibodies (α-AcH3(K9), Upstate #07-352; α-AcH3(K14), Upstate #07-353; α-AcH3(K9/K18), Upstate #07-593) were used at 1:1000 and secondary antibody (goat derived anti-rabbit; Amersham, #NA934), was used at 1:2500. Detection was performed using Amersham ECL reagent (#RPN2209) with High Sensitive Blue photographic film (RPS Imaging #33-0810). Markers are on left in kilodaltons (kD). FL = full length; ΔN_T = lacking the N-terminus; Sc = *Saccharomyces cerevisiae* GCN5. Tg = *Toxoplasma* GCN5; RH = wild type parasite lysate as a negative control

specific acetylated lysine residues in histone H3 (Figure 29-B). Given the activity of other GCN5 homologues (Kuo et al., 1996), antibodies recognizing acetylation on lysine 14, lysine 9, and on lysine 9+18 were used. The lysine 9+18 antibody was used because there is not a reliable antibody available for Western blot analysis that recognizes acetylation only on lysine 18 in H3.

The first antibody used in the Western blot analysis recognizes acetylation of lysine 14 on histone H3 (α-AcH3(K14); Upstate). In agreement with published literature (Kuo et al., 1996), ScGCN5 strongly acetylates histone H3 on lysine 14 (Figure 20B, middle panel). Surprisingly, α-AcH3(K14) did not react with recombinant histone H3 incubated with TgGCN5 suggesting that TgGCN5 does not acetylate lysine 14. Therefore, acetylation of lysines 9 and 18 were also examined. As expected, ScGCN5 is able to acetylate lysines 9 and 18, but to a lesser extent than lysine 14. TgGCN5 does not appear to have any activity on lysine 9 in histone H3 (Figure 29-B, top panel). However, histone H3 does show acetylation by TgGCN5 when probed for acetylation on lysines 9 and 18 (Figure 29-B, bottom panel). Given the lack of TgGCN5 activity on lysine 9 alone, (Figure 29-B, top panel) this suggests that TgGCN5 specifically acetylates lysine 18. ΔN_TTgGCN5 shows the same acetylation pattern as full-length TgGCN5. Therefore it does not appear that the N-terminal extension is involved in determining the lysines that are acetylated. There is a modest amount of background with anti-FLAG affinity purified RH (wild type) parasite lysate. However, the activity of TgGCN5 and ΔN_TTgGCN5 is much higher than background levels on lysine 18 in histone H3.

To date, TgGCN5 is the first GCN5 HAT incapable of acetylating lysine 14 of histone H3 *in vitro*. The bias for H3 lysine 18 is also an intriguing peculiarity, and begs the question if other HAT(s) are preset in *Toxoplasma* that can acetylate the other lysine residues of H3

CHAPTER 4: DISCUSSION AND FUTURE STUDIES

Phylum Apicomplexa is home to numerous intracellular protozoan parasites of tremendous medical and economical importance, including *Plasmodium* (malaria) and *Cryptosporidium* (cryptosporidiosis). *Toxoplasma gondii* is another apicomplexan, capable of causing congenital birth defects and serious complications in immunocompromised patients (Black and Boothroyd, 2000).

Apicomplexan parasites typically have complex life cycles and interconvert between various stages. The differentiation of *Toxoplasma* from its invasive form (tachyzoite) into an encysted form (bradyzoite) is central to its pathology and transmission (Weiss and Kim, 2000). Interfering with this process may lead to novel therapeutic strategies, but the processes orchestrating parasite differentiation have not been completely resolved. It is well established that stage-specific genes are present in *Toxoplasma*, indicating that transcriptional regulation participates in the control of differentiation (Singh et al., 2002). Furthermore, microarray analysis of differentiating tachyzoites has shown that several genes are expressed during the entire differentiation process and some are expressed only at certain time points during differentiation suggesting a highly orchestrated and controlled event (Cleary et al., 2002). There are currently no known transcription factors that are involved in governing the bradyzoite differentiation process. Searches of apicomplexan genome databases indicate a lack of conventional transcription factors conserved in other species (Saksouk et al., 2005; Templeton et al., 2004).

Tachyzoites develop into bradyzoites *in vitro* after being exposed to various stresses (e.g. pH, chemical, or heat shock; Soête et al., 1994), suggesting that stress response pathways may be involved in mediating signals to initiate differentiation. Chromatin remodeling proteins in other eukaryotic organisms play a key role in gene expression, some of which are instrumental in mediating stress responses (Huisinga and Pugh, 2004). Thus, chromatin

remodeling proteins should be investigated in *Toxoplasma* as they may participate in the differentiation process.

The GCN5 histone acetyltransferase (HAT) is involved in regulating approximately 10% of the yeast genome, most of which are induced by stress (Huisinga and Pugh, 2004). An unusual GCN5 HAT has been cloned in *Toxoplasma* (Sullivan and Smith, 2000), but no work has been done to characterize the novel aspects of this unique HAT. TgGCN5 contains a unique N-terminal "extension" that bears no similarity to known proteins and is devoid of known protein motifs. Elucidating the function of the N-terminus will facilitate our understanding of gene regulation in this parasite and other apicomplexans, and may shed new light on the bradyzoite differentiation process as well as other aspects of parasite biology. For example, the N-terminal extension of TgGCN5 has been used a tool to explore the understudied area of nuclear trafficking in *Toxoplasma*.

I. The N-terminus and TgGCN5

The TgGCN5 N-terminal extension is comprised of 820 amino acids. The large size of the N-terminal extension in TgGCN5 distinguishes it from GCN5 homologues found in other lower eukaryotic organisms. For example, the GCN5 homologue expressed in the related alveolate *Tetrahymena* possesses a very small N-terminal extension of 69 residues. Yeast GCN5 also possesses a short N-terminal extension composed of 118 amino acids. Although metazoan GCN5 homologues contain longer N-terminal extensions of approximately 500 residues, they are still smaller than TgGCN5. Furthermore, unlike the large N-terminal extensions present in metazoan GCN5 and the related HAT PCAF, the lengthy N-terminal extension present in TgGCN5 bears no similarity to the N-terminal extensions found in other GCN5 homologues. Our hypothesis is that this unique N-terminal extension is critical to the function of TgGCN5. The goal of this thesis was to ascertain functions for the novel N-terminus providing evidence that it is critical to the function of TgGCN5. Based on what is known about other GCN5 proteins, three roles for the N-terminal domain of TgGCN5 that were investigated

in this thesis include nuclear localization, protein-protein interactions, and modulation of enzymatic function.

A. Mapping the NLS of TgGCN

Discussion: The data reported in this thesis demonstrates that the TgGCN5 N-terminal extension contains a nuclear localization signal (NLS) that is required for targeting TgGCN5 to the parasite nucleus. Previous work in the Sullivan Lab had established that the N-terminal extension was essential for the nuclear translocation of TgGCN5. Therefore, TgGCN5 made an excellent model to study how a GCN5 protein is transported into the nucleus. Before our work on elucidating the nuclear localization of TgGCN5, no NLS had been reported for any GCN5 homologue. Furthermore, no NLS had ever been mapped in a *Toxoplasma* protein. Additional studies had narrowed the region involved in nuclear localization to between amino acids 58 and 260.

Stretches of 5-6 basic residues have been discovered to function as nuclear targeting sequences in many nuclear proteins (Boulikas, 1994). The mostly likely candidate for an NLS between amino acids 58 and 260 was the hexapeptide RKRVKR. Truncations were made to the N-terminal extension isolating this stretch of basic residues within the N-terminus. Using immunocytochemistry, it was determined that the RKRVKR sequence was required for the nuclear localization of TgGCN5. To demonstrate that the hexapeptide was a complete NLS, it was fused to the cytoplasmic protein, beta galactosidase. The fusion protein of beta galactosidase and the hexapeptide was able to enter the parasite nucleus. The ability for the hexapeptide to target a non-nuclear protein to the nucleus provides strong evidence that the hexapeptide is a complete NLS recognized by the parasite's nuclear trafficking machinery.

The sequence RKRVKR is the first NLS to be described in any GCN5 protein. Searches for the TgGCN5 NLS in other GCN5 proteins revealed no matches including metazoan GCN5 and PCAF homologues which possess long N-terminal extensions. However, there are clusters of 4–5 basic residues present in the N-termini of metazoan GCN5 and PCAF homologues that are

potential candidates for an NLS. The putative maize GCN5 NLS, RKRK, is similar to the NLS isolated in TgGCN5 (Bhat et al., 2003). However, it is not interrupted by a non-basic residue. The divergence between the maize and TgGCN5 NLS suggests that there may not be a conserved NLS between GCN5 homologues. Furthermore, yeast GCN5 (ScGCN5) does not possess any stretches of basic residues within its relatively short N-terminus. The signal sequenced responsible for targeting yeast GCN5 to the nucleus is unknown, but it is doubtful that the N-terminal extension is involved. There is a stretch of basic residues, RRKIR, within the Ada2 domain of ScGCN5 that is very similar to the NLS motif of TgGCN5 (Boulikas, 1994). The hexapeptide RRKIR may be the putative NLS for ScGCN5. Interestingly, the beginning four residues, RRKI are conserved in the Ada2 domain of TgGCN5. However, these residues within TgGCN5 do not appear to be involved in targeting the protein to the parasite nucleus. Furthermore, the stretch of basic residues RRKIR is not conserved within the Ada2 domains of other GCN5 homologues. Additional studies determining if this stretch is involved in targeted yeast GCN5 are warranted as it may provide further evidence that there is not a conserved NLS among GCN5 homologues.

Small proteins less than 40kD are capable of passively diffusing through the nuclear envelope (Pante and Kann, 2002). The calculated size of yeast GCN5 is approximately 51kD. The size of yeast GCN5 exceeds the threshold for passive diffusion through the nuclear envelope, and therefore requires a form of active transport in order to be present in the nucleus.

The hexapeptide RKRVKR is the first NLS that has ever been mapped for any *Toxoplasma* protein. The RKRVKR sequence does fit one of the nine motifs described by Boulikas, but this particular NLS motif is the rarest and is not commonly utilized by nuclear proteins (Boulikas, 1994). The presence of a rare NLS motif suggests that *Toxoplasma* and perhaps other apicomplexans may utilize NLSs that differ slightly in composition from those utilized by yeast and multicellular animals. Differences in NLS composition may mean differences in nuclear trafficking pathways which could be exploited in drug design. However,

additional NLS motifs will need to be elucidated in *Toxoplasma* and other apicomplexans to determine if this is true.

Future Studies: While RKRVKR fits a recognized NLS motif pattern (Boulikas, 1999), it may not be the minimal requirement for targeting TgGCN5 to the parasite nucleus. It is plausible that the three basic residues in front of the valine, RKR, could be the entire NLS. Although most NLSs are composed of 4 or more basic residues, there are rare NLSs composed of only three basic residues (Boukilas, 1999). Fusing only RKR to β-galactosidase would confirm whether or not three basic residues are enough to target a protein to the parasite nucleus. Alternatively, site-directed mutagenesis could be performed on the NLS present in TgGCN5 changing the basic residues to alanine. The results would narrow the NLS down to its minimal size and may indicate which residues within the hexapeptide are the most important in targeting the HAT to the parasite nucleus.

It would be interesting to ascertain if TgGCN5 is localized to the nucleus when expressed in mammalian cells. The purpose of such a study would be to determine if mammalian nuclear trafficking proteins could recognize the hexapeptide NLS and transport TgGCN5 into the nucleus. Several transfection and expression methods for mammalian cells have been developed (Recillas-Targa, 2004). If TgGCN5 is transported into the nucleus, it would argue that nuclear localization via the importin pathway is a highly conserved phenomenon. However, if the TgGCN5 is not recognized by the nuclear transport machinery it would indicate that differences exist between mammalian and apicomplexan nuclear import.

B. Apicomplexan database searches with the TgGCN5 NLS

Discussion: Defining motifs such as an NLS has the potential to assist in the annotation efforts of apicomplexan genomics. The NLS of TgGCN5 is one of the first fully mapped nuclear targeting sequences in an apicomplexan. The online genomic databases for *Toxoplasma*, *Plasmodium*, and *Cryptosporidium* were searched for proteins containing identical and similar motifs. If the

RKRVKR sequence was able to identify other proteins with a high probability of possessing nuclear function based on homology or other domains present in the protein than it provided additional evidence that the hexapeptide was a nuclear targeting sequence and had the capability of identifying other nuclear proteins. The NLS could also aid in characterizing parasite proteins without any known function. It is estimated that over 50% of the predicted proteins in the *Toxoplasma* and *Plasmodium* databases are parasite specific and possess no discernable homology to any known proteins (Li et al., 2003). Proteins that have no readily identifiable function but possess the sequence RKRVKR may indicate a function within the parasite nucleus. Additionally, searching the databases may provide further insight into nuclear proteins that have already been characterized suggesting a method for how these proteins are targeted to the parasite nucleus.

Interestingly, searches of the predicted proteins present in the *Toxoplasma* genomic database revealed no exact matches to the TgGCN5 NLS sequence except for TgGCN5 itself. Additionally, no proteins possessing the hexapeptide were found in the *Plasmodium* or *Cryptosporidium* databases. The lack of additional proteins possessing the exact same NLS suggests that the hexapeptide is a rare NLS not commonly used by apicomplexans. However, searching for an exact match to the RKRVKR hexapeptide is not the most appropriate search. Within an NLS motif, arginine can be substituted for lysine and vice versa and the non-basic residue can be substituted for any amino acid and nuclear targeting can still be achieved (Boulikas, 1994; Hodel et al., 2001). Therefore, the searches of the three apicomplexan databases were expanded to include these permutations.

The first variation was to search for the sequence RKRxKR where “x” can be any amino acid. A search of the *Toxoplasma* database of predicted proteins revealed a total of 19 potential hits. Eight of the proteins were parasite-specific and had no discernable homology to any known proteins nor contained any recognizable protein motifs. Of the remaining 11 proteins, nine contained additional domains or homology indicating a nuclear function and two had additional elements indicating a function within the cytoplasm. Therefore, over

80% of the proteins containing additional motifs or homology appear to have a nuclear function indicating the motif RKRxKR has predictive function in identifying nuclear proteins in *Toxoplasma*.

The nine proteins containing additional domains or homology implying a nuclear function included a RNA recognition protein, tRNA synthetase, a DNA polymerase, and several potential transcription factors. The tRNA synthase (38.m01112, Table II, page 100) belongs to a group of proteins that function in both the cytoplasm and nucleus (Schimmel and Wang, 1999). Several studies have implicated an NLS enriched with basic residues in targeting tRNA synthetases to the nucleus (Mucha, 2002). The NLS-like motif present in 38.m01112, RKRGKR, does not match the proposed NLS motifs in other tRNA synthetases homologues (Mucha, 2002). Nonetheless, the ability of the TgGCN5 NLS to identify other nuclear proteins including a tRNA synthetase that appears to enter the nucleus via an NLS composed of basic residues in other organisms provides additional evidence that RKRxKR is a nuclear targeting sequence and has the potential to assist in further characterizing known proteins.

The ability of RKRxKR motif to identify other nuclear proteins indicates that it possesses value in predicting the localization of proteins. Therefore, the NLS can assist in deciphering the functions of proteins with unknown function. The eight *Toxoplasma* proteins with unknown function containing RKRxKR motifs should be targeted to the parasite nucleus. However, two of the 19 proteins from the database search contained ribosomal motifs in addition to an NLS-like motif. The presence of a ribosomal domain indicates a function outside of the nucleus which illustrates the limitation of database searches. The presence of an NLS like-motif does not guarantee that the protein is nuclear and must be taken into context with other domains expressed in the protein. Ultimately, localization studies will need to be performed to confirm the distribution of these 19 proteins within the parasite.

Similar searches of the *Plasmodium* and *Cryptosporidium* databases were also performed looking for permutations of the TgGCN5 NLS. Searching the *P. falciparum* annotated protein data set with the RKRxKR motif returned over 10

proteins. One protein contained multiple zinc finger-like motifs suggesting it may be involved in transcriptional regulation and function within the nucleus (Aasland et al., 1995). A second search of the *P. falciparum* database was performed with less stringent parameters by allowing the five basic residues in the TgGCN5 NLS to be either lysine or arginine and the non-basic valine to be any amino acid. The search produced 1500 hits from the annotated protein data set. Preliminary analysis of some of these potential nuclear proteins included a putative histone deacetylase, a tRNA synthase, and a putative DNA polymerase subunit. The same criteria was also used to search the *Cryptosporidium parvum* expressed sequence tag database, which identified nine proteins with an NLS-like motif including the histone H2B and a putative nucleolar protein.

The ability of permutations in the TgGCN5 NLS hexapeptide to identify nuclear proteins in other apicomplexans suggests that there may be shared NLS sequences between *Toxoplasma*, *Plasmodium* and *Cryptosporidium* that target nuclear proteins to their respective nuclei. However, the ability of these NLS like-motifs to target a protein to the nucleus in *Plasmodium* and *Cryptosporidium* remains to be confirmed *in vivo*.

Future Studies: Since the completion of this thesis, additional apicomplexan online genomic databases have been established. *Theileria parva*, a tick-borne apicomplexan, is responsible for causing East Coast fever, a highly fatal lymphoproliferative disease of cattle in Africa (Graham et al., 2006). East Coast fever kills over 1 million cattle each year in sub-Saharan Africa, resulting in economic losses exceeding $200 million annually (Gardner et al., 2005). Results from sequencing efforts of the *Theileria parvum* genome are now available online (http://www.tigr.org/tdb/e2k1/tpa1/). *Eimeria tenella* is another apicomplexan that causes avian coccidiosis (Allen and Fetterer, 2002). *E. tenella* is a constant threat to the poultry industry with an annual worldwide cost estimated at about $800 million (Allen and Fetterer, 2002). An online genomic database is also available for *Eimeria tenella* (www.genedb.org/genedb/etenella/). As new apicomplexan genomes are

sequence and made available online, the TgGCN5 NLS will be an important tool to begin searching for nuclear proteins. Searches with the RKRVKR sequence and its appropriate permutations should be performed on these new databases. Proteins that possessing sequences similar to RKRVKR NLS should be investigated further as potential nuclear proteins.

While it has been demonstrated that RKRVKR NLS has the potential to assist in the identification of putative nuclear protein *in silico*, these results need to be confirmed *in vivo*. Expression of recombinant proteins and localization studies with putative nuclear proteins will need to be performed to determine the true potential of the TgGCN5 NLS to identify other nuclear proteins. Expression systems and techniques are well established in *Toxoplasma* for this purpose (Soldati and Boothroyd, 1993). Other members of the phylum Apicomplexa are not as readily amenable to genetic manipulation in the laboratory (Kim and Weiss, 2004). Therefore, *Toxoplasma* could be utilized as a model system to determine the cellular localization of putative nuclear proteins from other apicomplexans. Following heterologous expression in *Toxoplasma*, immunofluorescence assays could be employed to ascertain if the protein is targeted to the nucleus. However, it is possible that the expression of certain heterologous apicomplexan proteins may be lethal in *Toxoplasma* preventing localization experiments.

C. Importin pathway in Toxoplasma

Discussion: Proteins that possess an NLS composed of basic residues are transported through the nuclear envelop by interacting with two proteins: importin α and importin β (Gorlich et al., 1995). The role of importin α is to recognize and bind to NLSs (Gorlich et al., 1994). Importin α is a highly conserved protein present in virtually all eukaryotic organisms (Chook and Blobel, 2001). In fact, the *Giardia* genomic database (www.mbl.edu/Giardia) contains an annotated entry of an importin α homologue indicating that the use of importin α to transport proteins into the nucleus occurred early in eukaryotic evolution.

Despite the extensive ongoing research in apicomplexans, nuclear trafficking is a greatly understudied area of parasite biology. Although homologues of importin α and importin β have been identified in *Plasmodium falciparum* (Mohmmed et al., 2003), no functional studies have been performed on either of these nuclear trafficking proteins. Furthermore, there are no reports on the identification of any nuclear transport proteins in *Toxoplasma*. Thus, the newly mapped NLS in the N-terminal extension of TgGCN5 provided an excellent tool to studying the nuclear trafficking pathways in *Toxoplasma*.

As the NLS of TgGCN5 is composed of highly basic residues, it was logical to theorize that *Toxoplasma* possessed a homologue of importin α. The *Toxoplasma* protein database was screened using the BLAST algorithm with the *P. falciparum* importin α sequence. One search result revealed a high degree of similarity to importin α. Using the genomic database information, the gene was subsequently cloned. The *Toxoplasma* homologue of importin α (TgIMPα) is composed of 545 amino acids and possesses two conserved domains present in all known importin α proteins. The importin β binding (IBB) domain is near the N-terminus and is the region bound by importin β (Goldfarb, et al., 2004). A long series of armadillo (ARM) repeats in the C-terminal half of the protein recognizes and binds the basic NLS of nuclear proteins (Teh et al., 1999).

In contrast to multicellular organisms, it appears that both *Toxoplasma* and *P. falciparum* possess only one importin α (Stewart and Rhodes, 1999).

Future Studies: It is important to continue elucidating how GCN5 and other proteins enter the parasite nucleus. Discovering how to block chromatin remodeling proteins and transcription factors from entering the parasite nucleus may subvert parasite differentiation and other processes essential for its survival. Novel elements found in the nuclear import pathway may be exploited in the design of more selective therapeutic agents. Additionally, the study of these pathways in protozoa provides a unique perspective on how these systems evolved in early eukaryotic cells.

The presence of an importin α homologue in *Toxoplasma* suggests that other components of the importin trafficking pathway should also be expressed by the parasite. Bioinformatic searches using keywords or sequence data from other species reveals that *Toxoplasma* possesses homologues of other importin pathway components including: importin β, Ran, RCC1, CAS, and RanGAP1. Similar components of the importin pathway have been cloned previously in *P. falciparum*, including importin α, importin β (Mohammed et al., 2003), Ran (Dontfraid and Chakrabarti, 1994), and RCC1 (Ji et al., 1998). These observations are consistent with the notion that the importin-mediated nuclear transport pathway is ancient in origin (Chook and Blobel, 2001).

A screen of parasites that were treated with a non-specific mutagen revealed several clones defective in bradyzoite cyst formation in mice. One of the mutations disrupted the genomic loci of the *Toxoplasma* homologue of RCC1 (Frankel et al., 2006). RCC1 is a GTPase activating protein (GAP) involved in the importin trafficking pathway that activates the GTPase activity of Ran to hydrolyze GTP to GDP (Bischoff and Ponsting, 1991; see Figure 7, page 47). It has been hypothesized that the clone is defective in nuclear trafficking rendering the parasite unable form cysts making it susceptible to the host immune system (Frankel et al., 2006). Additional studies have shown that the cyst formation is restored when recombinant *Toxoplasma* RCC1 is expressed in the defective clone (Frankel et al., 2006). The loss of RCC1 resulting in the defective to bradyzoite formation in mice demonstrates the importance of continuing to characterize nuclear trafficking pathways in apicomplexans.

D. Interaction of TgGCN5 with TgIMPα

Discussion: Once it had been established that *Toxoplasma* possessed an importin α homologue, the next logical step was to determine if TgGCN5 interacted with TgIMPα and if that interaction occurred through the hexapeptide RKRVKR. Both proteins were generated using an *in vitro* translation system. The expression vectors used by the translation system incorporated an N-terminal cMyc tag on TgGCN5 and an N-terminal hemagluttin (HA) tag on

TgIMPα. Affinity resins directed to each tag were used in co-immunoprecipitation experiments to determine if the two proteins interacted. Unfortunately, the two proteins did not co-immunoprecipitate indicating that TgGCN5 and TgIMPα did not interact.

However, it was theorized that an interaction between the two proteins could mask the N-terminal fusion tags on both proteins and prevent the two interacting proteins from being co-immunoprecipitated. Therefore, additional co-immunoprecipitations were performed using an antiserum directed towards the C-terminal end of TgGCN5 that had been previously developed in the Sullivan Lab (Chapter 2, Section IV-B, page 71). The antiserum targeted a region of TgGCN5 that is away from the region involved in a potential interaction with TgIMPα alleviating the problem of masked fusion tags. Co-immunoprecipitation experiments using the antiserum were able to pull down both proteins indicating they interacted with each other. Control experiments were performed indicating that the antiserum is incapable of interacting with TgIMPα. Therefore, the only way TgIMPα could have been isolated during the co-immunoprecipitation was through an interaction with TgGCN5. The fact that the co-immunoprecipitation failed with the HA and cMyc affinity resins but was successful with the TgGCN5 antiserum supports the theory that the N-terminal tags were masked during an interaction between the two proteins. Therefore, the concealment of the N-terminal fusion tags provides a second line of evidence supporting an interaction between TgGCN5 and TgIMPα.

After it had been established that TgGCN5 and TgIMPα interacted with each other, additional experiments were employed to determine if the NLS hexapeptide was involved in mediating the interaction. A co-immunoprecipitation was performed using a truncated form of TgGCN5 lacking the first 99 amino acids (Δ99TgGCN5), which is just down stream from the NLS. In co-immunoprecipitation experiments, Δ99TgGCN5 was unable to pull down TgIMPα. To further isolate the NLS as the element responsible for interacting with TgGCN5, a deletion mutant was generated (ΔNLSTgGCN5) removing the NLS and replacing it with an alanine-serine dipeptide. ΔNLSTgGCN5 was also

incapable of pulling down TgIMPα indicating that the RKRVKR sequence is responsible for the mediating interaction with TgIMPα.

It is important to note that the apicomplexan versions of importin α are more plant-like, which may have relevant functional and pharmacological implications. In order for most proteins to be translocated into the nucleus, importin α must first be bound by importin β through an interaction with the IBB domain (Kobe, 1999). Once bound by importin β, importin α can interact with the NLS of a nuclear protein. However, importin α homologues in plants are able to translocate their cargo into the nucleus without being bound by importin β (Hubner et al., 1999). TgIMPα may also possess the ability to translocate proteins to the nucleus without interacting with importin β. Like all importin α homologues known to date, TgIMPα possess an IBB domain in its N-terminus. The IBB domain contains an autoinhibitory sequence which prevents importin α from binding an NLS (Kobe, 1999). When the IBB domain is bound by importin β, the autoinhibiton is abolished and importin α is able to bind to an NLS. The co-immunoprecipitation performed between TgIMPα and TgGCN5 used the entire native TgIMPα sequence which included the IBB domain. Despite the IBB domain being present, TgIMPα was able to bind the TgGCN5 *in vitro*. One explanation is that TgIMPα does not require importin β to bind an NLS as seen in plants. Another explanation is the presence of an importin β protein in the translation system used to generate the TgIMPα and TgGCN5 proteins. The *in vitro* translation system uses a cell fraction from rabbit reticulocytes, so it is possible that the cell fraction contains the rabbit homologue of importin β. If present, the rabbit importin β may be facilitating the interaction between TgIMPα and TgGCN5. Co-immunoprecipitation experiments performed with purified recombinant protein would provide a more accurate assessment of TgIMPα's ability to bind TgGCN5 in the absence of importin β.

In summary, the nuclear trafficking protein TgIMPα, interacts with TgGCN5 *in vitro* by binding the hexapeptide RKRVKR. The sequence RKRVKR has also been shown to be required to target TgGCN5 to the parasite nucleus. Therefore, it follows that TgIMPα is involved in trafficking TgGCN5 into the

parasite nucleus via an interaction with the hexapeptide RKRVKR. In addition, the interaction of TgIMPα with TgGCN5 illustrates that the N-terminal extension participates in protein-protein interactions. However, the interaction between TgIMPα and TgGCN5 has only been demonstrated through *in vitro* studies and additional *in vivo* studies will need to be performed to confirm if TgIMPα is truly involved in trafficking TgGCN5 into the parasite nucleus.

It is important to mention that these are the first studies elucidating how a nuclear protein may be transported into the nucleus in apicomplexan parasites. Furthermore, it is also the first study characterizing how a GCN5 protein is transported into the nucleus.

Future Studies: An interesting finding in demonstrating the interaction between TgIMPα and TgGCN5 was that TgIMPα did not require a prior interaction with an importin β protein to alleviate the autoinhibition of the IBB domain. It has been proposed that the cell lysate used in the *in vitro* translation system may contain a homologue of importin β that was facilitating the interaction between TgIMPα and TgGCN5. In order to remove potential artifact from additional unknown proteins present during immunoprecipitation, pure recombinant proteins could be employed in future co-immunoprecipitation experiments. TgIMPα could be expressed in *Toxoplasma* and purified by virtue of an inframe fusion tag. TgGCN5 would be purified from the transgenic parasite line used to obtain protein for the *in vitro* HAT assays (Chapter 2, Section IV-E, page 80). Purified recombinant TgGCN5 and TgIMPα could be mixed together and co-immunoprecipitated without the interference of other proteins. If TgGCN5 and TgIMPα interacted in the absence of any other proteins, it would provide fairly conclusive evidence that TgIMPα does not require importin β to bind an NLS indicating a sharp difference between *Toxoplasma* and other animal importin α homologues. However, if the proteins do not interact, it would indicate that some element present in the *in vitro* translation system was facilitating the interaction between TgGCN5 and TgIMPα.

The putative importin β homologue discovered in our bioinformatic searches of the *Toxoplasma* database could be cloned and overexpressed in the parasites. Following purification from parasite lysate, the *Toxoplasma* importin β homologue (TgIMPβ) would be added to co-immunoprecipitation experiments determining if it interacted with importin α through its IBB domain. Additionally, performing co-immunoprecipitations comparing the ability of TgIMPα to pulll down TgGCN5 in the presence and absence of TgIMPβ would be the definitive experiment to determine if TgIMPβ is required for TgIMPα to interact with TgGCN5. If co-immunoprecipitations showed an increase in the amount of protein pulled down in the presence TgIMPβ, it would argue that TgIMPβ increases the strength of the interaction between TgIMPα and TgGCN5. Furthermore, it would indicate that TgIMPβ is probably required for TgIMPα to interact with an NLS *in vivo*.

Once a stable transgenic parasite line overexpressing TgIMPα has been generated, the thought of performing co-immunoprecipitations attempting to pull down TgGCN5 is tempting. However, co-immunoprecipitations of TgIMPα using parasite lysate would pull down a multitude of nuclear proteins hindering detection of TgGCN5. Furthermore, the interaction between TgGCN5 and TgIMPα is transient resulting in low yields of co-immunoprecipitated TgGCN5 making detection even more difficult. The converse is also true. Attempting to perform co-immunoprecipitations using lysate from parasites overexpressing TgGCN5 would be plagued with similar limitations.

Another important aspect in studying the nuclear localization of TgGCN5 is determining if its interaction with TgIMPα is sufficient to transport the protein into the nucleus. Treating living cells with digitonin permeabilizes the cholesterol rich plasma membrane, but leaves the cholesterol poor nuclear membrane intact (Liu et al., 1999). Digitonin permeabilization depletes the cytoplasm of factors and proteins required for efficient nuclear import (Adam and Gerace, 1991). Therefore, the import of nuclear substrates in digitonin-permeabilized cells requires the addition of purified transport factors (Adam and Gerace, 1991). The advantage of this system is it allows for the selective addition of transport

proteins to determine which are involved in translocating TgGCN5 into the nucleus. The only limitation is the ability to generate pure transport factors to add back to the permeabilized cells. If parasites overexpressing TgIMPα and TgIMPβ are successfully generated, they would provide an excellent source of recombinant transport factors. Once cells are permeabilized, the proteins of interested are added to a HEPES-buffered acetate solution containing an ATP-regenerating system (1mM ATP, 5mM creatine phosphate, 20U/ml creatine phosphokinase) to provide energy necessary to facilitate active transport (Liu et al., 1999). Immunofluorescent assays would be employed to determine if nuclear transport occurred. A method has been reported for permeabilizing cells in suspension (Liu et al., 1999) that may work on freshly harvested extracellular tachyzoites. If the permeabilization of parasites fails, the experiment may work in a mammalian cell line. However, obtaining results using *Toxoplasma* parasites would be more meaningful as the nuclear pore complex may be divergent between parasite and mammal. The basic experimental outline would be to assess the translocation of TgIMPα and TgGCN5 with and without the hexapeptide NLS and to determine if the addition of TgIMPβ is necessary to localize TgGCN5 to the nucleus. If TgIMPα can transport TgGCN5 into the parasite nucleus without TgIMPβ it provides additional evidence that nuclear trafficking pathways in *Toxoplasma* are more similar to plants than animals. Establishing a digitonin permeabilization method in *Toxoplasma* will also assist in characterizing the trafficking of other parasite nuclear proteins.

E. Identifying members of TgGCN5 HAT complexes

Discussion: In addition to TgIMPα, we wanted to identify other proteins that could be interacting with TgGCN5. GCN5 in other organisms does not function singly to remodel chromatin, but rather as a member of large multi-subunit complexes (Roth et al., 2001). The most likely candidate proteins interacting with TgGCN5 are those forming HAT complexes analogous to the SAGA and ADA complexes found in yeast.

Both the SAGA and ADA complexes possess GCN5 as their catalytic core (Grant et al., 1997). The ADA complex is much smaller than SAGA, and is composed of three different proteins: Ada2, Ada3, and Ahc1 (Eberharter et al., 1999). SAGA is a very large complex, approximately 1.8MDa in size and has over 15 different subunits including four Ada proteins, several Spt and TAF proteins and Tra1 (Sterner and Berger, 2000). In addition to proteins known to exist in the SAGA and ADA complexes, TgGCN5 may interact with parasite specific proteins not present in other GCN5 HAT complexes. Novel proteins present in TgGCN5 complexes may interact with the unique N-terminus of TgGCN5.

To identify proteins interacting with the N-terminus, an attempt was made to overexpress the N-terminal extension of TgGCN5 in *Toxoplasma* fused to a FLAG tag. Co-immunoprecipitations performed on parasite nuclear extract from the transgenic parasites using FLAG affinity resin would identify interacting proteins. However, attempts at trying to generate a transgenic parasite line expressing just the N-terminal extension of TgGCN5 were unsuccessful. Overexpression of only the N-terminus of TgGCN5 appears to be deleterious to the parasite. The deleterious nature of expressing the N-terminus of TgGCN5 may be due to sequestering interacting proteins away from fully functional HAT complexes causing a depletion of complex members resulting in a dysregulation of histone acetylation and parasite death. Despite the failure of trying to overexpress the N-terminal extension of TgGCN5, it does provide preliminary evidence that the N-terminal extension is involved in mediating protein interactions.

A second option to identifying proteins interacting with TgGCN5 was to employ co-immunoprecipitation experiments using anti-FLAG affinity resin on parasites overexpressing recombinant full length TgGCN5 fused to FLAG tag ($_{FLAG}$TgGCN5) and TgGCN5 lacking the N-terminal extension also fused to a FLAG tag ($_{FLAG}\Delta N_T$TgGCN5). Proteins isolated from $_{FLAG}$TgGCN5 that are not present in pull down assays from $_{FLAG}\Delta N_T$TgGCN5 would be judged as interacting with the N-terminal extension. Dr. Mohamed-Ali Hakimi's laboratory

at the French National Centre for Scientific Research has had success in purifying protein complexes from *Toxoplasma*. Therefore, a collaboration was established with Dr. Hakimi's lab to biochemically purify proteins that may interact with full length $_{FLAG}$TgGCN5 and $_{FLAG}\Delta N_T$TgGCN5.

The co-immunoprecipitation from parasites overexpressing full length $_{FLAG}$TgGCN5 revealed several interesting results. The most interesting was the identification of two different Ada2 homologues in *Toxoplasma*. Ada2 is a well characterized protein known to associate with GCN5 in acetylation complexes (Eberharter et al., 1999). Ada2 proteins associate with GCN5 acting as coactivators bridging the HAT complex to the acidic activation domains of DNA-binding transcription factors and TATA-binding protein (Barlev et al., 1995). It was unexpected to find two novel Ada2 homologues in an early branching eukaryotic cell like *Toxoplasma*. While most metazoans and land plants have two Ada2s, our search results indicate that most protozoa, like yeast, have only one. *T. gondii* appears to be an exception to this observation. The related apicomplexan *Plasmodium falciparum* expresses a single Ada2 protein, termed PfAda2 (Fan et al., 2004). It is unknown why *Toxoplasma* would possess two Ada2 proteins when all other similar eukaryotes evidently only express one. It is possible that one of the TgAda2 homologues operates independently of TgGCN5. *Drosophila* has two Ada2 homologues, but only one (dAda2b) is incorporated into the SAGA complex (Kusch et al., 2003). Data argues that dAda2a may have novel roles in transcription that are independent of GCN5 (Kusch et al., 2003; Muratoglu et al., 2003). Two different Ada2 proteins generates the possibility that at least two different HAT complexes containing TgGCN5 exist in *Toxoplasma*. The different complexes may have independent roles in regulating chromatin structure. The apicomplexan Ada2 proteins PfAda2, TgAda2-A, TgAda2-B are much longer than the Ada2 homologues present in other species, containing large regions flanking the Ada2 domain that have no similarity to other protein sequences or known motifs (Bhatti et al., 2006). The function of these unique domains remains to be ascertained. The

presence of unique domains suggests that apicomplexan Ada2 proteins may possess parasite specific functions not seen in other Ada2 proteins.

In addition to the two Ada2 proteins, the full length $_{FLAG}$TgGCN5 also pulled down four histone proteins: H3.3, H2A.1, H2AF/Z, and H4. The histone proteins H3.3, H2A.1, and H2A F/Z, are not the canonical core histones, but variant histones. Histone variants are expressed in eukaryotic cells, and can replace the four canonical histones that comprise the nucleosome (Henikoff et al., 2004). The existence of histone variants for histones H3 and H2A has been previously reported in *Toxoplasma* and *Plasmodium* (Sullivan, 2003; Sullivan et al., 2006; Miao et al., 2006). Variant histones are associated with restructuring transcriptionally active chromatin (Kamakaka and Biggins, 2005). Finding variant histones interacting with TgGCN5 agrees with the TgGCN5's predicted role in remodeling chromatin. The mechanisms involved in incorporating histone variants into the genome are poorly understood (Pusarla and Bhargava, 2005). It is theoretically possible the HATs participate in depositing variant histones into areas of the genome where active transcription is taking place (Kusch et al., 2004).

As the N-terminus is parasite-specific, proteins interacting with this region may also be unique to the parasite. Parasite-specific proteins present in acetylation complexes may provide novel drug targets to combat apicomplexan disease. Two candidates, TgTwinScan_2624 and TgTigrScan_8186, were present in the co-immunoprecipitation of full length $_{FLAG}$TgGCN5. However, further experiments are required to determine if these two proteins interact with N-terminus of TgGCN5.

The co-immunoprecipitations of $_{FLAG}\Delta N_T$TgGCN5 were not as fruitful as full length TgGCN5. In hindsight, the experimental design may have been limited because $_{FLAG}\Delta N_T$TgGCN5 is cytoplasmic (Chapter 3, Figure 11, page 91). As $_{FLAG}\Delta N_T$TgGCN5 is not localized to the nucleus, a proper comparative analysis between the two co-immunoprecipitations cannot be adequately performed. It is possible that HAT complexes may be completely or partially assembled in the cytoplasm before being transported into the nucleus. Once assembled, an

interaction between one of the complex members (e.g. TgGCN5) and nuclear trafficking proteins would result in transport through the nuclear pore into the nucleus. The nuclear pore complex has a rather large diameter pore capable of translocating macromolecules and protein complexes exceeding 1.0MDa (Fahrenkrog and Aebi, 2003). A “piggy back mechanism” has the advantage of enabling the cell to provide equimolar concentrations of each subunit to the nucleus (Steidl et al., 2004). Such a mechanism has been observed for the translocation of a heterotrimeric CCAAT-binding complex to the nucleus of *Aspergillus nidulans*. A single subunit (HapB) within in the complex is responsible for nuclear localization of the entire complex (Steidl et al., 2004). It is plausible that HAT complexes in *Toxoplasma* are assembled in the cytoplasm, with TgGCN5 serving as the critical member responsible for directing the complex into the nucleus. Therefore, co-immunoprecipitations performed on parasites overexpressing $_{FLAG}\Delta N_T$TgGCN5 may still provide a way to identify complex members. However, because $_{FLAG}\Delta N_T$TgGCN5 is not present in the nucleus, whole cell extract rather than nuclear extract was used in the co-immunoprecipitations experiments. The disadvantage of using whole cell extract versus nuclear extract is the large amount of background proteins present in the lysate. The large amount of background proteins resulted in difficulty resolving the elutions following co-immunoprecipitations with $_{FLAG}\Delta N_T$TgGCN5 parasite lystate. Furthermore, the high amount of background proteins may have masked potential complex members from being identified. Thus, the interpretation of the results from the $_{FLAG}\Delta N_T$TgGCN5 co-immunoprecipitation is limited by artifact caused by signification background proteins and lack of proper localization.

Despite the limitations mentioned, co-immunoprecipitation with $_{FLAG}\Delta N_T$TgGCN5 indicated an interaction with the heat shock proteins Hsp60 and Hsp 70. Full length $_{FLAG}$TgGCN5 also revealed potential interactions with Hsp60 and Hsp 70 as well as Hsp90. Bradyzoite differentiation in *Toxoplasma* can be generated *in vitro* under thermal and pH stress conditions (Soête et al., 1994). The three heat shock proteins, Hsp60, Hsp 70, and Hsp90, have been shown to be involved in the transition between the two different parasite stages

(Weiss et al., 1998; Weiss and Kim, 2000). In other organisms, GCN5 is involved in stress remediation. GCN5 has been linked to activating stress response pathways during amino acid and phosphate starvation in yeast (Georgakopoulos and Thireos, 1992; Gregory et al., 1998). In *Arabidopsis*, GCN5 appears to be involved in the activation of stress response proteins during exposure to low ambient temperatures (Stockinger et al., 2001). Evidence of an interaction between heat shock proteins and TgGCN5 indicates that TgGCN5 may also be involved in response to stressful stimuli in *Toxoplasma*. Therefore, TgGCN5 may be involved in the bradyzoite differentiation process. The mechanisms involved in recruiting GCN5-containing complexes to remodel areas of the chromatin during stress are poorly understood. Heat shock proteins which are expressed in higher levels during stress events (Schlesinger, 1990) may participate in directing the chromatin remodeling activity of TgGCN5-containing complexes leading to the upregulation and expression of genes involved in bradyzoite conversion. The same process may occur with GCN5-containing complexes in other organisms during stressful stimuli.

Another reason for the interaction between heat shock proteins and TgGCN5 may be their involvement in the assembly of HAT complexes. Heat shock proteins function as chaperones in the construction of large protein complexes (Kimmins and MacRae, 2000). Hsp70 and hsp90 have been shown to aid in the assembly of steroid receptors (Kimmins and MacRae, 2000) and Hsp 90 is involved in the initial formation of the kinetochore complex (Bansal et al., 2004). Heat shock proteins may play a similar role in the assembly of HAT complexes in *Toxoplasma*. It is theoretically possible that TgGCN5 complexes change composition under bradyzoite conditions. Heat shock proteins may facilitate the change in composition TgGCN5 preceding bradyzoite conversion.

Future Studies: Bioinformatics could aid significantly in the pursuit of identifying potential proteins interacting with TgGCN5. As outlined in the introduction, many of the proteins present in GCN5-containing complexes have been identified in other species. Table III is the result of a BLAST search of the

Toxoplasma database predicted proteins based on the Draft3 algorithm with known members of yeast SAGA and ADA complexes.

Interestingly, the database contains no homologues for the members of the yeast ADA complex (Ahc1, Ada1, Ada3, or Ada5) besides Ada2. The BLAST search did confirm the presence of two Ada2 proteins in *Toxoplasma*. The lack of additional ADA complex members suggests that either *Toxoplasma* does not possess an ADA complex or an analogous complex contains divergent members not present in yeast.

None of the characterized Spt or SGF proteins found in the yeast SAGA complex appear to present in *Toxoplasma*. *Toxoplasma* is an ancient eukaryote and its GCN5-containing complexes may be more streamlined lacking many of the proteins in yeast or metazoans. Alternatively, additional HAT components lack sufficient homology to be detected by the BLAST search. The only proteins for which putative homologues could be identified were Taf90, Taf60, and Tra1. These proteins were not isolated in the full length TgGCN5 co-immunopreciptiation. The Taf90, Taf60, and Tra1 proteins may have been in too low in abundance to be isolated in the co-immunoprecipitation. Additionally, the size of the putative Tra 1 in *Toxoplasma*, 59.m03383, is predicted to be 909 kilodaltons, which exeeds the resolution capacity of the gels used to resolve the co-immunoprecipitated proteins.

In fact, the sizes of the putative SAGA subunit homologues in *Toxoplasma* are larger than their yeast counterparts. As shown in Table IV, the *Toxoplasma* homologues for Tra1, Ada2-B, and GCN5 are over twice the size of their corresponding homologues. The large size of the complex members in *Toxoplasma* suggests they have functions beyond their yeast counterparts. The additional function may compensate for the limited number of complex members. As these HAT complexes evolved, the larger *Toxoplasma* subunit homologues such as Tra1 and Ada2-B may have fragmented into the smaller subunits present in yeast, each with their own specialized function.

If additional complex members exist, they are either novel members not seen in other GCN5 containing complexes, or they are too divergent to be

Table III: Results of BLAST search of *Toxoplasma* database for ADA and SAGA complex members

Protein Name	Complex Found	Blast With	Potential Hits	Identity Score
Ahc1	ADA	YOR023C	NONE	
Ada1	SAGA	YPL254W	NONE	
Ada2	Both	AAA34393.1	35.m00936 - A	<1e-100
			55.m04988 - B	4.1e-67
Ada3	Both	YDR176W	NONE	
Ada5/Spt20	SAGA	AAB07900	NONE	
Spt 3	SAGA	YDR392W	NONE	
Spt 7	SAGA	YBR081C	NONE	
Spt8	SAGA	YLR055C	NONE	
Taf5/Taf90	SAGA	YBR198C	641.m01468	1.7e-13
Taf6/Taf 60	SAGA	YGL112C	86.m00838	1.8e-6
Taf9/17	SAGA	YMR236W	NONE	
Taf10/Taf25	SAGA	YDR167W	NONE	
Taf12/Taf68	SAGA	YDR145W	NONE	
Tra1	SAGA	AF076974	59.m03383	3.6e-66
Sin4	SAGA	AAA35044.1	NONE	
SGF11	SAGA	YPL047W	NONE	
SGF29	SAGA	YCL010C	NONE	
SGF73	SAGA	YGL066W	NONE	

list of complex members adapted from Roth et al., 2001

Table IV: Size comparison of *Toxoplasma* and *Saccharomyces cerevisiae* SAGA complex members

Protein Name	*S. cerevisiae* Identifier	Estimated Size (kD)	*Toxoplasma* Identifier	Estimated Size (kD)	Percent Difference
GCN5	YGR252W	51.1	AAF29981	126.8	248.1%
Ada2	AAA34393.1	101.5	35.m00936 (A)	109.9	108.3%
			55.m04988 (B)	248.9	245.2%
Taf5/Taf90	YBR198C	57.9	641.m01468	89.3	154.2%
Taf6/Taf 60	YGL112C	88.9	86.m00838	117.9	132.6%
Tra1	AF076974	433.2	59.m03383	909.0	209.8%

kD = kilodaltons

detected using bioinformatic techniques. Therefore, a large screening technique such as additional co-immunoprecipitations, a yeast two hybrid screen, or a GST pull down assay could be employed to identify additional complex members. Assays pertinent to this thesis would focus on identifying complex members that interact with the unique N-terminus of TgGCN5.

As mentioned in the discussion, the co-immunoprecipitations from parasite overexpressing $_{FLAG}\Delta N_T$TgGCN5 had limited utility because $_{FLAG}\Delta N_T$TgGCN5 does not localize to the parasite nucleus. $_{FLAG}\Delta N_T$TgGCN5 is cytoplasmic, hence it prevented the use of nuclear extract resulting in a higher amount of background proteins and subsequent poor resolution of elutions using gel electrophoresis. To better ascertain the involvement of the N-terminal extension in protein-protein interactions, a different protein could be expressed in the parasites. For example, fusing the TgGCN5 NLS sequence, RKRVKR, to $_{FLAG}\Delta N_T$TgGCN5 should localize it to the parasite nucleus. A protein that is targeted to the nucleus would allow for the use of nuclear extract in co-immunoprecipitations, lowering the amount of background interference from cytoplasmic proteins. Co-immunoprecipitations with $_{FLAG}\Delta N_T$TgGCN5 fused to the TgGCN5 NLS would provide a better comparison to $_{FLAG}$TgGCN5 in identifying proteins associating with the N-terminal extension of TgGCN5. In addition, elutions from additional co-immunoprecipitation experiments could be resolved using 2D gel electrophoresis to provide better resolution and assist in identifying more potential interacting proteins.

Other biochemical approaches that could be used include a GST pull down assay. The N-terminal extension of TgGCN5 would be expressed in bacteria or yeast fused to glutathione S-transferase (GST) tag. The recombinant N-terminus is purified, concentrated, and immobilized on a glutathione resin. The glutathione resin containing the N-terminus would be incubated with nuclear extract from wild-type parasites. The immobilized N-terminus would be bound by various proteins present in the extract. The N-terminus protein complexes would be eluted off the GST affinity resin and resolved using 2D gel electrophoresis. Spots present on the 2D gel could be excised and sent for sequencing analysis.

A yeast two-hybrid screen is currently being performed using the N-terminus as bait to screen a tachyzoite cDNA library for potentially interacting proteins. The advantage of yeast-two-hybrid is a large number of proteins can be rapidly screened searching for potential complex members. However, yeast two hybrid screens are often limited by the potential for false positives. Furthermore, some parasite proteins may be toxic to yeast and expressed in either low at levels or not at all. Depending on how the library is generated, proteins expressed in low abundance might not be adequately represented in such a screen. Nevertheless, a yeast two hybrid screen has the potential to provide a starting point from which to begin elucidating additional complex members.

There are variants of the yeast SAGA complex that have been identified and appear to be more involved in stress remediation than the SAGA complex. A variant of the SAGA complex, termed SALSA (SAGA altered, Spt8 absent) lacks the Spt8 subunit (Sterner et al., 2002). The SALSA complex has been reported to be more actively involved in the upreglation of histidine and tryptophan biosynthetic pathways during amino acid starvation over the traditional SAGA complex (Sterner et al., 2002). The SLIK (SAGA-like) complex is another SAGA variant that also lacks Spt8, contains an altered form of Spt7, and potentially includes a number of additional components not found in SAGA (Pray-Grant et al., 2002). The SLIK complex appears to be involved in the expression of certain nuclear genes during stress caused by mitochondrial dysfunction (Pray-Grant et al., 2002). The HAT complexes in *Toxoplasma* may also alter their compositions during stress (i.e. bradyzoite differentiation). Therefore, it would be important to include bradyzoite nuclear extract in the GST pull down assays or a bradyzoite cDNA library in a yeast two hybrid screen searching for potential GCN5-containing complex members.

F. Histone acetylation by TgGCN5

Discussion: All GCN5 homologues known to date selectively acetylate histone H3 in nucleosomal substrates *in vivo* (Ruiz-Garcia et al., 1997). Histone

H3 possesses several conserved lysine residues that can be acetylated, and GCN5 preferentially acetylates lysine 14 (Kuo et al., 1996). Recombinant GCN5 also weakly acetylates lysine 8 and lysine 16 of histone H4 *in vitro* but it is doubtful GCN5 mediates modifications of histone H4 *in vivo* once it has been incorporated into the SAGA complex (Grant et al., 1999). The unique N-terminal extension of TgGCN5 raised the possibility that it may influence the enzymatic activity of TgGCN5, making it different from previous study of GCN5 homologues. However, the catalytic domain of TgGCN5 is well conserved, sharing approximately ~60% with the yeast GCN5 catalytic domain (Sullivan and Smith, 2000).

The relatively long N-terminal extensions present in mouse and human GCN5 and PCAF are involved in substrate recognition. When the N-terminus is deleted in mouse and human GCN5 and PCAF, the ability to recognize and acetylate nucleosomal histones is lost (Xu et al., 1998). However, deletion of the mouse and human GCN5 and PCAF N-terminal extension has no effect on the ability of these GCN5 enzymes to recognize and acetylate histones that are not incorporated into nucleosomes. Even though the N-terminal extensions are not homologous, the unusual N-terminal domain of TgGCN5 may also have a role in regulating substrate specificity and/or acetylase function. It was of interest to determine if the long and divergent N-terminus of TgGCN5 was required for the acetylation of non-nucleosomal histones in solution and the specific histone and lysine residue(s) acetylated by TgGCN5.

To assess the enzymatic activity of TgGCN5, a positive control and *in vitro* HAT assay were developed based on previously described assays in the literature (Brownell et al., 1996). Recombinant *Saccharomyces cerevisiae* GCN5 (ScGCN5) was express in *E. coli* bacteria and purified by virtue of an in-frame N-terminal 6X histidine fusion tag. ScGCN5 was used in an *in vitro* radioactive HAT assay measuring HAT activity using tritiated acetyl CoA and core histones. In our *in vitro* assay, ScGCN5 showed a strong acetylation of histone H3 and minor activity on histone H4, consistent with activities reported previously in the literature (Kuo et al., 1996). ScGCN5 also appears to acetylate histone H1 but

additional studies indicated that GCN5 homologues do not acetylate H1 *in vivo* (Herrera et al., 1997). Thus, we successfully recreated the *in vitro* assay published in the literature and generated an in-house positive control allowing for the analysis of TgGCN5 and any additional *Toxoplasma* HATs identified in the future.

Since enzymatically active ScGCN5 was able to be expressed and purified from bacteria, an attempt was made to express TgGCN5 in bacteria. Despite our best efforts, we were unable to elicit significant acetylase activity in our HAT assay from recombinant TgGCN5 produced in bacteria. The most likely explanation for the lack of activity is the inability of bacteria to perform the necessary posttranslational modifications required by TgGCN5 to generate full acetylase activity. Bacteria have a limited capacity to generate disulfide bonds, which restricts their ability to generate functional eukaryotic proteins. Additionally, bacterial do not possess chaperone proteins that may be required to assist in folding TgGCN5 into its proper conformation (Makrides, 1996).

Multiple methods are available for improving the odds of obtaining functional protein from bacteria (Derman et al., 1993; Cole, 1996; Terpe, 2003). Alternatively, eukaryotic expression systems such as yeast and baculovirus are capable of generating active recombinant eukaryotic proteins (Schmidt and Hoffman, 2002). Any one of these methods could have been successful in producing enzymatically active TgGCN5, but a significant amount of time would have been spent optimizing expression conditions without any guarantee of success.

To obtain active TgGCN5 with minimal time devoted to optimizing protein expression, we purified protein from the transgenic parasites overexpressing $_{FLAG}$TgGCN5 and $_{FLAG}\Delta N_T$TgGCN5. The advantage of obtaining recombinant protein from *Toxoplasma* is that it was very likely to have proper folding and posttranslational modifications required for HAT activity. Using commercially available anti-FLAG affinity resin, the recombinant proteins were purified from parasite lysate. However, difficulty arose when trying to elute the proteins from the affinity resin. Therefore, *in vitro* HAT assays were performed with

$_{FLAG}$TgGCN5 or $_{FLAG}$$\Delta N_T$TgGCN5 still bound to the affinity resin. Fortunately, both proteins retained enzymatic activity and acetylated core histones while still bound to the affinity resin

Both TgGCN5 and ΔN_TTgGCN5 are able to acetylate free histones in solution suggesting that the N-terminal extension is dispensable for the acetylation of non-nucleosomal histones. As expected for a GCN5 HAT, TgGCN5 preferentially acetylates histone H3 and to a lesser extent H4 *in vitro*. Given the homology of the TgGCN5 catalytic domain, the preferential acetylation of histone H3 is not surprising. Loss of the N-terminal extension does not appear to affect the histones that are acetylated by TgGCN5 suggesting that the N-terminal extension does not play a role in the recognition with non-nucleosomal histones. The dispensability of the N-terminal extensions in TgGCN5 for acetylation of non-nucleosomal histones parallels observations of the N-terminal extensions present in metazoan GCN5 homologues.

At first glance, the HAT assay results indicate that full length TgGCN5 has a stronger acetylation activity than the truncated form (Figure 28-B, page 146). Furthermore, the Western blot performed in parallel suggests that there is more recombinant HAT protein present in the ΔN_TTgGCN5 HAT assay versus the full length reaction (Figure 28-A, page 146). The N-terminal extension may stabilize TgGCN5 generating stronger catalytic activity. There is evidence that human PCAF appears to transacetylate its own N-terminus, potentially stabilizing the enzyme (Herrera, 1997). The TgGCN5 N-terminal extension may undergo similar transacetylation. Loss of the N-terminal extension and subsequent modification would result in a less stable form of TgGCN5 and an overall lower activity. However, it may also be an artifact caused by using recombinant protein still bound to the affinity resin. The FLAG tag in both recombinant proteins is present on the amino-terminal end. Therefore, the FLAG tag places the catalytic domain of $_{FLAG}$$\Delta N_T$TgGCN5 in close proximity to the affinity resin, potentially impeding its HAT activity. Finding a method to effectively elute the proteins off the affinity resin or moving the FLAG tag to the C-terminus would give a meaningful comparison.

As with ScGCN5, TgGCN5 appears to be able to acetylate H1. Unlike other eukaryotic organisms, apicomplexan parasites do not appear to possess a histone H1 homologue (Sullivan et al., 2006). It is possible that the H1 orthologue present in apicomplexa is too divergent to detect. A histone H1 knockout in *Tetrahymena*, a protist related to *Toxoplasma*, grows normally, indicating that H1 may not be necessary in protozoa (Shen et al., 1995). Nonetheless, the ability to acetylate histone H1 *in vitro* is conserved in TgGCN5 despite the fact that *Toxoplasma* does not express a canonical histone H1 protein.

Future Studies: The N-terminus does not appear to play a role in the acetylation of non-nucleosomal histones. It still remains possible for the N-terminus of TgGCN5 to be involved in the recognition of nucleosomal histones. As mentioned previously, the long N-terminal extensions present in human and mouse GCN5 are required for the recognition and acetylation of nucleosomal histones. Therefore future studies should assess the enzymatic activity of full length TgGCN5 and ΔN_TTgGCN5 on nucleosomal substrates. Measuring activity on nucleosomal histones is desirable as these are the *in vivo* substrate of GCN5 proteins.

G. Determining the lysine(s) acetylated by TgGCN5

Discussion: Once it had been established that TgGCN5 preferentially acetylated histone H3, we wanted to identify which of the conserved lysine residues were acetylated by TgGCN5 and if the unique N-terminal extension influenced which residues were acetylated. To date, all known GCN5 homologues preferentially acetylate lysine 14 in histone H3 (Kuo et al., 1996).

To determine the acetylation target of TgGCN5, *in vitro* HAT assays were performed with TgGCN5 and $_{FLAG}\Delta N_T$TgGCN5 using commercially available recombinant histone H3 instead of core histones. The HAT assays were subsequently transferred to a Western blot and stained with antibodies specific to acetylated lysine residues. Given the acetylation patterns of previously studied GCN5 homologues (Kuo et al., 1996), antibodies detecting acetylation on lysine

14 were initially used. As predicted, the recombinant ScGCN5 (positive control) generated in bacteria strongly acetylated histone H3 on lysine 14. However, both TgGCN5 and $_{FLAG}\Delta N_T$TgGCN5 failed to show significant acetylation of lysine 14 in histone H3. The lack of acetylation by TgGCN5 on lysine 14 of histone H3 was a very surprising and unprecedented result. TgGCN5 represents the first GCN5 protein that does not acetylate lysine 14. However, our previous HAT assays performed with core histones clearly indicate that TgGCN5 acetylates histone H3. Therefore, acetylation of lysines 9 and 18 in histone H3 were also examined.

As expected, ScGCN5 showed strong acetylation on lysines 9 and 18, but TgGCN5 and $_{FLAG}\Delta N_T$TgGCN5 appeared to only acetylate lysine 18. Of the three lysine residues examined, lysine 18 of histone H3 appears to be the only one acetylated by TgGCN5, which is highly unusual for a GCN5 protein. However, despite the unusual acetylation pattern, removal of the N-terminus did not have any effect on which lysines were acetylated. Therefore, the N-terminal extension does not appear to play a role in dictating which lysine residues are acetylated by TgGCN5 in non-nucleosomal histone H3.

Selective acetylation of lysine 18 has not been reported for any other GCN5 HAT. As seen with ScGCN5, it was expected that TgGCN5 would be capable of acetylating lysine resides 9, 14, and 18, but with a stronger preference for lysine 14. In humans, acetylation of lysine 18 in histone H3 has been linked to subsequent methylation at arginine 17. In humans, the acetylation of lysine 18 *in vivo* is catalyzed by the HAT CBP followed by arginine 17 methylation by CARM1 (cofactor-associated arginine [R] methyltransferase 1; Daujat et al., 2002). A CBP homologue is absent in Apicomplexa and is not present in other early eukaryotes (Sullivan et al., 2006). It is plausible that TgGCN5 has adopted a CBP-like function in acetylating lysine 18 of histone H3 in *Toxoplasma*. Recently, a homologue of CARM1 has been cloned in *Toxoplasma* (TgCARM1) that methylates arginine 17 (Saksouk et al., 2005). Chromatin immunoprecipitation (ChIP) experiments revealed a strong correlation between lysine 18 acetylation and arginine 17 methylation suggesting that TgGCN5 and

TgCARM1 may work in concert to facilitate chromatin remodeling in *Toxoplasma* (Saksouk et al., 2005).

Future Studies: The selective acetylation of lysine 18 and not 14 in histone H3 by TgGCN5 was a most unexpected result. The selective acetylation of only one lysine seems to be highly unsual when compared to other HATs that possess the ability to acetylate multiple residues. The *Plasmodium* homologue of GCN5 (PfGCN5) has been demonstrated to acetylate lysines 9 and 14 of histone H3 (Miao et al., 2006). It is important to note that the *in vitro* HAT assays with PfGCN5 were performed using recombinant *Plasmodium* histone H3 (PfH3; Miao et al., 2006). The homology between the catalytic domain of PfGCN5 and TgGCN5 is very high (~ 85%). The high degree of homology suggests that the acetylase activity of both proteins should be comparable. The only difference between the PfGCN5 HAT assay and the assay used to assess TgGCN5 assay was the source of histones. Therefore, it may be worthwhile to use *Toxoplasma* histones in future enzymatic assays. The four canonical core histones have been identified and characterized in *Toxoplasma* (Sullivan et al., 2006). *Toxoplasma* histone H3 differs very little from the chicken erythrocyte histone H3 used in the radioactive assay or the recombinant *Xenopus laevis* histone H3 used in the non-radioactive assay (Sullivan et al., 2006). Nevertheless, enzymatic assays should be performed to confirm that TgGCN5 does not acetylate endogenous histones differently. The divergent N-terminal extension may interact with endogenous histones differently than heterologous histones resulting in a different enzymatic activity for TgGCN5.

When core histones were used as substrates, TgGCN5 showed modest activity on histone H4. The activity was dismissed as an *in vitro* artifact that has been seen in other GCN5 proteins that is not believed to occur *in vivo*. However, given the selective acetylation of TgGCN5 on histone H3, its activity on histone H4 may not be artifact and should be investigated further. Commercial antibodies recognizing residue specific acetylation of histone H4 are available and could be used to determine which lysines in histone H4 are being acetylated

by TgGCN5. However, the western blot approach is limited by the availability of antibodies to recognize specific histone modifications. An alternative detection method is to use MALDI/TOF (matrix-assisted laser desorption/ionization time-of-flight) and tandem mass spectrometry to identify the acetylation sites (Cocklin and Wang, 2003). The advantage of using MALIDI/TOF is it allows for a full post-translational modification analysis on an entire histone protein assessing any place that TgGCN5 acetylates the protein. It would be important to use recombinant histone proteins produces in bacteria to limit the amount of background acetylation present on the histones.

It should be noted that two MYST homologues, TgMYST-A and –B, have been cloned in *Toxoplasma* (Smith et al., 2005). In other organisms, MYST HATs are the proteins typically responsible for acetylating histone H4. *In vitro* HAT assays with TgMYST -A demonstrate it strongly acetylates histone H4 (Smith et al., 2005). The presence of MYST HATs in *Toxoplasma* suggests that the parasite already possesses proteins capable of acetylating histone H4, reducing the likelihood that TgGCN5 acetylates histone H4 *in vivo*. It is plausible that TgGCN5's primary function is to acetylate other transcription factors or proteins in the parasite. Future studies to assess this possibility will be addressed in a latter section.

As TgGCN5-containing complexes are elucidated, their enzymatic activity should be assessed as well. TgGCN5 may have an expanded acetylation profile once it is incorporated into a complex. Recombinant yeast GCN5 is capable of acetylating multiple lysine residues, but has a strong preference for lysine 14 in histone H3 (Grant et al., 1999). *In vitro* HAT with yeast ADA and SAGA complexes indicated that the ADA complex acetylates both lysines 14 and 18, and SAGA acetylates to all four lysines in H3. These results indicate that the association of yeast GCN5 with different proteins within HAT complexes may alter the substrate specificity of GCN5.

The unusual acetylation specificity of TgGCN5 may have specific functions in the parasite. Chromatin immunoprecipitation (ChIP) may yield some insight on the promoters that are remodeled through acetylation at lysine 18 in

histone H3. The transgenic parasite line overexpressing $_{FLAG}$TgGCN5 has been used in preliminary ChIP assays showing TgGCN5 present at the SAG2A (surface antigen 2-A) promoter (Saksouk et al., 2005). SAG2A is a tachyzoite specific gene (Lekutis et al., 2001) suggesting that TgGCN5 may be involved regulating tachyzoite gene expression. Additional ChIP assays are warranted to determine if TgGCN5 is associated with other tachyzoite promoters. It is possible that TgGCN5 may be involved in the genetic switching that occurs during reactivation of bradyzoites back into tachyzoites. Recrudescence of bradyzoites is the major cause of illness in humans with suppressed immune systems (Reiter-Owona et al., 2000). The possibility that TgGCN5 may be involved in this process is exciting as it would represent the first factor associated with this process.

The fact that TgGCN5 does not acetylate histone H3 at lysine 14 suggests that either lysine 14 is not acetylated in *Toxoplasma* or another protein is responsible for the acetylation. Searches of the *Toxoplasma* database revealed the presence of another GCN5 homologue (Bhatti et al., 2006). Two GCN5 homologues present in a lower eukaryote is unprecedented and will be discussed in a following section.

II. A TgGCN5 Deletional Mutant

Experiments performed by others in the Sullivan Lab lead to the development of a stable deletional mutant of TgGCN5 (ΔTgGCN5; Bhatti et al., 2006). No observable phenotype difference is noted for the ΔTgGCN5 clone. When the expression of GCN5 is disrupted in yeast, there is no readily observable phenotype, either (Georgakopoulos and Thireos, 1992). Further analysis revealed no appreciable difference in parasite growth rates between the wild type and ΔTgGCN5 (Bhatti et al., 2006). In collaboration with Dr. David Sibley (Washington University, St. Louis, MO), mice were infected with ΔTgGCN5 to determine if loss of TgGCN5 impaired virulence *in vivo*. Loss of TgGCN5 did not have any effect on parasite virulence suggesting, that TgGCN5

does not play a significant role during the acute stage of infection (Bhatti et al., 2006).

While a viable knockout of TgGCN5 suggests the protein is not essential in tachyzoites, it offers the opportunity to identify genes that may be controlled by this HAT in the parasite. Others in the Sullivan Lab have employed 2D gel analysis to assess the differential expression between wild-type and ΔTgGCN5 parasites. The 2D-gel analysis revealed several proteins down-regulated in the ΔTgGCN5 that are associated with host cell invasion including MIC4, MIC5, and MIC11 (Bhatti et al., 2006). The MIC proteins are found within the micronemes, a secretory organelle apart of the apical complex that facilitates invasion of host cells (Achbarou et al., 1991). Despite the apparent reductions in the expression and secretion of these microneme proteins, any difference in parasite attachment or invasion was undetectable (Bhatti et al., 2006). It is possible that the observed deficiencies on the 2D gel are not profound enough to alter the invasion process or other invasion proteins are compensating for the reduced amounts of MIC4, MIC5, and MIC11.

Interestingly, one of the enzymes that was upregulated in the ΔTgGCN5 mutant was enolase I. There are two paralogues of enolase in *Toxoplasma* that are stage specific (Singh et al., 2002). Enolase 2 (ENO2) is expressed only in tachyzoites (Dzierszinski et al., 2001). Enolase 1 (ENO1) is bradyzoite-specific and possesses different enzyme kinetics that are believed to slow the metabolic rate of bradyzoites (Dzierszinski et al., 2001). During the tachyzoite stage, enolase 1 expression is repressed (Kibe et al., 2005). The increased levels of ENO1 suggest that TgGCN5 may play a role in maintaining the repression of the ENO1 promoter during the tachyzoite stage. The dysregulation of a bradyzoite specific-gene provides evidence that TgGCN5 is linked to stage differentiation.

Comprehensive microarrays for *Toxoplasma* are in the final stages of development. When they are complete, it will be important to compare and contrast mRNA expression between wild-type and the TgGCN5 knockout. It would provide another tool to search for genes regulated by TgGCN5.

Microarray analysis would also determine if the differences detected by 2D gel results are reflected in the mRNA pool.

There are three different strains of *Toxoplasma* (Howe and Sibley, 1995). The type I strains, like the RH strain, are termed virulent strains because they are fast growing and uniformly lethal in mice (Sibley et al., 2002). Types II and III grow much slower and mice often survive infection with these parasite lines and remain chronically infected and seropositive (Sibley et al., 2002). The majority of human toxoplasmosis cases are caused by a member of the type II strain (Howe and Sibley, 1995). Interestingly, the three lineages of *T. gondii* differ by only 1–2% at the DNA sequence level (Su et al., 2002). The ΔTgGCN5 clone was generated in the RH strain which belongs to the type I group and used in routine molecular experiments because the fast growing nature of the parasite makes it more amenable to genetic manipulation (Bhatti et al., 2006). However, the virulent, type I, RH strain of *Toxoplasma*, has largely lost its ability to convert to the bradyzoite cyst form (Dubey et al., 1999). Therefore, it is not the ideal strain to assess the deletional effect of TgGCN5 on bradzyoite differentiation. A more appropriate study would be to generate a deletional mutant in a type II strain such as ME49 or Pru. The type II strains are much slower growing and therefore easier to differentiate into bradyzoites *in vitro* (Howe and Sibley, 1995). Since the type II parasites are more commonly found human infections, making a TgGCN5 deletional mutant in these strains is more clinically relevant. However, generating a deletional mutant in a type II strain is much more challenging because of the slow growing nature of the parasites. Nonetheless, a deletional mutant of TgGCN5 in a type II strain would allow for a better assessment of the potential involvement of TgGCN5 in differentiation (Dubey et al., 1998). Additionally, a deletional mutant in a type II strain provides the ability to assess bradyzoite differentiation *in vivo* using animal models and can also be fed to felines under laboratory conditions to assess the loss of TgGCN5 on the development of micro- and macrogametes and oocysts (Dubey et al., 1970).

Continued characterization of the ΔTgGCN5 strain will provide further information about role of TgGCN5 histone and protein acetylation in the parasite

and may provide a better understanding of bradyzoite differentiation and host cell invasion.

III. Additional HATs in Apicomplexan Parasites

A. A second homologue of GCN5 in Toxoplasma

From the studies conducted during this thesis, we have concluded that TgGCN5 is unable to acetylate histone H3 at lysine 14. In addition, when the genomic loci for TgGCN5 is disrupted, there is no observable phenotype even under stress conditions (Bhatti et al., 2006). These two pieces of data argue that there must be another protein expressed in *Toxoplasma* capable of acetylating histone H3 at lysine 14 with the potential to compensate for the loss for TgGCN5.

Subsequent searches of the *Toxoplasma* database revealed the presence of a second GCN5 homologue (Bhatti et al., 2006). Mammals possess a second GCN5 family member, PCAF, but there are no reports of multiple GCN5 HATs in other animals or plants. Our lab subsequently cloned the second homologue of GCN5 (termed TgGCN5-B) and the TgGCN5 protein previously reported is now referred to as TgGCN5-A. Thus, *Toxoplasma* appears to be unique among invertebrates in harboring two independent GCN5 HATs, and the relevance to parasite physiology is subject to future investigation.

The TgGCN5-B ORF encodes a predicted protein of 1032 amino acids, which is 137 residues shorter than TgGCN5-A. The HAT domains are nearly identical, but the putative Ada2-binding domain and bromodomain are less conserved. Most striking is the lack of homology in the N-terminal extensions of both proteins. Neither one of these N-terminal extensions has homology to protein sequences in other species, nor are any known protein motifs evident. Furthermore, the N-termini have no homology to each other. The studies outlined in this thesis will serve as a template to begin investigating TgGCN5-B, particularly the role of its unique N-terminal extension.

Preliminary enzymatic data produced in our lab indicates that TgGCN5-B acetylates histone H3 at the expected lysine 14, indicating it may be the true GCN5 orthologue in *Toxoplasma* (Bhatti et al., 2006). Furthermore, TgGCN5-B

is able to acetylate lysine residues 9 and 18 suggesting that it can compensate for the loss of TgGCN5-A explaining the lack of an observable phenotype in the TgGCN5-A knock out (Bhatti et al., 2006). TgGCN5-A may be the result of a gene duplication event that has evolved a highly specific function in chromatin remodeling in these parasites, as suggested by its proclivity to acetylate lysine 18 of H3. Finally, elucidating the proteins TgGCN5-B interacts with will provide additional insight into the overall function of this HAT in the parasite.

Repeated attempts by our lab to disrupt the TgGCN5-B genomic loci have been unsuccessful. The inability to knock out TgGCN5-B suggests that the protein is essential in tachyzoites. Furthermore, it argues that TgGCN5-A is unable to compensate for the loss of TgGCN5-B. Given the completely different acetylation patterns between the two HATs, the inability for TgGCN5-A to compensate for the loss of TgGCN5-B is not surprising. The similarities and differences between the two GCN5 proteins in *Toxoplasma* is akin to the relationship between GCN5 and PCAF in mammals.

When murine GCN5 (mGCN5) is deleted in mice, the embryos do not survive past day 10 post conception due to developmental defects (Xu et al., 2000). Thus a deletion of mGCN5 is lethal and the same appears to be true of TgGCN5-B in *Toxoplasma* tachyzoites. Mice that are homozygous null for murine PCAF (mPCAF) are developmentally normal without any distinct phenotype (Yamauchi et al., 2000). It has been proposed that mGCN5 can functionally compensate for the loss of mPCAF (Yamauchi et al., 2000). The deletional mutant of TgGCN5-A possess no phenotypic anomaly and it is likely TgGCN5-B is able to compensate for the loss of TgGCN5-A. The lethality of a mGCN5 knockout suggests that mPCAF is unable to compensate for the loss mGCN5. Given the inability to obtain a TgGCN5-B knockout, argues that TgGCN5-A is unable to compensate for the loss of TgGCN5-B. Therefore, TgGCN5-A may be analogous to mPCAF and TgGCN5-B is a functional equivalent to mGCN5. However, the acetylation pattern of PCAF is highly similar to GCN5 showing strong activity on lysine 14 of histone H3 with modest activity

on lysines 9 and 18 (Schiltz et al., 1999). Thus, the analogy between TgGCN5-A and PCAF is not flawless.

B. GCN5 homologue in Plasmodium falciparum

A GCN5 homologue was also identified and cloned in *Plasmodium falciparum* (PfGCN5; Fan et al., 2004a). PfGCN5 also contains a long N-terminal extension upstream from its catalytic domain, currently the longest one known to date (1126 amino acids). The N-terminus in PfGCN5 possesses no homology to any known protein and it is completely divergent from the N-termini in TgGCN5-A and -B. PfGCN5 contains the conserved domains present in other GCN5 proteins: the catalytic HAT domain, Ada2 binding domain, and the bromodomain. Enzymatic studies have shown that PfGCN5 acetylates histone H3 at lysines 9 and 14 of *Plasmodium* histone H3 *in vitro* (Miao et al., 2006). *Plasmodium* also possesses a single Ada2 protein, (PfAda2; Fan et al., 2004b). Unlike *Toxoplasma*, searches of the *Plasmodium* genome indicate that *Plasmodium* expresses only one Ada2 and GCN5 protein. The lack of additional copies of Ada2 and GCN5 in *Plasmodium* suggests that the *Toxoplasma* has generated duplications of Ada2 and GCN5. The purpose of these gene duplications in *Toxoplasma* is unknown and warrants additional studies.

C. A MYST homologue in Toxoplasma

Two MYST HATs have been isolated in *Toxoplasma* termed TgMYST-A and -B (Smith et al., 2005). The TgMYST-A genomic locus generates a transcript of approximately 3.5-kb that can encode two proteins of 411 or 471 amino acids (Smith et al., 2005). The longer protein has been termed TgMYST-A(L) and the shorter one TgMYST-A(S). TgMYST-B mRNA is approximately 7.0kb and encodes a second MYST homologue (Smith et al., 2005). Interestingly, there is a decrease of in the expression of both forms of TgMYST-A during *in vitro* bradyzoite induction (Smith et al., 2005). The decrease in expression suggests that both forms of TgMYST-A are predominantly expressed in tachyzoites.

Both the long and short forms of recombinant TgMYST-A acetylate histone H4 *in vitro* at lysines 5, 8, 12, and 16 (Smith et al., 2005). While GCN5 family members show a bias towards acetylating H3, the MYST family of HATs display a preference for acetylating H4 (Utley and Cote, 2003). The ability for TgMYST-A to acetylate all for residue in histone H4 suggests that TgMYST-A may be an orthologue of Esa1/Tip60 since it can target every lysine in the H4 tail.

Similar studies assessing the expression of TgMYST-B in tachyzoites versus bradyzoites and determining the acetylation pattern have yet to be completed. It is conceivable that the second MYST may be expressed during the bradyzoite stage.

IV. Additional Future Studies for TgGCN5-A

A. Acetylation of non-histone substrates

The unusual acetylation pattern of TgGCN5-A and the presence of a second homologue capable of acetylating histone H3 at all the lysine residues suggests that TgGCN5-A may involved in additional processes beyond chromatin remodeling. It has been suggested that TgGCN5-A may represent a functional equivalent to mammalian PCAF. In addition to acetylating histones, PCAF possesses the ability to acetylate non-histone proteins. Therefore, it is possible that TgGCN5-A may also acetylate non-histone proteins.

The process of acetylating non-histone proteins is referred to as factor acetyltransferase or FAT activity (Roth et al., 2001). The cell cycle regulator p53 is one the of most widely study targets of the FAT activity of PCAF. The acetylation of p53 by PCAF appears to be triggered in response to the stress caused by DNA damage due to UV or ionizing radiation (Liu et al., 1999). The acetylation of p53 by PCAF increases the affinity of p53 to bind DNA, augmenting its ability to upregulate stress response proteins (Liu et al., 1999). Therefore, PCAF is involved in stress remediation pathways through its FAT activity.

PCAF also acetylates the chromatin associating protein HMG-17 (Herrera et al., 1999). HMG (high mobility group) proteins are a group of non-histone

proteins which modulate the structure chromatin altering the accessibility of DNA (Bustin et al., 1995). HMG-17 specifically binds to the 146 base pair nucleosomal core particle and modulates the chromatin structure of active genes (Bustin et al., 1995). The acetylation of HMG-17 by PCAF reduces the affinity of HMG-17 for nucleosome cores. Therefore, PCAF acetylation can affect chromatin structure through the acetylation of non-histone proteins. The overall affect of acetylating HMG-17 on gene transcription is still controversial. The binding of HMG-17 to nucleosomes unfolds the higher-order chromatin fiber and enhances the accessibility of various factors to the chromatin (Herrera et al., 1999). Acetylation appears to weaken the interaction of HMG-17 with DNA suggesting it may cause the chromatin to refold generating a repressive effect to transcription. Alternatively, the acetylation of HMG-17 may have the same effect as the acetylation as histones, reducing the affinity for DNA and further opening chromatin allowing for an increase in transcription (Herrera et al., 1999).

If TgGCN5-A possess FAT activity, it may also acetylated non-histone proteins involved in stress remediation. Therefore, it may acetylate DNA-binding transcription factors that are involved in bradyzoite differentiation. Searches of the *Toxoplasma* database indicate that the parasite does not containing a homologue of p53. However, HMG proteins have been identified in *Plasmodium* and similar proteins may exist in *Toxoplasma* (Nambiar et al., 1997). The presence of HMG proteins presents another level of chromatin remodeling present in apicomplexan parasites. Apicomplexan HMG proteins may be acetylated by GCN5 proteins such as TgGCN5-A. *In vitro* HAT assays with TgGCN5-A and recombinant non-histone substrates such as HMG would determine if TgGCN5-A possesses the ability to acetylate non-histone substrates. If activity on non-histone substrates exists, HAT assays with $_{F}\Delta N_{T}$TgGCN5 would determine if the N-terminus is required to acetylate non-histone proteins.

A screen could be established to further identify non-histone substrates. A protocol has been established for monitoring acetylation by incubating cells with with tritiated acetate (Zhang and Nelson, 1988). The tritated acetate is

incorporated into the cellular acetyl CoA pool and is utilized by HATs to acetylate proteins. Both wild-type and the TgGCN5-A knock out parasites could be incubated in tritiated acetate. Any proteins capable of being acetylated would be marked with a radiolabeled acetyl group in the parasites. The parasites would be fractionated between nuclear and cytoplasmic fractions. The nuclear fraction would be resolved via 2D gel electrophoresis. The gel would be processed for autoradiography and exposed to film. Dots on the exposed film would indicate the presence of radioactive acetylation. The knock out would provide a necessary comparison to establish which proteins were aceylated by TgGCN5-A. Dots present on film from wild type lysate that are missing on the film from the knock out would indicate a loss of aceyltation. Proteins that appear to be differentially acetylated between wild-type and knock out would be sent for sequencing analysis. Proteins that were acetylated could be eventually cloned and recombinant protein expressed. The recombinant protein would be used as substrate in an *in vitro* acetylation assay with TgGCN5-A to confirm that TgGCN5-A is responsible for the acetylation of the protein.

The differential expression of proteins revealed in the 2D gel analysis of ΔTgGCN5-A may have been caused by decreased protein acetylation. The decreased acetylation may have caused proteins such as MIC4, MIC5, and MIC11 to migrate differently in the ΔTgGCN5-A 2D gel versus the wild type 2D gel. Therefore, it would be interesting to see if the differences in the 2D gel analysis are paralleled in the acetyl Co-A labeling experiment. No studies have been completed determining if the MIC (microneme) proteins are acetylated.

B. Posttranslational modification of TgGCN5-A

Like other proteins, GCN5 has the potential to be post-translationally modified. The co-immunoprecipitations performed on parasites overexpressing $_{FLAG}$TgGCN5 and $_{FLAG}\Delta N_T$TgGCN5 through our collaboration with Dr. Mohamed-Ali Hakimi support this possibility. The SDS-PAGE analysis of co-immunprecipitations with both proteins produced multiple bands migrating at sizes larger than expected for $_{FLAG}$TgGCN5 and $_{FLAG}\Delta N_T$TgGCN5 (Figures 21

and 22 on pages 121 and 126, respectively). The slower migrating versions of $_{FLAG}$TgGCN5 and $_{FLAG}\Delta N_T$TgGCN5 can be explained by the presence of post translational modifications that are retarding their migration through the gel. A single phosphorylation can impede the SDS-PAGE migration of a protein by up to 5 kilodaltons (Smith et al., 1989). As more modifications are added, the retardation of migration becomes more pronounced. There is evidence that GCN5 can be phosphorylated (Barlev et al, 1998) and the related HAT PCAF can be transacetylated (Herrera et al., 1997). These same modifications may occur to TgGCN5-A and may impact the activity of TgGCN5-A.

Phosphorylation of human GCN5 (hGCN5) by the DNA-dependent kinase Ku70 attenuates the HAT activity of hGCN5 (Barlev et al, 1998). The kinase Ku70 interacts with the bromodomain of hGCN5 and through this interaction is believed to phosphorylate the amino terminus of recombinant hGCN5. The precise residues acetylated by Ku70 are still being resolved. Using siRNA, the expression of Ku70 was knocked down in human cells resulting in an increased sensitivity to DNA damage caused by ionizing radiation suggesting the kinase has a role in DNA repair (Ayene et al., 2005). The purpose of down regulating the HAT activity of GCN5 during DNA repair is unknown however it may be to prevent the remodeling of chromatin containing damaged DNA.

Additionally, human PCAF appears to transacetylate its N-terminus, which is essential for it to be transported into the nucleus and may potentially stabilize the enzymatic function of the enzyme (Santos-Rosa et al., 2003; Herrera 1997). TgGCN5-A may be capable of similar transacetylation. There are five lysines in the PCAF N-terminus that are acetylated; lysines 416, 428, 430, 441, and 442 (Santos-Rosa et al., 2003). Based on the alignment between TgGCN5-A and human PCAF using the Clustal W algorithm (http://www.ch.embnet.org/software/ClustalW.html), TgGCN5-A only possess one of these lysine residues, lysine 430, which corresponds to lysine 712 in TgGCN5-A. The lack of conserved lysine residues suggests that TgGCN5-A may not be transacetylated. Human GCN5 does not appear to acetylate its N-terminus (Herrera et al., 1997). Only two of the five lysines that are acetylated in

human PCAF are conserved in human GCN5 (lysines 428 and 441). However, there are 25 other lysine resides present in the N-terminal extension of TgGCN5-A that are potential acetylation sites.

The acetylation mechanism of the N-terminus occurs via an interaction with a second PCAF protein (Santos-Rosa et al., 2003). The *in vitro* HAT assays used in this thesis could be modified using recombinant TgGCN5-A N-termininal peptide as a substrate instead of histones. Radioactive assays could be performed by comparing the acetylase activity of full length TgGCN5-A and a catalytically inactive mutant TgGCN5-A on the N-terminal peptide. Following the *in vitro* assay, the proteins would be resolved via SDS-PAGE and the gel processed for autoradiography. If a band corresponding to the size of the N-terminal peptide developed on film after being exposed to the gel, it would indicate that the N-terminal extension is acetylated by TgGCN5-A. The catalytically inactive mutant TgGCN5-A would serve as a negative control and should be unable to acetylate the N-terminal peptide. In addition, non-radioactive acetylation assays could be performed and the N-terminal peptide excised from the gel, and sent for MALDI/TOF analysis to determine which, if any, lysine residue(s) are being acetylated.

If it is determined that lysines in the N-terminus are being acetylated, site directed mutagenesis switching the lysine for arginine would provide information about the importance of acetylating these residues. Arginine is not susceptible to acetylation but is positively charged like lysine and therefore should not generate significant change it the conformation of TgGCN5-A when exchanged for lysine (Santos-Rosa et al., 2003). Immunofluorescence assays performed on parasites expressing mutated TgGCN5-A with the lysine exchanged for arginine would determine if acetylation has an effect on nuclear localization as it does in human PCAF. The enzymatic function of recombinant protein purified from parasites expressing the same mutated versions of TgGCN5-A could be assessed via *in vitro* HAT assays to determine if acetylation has any effect of the ability of TgGCG5-A to acetylate histone. The enzymatic activity of human PCAF is increased when it is acetylated (Santos-Rosa et al., 2003).

V. HAT and HDAC Inhibitors

Histones were originally thought to be no more than simple scaffolding proteins, providing a mechanism for compacting DNA into the nucleus. However, in the past few decades it has become abundantly clear that histones can be posttranslationally modified altering the structure of chromatin. The resulting changes in chromatin structure have profound impact on the expressed genome resulting differentiation, activation of stress response pathways, apoptosis, and cell-cycle regulation. Acetylation of histones is the most well studied modification and is carried about by histone acetyltransferases (HATs) and histone deacetyl transferases (HDACs). HATs and HDACs work cooperatively to provide the proper balance of histone and protein acetylation to maintain a functional cellular state (Kurdistani and Grunstein, 2003). Disruption in the balance of histone and protein acetylation is deleterious to cells resulting in cell death or carcinogenesis (Timmermann et al., 2001). HAT and HDAC inhibitors hold the promise of being able to restore the acetylation balance and may be effective antitumor agents. Alternatively, HAT and HDAC inhibitors can be used to unbalance acetylation in pathogenic organisms providing potential therapies for HIV and apicomplexan infections.

A. HDAC Inhibitors

Interestingly, the involvement of histone modification in regulating global processes was originally discovered through screening compounds to aid in the treatment of cancer (Kim et al., 2003). Trichostatin A (TSA), a well-known HDAC inhibitor, was originally reported as a fungistatic antibiotic (Tsuji and Kobayashi, 1978). The treatment of a leukemia cell line with TSA resulted in differentiation and inhibition of the cell proliferation at very low concentrations (Yoshida et al., 1987). Additional analysis revealed that cells treated with TSA caused an accumulation of acetylated histones in a variety of mammalian cell lines. Furthermore, TSA strongly inhibits the activity of the partially purified histone deacetylases *in vitro*. (Yoshida et al., 1990). The same story is true for

depudecin which was isolated from a screening project for agents capable of reverting malignant carcinogenic cells to a normal phenotype and was later identified as targeting HDACs (Kim et al., 2003).

The mechanism of HDAC inhibition leading to reversion of cancerous cells is not entirely understood. Direct alterations in the HDAC genes have not been demonstrated in human cancers (Cress and Seto, 2000). Typically HDACs are seen as repressive proteins causing the deacetylation of histones restoring chromatin into its condensed state. Furthermore, cell cycle regulators such as Mad (mitotic arrest deficient) and the retinoblastoma (Rb) protein associate with HDAC complexes to mediate transcriptional repression leading to differentiation and suppressing proliferation (Vigushin and Coombes, 2004). Thus, it would seem that inhibiting HDACs would push a cell further into carcinogensis, but paradoxically, the opposite occurs with HDAC inhibitor exposure. It has been theorized that the HDAC proteins act upstream of the dysregulated proteins allowing for histone hyperacetylation at promoters of cell regulatory proteins. The tumor suppressor protein p21 is a cyclin-dependent kinase (CDK) inhibitor which binds defective cell cycle regulators and inhibits their activity, leading to suppression of cell proliferation (Vigushin and Coombes, 2004). Inhibition of HDACs results in hyperacetylated histones, markedly increasing transcription of p21. In fact, treatment of colon cancer cells with HDAC inhibitors results in an upregulation of p21 expression and subsequent growth arrest (Timmerman et al., 2001). Given the impressive antitumor activity of HDAC, several inhibitors are currently in phase I clinical trials for the treatment of the treatment of cancer (Vigushin and Coombes, 2002).

The inhibition of HDACs may be beneficial to other disease states besides cancer. Huntington's disease (HD) is an autosomal dominant neurodegenerative disorder caused by a polyglutamine repeat expansion in the huntingtin gene (Davies and Ramsden, 2001). Recent studies have indicated transcriptional dysregulation as a mechanism of HD pathogenesis through disruption of histone acetylation (Sadri-Vakili and Cha, 2006). Several HATs including CBP and PCAF are depleted by the sequestration into the aggregates of mutant huntingtin

resulting in a globally hypoacetylated state in a cell culture model (Butler and Bates, 2006). The inhibition of histone deacetylases is theorized to compensate for the reductions in HAT protein and restore the proper balance of acetylated histones in HD models (Butler and Bates, 2006). The use of HDAC inhibitors in HD animal models has shown promise as potential therapeutic agents (Sadri-Vakili and Cha, 2006).

The inhibition of HDACs has also been linked to the repression of proinflammatory cytokine production (Barnes et al., 2005). Therefore, inhibition of HDACs may provide novel therapeutic treatments for inflammatory disease such as asthma, inflammatory bowel disease, and systemic lupus erythematosus (Huang, 2006).

B. HAT inhibitors

Aberrant lysine acetylation has also been implicated in carcinogenesis. As mentioned previously, the HAT protein p300/CBP has the ability to acetylate non-histone proteins like p53 (Liu et al., 1999). The acetylation of p53 augments its ability to bind DNA. p53 is often referred to as the “guardian of the genome” and its importance is emphasized by the fact that 50% of all human cancers possess mutations in p53 (Balasubramanyam et al., 2004b). Acetylation by p300/CBP and PCAF is key in activating p53 when DNA damage occurs (Liu et al., 1999). Loss or a defect in the ability to acetylate p53 may lead to carcinogenesis. A disruption in the p300 gene has been demonstrated to be common in colorectal and gastric cancer samples (Muraoka et al., 1996). It is estimated that 80% of glioblastoma, the most aggressive form of primary brain tumors) cases have been associated with the loss of heterozygosity of the p300 gene (Balasubramanyam et al., 2004b).

Chromosomal translocations resulting in the three different fusions between p300/CBP and a MYST member are associated with leukemia (Dash and Gilliland, 2001). The three fusion proteins are MLL/CBP, MLL/p300 and MOZ/CBP and incorporate the well characterized transcriptional co-activators CBP/p300 to either the monocytic leukemia zinc finger (MOZ) gene or to MLL

(mixed linage leukemia; Timmerman et al., 2001). The acetyltransferase domains of the MYST member and p300/CBP remain intact in the fusion proteins resulting in unregulated HAT activity and aberrant histone acetylation leading to leukemia (Yang, 2004).

Mutations in human CBP have also been associated with the developmental disorder, Rubinstein-Taybi syndrome (RTS; Timmerman et al., 2001). RTS occurs in 1/125,000 births and is characterized by growth retardation and psychomotor developmental delay (Coupry et al., 2002). Interestingly, patients with RTS have a 5% increased incidence in childhood tumors (Miller and Rubinstein, 1995).

Protein acetylation has also been implicated in the regulation of HIV-1 gene transcription. The histone deacetylase inhibitor trichostatin A (TSA) is a potent inducer of HIV-1 transcription in latently infected T-cell lines through hyperacetylation of histones (Sheridan et al., 1997). Furthermore, the acetylation of histone H3 and H4 by HAT complexes stimulate HIV-1 transcription from preassembled nucleosomal templates (Steger et al., 1998). The HIV viral protein Tat (trans-activator protein) appears to highjack various acetyltransferases including Tip60, hGCN5 , p300/CBP and PCAF (Quivy and Van Lint, 2002). Tat interacts with Tip60 (Tat-interactive protein) to hinder the expression of cellular genes which normally interfere with the replication and propagation of the virus (Creaven et al., 1999). p300/CBP, hGCN5, and PCAF appear to each acetylate Tat at specific lysine residues within the protein, stimulating the transcriptional activity of Tat (Col et al., 2001). It has been suggested that the requirement of histone and Tat acetylation in the replication of HIV defines HATs as new targets for HIV drug design (Quivy and Van Lint, 2002). HAT inhibitors could maintain infected cells in the latent state and may augment highly active anti-retroviral therapy (HAART) in controlling HIV infection (Quivy and Van Lint, 2002).

Although significant progress has been made in the field of histone deacetylase inhibitors, there are only a few HAT inhibitors (Balasubramanyam et al., 2003). Two synthetic peptide-CoA conjugates have been generated that are specific inhibitors of PCAF and p300. Lysyl CoA (Lys-CoA) specifically inhibits

p300 and PCAF is inhibited by H3-CoA-20 (Lau et al., 2000). Interestingly, Lys-CoA does not effectively inhibit PCAF and H3-CoA-20 only affects p300 at very high concentrations (Lau et al., 2000). The difference between these two inhibitors demonstrates that different HATs can be selectively inhibited despite homology between their catalytic domains. Unfortunatley, Lys-CoA is unable to permate cell membranes limiting its usefulness to only *in vitro* assays (Balasubramanyam et al., 2003).

Newer HAT inhibitors that are permeable to cell membranes are being derived from naturally occurring substances. Anacardic acid is derived from cashew nut shell liquid, and has been shown to exhibit antitumor activity (Balasubramanyam et al., 2003). Further *in vitro* analysis demonstrates that it inhibits the HAT activity of p300 and PCAF (Balasubramanyam et al., 2003). A p300 selective inhibitor curcumin (diferuloylmethane) is derived from the spice turmeric (Kang et al., 2005). Curcumin is selective for p300 and does not have any effect on the activity of PCAF (Balasubramanyam et al., 2004b). Treatment of HIV infected cells with curcumin significally repressed the multiplication of the virus and inhibited the acetylation of HIV-Tat protein (Balasubramanyam et al., 2004b). Garcinol, a polyisoprenylated benzophenone derivative from the rind of *Garcinia indica* berries, is capable of inhibiting both p300 and PCAF (Balasubramanyam et al., 2004a). Garcinol and curcumin displayed strong growth inhibitory effects and induction of apoptosis in human leukemia cells (Pan et al., 2001). Furthermore, curcumin also appears to selectively induce apoptosis in mouse and human cancer cell lines but has minimal effect upon exposure to healthy primary mouse and human cell cultures (Jiang et al., 1996). However, it is still unknown if the apoptotic effects of garcinol and curcumin are caused by acetylase inhibition or through different mechanisms.

Work has been done using natural occurring HAT inhibitors to derive more selective inhibitors as potential theraputics and tools to further characterize HAT proteins. Compound MC1626 (2-methyl-3-carbethoxyquinoline) is an analog of anacardic acid (Ornaghi et al., 2005). It is able to permeate cell membranes and treatment of yeast cells with MC1626 inhibits cell growth in a Gcn5p-dependent

way (Ornaghi et al., 2005). Cell growth inhibition was significantly diminished in a null gcn5Δ and a HAT catalytic mutant strain (Ornaghi et al., 2005). MC1626 represents the first membrane permeable inhibitor of GCN5. While MC1626 may have potential as a therapeutic compound, it may be an even more important as molecular tool facilitating continued dissection of GCN5-mediated expression and chromatin remodeling. It may also provide a starting point from which additional GCN5 inhibitors or activators can be derived.

C. Apicomplexan HDAC and HAT inhibitors

Compounds able to inhibit HDAC and HAT proteins eventually made its way into the apicomplexan world led by the discovery of apicidin. Apicidin is a fugal metabolite isolated from a Costa Rican fungus (*Fusarium* spp.) and possesses a tetrapeptide structure (Darkin-Rattray et al., 1996). When cultured parasites were treated with apicidin, it inhibited the growth of drug-resistant human malaria, *C. parvum*, and *Toxoplasma*, demonstrating it as a broad spectrum antiprotozoal (Darkin-Rattray et al., 1996). Further analysis revealed that apicidin inhibits the histone deacetylases in apicomplexan parasites (Singh et al., 2002). Unfortunately, apicidin has a narrow therapeutic window and inhibits human HDACs resulting in an anti-proliferative effect on human cells, limiting its clinical potential to treat apicomplexan diseases. Attempts have been made by Merck Research Laboratories at modifying the parent compound to increase the selectivity for parasite HDACs. However, none of the related compounds were active against parasites *in vivo* (Tucker-Samaras, 2003). Nevertheless, the potent and broad spectrum antiprotozoal effect of apicidin illustrates the importance of pursuing chromatin remodeling enzymes as potential therapeutic targets to combat apicomplexan diseases (Darkin-Rattray et al., 1996).

Disruption of HATs in other systems is often deleterious to the organism (Timmerman et al., 2001). HATs in *Toxoplasma* and other apicomplexans are involved in regulating gene expression. In fact, progression of malarial parasites through different life cycle stages has been linked to changes in histone

acetylation (Miao et al., 2006). Therefore, disrupting HAT and HDAC complexes could influence stage differentiation in apicomplexans. Evidence exists suggesting that TgMYST-A and TgGCN5-A HATs are involved in regulating tachyzoite expression (Saksouk et al., 2005; Smith et al., 2005). Inhibition of these proteins may lock the parasite in the tachzoite stage or prevent recrudescence from bradyzoites back into tachyzoites. The inability to successfully generate a TgGCN5-B and TgMYST-A deletional mutants suggests that these two chromatin remodelers are essential to the tachyzoite state of *Toxoplasma* (Bhatti et al., 2006; Smith et al., 2005). Therefore, HAT inhibitors able to target these HATs may inhibit parasite growth. A limiting step in the inhibition of HATs and HDACs in parasites is finding an inhibitor that is selective for parasites and has little to no effect on human tissues.

A screening of several HAT inhibitors identified a candidate compound, Compound 1, with good parasite selectivity and broad-spectrum antiprotozoal effects (Tucker-Samaras, 2003). Unfortunately, Compound 1 was extremely toxic *in vivo* but it provides evidence that parasite selective inhibitors are obtainable (Tucker-Samaras, 2003). The selective nature of Compound 1 still makes it an excellent investigational tool for further characterizing parasite HAT proteins. Compound 1 may also provide a starting point upon which to design further compounds that maintain its parasite selectivity but lower its *in vivo* toxicity (Tucker-Samaras, 2003). The *in vivo* toxicity of Comound 1 was very rapid suggesting that it was mediated by a mechanism other than HAT inhibition (Tucker-Samaras, 2003). Additional studies indicate that Compound 1 targets one of the MYST acetyltransferases in apicomplexans.

Future experiments testing the effect of anacardic acid, curcumin, garcinol, and compound MC1626 on parasite growth are warranted. *Toxoplasma* does not possess a homologue of p300 (Nallani and Sullivan 2005). Therefore, anacardic acid, curcumin, and garcinol may have no effect on the growth of *Toxoplasma*. However, it has been suggested that TgGCN5-A may be a functional equivalent of PCAF. The ability of garcinol and anacardic acid to inhibit its HAT activity would provide further evidence of a PCAF-like function.

However, all four of these compounds may possess activity against *Toxoplasma* HATs. As compound MC1626 has been demonstrated to inhibit yeast GCN5, it would be of particular interest to demonstrate its ability to inhibit parasite growth. If compound MC1626 is able to inhibit parasite growth, it is highly probable that it is targeting TgGCN5-B and not TgGCN5-A, because a knock out of TgGCN5-A is viable (Bhatti et al., 2006).

In vitro HAT assays could also be employed to determine the targets of HAT inhibitors in *Toxoplasma*. Different inhibitors may display selectivity against different HATs expressed in *Toxoplasma*. Additionally, as new HAT inhibitors are discovered they should be tested against known and newly cloned apicomplexan HATs. It is important to investigate compounds that are ineffective or weak inhibitors of mammalian HATs. These inhibitors may selectively inhibit apicomplexan HATs. Thus, when they are added to parasites in culture, they will injure the parasites but have no effect on the host cells. Inhibitors with parasite selectively have the potential to be novel therapeutics in treatment of *Toxoplasma* and other apicomplexan infections.

Novel inhibitors of apicomplexan HATs and HDACs may also be applicable to other disease caused by dysregulated histone acetylation. Although originally discovered to thwart apicomplexan disease, apicidin is now being investigated as a potential anticancer agent. Treatment of v-ras-transformed mouse fibroblasts with apicidin induced morphological alteration towards differentiation due to hyperacetylation of histone H4 (Kim et al., 2004). Furthermore, apicidin strongly inhibited angiogenesis of vascular endothelial cells (Kim et al., 2004). Angiogensis is the formation of new vessels and is a key step in the spread of metastatic cancer and has been linked to histone acetylation (Kim et al., 2003). Therefore, the anti-angiogenic and anti-oncogenic potential of apicidin may lead to its use as a new type of drug to treat cancer.

Clearly, the continued characterization of *Toxoplasma* HATs and other chromatin remodeling proteins will aid in understanding parasite biology and the bradyzoite differentiation process. Furthermore, it will assist in the development of development of HAT and HDAC inhibitors which have the therapeutic potential

to treat not only in apicomplexan infections but many other disorders including cancer, HIV infection, genetic diseases, and inflammatory conditions.

CHAPTER 5: SUMMARY

- The N-terminal extension of TgGCN5 is required for the nuclear localization of TgGCN5-A.
- The motif RKRVKR present in the N-terminus between residues 93 and 99 is required for TgGCN5-A nuclear localization.
- The motif RKRVKR is capable of targeting a cytoplasmic protein to the parasite nucleus.
- RKRVKR is the first NLS to be described in any apicomplexan.
- The RKRVKR NLS is not present in any other known GCN5 proteins.
- Permutations of RKRVKR have predictive value in bioinformatic searches of apicomplexan sequence databases.
- A homologue of importin α was identified and cloned in *Toxoplasma* (TgIMPα).
- TgGCN5-A interacts with TgIMPα via the RKRVKR NLS *in vitro*, demonstrating that the N-terminus can participate in protein-protein interactions.
- In collaboration with Dr. Mohamed-Ali Hakimi's laboratory at the French National Centre for Scientific Research in Grenoble, France, proteins associating with TgGCN5-A were identified in co-immunoprecipitation experiment suggesting that HAT complexes exist in *Toxoplasma*.
- TgGCN5-A preferentially acetylates histone H3 only at lysine 18 *in vitro*. The selective acetylation of lysine 18 in histone H3 has not been described for any other GCN5 homologue.
- Removal of the TgGCN5-A N-terminal extension does not affect the acetylation of free histones *in vitro*.

REFERENCES

Aasland R, Gibson TJ, and Stewart AF. (1995) The PHD finger: implications for chromatin-mediated transcriptional regulation. *Trends Biochem Sci.* 20:56-9.

Achbarou A, Mercereau-Puijalon O, Autheman JM, Fortier B, Camus D, and Dubremetz JF. (1991) Characterization of microneme proteins of *Toxoplasma gondii*. *Mol Biochem Parasitol.* 47:223-33.

Adam SA and Gerace L. (1991) Cytosolic proteins that specifically bind nuclear location signals are receptors for nuclear import. *Cell.* 66:837-47.

Ahmad K and Henikoff S. (2002) Histone H3 variants specify modes of chromatin assembly. *Proc Natl Acad Sci.* 99 Suppl 4:16477-84.

Alberts B, Bray D, Lewis J, Raff M, Roberts K, and Watson JD. (1994) *Molecular Biology of the Cell.* 3rd ed. New York: Garland Publishing. pp. 335-99.

Allen PC and Fetterer RH. (2002) Recent advances in biology and immunobiology of *Eimeria* species and in diagnosis and control of infection with these coccidian parasites of poultry. *Clin Microbiol Rev.* 15:58-65.

Allfrey VG, Faulkner R, and Mirsky AE. (1964) Acetylation and methylation of histones and their possible role in the regulation of RNA synthesis. *Proc Natl Acad Sci.* 51:786-94.

Allfrey VG, Littau VC, and Mirsky, AE. (1963) On the role of histones in regulation ribonucleic acid synthesis in the cell nucleus. *Proc Natl Acad Sci.* 49:414-21.

Anderson BC. (1998) Cryptosporidiosis in bovine and human health. *J Dairy Sci.* 81:3036-41.

Ayene IS, Ford LP, and Koch CJ. (2005) Ku protein targeting by Ku70 small interfering RNA enhances human cancer cell response to topoisomerase II inhibitor and gamma radiation. *Mol Cancer Ther.* 4:529-36.

Bahl A, Brunk B, Crabtree J, Fraunholz MJ, Gajria B, Grant GR, Ginsburg H, Gupta D, Kissinger JC, Labo P, Li L, Mailman MD, Milgram AJ, Pearson DS, Roos DS, Schug J, Stoeckert CJ Jr, and Whetzel P. (2003) PlasmoDB: the *Plasmodium* genome resource. A database integrating experimental and computational data. *Nucleic Acids Res.* 31(1):212-5.

Balasubramanyam K, Swaminathan V, Ranganathan A, and Kundu TK. (2003) Small molecule modulators of histone acetyltransferase p300. *J Biol Chem.* 278:19134-40.

Balasubramanyam K, Altaf M, Varier RA, Swaminathan V, Ravindran A, Sadhale PP, and Kundu TK. (2004a) Polyisoprenylated benzophenone, garcinol, a natural histone acetyltransferase inhibitor, represses chromatin transcription and alters global gene expression. *J Biol Chem.* 279:33716-26.

Balasubramanyam K, Varier RA, Altaf M, Swaminathan V, Siddappa NB, Ranga U, and Kundu TK. (2004b) 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.* 279:51163-71.

Bansal PK, Abdulle R, and Kitagawa K. (2004) Sgt1 associates with Hsp90: an initial step of assembly of the core kinetochore complex. *Mol Cell Biol.* 24:8069-79.

Barlev NA, Candau R, Wang L, Darpino P, Silverman N, and Berger SL. (1995) Characterization of physical interactions of the putative transcriptional adaptor, ADA2, with acidic activation domains and TATA-binding protein. *J Biol Chem.* 270:19337-44.

Barlev NA, Emelyanov AV, Castagnino P, Zegerman P, Bannister AJ, Sepulveda MA, Robert F, Tora L, Kouzarides T, Birshtein BK, and Berger SL. (2003) A novel human Ada2 homologue functions with Gcn5 or Brg1 to coactivate transcription. *Mol Cell Biol.* 23:6944-57.

Barlev NA, Liu L, Chehab NH, Mansfield K, Harris KG, Halazonetis TD, and Berger SL. (2001) Acetylation of p53 activates transcription through recruitment of coactivators/histone acetyltransferases. *Mol Cell.* 8:1243-54.

Barlev NA, Poltoratsky V, Owen-Hughes T, Ying C, Liu L, Workman JL, and Berger SL. (1998) Repression of GCN5 histone acetyltransferase activity via bromodomain-mediated binding and phosphorylation by the Ku-DNA-dependent protein kinase complex. *Mol Cell Biol.* 18:1349-58.

Barnes PJ, Adcock IM, and Ito K. (2005) Histone acetylation and deacetylation: importance in inflammatory lung diseases. *Eur Respir J.* 25:552-63.

Beckers CJ, Dubremetz JF, Mercereau-Puijalon O, and Joiner KA (1994) The *Toxoplasma gondii* rhoptry protein ROP2 is inserted into the parasitophorous vacuole membrane, surrounding the intracellular parasite, and is exposed to the host cell cytoplasm. *J Cell Biol* 127:947–961.

Berger SL, Pina B, Silverman N, Marcus GA, Agapite J, Regier JL, Triezenberg SJ, and Guarente L. (1992) Genetic isolation of ADA2: a potential transcriptional adaptor required for function of certain acidic activation domains. *Cell.* 70(2):251-65.

Bhat RA, Riehl M, Santandrea G, Velasco R, Slocombe S, Donn G, Steinbiss HH, Thompson RD, and Becker HA (2003) Alteration of GCN5 levels in maize reveals dynamic responses to manipulating histone acetylation. *Plant J.* 33:455-69.

Bhatti MM and Sullivan WJ Jr. (2005) Histone acetylase GCN5 enters the nucleus via importin-alpha in protozoan parasite *Toxoplasma gondii. J Biol Chem.* 280:5902-8.

Bhatti MM, Livingston M, Mullapudi N, and Sullivan WJ Jr. (2006) Pair of unusual GCN5 histone acetyltransferases and ADA2 homologues in the protozoan parasite *Toxoplasma gondii*. *Eukaryot Cell.* 5:62-76.

Binder EM and Kim K. (2004) Location, Location, Location: Trafficking and Function of Secreted Proteases of *Toxoplasma* and *Plasmodium*. *Traffic.* 5:914–924.

Bischoff FR and Ponstingl H. (1991) Catalysis of guanine nucleotide exchange on Ran by the mitotic regulator RCC1. *Nature.* 354:80-2.

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

Blackman MJ and Bannister LH. (2001) Apical organelles of Apicomplexa: biology and isolation by subcellular fractionation. *Mol Biochem Parasitol.* 117:11-25.

Blander G and Guarente L. (2004) The Sir2 family of protein deacetylases. *Annu. Rev. Biochem.* 73:417–435.

Bohne W, Gross U, Ferguson DJ, and Heesemann J. (1995) Cloning and characterization of a bradyzoite-specifically expressed gene (hsp30/bag1) of *Toxoplasma gondii*, related to genes encoding small heat-shock proteins of plants. *Mol Microbiol.* 16:1221-30.

Borrow J, Stanton VP Jr, Andresen JM, Becher R, Behm FG, Chaganti RS, Civin CI, Disteche C, Dube I, Frischauf AM, Horsman D, Mitelman F, Volinia S, Watmore AE, and Housman DE. (1996) The translocation t(8;16)(p11;p13) of acute myeloid leukaemia fuses a putative acetyltransferase to the CREB-binding protein. *Nat Genet.* 14:33-41.

Bottomley MJ. (2004) Structures of protein domains that create or recognize histone modifications. *EMBO Rep* 5:464-9.

Boulikas T. (1994) Putative nuclear localization signals (NLS) in protein transcription factors. *J Cell Biochem.* 55:32-58.

Boyer KM. (1996) Diagnosis and treatment of congenital toxoplasmosis. *Adv Pediatr Infect Dis.* 11:449-67.

Brownell JE and Allis CD. (1995) An activity gel assay detects a single, catalytically active histone acetyltransferase subunit in *Tetrahymena* macronuclei. *Proc Natl Acad Sci.* 92:6364-8.

Brownell JE, Zhou J, Ranalli T, Kobayashi R, Edmondson DG, Roth SY, and Allis CD. (1996) *Tetrahymena* histone acetyltransferase A: a homolog to yeast Gcn5p linking histone acetylation to gene activation. *Cell.* 84:843-51.

Bustin M, Trieschmann L, and Postnikov YV. (1995) The HMG-14/-17 chromosomal protein family: architectural elements that enhance transcription from chromatin templates. *Semin Cell Biol.* 6:247-55.

Butler R and Bates GP. (2006) Histone deacetylase inhibitors as therapeutics for polyglutamine disorders. *Nat Rev Neurosci.* 7:784-96.

Byrne BM and Welsh J. (2005) Altered thioredoxin subcellular localization and redox status in MCF-7 cells following 1,25-dihydroxyvitamin D3 treatment. *J Steroid Biochem Mol Biol.* 97:57-64.

Candau R, Moore PA, Wang L, Barlev N, Ying CY, Rosen CA, and Berger SL. (1996) Identification of human proteins functionally conserved with the yeast putative adaptors ADA2 and GCN5. *Mol Cell Biol.* 16:593-602.

Candau R, Zhou JX, Allis CD, and Berger SL. (1997) Histone acetyltransferase activity and interaction with ADA2 are critical for GCN5 function *in vivo*. *EMBO J.* 16:555-65.

Carruthers VB. (2002) Host cell invasion by the opportunistic pathogen *Toxoplasma gondii*. *Acta Trop.* 81:111-22.

Carruthers VB, Moreno SN, and Sibley LD. (1999) Ethanol and acetaldehyde elevate intracellular [Ca^{2+}] and stimulate microneme discharge in *Toxoplasma gondii*. *Biochem J.* 342:379-86.

Carruthers VB and Sibley LD. (1999) Mobilization of intracellular calcium stimulates microneme discharge in *Toxoplasma gondii*. *Mol Microbiol.* 31:421-8.

Catimel B, Teh T, Fontes MR, Jennings IG, Jans DA, Howlett GJ, Nice EC, and Kobe B. (2001) Biophysical characterization of interactions involving importin-alpha during nuclear import. *J Biol Chem.* 276(36):34189-98.

Chelikani P, Fita I, and Loewen PC. (2004) Diversity of structures and properties among catalases. *Cell Mol Life Sci.* 61:192-208.

Chen T, Sun H, Lu J, Zhao Y, Tao D, Li X, and Huang B. (2002) Histone acetylation is involved in hsp70 gene transcription regulation in *Drosophila melanogaster*. *Arch Biochem Biophys*. 408:171-6.

Chook YM and Blobel G. (2001) Karyopherins and nuclear import. *Curr Opin Struct Biol.* 11:703-15.

Cleary MD, Singh U, Blader IJ, Brewer JL, and Boothroyd JC. (2002) *Toxoplasma gondii* asexual development: identification of developmentally regulated genes and distinct patterns of gene expression. *Eukaryot Cell.* 1:329-40.

Cocklin RR and Wang M. (2003) Identification of methylation and acetylation sites on mouse histone H3 using matrix-assisted laser desorption/ionization time-of-flight and nanoelectrospray ionization tandem mass spectrometry. *J. Protein Chem.* 22:327-34.

Col E, Caron C, Seigneurin-Berny D, Gracia J, Favier A, and Khochbin S. (2001) The histone acetyltransferase, hGCN5, interacts with and acetylates the HIV transactivator, Tat. *J Biol Chem.* 276:28179-84.

Cole PA. (1996) Chaperone-assisted protein expression. *Structure.* 4:239-42.

Coupry I, Roudaut C, Stef M, Delrue MA, Marche M, Burgelin I, Taine L, Cruaud C, Lacombe D, and Arveiler B. (2002) Molecular analysis of the CBP gene in 60 patients with Rubinstein-Taybi syndrome. *J Med Genet.* 39:415-21.

Creaven M, Hans F, Mutskov V, Col E, Caron C, Dimitrov S, and Khochbin S. (1999) Control of the histone-acetyltransferase activity of Tip60 by the HIV-1 transactivator protein, Tat. *Biochemistry.* 38:8826-30.

Cress WD and Seto E. (2000) Histone deacetylases, transcriptional control, and cancer. *J Cell Physiol.* 184:1-16.

Daniel JA, Torok MS, Sun ZW, Schieltz D, Allis CD, Yates JR 3rd, and Grant, PA. (2004) Deubiquitination of histone H2B by a yeast acetyltransferase complex regulates transcription. *J Biol Chem.* 279:1867-71.

Darkin-Rattray SJ, Gurnett AM, Myers RW, Dulski PM, Crumley TM, Allocco JJ, Cannova C, Meinke PT, Colletti SL, Bednarek MA, Singh SB, Goetz MA, Dombrowski AW, Polishook JD, and Schmatz DM. (1996) Apicidin: a novel antiprotozoal agent that inhibits parasite histone deacetylase. *Proc Natl Acad Sci.* 93:13143-7.

Dash A and Gilliland DG. (2001) Molecular genetics of acute myeloid leukaemia. *Best Pract Res Clin Haematol.* 14:49-64.

Daujat S, Bauer UM, Shah V, Turner B, Berger S, and Kouzarides T. (2002) Crosstalk between CARM1 methylation and CBP acetylation on histone H3. *Curr Biol.* 12:2090-7.

Davies S and Ramsden DB. (2001) Huntington's disease. *Mol Pathol.* 54:409-13.
de Ruijter AJ, van Gennip AH, Caron HN, Kemp S, and van Kuilenburg AB. (2003) Histone deacetylases (HDACs): characterization of the classical HDAC family. *Biochem J.* 370(Pt 3):737-49.
Derman AI, Prinz WA, Belin D, and Beckwith J. (1993) Mutations that allow disulfide bond formation in the cytoplasm of *Escherichia coli. Science.* 262:1744-7.
Derouin F. (1992) Pathogeny and immunological control of toxoplasmosis. *Braz J Med Biol Res.* 25:1163-9.
Dhalluin C, Carlson JE, Zeng L, He C, Aggarwal AK, and Zhou MM. (1999) Structure and ligand of a histone acetyltransferase bromodomain. *Nature.* 399:491-6.
Ding M, Kwok LY, Schluter D, Clayton C, and Soldati D. (2004) The antioxidant systems in *Toxoplasma gondii* and the role of cytosolic catalase in defence against oxidative injury. *Mol Microbiol.* 51:47-61.
Dingwall C and Laskey RA. (1991) Nuclear targeting sequences--a consensus? *Trends Biochem Sci.* 16:478-81.
Dobrowolski JM and Sibley LD. (1996) *Toxoplasma* invasion of mammalian cells is powered by the actin cytoskeleton of the parasite. *Cell.* 84:933-9.
Donald RG, Carter D, Ullman B, and 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.* 271:14010-9.
Donald RG and Liberator PA. (2002) Molecular characterization of a coccidian parasite cGMP dependent protein kinase. *Mol Biochem Parasitol.* 120:165-75.
Donald RG and 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.* 90:11703-7.
Dontfraid FF and Chakrabarti D. (1994) Cloning and expression of a cDNA encoding the homologue of Ran/TC4 GTP-binding protein from *Plasmodium falciparum*. *Biochem Biophys Res Commun.* 201:423-9.
Driessen CA, Winkens HJ, Kuhlmann LD, Janssen BP, van Vugt AH, Deutman AF, and Janssen JJ. (1997) Cloning and structural analysis of the murine *GCN5L1* gene. *Gene.* 203:27-31.
Dubey JP, Lindsay DS, and Speer CA. (1998) Structures of *Toxoplasma gondii* tachyzoites, bradyzoites, and sporozoites and biology and development of tissue cysts. *Clin Microbiol Rev.* 11:267-99.
Dubey JP, Miller NL, and Frenkel JK (1970) The *Toxoplasma gondii* oocyst from cat feces. *J Exp Med.* 132:636-62.
Dubey JP, Shen SK, Kwok OC, and Frenkel JK. (1999) Infection and immunity with the RH strain of *Toxoplasma gondii* in rats and mice. *J. Parasitol.* 85:657-62.

Dubremetz JF, Achbarou A, Bermudes D, and Joiner KA. (1993) Kinetics and pattern of organelle exocytosis during *Toxoplasma gondii* host-cell interaction. *Parasitol Res.* 79:402-8.

Dudley AM, Rougeulle C, and Winston F. (1999) The Spt components of SAGA facilitate TBP binding to a promoter at a post-activator-binding step *in vivo*. *Genes Dev.* 13:2940-5.

Dzierszinski F, Mortuaire M, Dendouga N, Popescu O, and Tomavo S. (2001) Differential expression of two plant-like enolases with distinct enzymatic and antigenic properties during stage conversion of the protozoan parasite *Toxoplasma gondii*. *J Mol Biol.* 309:1017-27.

Eberharter A, Sterner DE, Schieltz D, Hassan A, Yates JR 3rd, Berger SL, and Workman JL. (1999) The ADA complex is a distinct histone acetyltransferase complex in *Saccharomyces cerevisiae*. *Mol Cell Biol.* 19:6621-31.

Eberharter A, John S, Grant PA, Utley RT, and Workman JL. (1998) Identification and analysis of yeast nucleosomal histone acetyltransferase complexes. *Methods.* 15:315-21.

Echeverria PC, Matrajt M, Harb OS, Zappia MP, Costas MA, Roos DS, Dubremetz JF, and Angel SO. (2005) *Toxoplasma gondii* Hsp90 is a potential drug target whose expression and subcellular localization are developmentally regulated. *J Mol Biol.* 350:723-34.

Ekwall K. (2005) Genome-wide analysis of HDAC function. *Trends Genet.* 21:608-15.

Elmendorf HG, Singer SM, and Nash TE. (2000) Targeting of proteins to the nuclei of *Giardia lamblia. Mol Biochem Parasitol.* 106:315-9.

Fahrenkrog B and Aebi U. (2003) The nuclear pore complex: nucleocytoplasmic transport and beyond. *Nat Rev Mol Cell Biol.* 4:757-66.

Fan Q, An L, and Cui L. (2004a) *Plasmodium falciparum* histone acetyltransferase, a yeast GCN5 homologue involved in chromatin remodeling. *Eukaryot Cell.* 3:264-76.

Fan Q, An L, and Cui L. (2004b) PfADA2, a *Plasmodium falciparum* homologue of the transcriptional coactivator ADA2 and its *in vivo* association with the histone acetyltransferase PfGCN5. *Gene.* 336:251-61.

Ferguson DJ, Hutchison WM, and Siim JC. (1975) The ultrastructural development of the macrogamete and formation of the oocyst wall of *Toxoplasma gondii*. *Acta Pathol Microbiol Scand [B].* (5):491-505.

Ferro M, Salvi D, Brugiere S, Miras S, Kowalski S, Louwagie M, Garin J, Joyard J, and Rolland N. (2003) Proteomics of the chloroplast envelope membranes from *Arabidopsis thaliana. Mol Cell Proteomics.* 2:325-45.

Fichera ME, Bhopale MK, and Roos DS. (1995) *In vitro* assays elucidate peculiar kinetics of clindamycin action against *Toxoplasma gondii*. *Antimicrob Agents Chemother.* 39:1530-7.

Foth BJ and McFadden GI. (2003) The apicoplast: a plastid in *Plasmodium falciparum* and other Apicomplexan parasites. *Int Rev Cytol.* 224:57-110.

Fox BA and Bzik DJ. (2002) De novo pyrimidine biosynthesis is required for virulence of *Toxoplasma gondii*. *Nature.* 415:926-9.

Fox BA, Gigley JP, and Bzik DJ. (2004) *Toxoplasma gondii* lacks the enzymes required for de novo arginine biosynthesis and arginine starvation triggers cyst formation. *Int J Parasitol.* 34:323-31.

Frankel MB, Mordue DG, and Knoll LJ. (2006) A disruption in the promoter of the *Toxoplasma gondii* RCC1 protein renders parasites defective in establishing a chronic infection. Poster Presentation #219B. Molecular Parasitology Meeting, Woods Hole, MA.

Fritsch PO and Sidoroff A. (2000) Drug-induced Stevens-Johnson syndrome/toxic epidermal necrolysis. *Am J Clin Dermatol.* 1:349-60.

Gardner MJ, Bishop R, Shah T, de Villiers EP, Carlton JM, Hall N, Ren Q, Paulsen IT, Pain A, Berriman M, Wilson RJ, Sato S, Ralph SA, Mann DJ, Xiong Z, Shallom SJ, Weidman J, Jiang L, Lynn J, Weaver B, Shoaibi A, Domingo AR, Wasawo D, Crabtree J, Wortman JR, Haas B, Angiuoli SV, Creasy TH, Lu C, Suh B, Silva JC, Utterback TR, Feldblyum TV, Pertea M, Allen J, Nierman WC, Taracha EL, Salzberg SL, White OR, Fitzhugh HA, Morzaria S, Venter JC, Fraser CM, and Nene V. (2005) Genome sequence of *Theileria parva*, a bovine pathogen that transforms lymphocytes. *Science.* 309:134-7.

Germain RN and Stefanova I. (1999) The dynamics of T cell receptor signaling: complex orchestration and the key roles of tempo and cooperation. *Annu Rev Immunol.* 17:467-522.

Georgakopoulos T and Thireos G. (1992) Two distinct yeast transcriptional activators require the function of the GCN5 protein to promote normal levels of transcription. *EMBO J.* 11:4145-52.

Glotzer M. (1996) Mitosis: don't get mad, get even. *Curr Biol.* 6:1592-4.

Goldfarb DS, Corbett AH, Mason DA, Harreman MT, and Adam SA. (2004) Importin alpha: a multipurpose nuclear-transport receptor. *Trends Cell Biol.* 14:505-14.

Goodman RH and Smolik S. (2000) CBP/p300 in cell growth, transformation, and development. *Genes Dev.* 14:1553-77.

Gorlich D, Kostka S, Kraft R, Dingwall C, Laskey RA, Hartmann E, and Prehn S. (1995) Two different subunits of importin cooperate to recognize nuclear localization signals and bind them to the nuclear envelope. *Curr Biol.* 5:383-92.

Gorlich D, Prehn S, Laskey RA, and Hartmann E. (1994) Isolation of a protein that is essential for the first step of nuclear protein import. *Cell.* 79:767-78.

Graham SP, Pelle R, Honda Y, Mwangi DM, Tonukari NJ, Yamage M, Glew EJ, de Villiers EP, Shah T, Bishop R, Abuya E, Awino E, Gachanja J, Luyai AE, Mbwika F, Muthiani AM, Ndegwa DM, Njahira M, Nyanjui JK, Onono FO, Osaso J, Saya RM, Wildmann C, Fraser CM, Maudlin I, Gardner MJ, Morzaria SP, Loosmore S, Gilbert SC, Audonnet JC, van der Bruggen P, Nene V, and Taracha EL. (2006) *Theileria parva* candidate vaccine antigens recognized by immune bovine cytotoxic T lymphocytes. *Proc Natl Acad Sci.* 103:3286-91.

Grant PA. (2001) A tale of histone modifications. *Genome Biol.* 2: reviews0003.1-reviews0003.6.

Grant PA, Duggan L, Cote J, Roberts SM, Brownell JE, Candau R, Ohba R, Owen-Hughes T, Allis CD, Winston F, Berger SL, and Workman JL. (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.* 11:1640-50.

Grant PA, Eberharter A, John S, Cook RG, Turner BM, and Workman JL. (1999) Expanded lysine acetylation specificity of Gcn5 in native complexes. *J Biol Chem.* 274:5895-900.

Grant PA, Schieltz D, Pray-Grant MG, Steger DJ, Reese JC, Yates JR 3rd, and Workman JL. (1998) A subset of TAF(II)s are integral components of the SAGA complex required for nucleosome acetylation and transcriptional stimulation. *Cell.* 94:45-53.

Gregory PD, Schmid A, Zavari M, Lui L, Berger SL, and Horz W. (1998) Absence of Gcn5 HAT activity defines a novel state in the opening of chromatin at the PHO5 promoter in yeast. *Mol Cell.* 1:495-505.

Gregory PD, Schmid A, Zavari M, Munsterkotter M, and Horz W. (1999) Chromatin remodelling at the PHO8 promoter requires SWI-SNF and SAGA at a step subsequent to activator binding. *EMBO J.* 18:6407-14.

Grunstein M. (1992) Histones as regulators of genes. *Sci Am.* 267:68-74B.

Gu W and Roeder RG. (1997) Activation of p53 sequence-specific DNA binding by acetylation of the p53 C-terminal domain. *Cell.* 90:595-606.

Guerina NG, Hsu HW, Meissner HC, Maguire JH, Lynfield R, Stechenberg B, Abroms I, Pasternack MS, Hoff R, and Eaton RB. (1994) Neonatal serologic screening and early treatment for congenital *Toxoplasma gondii* infection. The New England Regional *Toxoplasma* Working Group. *N Engl J Med.* 330:1858-63.

Gustafsson C, Govindarajan S, and Minshull J. (2004) Codon bias and heterologous protein expression. *Trends Biotechnol.* 22:346-53.

Hammouda NA, Abo el-Naga I, Hussein ED, and Rashwan EA. (1995) Opsonization and intracellular killing of *Toxoplasma gondii* by human mononuclear phagocytes. *J Egypt Soc Parasitol.* 25:11-7.

Herold A, Truant R, Wiegand H, and Cullen BR. (1998) Determination of the functional domain organization of the importin alpha nuclear import factor. *J Cell Biol.* 143:309-18.

Hassan AH, Prochasson P, Neely KE, Galasinski SC, Chandy M, Carrozza MJ, and Workman JL. (2002) Function and selectivity of bromodomains in anchoring chromatin-modifying complexes to promoter nucleosomes. *Cell.* 111:369-79.

Helmlinger D, Hardy S, Sasorith S, Klein F, Robert F, Weber C, Miguet L, Potier N, Van-Dorsselaer A, Wurtz JM, Mandel JL, Tora L, and Devys D. (2004) Ataxin-7 is a subunit of GCN5 histone acetyltransferase-containing complexes. *Hum Mol Genet.* 13:1257-65.

Henikoff S, Furuyama T, and Ahmad K. (2004) Histone variants, nucleosome assembly and epigenetic inheritance. *Trends Genet.* 20:320-6.

Hermanns B, Brunn A, Schwarz ER, Sachweh JS, Seipelt I, Schroder JM, Vogel U, Schoendube FA, and Buettner R. (2001) Fulminant toxoplasmosis in a heart transplant recipient. *Pathol Res Pract.* 197:211-5.

Herrera JE, Bergel M, Yang XJ, Nakatani Y, and Bustin M. (1997) The histone acetyltransferase activity of human GCN5 and PCAF is stabilized by coenzymes. *J Biol Chem.* 272:27253-8.

Herrera JE, Sakaguchi K, Bergel M, Trieschmann L, Nakatani Y, and Bustin M. (1999) Specific acetylation of chromosomal protein HMG-17 by PCAF alters its interaction with nucleosomes. *Mol Cell Biol.* 19:3466-73.

Hershko A and Ciechanover A. (1998) The ubiquitin system. *Annu Rev Biochem.* 67:425-79.

Hettmann C and Soldati D. (1999) Cloning and analysis of a *Toxoplasma gondii* histone acetyltransferase: a novel chromatin remodelling factor in Apicomplexan parasites. *Nucleic Acids Res.* 27:4344-52.

Hilfiker A, Hilfiker-Kleiner D, Pannuti A, and Lucchesi JC. (1997) mof, a putative acetyl transferase gene related to the Tip60 and *MOZ* human genes and to the *SAS* genes of yeast, is required for dosage compensation in *Drosophila*. *EMBO J.* 16:2054-60.

Hilfiker A, Yang Y, Hayes DH, Beard CA, Manning JE, and Lucchesi JC. (1994) Dosage compensation in Drosophila: the X-chromosomal binding of MSL-1 and MLE is dependent on Sxl activity. *EMBO J.* 13:3542-50.

Hill D and Dubey JP. (2002) *Toxoplasma gondii*: transmission, diagnosis and prevention. *Clin Microbiol Infect.* 8:634-40.

Hinnebusch AG. (1985) A hierarchy of trans-acting factors modulates translation of an activator of amino acid biosynthetic genes in *Saccharomyces cerevisiae*. *Mol Cell Biol.* 5:2349-60.

Ho-Yen DO. (2001) Infection in the immunocompetent. *Toxoplasmosis: A Comprehensive Clinical Guide.* Cambridge University Press: London. pp. 125-146.

Hodel MR, Corbett AH, and Hodel AE. (2001) Dissection of a nuclear localization signal. *J Biol Chem.* 276(2):1317-25.

Hope IA and Struhl K. (1985) GCN4 protein, synthesized *in vitro*, binds *HIS3* regulatory sequences: implications for general control of amino acid biosynthetic genes in yeast. *Cell.* 43:177-88.

Hope IA and Struhl K. (1987) GCN4, a eukaryotic transcriptional activator protein, binds as a dimer to target DNA. *EMBO J.* 6:2781-4.

Howe DK and Sibley LD. (1995) *Toxoplasma gondii* comprises three clonal lineages: correlation of parasite genotype with human disease. *J Infect Dis.* 172:1561-6.

Hu K, Mann T, Striepen B, Beckers CJ, Roos DS, and Murray JM. (2002) Daughter cell assembly in the protozoan parasite *Toxoplasma gondii*. *Mol Biol Cell.* 13:593-606.

Huang L. (2006) Targeting histone deacetylases for the treatment of cancer and inflammatory diseases. *J Cell Physiol.* 209:611-6.

Huang J, Mullapudi N, Sicheritz-Ponten T, and Kissinger JC. (2004) A first glimpse into the pattern and scale of gene transfer in Apicomplexa. *Int J Parasitol.* 34:265-74.

Hubner S, Smith HM, Hu W, Chan CK, Rihs HP, Paschal BM, Raikhel NV, and Jans DA. (1999) Plant importin alpha binds nuclear localization sequences with high affinity and can mediate nuclear import independent of importin beta. *J Biol Chem.* 274:22610-7.

Huisinga KL and Pugh BF. (2004) A genome-wide housekeeping role for TFIID and a highly regulated stress-related role for SAGA in *Saccharomyces cerevisiae*. *Mol Cell.* 13:573-85.

Hunter CA and Reichmann G. (2001) Immunology of *Toxoplasma* infection *Toxoplasmosis: A Comprehensive Clinical Guide.* Cambridge University Press: London. pp. 43-57.

Imhof A, Yang XJ, Ogryzko VV, Nakatani Y, Wolffe AP, and Ge H. (1997) Acetylation of general transcription factors by histone acetyltransferases. *Curr Biol.* 7:689-92.

Ingvarsdottir K, Krogan NJ, Emre NC, Wyce A, Thompson NJ, Emili A, Hughes TR, Greenblatt JF, and Berger SL. (2005) H2B ubiquitin protease Ubp8 and Sgf11 constitute a discrete functional module within the *Saccharomyces cerevisiae* SAGA complex. *Mol Cell Biol.* 25:1162-72.

Jiang MC, Yang-Yen HF, Yen JJ, and Lin JK. (1996) Curcumin induces apoptosis in immortalized NIH 3T3 and malignant cancer cell lines. *Nutr Cancer.* 26:111-20.

Ji DD, Sultan AA, Chakrabarti D, Horrocks P, Doerig C, and Arnot DE. (1998) An RCC1-type guanidine exchange factor for the Ran G protein is found in the *Plasmodium falciparum* nucleus. *Mol Biochem Parasitol.* 95:165-70.

Jones J, Lopez A, and Wilson M. (2003) Congenital toxoplasmosis. *Am Fam Physician.* 67:2131-8.

Jones JL, Sehgal M, and Maguire JH. (2002) Toxoplasmosis-associated deaths among Human Immunodeficiency Virus-infected persons in the United States, 1992-1998. *Clin Infect Dis.* 34:1161.

Jones TC, Bienz KA, and Erb P. (1986) *In vitro* cultivation of *Toxoplasma gondii* cysts in astrocytes in the presence of gamma interferon. *Infect Immun.* 51:147-56.

Juan LJ, Shia WJ, Chen MH, Yang WM, Seto E, Lin YS, and Wu CW. (2000) Histone deacetylases specifically down-regulate p53-dependent gene activation. *J Biol Chem.* 275:20436-43.

Kalderon D, Richardson WD, Markham AF, and Smith AE. (1984a) Sequence requirements for nuclear location of simian virus 40 large-T antigen. *Nature.* 311:33-8.

Kalderon D, Roberts BL, Richardson WD, and Smith AE. (1984b) A short amino acid sequence able to specify nuclear location. *Cell.* 39:499-509.

Kamakaka RT and Biggins S. (2005) Histone variants: deviants? *Genes Dev.* 19:295-310.

Kamine J, Elangovan B, Subramanian T, Coleman D, and Chinnadurai G. (1996) Identification of a cellular protein that specifically interacts with the essential cysteine region of the HIV-1 Tat transactivator. *Virology.* 216:357-66.

Kang J, Chen J, Shi Y, Jia J, and Zhang Y. (2005) Curcumin-induced histone hypoacetylation: the role of reactive oxygen species. *Biochem Pharmacol.* 69:1205-13.

Karimpour S, Lou J, Lin LL, Rene LM, Lagunas L, Ma X, Karra S, Bradbury CM, Markovina S, Goswami PC, Spitz DR, Hirota K, Kalvakolanu DV, Yodoi J, and Gius D. (2002) Thioredoxin reductase regulates AP-1 activity as well as thioredoxin nuclear localization via active cysteines in response to ionizing radiation. *Oncogene.* 21:6317-27.

Kibe MK, Coppin A, Dendouga N, Oria G, Meurice E, Mortuaire M, Madec E, and Tomavo S. (2005) Transcriptional regulation of two stage-specifically expressed genes in the protozoan parasite *Toxoplasma gondii*. *Nucleic Acids Res.* 33:1722-36.

Kim SH, Ahn S, Han JW, Lee HW, Lee HY, Lee YW, Kim MR, Kim KW, Kim WB, and Hong S. (2004) Apicidin is a histone deacetylase inhibitor with anti-invasive and anti-angiogenic potentials. *Biochem Biophys Res Commun.* 315:964-70.

Kim DH, Kim M, and Kwon HJ. (2003) Histone deacetylase in carcinogenesis and its inhibitors as anti-cancer agents. *J Biochem Mol Biol.* 36:110-9.

Kim K and Weiss LM. (2004) *Toxoplasma gondii*: the model apicomplexan. *Int J Parasitol.* 34:423-32.

Kimmins S and MacRae TH. (2000) Maturation of steroid receptors: an example of functional cooperation among molecular chaperones and their associated proteins. *Cell Stress Chaperones.* 5:76-86.

Kimura S, Oka S, Mohri H, Mitamura K, and Shimada K. (1991) Three cases of acquired immunodeficiency syndrome complicated with toxic epidermal necrolysis. *Jpn J Med.* 30:553-8.

Kissinger JC, Gajria B, Li L, Paulsen IT, and Roos DS. (2003) ToxoDB: accessing the *Toxoplasma gondii* genome. *Nucleic Acids Res.* 31:234-6.

Kobe B. (1999) Autoinhibition by an internal nuclear localization signal revealed by the crystal structure of mammalian importin alpha. *Nat Struct Biol.* 6:388-97.

Kohler M, Speck C, Christiansen M, Bischoff FR, Prehn S, Haller H, Gorlich D, and Hartmann E. (1999) Evidence for distinct substrate specificities of importin alpha family members in nuclear protein import. *Mol Cell Biol.* 19:7782-91.

Kotani T, Zhang X, Schiltz RL, Ogryzko VV, Howard T, Swanson MJ, Vassilev A, Zhang H, Yamauchi J, Howard BH, Qin J, and Nakatani Y. (1998) TBP-associated factors in the PCAF histone acetylase complex. *Cold Spring Harb Symp Quant Biol.* 63:493-9.

Kozak M. (1991) An analysis of vertebrate mRNA sequences: intimations of translational control. *J Cell Biol.* 115:887-903.

Kuo MH, Brownell JE, Sobel RE, Ranalli TA, Cook RG, Edmondson DG, Roth SY, and Allis CD. (1996) Transcription-linked acetylation by Gcn5p of histones H3 and H4 at specific lysines. *Nature.* 383:269-72.

Kurdistani SK and Grunstein M. (2003) Histone acetylation and deacetylation in yeast. *Nat Rev Mol Cell Biol.* 4:276-84.

Kusch T, Florens L, Macdonald WH, Swanson SK, Glaser RL, Yates JR 3rd, Abmayr SM, Washburn MP, and Workman JL. (2004) Acetylation by Tip60 is required for selective histone variant exchange at DNA lesions. *Science.* 306:2084-7.

Kusch T, Guelman S, Abmayr SM, and Workman JL. (2003) Two *Drosophila* Ada2 homologues function in different multiprotein complexes. *Mol Cell Biol.* 23:3305-19.

LaCasse EC and Lefebvre YA. (1995) Nuclear localization signals overlap DNA- or RNA-binding domains in nucleic acid-binding proteins. *Nucleic Acids Res.* 23:1647-56.

Lane HC, Laughon BE, Falloon J, Kovacs, JA, Davey RT Jr, Polis MA, and Masur H. (1994) NIH conference. Recent advances in the management of AIDS-related opportunistic infections. *Ann Intern Med.* 120:945-55.

Langermans JA, Van der Hulst ME, Nibbering PH, Hiemstra PS, Fransen L, and Van Furth R. (1992) IFN-gamma-induced L-arginine-dependent toxoplasmastatic activity in murine peritoneal macrophages is mediated by endogenous tumor necrosis factor-alpha. *J Immunol.* 148:568-74.

Lau OD, Kundu TK, Soccio RE, Ait-Si-Ali S, Khalil EM, Vassilev A, Wolffe AP, Nakatani Y, Roeder RG, and Cole PA. (2000) HATs off: selective synthetic inhibitors of the histone acetyltransferases p300 and PCAF. *Mol Cell.* 5:589-95.

Lekutis C, Ferguson DJ, Grigg ME, Camps M, and Boothroyd JC. (2001) Surface antigens of *Toxoplasma gondii*: variations on a theme. *Int J Parasitol.* 31:1285-92.

Leirião P, Rodrigues CD, Albuquerque SS, and Mota MM. (2004) Survival of protozoan intracellular parasites in host cells. *EMBO Rep.* 5:1142-7.

Lewis JD, Abbott DW, and Ausio J. (2003) A haploid affair: core histone transitions during spermatogenesis. *Biochem Cell Biol.* 81:131-40.

Li L, Brunk BP, Kissinger JC, Pape D, Tang K, Cole RH, Martin J, Wylie T, Dante M, Fogarty SJ, Howe DK, Liberator P, Diaz C, Anderson J, White M, Jerome ME, Johnson EA, Radke JA, Stoeckert CJ Jr, Waterston RH, Clifton SW, Roos DS, and Sibley LD. (2003) Gene discovery in the apicomplexa as revealed by EST sequencing and assembly of a comparative gene database. *Genome Res.* 13:443-54.

Liscum E and Reed JW. (2002) Genetics of Aux/IAA and ARF action in plant growth and development. *Plant Mol Biol.* 49:387-400.

Liu L, Scolnick DM, Trievel RC, Zhang HB, Marmorstein R, Halazonetis TD, and Berger SL. (1999) p53 sites acetylated *in vitro* by PCAF and p300 are acetylated *in vivo* in response to DNA damage. *Mol Cell Biol.* 19:1202-9.

Liu J, Xiao N, and DeFranco DB. (1999) Use of digitonin-permeabilized cells in studies of steroid receptor subnuclear trafficking. *Methods.* 19:403-9.

Lodish H, Berk A, Zipusky SL, Matsudaira P, Balitmore D, and Darnell J. (2000) Section 11.4. signal-mediated transport through nuclear pore complexes. *Molecular Cell Biology.* 4th ed. New York: W.H. Freeman and Company. pp. 427-36.

Loidl P. (2001) Introduction: assembly, remodeling and modification of chromatin. *Cell Mol Life Sci.* 58:663-4.

Lopez A, Dietz VJ, Wilson M, Navin TR, and Jones JL. (2000) Preventing congenital toxoplasmosis. *MMRW.* 49:57-75.

Lucchesi JC, Kelly WG, and Panning B. (2005) Chromatin remodeling in dosage compensation. *Annu Rev Genet.* 39:615-51.

Luft BJ and Remington JS. (1988) AIDS commentary. Toxoplasmic encephalitis. *J Infect Dis.* 157:1-6.

Luft BJ, Hafner R, Korzun AH, Leport C, Antoniskis D, Bosler, EM, Bourland DD, Uttamchandani R, Fuhrer J, and Jacobson J. (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.

Lyons RE, McLeod R, and Roberts CW. (2002) *Toxoplasma gondii* tachyzoite-bradyzoite interconversion. *Trends Parasitol.* 18:198-201.

MacKenzie WR, Neil MD, Hoxie NJ, Mary MS, Proctor ME, Gradus MS, Blair KA, Peterson DE, Kasmierczak JJ, Addiss DG, Fox KR, Rose JB, and Davis JP. (1994) A massive outbreak in Milwaukee of *Cryptosporidium* infection transmitted through the public water supply. *New England J. Med.* 331:161–167.

Makrides SC. (1996) Strategies for achieving high-level expression of genes in *Escherichia coli. Microbiol Rev.* 60:512-38.

Malik HS and Henikoff S. (2003) Phylogenomics of the nucleosome. *Nat Struct Biol.* 10:882-91.

Many A and Koren G. (2006) Toxoplasmosis during pregnancy. *Can Fam Physician.* 52:29-30, 32.

Marcus GA, Horiuchi J, Silverman N, and Guarente L. (1996) *ADA5/SPT20* links the *ADA* and *SPT* genes, which are involved in yeast transcription. *Mol Cell Biol.* 16:3197-205.

Marechal E and Cesbron-Delauw MF. (2001) The apicoplast: a new member of the plastid family. *Trends Plant Sci.* 6:200-5.

Mariuz P and Steigbigel RT. (2001) *Toxoplasma* infection in HIV-infected patients. *Toxoplasmosis: A Comprehensive Clinical Guide.* Cambridge University Press: London. pp. 147-177.

Maris C, Dominguez C, and Allain FH. (2005) The RNA recognition motif, a plastic RNA-binding platform to regulate post-transcriptional gene expression. *FEBS J.* 272:2118-31.

Marmorstein R. (2001) Structure of histone acetyltransferases. *J Mol Biol.* 311:433-44.

Marmorstein R and Berger SL. (2001) Structure and function of bromodomains in chromatin-regulating complexes. *Gene.* 272:1-9.

McCabe RE. (2001) Antitoxoplasma chemotherapy. *Toxoplasmosis : A Comprehensive Clinical Guide.* Cambridge University Press: London. pp. 319-59.
McMahon SB, Van Buskirk HA, Dugan KA, Copeland TD, and Cole MD. (1998) The novel ATM-related protein TRRAP is an essential cofactor for the c-Myc and E2F oncoproteins. *Cell.* 94:363-74.
Miao J, Fan Q, Cui L, and Li J. (2006) The malaria parasite *Plasmodium falciparum* histones: organization, expression, and acetylation. *Gene.* 369:53-65.
Miller RW and Rubinstein JH. (1995) Tumors in Rubinstein-Taybi syndrome. *Am J Med Genet.* 56:112-5.
Mohmmed A, Kishore S, Dasaradhi PV, Patra K, Malhotra P, and Chauhan VS. (2003) Cloning and characterization of *Plasmodium falciparum* homologues of nuclear import factors, karyopherin alpha and karyopherin beta. *Mol Biochem Parasitol.* 127:199-203.
Mucha P. (2002) Aminoacyl-tRNA synthetases and aminoacylation of tRNA in the nucleus. *Acta Biochim Pol.* 49:1-10.
Muraoka M, Konishi M, Kikuchi-Yanoshita R, Tanaka K, Shitara N, Chong JM, Iwama T, and Miyaki M. (1996) p300 gene alterations in colorectal and gastric carcinomas. *Oncogene.* 12:1565-9.
Muratoglu S, Georgieva S, Papai G, Scheer E, Enunlu I, Komonyi O, Cserpan I, Lebedeva L, Nabirochkina E, Udvardy A, Tora L, and Boros I. (2003) Two different *Drosophila* ADA2 homologues are present in distinct GCN5 histone acetyltransferase-containing complexes. *Mol Cell Biol.* 23:306-21.
Murray HW and Cohn ZA. (1979) Macrophage oxygen-dependent antimicrobial activity. I. Susceptibility of *Toxoplasma gondii* to oxygen intermediates. *J Exp Med.* 150:938-49.
Murray HW, Nathan CF, and Cohn ZA. (1980) Macrophage oxygen-dependent antimicrobial activity. IV. Role of endogenous scavengers of oxygen intermediates. *J Exp Med.* 152:1610-24.
Nambiar A, Kandala JC, Dolan SA, Jensen JB, and Guntaka RV. (1997) Molecular cloning and characterization of a cDNA for the highly conserved HMG-like protein (Pf16) gene of *Plasmodium falciparum*. *Biochem Biophys Res Commun.* 234:101-6.
Nallani KC and Sullivan WJ Jr. (2005) Identification of proteins interacting with *Toxoplasma* SRCAP by yeast two-hybrid screening. *Parasitol Res.* 95:236-42.
Nath A and Sinai AP. (2003) Cerebral toxoplasmosis. *Curr Treat Options Neurol.* 5:3-12.
Neuwald AF, Aravind L, Spouge JL, and Koonin EV. (1999) AAA+: A class of chaperone-like ATPases associated with the assembly, operation, and disassembly of protein complexes. *Genome Res.* 9:27-43.
Ornaghi P, Rotili D, Sbardella G, Mai A, and Filetici P. (2005) A novel Gcn5p inhibitor represses cell growth, gene transcription and histone acetylation in budding yeast. *Biochem Pharmacol.* 70:911-7.

Owen DJ, Ornaghi P, Yang JC, Lowe N, Evans PR, Ballario P, Neuhaus D, Filetici P, and Travers AA. (2000) The structural basis for the recognition of acetylated histone H4 by the bromodomain of histone acetyltransferase gcn5p. *EMBO J.* 19:6141-9.

Pan MH, Chang WL, Lin-Shiau SY, Ho CT, and Lin JK. (2001) Induction of apoptosis by garcinol and curcumin through cytochrome c release and activation of caspases in human leukemia HL-60 cells. *J Agric Food Chem.* 49:1464-74.

Pante N and Kann M. (2002) Nuclear pore complex is able to transport macromolecules with diameters of about 39 nm. *Mol Biol Cell.* 13:425-34.

Pattanapanyasat K and Thakar MR. (2005) $CD4^+$ T cell count as a tool to monitor HIV progression & anti-retroviral therapy. *Indian J Med Res.*121:539-49.

Petersen E and Dubey JP. (2001) Biology of toxoplasmosis. *Toxoplasmosis: A Comprehensive Clinical Guide.* Cambrdige University Press: London. pp. 1-42.

Penn MD, Galgoci B, and Greer H. (1983) Identification of AAS genes and their regulatory role in general control of amino acid biosynthesis in yeast. *Proc Natl Acad Sci.* 80:2704-8.

Pfefferkorn ER. (1984) Interferon gamma blocks the growth of *Toxoplasma gondii* in human fibroblasts by inducing the host cells to degrade tryptophan. *Proc Natl Acad Sci.* 81:908-12.

Pfefferkorn ER, Bzik DJ, and Honsinger CP. (2001) *Toxoplasma gondii*: mechanism of the parasitostatic action of 6-thioxanthine. *Exp Parasitol.* 99:235-43.

Phillips DM. (1963) The presence of acetyl groups of histones. *Biochem J.* 87:258-63.

Pina B, Berger S, Marcus GA, Silverman N, Agapite J, and Guarente L. (1993) *ADA3*: a gene, identified by resistance to GAL4-VP16, with properties similar to and different from those of *ADA2*. *Mol Cell Biol.* 13:5981-9.

Powell DW, Weaver CM, Jennings JL, McAfee KJ, He Y, Weil PA, and Link AJ. (2004) Cluster analysis of mass spectrometry data reveals a novel component of SAGA. *Mol Cell Biol.* 24:7249-59.

Pray-Grant MG, Schieltz D, McMahon SJ, Wood JM, Kennedy EL, Cook RG, Workman JL, Yates JR 3rd, and Grant PA. (2002) The novel SLIK histone acetyltransferase complex functions in the yeast retrograde response pathway. *Mol Cell Biol.* 22:8774-86.

Puiu D, Enomoto S, Buck GA, Abrahamsen MS, and Kissinger JC. (2004) CryptoDB: the *Cryptosporidium* genome resource. *Nucleic Acids Res.* 32:D329-31.

Pusarla RH and Bhargava P. (2005) Histones in functional diversification. Core histone variants. *FEBS J.* 272:5149-68.

Quimby BB and Corbett AH. (2001) Nuclear transport mechanisms. *Cell Mol Life Sci.* 58:1766-73.

Quivy V and Van Lint C. (2002) Diversity of acetylation targets and roles in transcriptional regulation: the human immunodeficiency virus type 1 promoter as a model system. *Biochem Pharmacol.* 64:925-34.

Reifsnyder C, Lowell J, Clarke A, and Pillus L. (1996) Yeast *SAS* silencing genes and human genes associated with AML and HIV-1 Tat interactions are homologous with acetyltransferases. *Nat Genet.* 14:42-9.

Reiter-Owona I, Seitz H, Gross U, Sahm M, Rockstroh JK, and Seitz HM. (2000) Is stage conversion the initiating event for reactivation of *Toxoplasma gondii* in brain tissue of AIDS patients? *J Parasitol.* 86:531-6.

Renold C, Sugar A, Chave JP, Perrin L, Delavelle J, Pizzolato G, Burkhard P, Gabriel V, and Hirschel B. (1992) *Toxoplasma* encephalitis in patients with the acquired immunodeficiency syndrome. *Medicine (Baltimore).* 71:224-39.

Robinson KA and Lopes JM. (2000) SURVEY AND SUMMARY: *Saccharomyces cerevisiae* basic helix-loop-helix proteins regulate diverse biological processes. *Nucleic Acids Res.* 28:1499-505.

Roos DS, Crawford MJ, Donald RG, Fohl LM, Hager KM, Kissinger JC, Reynolds MG, Striepen B, and Sullivan WJ Jr. (1999) Transport and trafficking: *Toxoplasma* as a model for *Plasmodium*. *Novartis Found Symp.* 226:176-95.

Roth SY, Denu JM, and Allis CD. (2001) Histone acetyltransferases. *Annu Rev Biochem.* 70:81-120.

Ruiz-Garcia AB, Sendra R, Pamblanco M, and Tordera V. (1997) Gcn5p is involved in the acetylation of histone H3 in nucleosomes. *FEBS Lett.* 403:186-90.

Rutkowska SM and Skowron PM. (1999) Rapid screening of elution conditions prior to immunopurification of proteins. *Biotechniques.* 27:930-2, 934.

Sacktor N. (2002) The epidemiology of human immunodeficiency virus-associated neurological disease in the era of highly active antiretroviral therapy. *J Neurovirol.* Suppl 2:115-21.

Sadri-Vakili G and Cha JH. (2006) Histone deacetylase inhibitors: a novel therapeutic approach to Huntington's disease (complex mechanism of neuronal death). *Curr Alzheimer Res.* 3:403-8.

Saksouk N, Bhatti MM, Kieffer S, Smith AT, Musset K, Garin J, Sullivan WJ Jr, Cesbron-Delauw MF, and Hakimi MA. (2005) Histone-modifying complexes regulate gene expression pertinent to the differentiation of the protozoan parasite *Toxoplasma gondii. Mol Cell Biol.* 25:10301-14.

Saleh A, Schieltz D, Ting N, McMahon SB, Litchfield DW, Yates JR 3rd, Lees-Miller SP, Cole MD, and Brandl CJ. (1998) Tra1p is a component of the yeast Ada-Spt transcriptional regulatory complexes. *J Biol Chem.* 273:26559-65.

Samuel R, Bettiker RL, and Suh B (2002) AIDS related opportunistic infections, going but not gone. *Arch Pharm Res.* 25(3):215-28.

Santos-Rosa H and Caldas C. (2005) Chromatin modifier enzymes, the histone code and cancer. *Eur J Cancer.* 41:2381-402.

Santos-Rosa H, Valls E, Kouzarides T, and Martinez-Balbas M. (2003) Mechanisms of P/CAF auto-acetylation. *Nucleic Acids Res.* 31:4285-92.

Schiltz RL, Mizzen CA, Vassilev A, Cook RG, Allis CD, and Nakatani Y. (1999) Overlapping but distinct patterns of histone acetylation by the human coactivators p300 and PCAF within nucleosomal substrates. *J. Biol. Chem.* 274:1189-1192.
Schimmel P and Wang CC. (1999) Getting tRNA synthetases into the nucleus. *Trends Biochem Sci.* 24:127-8.
Schlesinger MJ. (1990) Heat shock proteins. *J Biol Chem.* 265:12111-4.
Schmidt M and Hoffman DR. (2002) Expression systems for production of recombinant allergens. *Int Arch Allergy Immunol.* 128:264-70.
Schmidt-Westhausen A, Grunewald T, Reichart PA, and Pohle HD. (1998) Oral manifestations of toxic epidermal necrolysis (TEN) in patients with AIDS: report of five cases. *Oral Dis.* 4:90-4.
Schroeder SC and Weil PA. (1998) Biochemical and genetic characterization of the dominant positive element driving transcription of the yeast TBP-encoding gene, *SPT15*. *Nucleic Acids Res.* 26:4186-95.
Seeber F and Boothroyd JC. (1996) *Escherichia coli* beta-galactosidase as an *in vitro* and *in vivo* reporter enzyme and stable transfection marker in the intracellular protozoan parasite *Toxoplasma gondii*. *Gene.* 169:39-45.
Seigneurin-Berny D, Verdel A, Curtet S, Lemercier C, Garin J, Rousseaux S, and Khochbin S. (2001) Identification of components of the murine histone deacetylase 6 complex: link between acetylation and ubiquitination signaling pathways. *Mol Cell Biol.* 21:8035-44.
Shen X, Yu L, Weir JW, and Gorovsky MA. (1995) Linker histones are not essential and affect chromatin condensation *in vivo*. *Cell.* 82:47-56.
Sheridan PL, Mayall TP, Verdin E, and Jones KA. (1997) Histone acetyltransferases regulate HIV-1 enhancer activity *in vitro*. *Genes Dev.* 11:3327-40.
Shirano Y and Shibata D. (1990) Low temperature cultivation of *Escherichia coli* carrying a rice lipoxygenase L-2 cDNA produces a soluble and active enzyme at a high level. *FEBS Lett.* 271:128-30.
Sibley LD, Adams LB, Fukutomi Y, and Krahenbuhl JL. (1991) Tumor necrosis factor-alpha triggers antitoxoplasmal activity of IFN-gamma primed macrophages. *J Immunol.* 147:2340-5.
Sibley LD and Krahenbuhl JL. (1988) Modification of host cell phagosomes by *Toxoplasma gondii* involves redistribution of surface proteins and secretion of a 32 kDa protein. *Eur J Cell Biol.* 47:81–87.
Sibley LD, Mordue DG, Su C, Robben PM, and Howe DK. (2002) Genetic approaches to studying virulence and pathogenesis in *Toxoplasma gondii. Philos Trans R Soc Lond B Biol Sci.* 357:81-8.
Sibley LD, Weidner E, and Krahenbuhl JL. (1985) Phagosome acidification blocked by intracellular *Toxoplasma gondii. Nature.* 315:416–419.
Silva NM, Gazzinelli RT, Silva DA, Ferro EA, Kasper LH, and Mineo JR. (1998) Expression of *Toxoplasma gondii*-specific heat shock protein 70 during *in vivo* conversion of bradyzoites to tachyzoites. *Infect Immun.* 66:3959-63.
Silverman SJ and Fink GR. (1984) Effects of Ty insertions on *HIS4* transcription in *Saccharomyces cerevisiae*. *Mol Cell Biol.* 4:1246-51.

Singh SB, Zink DL, Liesch JM, Mosley RT, Dombrowski AW, Bills GF, Darkin-Rattray SJ, Schmatz DM, and Goetz MA. (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.* 67:815-25.

Singh U, Brewer JL, and Boothroyd JC. (2002) Genetic analysis of tachyzoite to bradyzoite differentiation mutants in *Toxoplasma gondii* reveals a hierarchy of gene induction. *Mol Microbiol.* 44:721-33.

Smith AT, Tucker-Samaras SD, Fairlamb AH, and Sullivan WJ Jr. (2005) MYST family histone acetyltransferases in the protozoan parasite *Toxoplasma gondii*. *Eukaryot Cell*. 4:2057-65.

Smith CL, Debouck C, Rosenberg M, and Culp JS. (1989) Phosphorylation of serine residue 89 of human adenovirus E1A proteins is responsible for their characteristic electrophoretic mobility shifts, and its mutation affects biological function. *J Virol.* 63:1569-77.

Smith ER, Belote JM, Schiltz RL, Yang XJ, Moore PA, Berger SL, Nakatani Y, and Allis CD. (1998a) Cloning of *Drosophila* GCN5: conserved features among metazoan GCN5 family members. *Nucleic Acids Res.* 26:2948-54.

Smith ER, Eisen A, Gu W, Sattah M, Pannuti A, Zhou J, Cook RG, Lucchesi JC, and Allis CD. (1998b) *ESA1* is a histone acetyltransferase that is essential for growth in yeast. *Proc Natl Acad Sci.* 95:3561-5.

Smith TF, Gaitatzes C, Saxena K, and Neer EJ. (1999) The WD repeat: a common architecture for diverse functions. *Trends Biochem Sci.* 24:181-5.

Snow RW, Craig MH, Deichmann U, and le Sueur D. (1999) A preliminary continental risk map for malaria mortality among African children. *Parasitol Today.* 15:99-104.

Soête M, Camus D, and Dubremetz JF. (1994) Experimental induction of bradyzoite-specific antigen expression and cyst formation by the RH strain of *Toxoplasma gondii in vitro*. *Exp Parasitol.* 78:361-70.

Soldati D and Boothroyd JC. (1993) Transient transfection and expression in the obligate intracellular parasite *Toxoplasma gondii*. *Science.* 260:349-52.

Sogin ML, Gunderson JH, Elwood HJ, Alonso RA, and Peattie DA. (1989) Phylogenetic meaning of the kingdom concept: an unusual ribosomal RNA from *Giardia lamblia*. *Science.* 243:75-7.

Speer CA, Clark S, and Dubey JP. (1998) Ultrastructure of the oocysts, sporocysts, and sporozoites of *Toxoplasma gondii*. *J Parasitol.* 84:505-12.

Speirs, GE, Hakim M, Calne, RY, and Wreghitt TG. (1988) Relative risk of donor-transmitted *Toxoplasma gondii* infection in heart, liver and kidney transplant recipients. *Clinical Tranlantation.* 2:257-60.

Steger DJ, Eberharter A, John S, Grant PA, and Workman JL. (1998) Purified histone acetyltransferase complexes stimulate HIV-1 transcription from preassembled nucleosomal arrays. *Proc Natl Acad Sci.* 95:12924-9.

Steidl S, Tuncher A, Goda H, Guder C, Papadopoulou N, Kobayashi T, Tsukagoshi N, Kato M, and Brakhage AA. (2004) A single subunit of a heterotrimeric CCAAT-binding complex carries a nuclear localization signal: piggy back transport of the pre-assembled complex to the nucleus. *J Mol Biol.* 342:515-24.

Sterner DE, Belotserkovskaya R, and Berger SL. (2002) SALSA, a variant of yeast SAGA, contains truncated Spt7, which correlates with activated transcription. *Proc Natl Acad Sci.* 99:11622-7.

Sterner DE and Berger SL. (2000) Acetylation of histones and transcription-related factors. *Microbiol Mol Biol Rev.* 64:435-59.

Sternglanz R and Schindelin H. (1999) Structure and mechanism of action of the histone acetyltransferase Gcn5 and similarity to other N-acetyltransferases. *Proc Natl Acad Sci.* 96:8807-8.

Stewart M and Rhodes D. (1999) Switching affinities in nuclear trafficking. *Nat Struct Biol.* 6:301-4.

Stockinger EJ, Mao Y, Regier MK, Triezenberg SJ, and 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.* 29:1524-33.

Stratagene. (2002) BL21-CodonPlus® Competent Cells. INSTRUCTION MANUAL Revision #042005b. www.stratagene.com.

Strahl BD and Allis CD. (2000) The language of covalent histone modifications. *Nature.* 403:41-5.

Striepen B, He CY, Matrajt M, Soldati D, and Roos DS. (1998) Expression, selection, and organellar targeting of the green fluorescent protein in *Toxoplasma gondii. Mol Biochem Parasitol.* 92:325-38.

Su C, Howe DK, Dubey JP, Ajioka JW, and Sibley LD. (2002) Identification of quantitative trait loci controlling acute virulence in *Toxoplasma gondii. Proc Natl Acad Sci.* 99:10753-8.

Sullivan, W.J. Jr. (2003) Histone H3 and H3.3 variants in the protozoan pathogens *Plasmodium falciparum* and *Toxoplasma gondii*. *DNA Seq.* 14:227-31.

Sullivan WJ Jr, Chiang CW, Wilson CM, Naguib FN, el Kouni MH, Donald RG, and Roos DS. (1999) Insertional tagging of at least two loci associated with resistance to adenine arabinoside in *Toxoplasma gondii*, and cloning of the adenosine kinase locus. *Mol Biochem Parasitol.* 103:1-14.

Sullivan WJ Jr and Hakimi MA. (2006) Histone mediated gene activation in *Toxoplasma gondii. Mol Biochem Parasitol.* 148:109-16.

Sullivan WJ Jr, Naguleswaran A, and Angel SO. (2006) Histones and histone modifications in protozoan parasites. *Cell Microbiol.* 8:1850-61.

Sullivan WJ Jr and Smith CK II. (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.

Taipale M and Akhtar A. (2005) Chromatin mechanisms in *Drosophila* dosage compensation. *Prog Mol Subcell Biol.* 38:123-49.
Tavaria M, Gabriele T, Kola I, and Anderson RL. (1996) A hitchhiker's guide to the human Hsp70 family. *Cell Stress Chaperones.* 1:23-8.
Teh T, Tiganis T, and Kobe B. (1999) Crystallization of importin alpha, the nuclear-import receptor. *Acta Crystallogr D Biol Crystallogr.* 55:561-3.
Templeton TJ, Iyer LM, Anantharaman V, Enomoto S, Abrahante JE, Subramanian GM, Hoffman SL, Abrahamsen MS, and Aravind L. (2004) Comparative analysis of apicomplexa and genomic diversity in eukaryotes. *Genome Res.* 14:1686-95.
Terpe K. (2003) Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems. *Appl Microbiol Biotechnol.* 60:523-33.
Timmermann S, Lehrmann H, Polesskaya A, and Harel-Bellan A. (2001) Histone acetylation and disease. *Cell Mol Life Sci.* 58:728-36.
Titz B, Thomas S, Rajagopala SV, Chiba T, Ito T, and Uetz P. (2006) Transcriptional activators in yeast. *Nucleic Acids Res.* 34:955-67.
Toursel C, Dzierszinski F, Bernigaud A, Mortuaire M, and Tomavo S. (2000) Molecular cloning, organellar targeting and developmental expression of mitochondrial chaperone HSP60 in *Toxoplasma gondii. Mol Biochem Parasitol.* 111:319-32.
Tracy JW and Webster LT Jr. (1996) Drugs used in the chemotherapy of protozoal infections. *Goodman and Gillman.* 9th ed. New York: McGraw-Hill. pp. 965-85.
Tsuji N and Kobayashi M. (1978) Trichostatin C, a glucopyranosyl hydroxamate. *J. Antibiot (Tokyo).* 31:939-44.
Tucker-Samaras S. (2003). Histone Acetyltransferase as an Apicomplexan Drug Target. Thesis Disertation.
Turner BM. (2000) Histone acetylation and an epigenetic code. *Bioessays.* 22:836-45.
Utley RT and Cote J. (2003) The MYST family of histone acetyltransferases. *Curr Top Microbiol Immunol.* 274:203-36.
Uhlmann F. (2003) Separase regulation during mitosis. *Biochem Soc Symp.* 70:243-51.
Vigushin DM and Coombes RC. (2004) Targeted histone deacetylase inhibition for cancer therapy. *Curr Cancer Drug Targets.* 4:205-18.
Vlachonasios KE, Thomashow MF, and Triezenberg SJ. (2003) Disruption mutations of ADA2b and GCN5 transcriptional adaptor genes dramatically affect *Arabidopsis* growth, development, and gene expression. *Plant Cell.* 15:626-38.
Wagner FM, Reichenspurner H, Uberfuhr P, Weiss M, Fingerle V, and Reichart B. (1994) Toxoplasmosis after heart transplantation: diagnosis by endomyocardial biopsy. *J Heart Lung Transplant.* 13:916-8.
Waller RF and McFadden GI. (2005) The apicoplast: a review of the derived plastid of apicomplexan parasites. *Curr Issues Mol Biol.* 7:57-79.

Wang Q, Song C, and Li CC. (2004) Molecular perspectives on p97-VCP: progress in understanding its structure and diverse biological functions. *J Struct Biol*. 146:44-57.
Weis K. (2003) Regulating access to the genome: nucleocytoplasmic transport throughout the cell cycle. *Cell.* 12:441-51.
Weiss LM and Kim K. (2000) The development and biology of bradyzoites of *Toxoplasma gondii. Front Biosci.* 5:D391-405.
Weiss LM, Ma YF, Takvorian PM, Tanowitz HB, and Wittner M. (1998) Bradyzoite development in *Toxoplasma gondii* and the hsp70 stress response. *Infect Immun.* 66:3295-302.
White DA, Belyaev ND, and Turner BM. (1999) Preparation of site-specific antibodies to acetylated histones. *Methods.* 19:417-24.
Winstanley P, Khoo S, Szwandt S, Edwards G, Wilkins E, Tija J, Coker R, McKane W, Beeching N, and Watkin S. (1995) Marked variation in pyrimethamine disposition in AIDS patients treated for cerebral toxoplasmosis. *J Antimicrob Chemother.* 36:435-9.
Wreghitt TG, Hakim M, Gray JJ, Balfour AH, Stovin PG, Stewart S, Scott J, English TA, and Wallwork J. (1989) Toxoplasmosis in heart and heart and lung transplant recipients. *J Clin Pathol.* 42:194-9.
Wreghitt TG and Joynson DHM. (2001) *Toxoplasma* infection in immunosuppressed (HIV-negative) patients. *Toxoplasmosis: A Comprehensive Clinical Guide.* Cambridge University Press: London. pp. 178-92.
Wong SY and Remington JS. (1993) Biology of *Toxoplasma gondii*. *AIDS.* 7:299-316.
Xu W, Edmondson DG, Evrard YA, Wakamiya M, Behringer RR, and Roth SY. (2000) Loss of *Gcn5l2* leads to increased apoptosis and mesodermal defects during mouse development. *Nat Genet.* 26:229-32.
Xu W, Edmondson DG, and Roth SY. (1998) Mammalian GCN5 and P/CAF acetyltransferases have homologous amino-terminal domains important for recognition of nucleosomal substrates. *Mol Cell Biol.* 18:5659-69.
Yamada M, Yamada M, Kiuchi Y, Nara K, Kanda Y, Morinobu S, Momose K, Oguchi K, Kamijima K, and Higuchi T. (1999) Identification of a novel splice variant of heat shock cognate protein 70 after chronic antidepressant treatment in rat frontal cortex. *Biochem Biophys Res Commun.* 261:541-5.
Yamauchi T, Yamauchi J, Kuwata T, Tamura T, Yamashita T, Bae N, Westphal H, Ozato K, and Nakatani Y. (2000) Distinct but overlapping roles of histone acetylase PCAF and of the closely related PCAF-B/GCN5 in mouse embryogenesis. *Proc Natl Acad Sci.* 97:11303-6.
Yang XJ. (2004) The diverse superfamily of lysine acetyltransferases and their roles in leukemia and other diseases. *Nucleic Acids Res.* 32:959-76.
Yang S and Parmley SF. (1997) *Toxoplasma gondii* expresses two distinct lactate dehydrogenase homologous genes during its life cycle in intermediate hosts. *Gene.* 184:1-12.

Yang XJ, Ogryzko VV, Nishikawa J, Howard BH, and Nakatani Y. (1996) A p300/CBP-associated factor that competes with the adenoviral oncoprotein E1A. *Nature.* 382:319-24.

Yolken RH, Bachmann S, Ruslanova I, Lillehoj E, Ford G, Torrey EF, and Schroeder J. (2001) Antibodies to *Toxoplasma gondii* in individuals with first-episode schizophrenia. *Clin Infect Dis.* 32:842-4.

Yoshida M, Kijima M, Akita M, and Beppu T. (1990) Potent and specific inhibition of mammalian histone deacetylase both *in vivo* and *in vitro* by trichostatin A. *J Biol Chem.* 265:17174-9.

Yoshida M, Nomura S, and Beppu T. (1987) Effects of trichostatins on differentiation of murine erythroleukemia cells. *Cancer Res.* 47:3688-91.

Zaidi SK, Young DW, Choi JY, Pratap J, Javed A, Montecino M, Stein JL, Lian JB, van Wijnen AJ, and Stein GS. (2004) Intranuclear trafficking: organization and assembly of regulatory machinery for combinatorial biological control. *J Biol Chem.* 279:43363-6.

Zhang W, Bone JR, Edmondson DG, Turner BM, and Roth SY. (1998) Essential and redundant functions of histone acetylation revealed by mutation of target lysines and loss of the Gcn5p acetyltransferase. *EMBO.* 17:3155-67.

Zhang DE and Nelson DA. (1988) Histone acetylation in chicken erythrocytes. Rates of deacetylation in immature and mature red blood cells. *Biochem J.* 250:241-5.

Zhou XW, Kafsack BF, Cole RN, Beckett P, Shen RF, and Carruthers VB. (2005) The opportunistic pathogen *Toxoplasma gondii* deploys a diverse legion of invasion and survival proteins. *J Biol Chem.* 280:34233-44.

CURRICULUM VITAE

MICAH M. BHATTI

EDUCATION

Doctorate of Philosophy
August 2002 - May 2007
Department of Pharmacology and Toxicology
Indiana University, Indianapolis, IN

Doctorate of Medicine
August 2000 - May 2007
Indiana University School of Medicine, Indianapolis, IN

Bachelors of Science
August 1996 - May 2000
Department of Chemistry
Butler University, Indianapolis, IN

HONORS AND AWARDS

K.K. Chen Fellowship
December 2005
Department of Pharmacology and Toxicology
Indiana University School of Medicine, Indianapolis, IN
Awarded in recognition of outstanding scholarship, innovative research, and exemplary dedication to the spirit of scientific investigation

First Place in Poster Competition
May 2004
Chicago Area Mycology & Parasitology Club Meeting
Northwestern University, Chicago, IL

Judy Boyd-White Award (second place)
April 2004
Sigma Xi Research Competition
Indiana University School of Medicine, Indianapolis, IN

William H. and Fern L. Groves Hardiman Research Scholarship
October 2001
Program in Academic Medicine Presentation Competition
Indiana University School of Medicine, Indianapolis, IN

Butler University Undergraduate Scholar Research Fellowship
June 1999
Butler University, Indianapolis, IN

MEMBERSHIP IN HONORARY/PROFESSIONAL SOCIETIES

Associate Membership in Sigma Xi
Member of the American Medical Association

EMPLOYMENT/TEACHING ACTIVITIES

Instructor for Summer Organic Chemistry Lab Section (CH351)
June 2000 - August 2000
Department of Chemistry
Butler University, Indianapolis, IN

Teaching Assistant for Summer Organic Chemistry Lab Section (CH351)
June 1999 - August 1999
Department of Chemistry
Butler University, Indianapolis, IN

Teaching Assistant for General Chemistry Lab Section (CH105 & CH106)
August 1997 - May 2000
Department of Chemistry
Butler University, Indianapolis, IN

Residential Assistant for Ross Hall Freshman Dormitory
August 1997 - May 2000
Department of Residence Life
Butler University, Indianapolis, IN

Gallery Interpreter
August 1996 - May 2000
Department of Special Exhibits
The Children's Museum of Indianapolis

RESEARCH EXPERIENCE

Doctoral Thesis Project
August 2002 - June 2005
Functions of the Unique N-terminus of a GCN5 Histone Acetylase in *Toxoplasma gondii*
Mentor: William Sullivan Jr., PhD
Department of Pharmacology & Toxicology
Indiana University School of Medicine, Indianapolis, IN

Research Rotation
July 2001 - August 2001
Worked on generating novel viral genomes for packaging in a lentiviral vector for gene therapy
Mentor: Kenneth Cornetta, MD
Department of Medical & Molecular Genetics
Indiana University School of Medicine, Indianapolis, IN

Summer Research Program in Academic Medicine
May 2001 - June 2001
Investigated chromatin remodeling proteins in the protozoan pathogen *Toxoplasma gondii*
Mentor: William Sullivan Jr., PhD
Department of Pharmacology & Toxicology
Indiana University School of Medicine, Indianapolis, IN

Research Rotation
May 2000 - August 2000
Investigated anti-sense RNA strategies to “knock-down” MGMT expression in cancer cell lines
Mentor: Leonard Erickson, PhD
Department of Pharmacology & Toxicology
Indiana University School of Medicine, Indianapolis, IN

Undergraduate Honors Thesis Project
May 1995 - December 1996
Novel Synthesis of Cryptands - Investigated the use of reductive amination as a novel synthetic route for cryptands, a class of polycyclic compounds
Mentors: Jim Ciszewski, MS and Shannon Lieb, PhD
Department of Chemistry
Butler University, Indianapolis, IN

EXTRACURRICULAR GROUPS AND ACTIVITIES

Member of Gold Humanism Honor Society Task Force
August 2004 - April 2005
Indiana University School of Medicine, Indianapolis, IN

Co-Chair of Pharmacology and Toxicology Graduate Student Organization
August 2004 - June 2005
Department of Pharmacology and Toxicology
Indiana University School of Medicine, Indianapolis, IN

Member of MD/PhD Combined Degree Student Advisory Committee
March 2004 - June 2005
Indiana University School of Medicine, Indianapolis, IN

Co-Chair for Evening of the Arts
August 2003 - May 2007
Indiana University School of Medicine, Indianapolis, IN

PUBLICATIONS

Bhatti MM, Livingston M, Mullapudi N, and Sullivan WJ Jr. (2006) Pair of unusual GCN5 histone acetyltransferases and ADA2 homologues in the protozoan parasite *Toxoplasma gondii. Eukaryotic Cell*. 5:62-76.

Saksouk N, **Bhatti MM**, Kieffer S, Smith AT, Musset K, Garin J, Sullivan WJ Jr, Cesbron-Delauw MF, and Hakimi MA. (2005) Histone-modifying complexes regulate gene expression pertinent to the differentiation of the protozoan parasite *Toxoplasma gondii*. *Molecular and Cellular Biology*. 25:10301-14.

Bhatti MM and Sullivan WJ Jr. (2005) Histone acetylase GCN5 enters the nucleus via importin-alpha in protozoan parasite *Toxoplasma gondii*. *The Journal of Biological Chemistry.* 280:5902-8.

Sullivan WJ Jr, Narasimhan J, **Bhatti MM**, and Wek RC. (2004) Parasite-specific eIF2 (eukaryotic initiation factor-2) kinase required for stress-induced translation control. *The Biochemical Journal*. 380:523-31.

Carlson MW, Ciszewski JT, **Bhatti MM**, Swanson WF, and Wilson AW. (2000) A Simple Secondary Amine Synthesis: Reductive Amination using Sodium Triacetoxyborohydride. *Journal of Chemical Education.* 77:270.

SELECTED PRESENTATIONS

Oral Presentations

Bhatti MM and Sullivan WJ Jr
April 2005
Nuclear Trafficking of a Histone Acetylase in *Toxoplasma gondii*
Chicago Area Mycology & Parasitology Club Meeting
Northwestern University, Chicago, IL

Bhatti MM and Sullivan WJ Jr
April 2004
Nuclear trafficking of a histone acetylase in *Toxoplasma gondii*
Sigma Xi research competition
Indiana University School of Medicine, Indianapolis, IN

Bhatti MM, Lou Q, and Sullivan WJ Jr
September 2003
A tale of two GCN5s: Deciphering the roles of a pair of novel histone acetyltransferases in *Toxoplasma gondii*
Molecular Parasitology Meeting
Woods Hole, MA

Posters

National/International Meetings

Bhatti MM and Sullivan WJ Jr
May 2005
Nuclear Trafficking of a Histone Acetylase in *Toxoplasma gondii*
8th International Congress on Toxoplasmosis
Corsica, France

Bhatti MM and Sullivan WJ Jr
September 2004
Toxoplasma gondii histone acetyltransferases: different HATs for different occasions
Molecular Parasitology Meeting
Woods Hole, MA

Local/Regional Meetings

Bhatti MM and Sullivan WJ Jr
May 2004
Nuclear Trafficking of a Histone Acetylase in *Toxoplasma gondii*
Chicago Area Mycology & Parasitology Club Meeting
Northwestern University, Chicago, IL

Bhatti MM and Sullivan WJ Jr
March 2004
Nuclear Trafficking of a Histone Acetylase in *Toxoplasma gondii*
New & Re-Emerging Infectious Diseases Conference
University of Illinois, Urbana, IL

Bhatti MM and Sullivan WJ Jr
March 2003
Characterization of TgGCN5: A novel histone acetyltransferase in the protozoan parasite *Toxoplasma gondii*
New & Re-Emerging Infectious Diseases Conference
University of Illinois, Urbana, IL

Bhatti MM, Nallani KC, and Sullivan WJ Jr
August 2002
A novel nuclear localization signal in AIDS opportunist *Toxoplasma gondii*
Midwest Microbial Pathogenesis Meeting
Indiana University, Indianapolis, IN

GRANTS

GSO Educational Enhancement Grants and Travel Fellowship
May 2004
Indiana University School of Medicine, Indianapolis, IN

IUSM Travel Grant
May 2004
Indiana University School of Medicine, Indianapolis, IN

GSO Educational Enhancement Grants and Travel Fellowship
May 2003
Indiana University School of Medicine, Indianapolis, IN