Factor Fiction? Identifying a Putative Toxoplasma gondii Transcriptional Complex

Julia M. Paquette
University of New Hampshire, Durham

Follow this and additional works at: https://scholars.unh.edu/honors

Part of the Molecular Biology Commons

Recommended Citation
https://scholars.unh.edu/honors/822

This Senior Honors Thesis is brought to you for free and open access by the Student Scholarship at University of New Hampshire Scholars' Repository. It has been accepted for inclusion in Honors Theses and Capstones by an authorized administrator of University of New Hampshire Scholars' Repository. For more information, please contact Scholarly.Communication@unh.edu.
Factor Fiction? Identifying a Putative *Toxoplasma gondii* Transcriptional Complex

Julia Paquette

Senior Honors Thesis

Advisors: Vicki Jeffers, PhD and Krista Fleck, PhD

College of Life Sciences and Agriculture

Department of Molecular, Cellular, and Biomedical Sciences, University of New Hampshire

Spring 2021
Abstract:

*Toxoplasma gondii* is a highly prevalent protozoan parasite that is estimated to infect 30-50% of the global population, though there is no treatment for chronic infection and current treatments for acute infection may have serious side effects. Transcription is a tightly regulated process in *T. gondii*, allowing the parasite to successfully invade and replicate within host cells, and it is thus a promising avenue to study gene regulation and to investigate possible novel therapeutics. In our lab’s previous research, a TFIID-like complex was identified in *T. gondii* and found to be associated with the parasite specific bromodomain protein BDP3. In all eukaryotes, the TFIID complex is composed of TAF proteins and is a highly conserved general transcription factor that has yet to be well characterized *T. gondii*. By comparing protein sequences and domain composition between predicted *T. gondii* TAF proteins and other eukaryotic organisms, I identified two proteins that are most likely to be core components of this TFIID-like complex, TgTAF5 and TgTAF6. Ongoing experiments to epitope tag TgTAF6 will allow localization and proteomic analysis of TgTAF6 interactors to determine the composition of possible TFIID complex variants in *T. gondii*. 
**Introduction:**

Toxoplasma gondii is an obligate intracellular protozoan parasite and part of the phylum Apicomplexa, a group that includes other clinically relevant human parasites such as Plasmodium, the causative agent of malaria.¹ T. gondii is the causative agent of toxoplasmosis and is present worldwide, and estimates of infection in some countries are upwards of 60+%, with more than 40 million US citizens estimated to be infected.² The primary hosts of T. gondii are felids such as the common house cat. The parasite reproduces sexually in the cat gut and parasites are shed as infectious oocysts in the feces.¹ T. gondii can infect most warm-blooded mammals, including humans, most commonly through ingestion of oocysts or tissue cysts in contaminated food or water.¹ Once ingested, the parasites begin the asexual stage of their life cycle. In the acute stage of infection, parasites exist as rapidly proliferating tachyzoites that spread through various body tissues, especially brain and muscle tissue.¹ In the chronic stage of infection, parasites are put under stress by the host immune system and differentiate into slowly replicating bradyzoites that are protected from host factors by a thick cyst wall.¹ Immunocompetent individuals are typically asymptomatic, though may experience a mild flu-like illness during the acute stage. Immunocompromised individuals, such as HIV/AIDS patients, can experience severe complications such as myocarditis, severe retinal inflammation, seizures, and encephalitis.³ Though healthy individuals will likely not experience severe acute infection, research shows that chronic infection may be linked to or a risk factor in various mental illnesses such as schizophrenia, depression, and anxiety.⁴ Chronically infected patients can also experience severe recurrent acute infection as a result of a compromised immune system, a problem in HIV/AIDS patients, organ transplant recipients, and those on immunosuppressant drugs.⁵ In addition, mothers who become infected during pregnancy can transmit the parasite to
their fetus, which can result in spontaneous abortion, miscarriage, and birth defects.³ Treatments for acute infection exist but can have toxic side effects, and there are currently no approved drugs for chronic infection.¹

Considering the global prevalence of T. gondii and its impact on human health, it is necessary to develop better novel therapeutics. A promising avenue for drug development is targeting transcription initiation, a vital yet understudied process in T. gondii, particularly in the context of tachyzoite to bradyzoite differentiation. Bromodomain proteins are a protein class of interest because of their role in epigenetic gene regulation. These proteins are well conserved across many organisms including humans, and they facilitate transcriptional activation via reading and binding acetyl lysine residues on histone tails, as shown in Figure 1.⁶ In a previous study, a T. gondii bromodomain protein unique to Apicomplexa, TgBDP3, was identified as a protein of interest because it was shown via co-immunoprecipitation (co-IP) to associate with conserved components of the TFIID complex.⁷ The TFIID complex is a well conserved transcription preinitiation complex and general transcription factor in eukaryotes that participates in the recruitment of RNA polymerase II.⁸ In a chromatin immunoprecipitation sequencing (ChIP-seq) study, TgBDP3 was found to associate with about 12% of all active tachyzoite gene promoters, indicating that there are likely multiple transcription initiation complexes in the parasite.⁷ The T. gondii TFIID complex has not been studied and may play a role in transcriptional activation in the parasite, and considering the transcriptional role of BDP3, there may be TFIID variant complexes present.

To investigate the T. gondii TFIID transcription complex, a bioinformatics study on the complex in other apicomplexans and in humans and yeast was carried out to identify the most well conserved proteins within the complex. Protein databases and bioinformatic tools were used
to compare the proteins in the TgBDP3 co-IP to other known TFIID components in other apicomplexans, humans, and yeast. Two structural proteins were identified as likely core components to the complex, TgTAF5 and TgTAF6. According to a CRISPR-Cas9 phenotype study, TgTAF6 and TgTAF6 are predicted to be essential to parasite viability and both proteins are the most highly conserved TFIID components among apicomplexans (Table 1). TgTAF6 is predicted to have a lower phenotype score than TgTAF5, indicating it is likely more essential, and thus TgTAF6 was selected to be used experimentally. Ongoing experiments are being performed to epitope tag TgTAF6. Once tagged, a co-IP will be performed to identify TgTAF6 protein interactors to better characterize the *T. gondii* TFIID complex and to identify any TFIID variants.

**Figure 1.** Schematic of bromodomain protein reading and binding of acetyl lysine residues on chromatin.
Materials and Methods:

Bioinformatic Analyses

To determine which TAF proteins in the BDP3 co-IP were integral complex members of a *T. gondii* TFIID complex, protein homology within and outside of Apicomplexa was considered. The TAF proteins present in three co-IP replicates were chosen as candidates because of their high likelihood to be in complex with BDP3. NCBI BLAST was used to look for orthologous proteins, and NCBI Conserved Domain Database (CDD) was used to compare domain architecture between orthologs in humans, yeast, and various apicomplexans. IUPred was used to investigate regions of order and disorder within proteins, and human and yeast orthologs were also compared. Finally, TAF protein sequences from multiple apicomplexans, human, and yeast were aligned using ClustalW and a maximum-likelihood phylogenetic tree was constructed using the alignments in MEGA-X with 100 bootstrap replications.

Epitope Tagging of TgTAF6

A C-terminal hemagglutinin (HA) epitope tag was inserted into the TAF6 coding region of the genome of RHΔHXΔKu80 *T. gondii* parasites. TgTAF6 gDNA sequence was amplified by PCR and ligated into a previously constructed plasmid vector that contained a DHFR drug selection cassette, an ampicillin resistance cassette, and a 3xHA epitope tag (pLIC-DHFR-3xHA). Specific primers retaining a unique restriction site were used to amplify the C-terminal end of the TgTAF6 gene: JP-TgTAF6_F1 (5’ TACTTCCAATCCAATTATAACGTATCGAACAGGCATAGCC 3’) and JP-TgTAF6_R1 (5’ CCTCCACTTCCAATTATTAAATGCTCCACGTACGCAAG 3’). Bolded sites on the primers indicate the region that retains a PacI restriction site and complementary DNA for integration into the plasmid backbone using PacI restriction enzyme.
PCR products were resolved via agarose gel electrophoresis, yielding a band around 1.5kb. TgTAF6 amplified gDNA was ligated into the PacI digested pLIC-DHFR-3xHA construct in a 3:1 insert to vector ratio following NEB HiFi DNA Assembly protocol.\textsuperscript{10} Resulting construct was transformed into NEB 5-alpha chemically competent \textit{E. coli} cells. \textit{E. coli} cells were plated on LB agar plates containing ampicillin for selection of plasmid positive clones. After bacteria grew and were subject to drug selection, clones were picked and cultured in 3mL LB broth and ampicillin. DNA was harvested from positive clones and digested with PacI to check for integration of pLIC-DHFR-TgTAF6-3xHA, with bands expected at 1600bp and 7100bp. A positive clone was sent for Sanger sequencing to confirm the construct was correct and had no mutations.

**Parasite Transfection:**

pLIC-DHFR-TgTAF6-3xHA construct was grown in NEB 5-alpha chemically competent \textit{E. coli} cells and purified via DNA extraction and phenol chloroform precipitation. Plasmid DNA was resuspended in cytomix. Half of a T-25 flask of fully lysed RH parasites was pelleted at 1500rpm for 10 minutes and resuspended in cytomix. DNA and resuspended parasites were gently mixed and electroporated as previously described.\textsuperscript{11} Contents were transferred to a T-25 flask of human foreskin fibroblast (HFF) cells in DMEM + 1% heat-inactivated FBS and parasites were left to recover for 24 hours. Pyrimethamine was added for drug selection after 24 hours.

**Immunofluorescence Assay:**

25uL of parasites were added to wells of a 24-well plate with coverslips affixed to wells. After a 24 hour incubation, media was removed and wells were rinsed with phosphate-buffered saline (PBS) twice, then parasites were fixed to coverslips using 0.5mL of 4% paraformaldehyde
for 15 minutes at room temperature. The coverslips were rinsed with PBS three times and then permeabilized using 0.5mL Triton X-100 in 3% BSA for 15 minutes at room temperature. The coverslips were washed again with PBS three times, 10 minutes each while mildly agitated using a platform shaker. A blocking buffer consisting of 0.5mL of 3% BSA (in PBS) was added after the last wash and incubated for 1 hour at room temperature. Primary antibody (anti-HA) was incubated overnight at 4°C, then wells were rinsed with PBS three times again. The remaining steps were performed in the dark to prevent fluorescent degradation: secondary antibody (anti-mouse conjugated Alexa fluor 594) and 4′,6-diamidino-2-phenylindole (DAPI) counterstain were added in 0.5mL of 3% BSA for 1 hour at room temperature. A final three rinses with PBS were done, and then a drop of mounting reagent was put on a slide. Coverslips were placed face-down on the slide with the mounting reagent and sealed with nail polish. Visualization and imaging of tagged protein were performed using fluorescence microscopy.

**Results:**

**TgTAF6 is likely an essential member of the *T. gondii* TFIID complex**

Bioinformatic protein analyses of TAF proteins pulled down in the BDP3 co-IP were carried out using NCBI CDD, BLAST, IUPred, and ClustalW. Of the proteins analyzed, TgTAF5 and TgTAF6 are the two most well-conserved proteins across Apicomplexa, indicating they are likely core components (Table 1). Though the conserved TAF1 protein in *T. gondii* is also highly conserved, the protein is likely too long to create a tagging construct for. TAF6 is an important structural component of the human TFIID complex, binding TAF5 and other proteins to effectively tether the complex together. TAF6 contains a N-terminal histone fold domain that primarily interacts with another TFIID member, TAF9, followed by a region of HEAT repeats. Protein sequence alignments and domain comparisons using NCBI CDD and InterPro indicate
that the C-terminal HEAT repeats are very well conserved between TgTAF6 and human and yeast orthologs (Fig. 2). Sequence homology is shared between a region in the N-terminus of TgTAF6 and the histone fold domain (HFD) of human and yeast orthologs (Fig. 3). TAF6 homologs are present in most apicomplexans and TAF6 in other coccidian parasites are closely related to TgTAF6 (Fig. 3). Additionally, data from a tachyzoite genome-wide CRISPR phenotype study indicates that TgTAF6 is an essential gene, with a CRISPR phenotype score of -5.64. These results indicate that TgTAF6 is likely an important structural member of the putative *T. gondii* TFIID complex.

**Table 1.** The CRISPR phenotype scores and conservation for each *T. gondii* TAF present in the BDP3 co-IP dataset.

<table>
<thead>
<tr>
<th>TAF</th>
<th>CRISPR Score</th>
<th>Neospora caninum</th>
<th>Hammondia hammondi</th>
<th>Eimeria acervulina</th>
<th>Cryptosporidium parvum</th>
<th>Cystoisospora suis</th>
<th>Plasmodium falciparum</th>
<th>Babesia ovata</th>
</tr>
</thead>
<tbody>
<tr>
<td>TAF1</td>
<td>-4.26</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TAF2</td>
<td>-3.58</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TAF5</td>
<td>-2.79</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TAF6</td>
<td>-5.64</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TAF7L</td>
<td>-2.66</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TAF8</td>
<td>-3.48</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TAF9</td>
<td>-3.64</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TAF10</td>
<td>-2.37</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TAF12</td>
<td>-2.03</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Black = Protein is conserved, grey = possible homolog identified in BLAST, white = no homologs found.

**Figure 2.** TgTAF6 (bottom) domain composition comparison to human and yeast TAF6 proteins. TgTAF6 contains a more disordered N-terminus than human and yeast TAF6.
Figure 3. Portion of sequence alignment with *T. gondii*, human, and yeast TAF6 protein sequences and phylogenetic tree including apicomplexans. (a) Sequence alignment at the N-terminus of *T. gondii*, human, and yeast TAF6 proteins using ClustalW. Portion highlighted by red boxes shows sequence homology. (b) Maximum likelihood tree constructed from ClustalW alignments in MEGA-X. Closely related apicomplexans such as *Neospora* and *Hammondia* have similar TAF6 proteins compared to *Cryptosporidium, Cystoisospora*, and *Babesia*.

An endogenous TgTAF6 C-terminal epitope tag does not integrate properly in *T. gondii*

A plasmid containing the C-terminal end of TgTAF6, a pyrimethamine drug selection cassette, and a 3x-HA tag was transfected into RHΔku80 parasites and inserted into the endogenous TAF6 locus via single crossover homologous recombination. Parasites were incubated for 24 hours and then subjected to drug selection with pyrimethamine for 48 hours. IFA protocol was followed and no HA signal was observed. A genomic DNA PCR of transfected parasites showed no amplification of the plasmid construct indicating that it was not being properly integrated. Since this method of tagging did not work, another tagging strategy will be employed. The current strategy being pursued is to insert an ectopic copy of TgTAF6 protein under a tubulin promoter with a 3xHA N-terminal tag.
Figure 4. Schematic for TFIID recruitment in humans and *T. gondii*. (a) Schematic of TAF1 binding of acetylated lysines on histone tails in humans via a double bromodomain. (b) Schematic of possible *T. gondii* TFIID recruitment mechanism involving BDP3.

**Discussion**

*T. gondii* is a parasite that is highly prevalent globally, able to establish chronic infection in immunocompetent hosts for a lifetime.\(^1\) Acute infection in immunocompromised individuals can be life-threatening and vertical transmission from mother to fetus is a serious concern.\(^2\)
There are currently no treatments for chronic infection, and current treatments for acute infection have toxic side effects.\textsuperscript{1} Transcription has not been well-studied in \textit{T. gondii} and could be a promising avenue to identify novel drug targets. A co-IP in a previous study showed a unique apicomplexan bromodomain protein, BDP3, was associated with conserved TFIID subunits.\textsuperscript{7} The TFIID complex is a widely conserved transcription factor and transcription pre-initiation complex in eukaryotes.\textsuperscript{8} The TFIID is not well characterized in \textit{T. gondii} or other apicomplexans, though there are conserved homologs identified across multiple genera. These conserved components are predicted to be essential to tachyzoite viability and may play an essential role in transcription initiation in \textit{T. gondii}.

TgTAF6 is one TFIID component identified in the BDP3 co-IP, and it is a core component of the TFIID humans and yeast as well as \textit{T. gondii}. Bioinformatic protein comparisons showed that TgTAF6 has similar domain architecture to human and yeast TAF6 proteins (Fig. 2), and TAF6 is a conserved protein across multiple apicomplexans (Fig. 3). These bioinformatic comparisons were mostly preliminary, and additional analyses examining the three dimensional structure of TgTAF6 could be useful in determining similarities between it and TAF6 orthologs.

Attempts to endogenously epitope tag TgTAF6 have been unsuccessful. This is likely due to the integral nature of TAF6 in the TFIID complex and introducing a tag could alter its binding capabilities within the complex as a result.\textsuperscript{12} Since a PCR showed that the tag construct was not integrating into parasite gDNA, it is also possible that the parasites simply did not integrate the entire construct. With these issues in mind, another tagging method is currently being used. By introducing an ectopic copy of TgTAF6, the endogenously expressed protein would not be interrupted by an epitope tag and thus should be tolerated better by the parasites. This gives the
tagged protein a better chance of integrating without disrupting the structural integrity of the TFIID complex. A previous study in the yeast *Saccharomyces cerevisiae* yielded successful TAF6 tagging with an N-terminal 3xHA epitope, so N-terminal tagging will be used instead of C-terminal tagging.\(^{13}\)

Though the function and structure of a *T. gondii* TFIID complex cannot currently be described and the presence of possible variants cannot be confirmed, this preliminary data and knowledge of the TFIID gives some insight into possible mechanisms for a *T. gondii* TFIID complex. Figure 4 outlines a possible mechanism for TFIID recruitment in *T. gondii* with BDP3 compared to the recruitment of the TFIID in humans. In humans, TAF1 is a key TFIID member that contains a double bromodomain region that can bind acetyl lysine residues to open chromatin and allow the TFIID to bind DNA via the TATA-binding protein.\(^{12}\) TAF1 is conserved in *T. gondii* as well, but it only contains one putative bromodomain, and it is unknown how the TFIID binds acetyl lysine residues and chromatin. Figure 4 details a possible binding mechanism for the TFIID in *T. gondii*, with BDP3 binding acetylated lysine residues to open chromatin and allow TFIID binding to DNA.

Future studies including a co-IP and ChIP-seq will elucidate the functions of the TFIID in *T. gondii*. A TgTAF6 co-IP will identify unique proteins that associate with the complex, including possible alternative epigenetic reader proteins. ChIP-seq will then identify where the TFIID and possible variants bind on the genome. These experiments will help to characterize the process of transcription in the parasite as well as serve as a way to identify possible novel drug targets.
References:


7. Jeffers et al., in preparation


