Specific Aims:

Apicomplexa are responsible for a number of important human diseases including malaria, toxoplasmosis, cryptosporidiosis and cyclosporidiosis. Management of these diseases rests heavily on chemotherapy but anti-parasitic drug treatment faces multiple challenges. These include poor overall potency, restriction to certain life-cycle stages, unwanted side effects, and rapidly emerging multiple drug resistance. A constant stream of new drugs and potential drug targets is required to stay abreast of the threat posed by these pathogens. One of the most promising sources of such parasite specific targets is the apicomplexan plastid or apicoplast. The apicoplast is unique to the parasite and its function is essential to parasite survival. This organelle is a holdover from a free-living photosynthetic past. The structure and biology of the apicoplast is remarkably complex as it is derived from the endosymbiotic marriage of two eukaryotes: a red alga and an auxotrophic protist. The goal of this application is to unravel the complexity of this biology in mechanistic detail. We hypothesize that the photosynthetic past of Apicomplexa and the continued presence of a plastid has profound and lasting implications for their current metabolism and cell biology. Further we believe that discovering and characterizing this biology in its molecular detail will lead us to important insights into the biology of Apicomplexa, the evolution of the eukaryotic cell, and ultimately to novel targets for anti-parasitic interference. In our current funding period we have conducted genetic studies on proteins involved in apicoplast replication, protein import, and metabolism that were identifiable as plastid proteins in part based on their similarity to plant chloroplast proteins. We did so in a gene-by-gene fashion characterizing a limited number of proteins in considerable depth using a genetic approach. This has been an excellent strategy that served us well, we will continue to use this approach to dig deeper into mechanism in the current application. However, we also feel that we might have harvested the lower hanging fruit of candidates with a lot of function left unassigned. We therefore will complement this approach with a broader effort to define a comprehensive set of plastid proteins to continue to feed our pipeline of hypothesis-driven mechanistic experiments with strong candidate genes.

Specific Aim1: Dissect the mechanism of apicoplast protein import. The bulk of the ~500 apicoplast proteins is nuclear encoded and post-translationally imported across four membranes. We (and others) have described three mechanistically distinct candidate protein translocons that reside in the three inner membranes of complex plastids. In the current funding period we will focus on a newly discovered mechanism that was derived from the ER-associated degradation system (ERAD) of the algal endosymbiont. We will use conditional gene disruptions and complementation assays to establish the importance of individual components and to define the energy source of the translocation process.

Specific Aim2: Understand the function of the apicoplast ubiquitination pathway. The ER-localized ERAD pathway goes hand in hand with the ubiquitination and subsequent proteasomal degradation of translocated proteins. Our preliminary data indicates that aspects of this protein modification pathway are still present in the apicoplast. What is the enzymatic machinery involved in this process? What are its substrates? And most importantly, what is the molecular function of apicoplast ubiquitination? A combination of genetic and biochemical approaches will be used to answer these important questions.

Specific Aim 3: Discover a comprehensive set of apicoplast proteins and characterize their function. Mining comparative and functional genomic information we have assembled an extensive list of proteins for which we hypothesize a role in apicoplast biology. We will establish the localization of their protein products for a comprehensive set of these candidate genes by epitope tagging. In the previous funding period we have found conditional null mutants to be highly informative to study apicoplast protein function and we have developed phenotypic assays to detect defects in apicoplast protein import, apicoplast division, and apicoplast metabolism. We will apply this genetic approach to a prioritized list of validated candidates. To increase the throughput of our analyses we will develop and test a new mutagenesis approach based on promoter replacement.

Research Strategy

(A) Significance

Apicomplexa are important human pathogens responsible for numerous severe diseases around the World. These include the various forms of malaria (1-3) as well as opportunistic infections associated with AIDS (4, 5). Several of these organisms have been included on the NIH/CDC appendix B list of pathogens considered potential bioterrorism threats (*Cryptosporidium, Toxoplasma, Cyclospora*). This is due to the marked resistance of infectious oocysts to conventional water treatment, which has led to large water borne outbreaks in the past (6-11). No effective vaccines are available for use in humans and significant challenges remain in the antimicrobial drug treatment for diseases caused by Apicomplexa. These challenges include widespread multiple drug resistance (malaria (12-14), coccidosis (15-17)), lack of efficacy of current treatment against chronic stages (toxoplasmosis, with particular clinical importance in the case of ocular infection (18-21)), or absence of fully effective treatment (cryptosporidiosis (21, 22)). New treatment options with independent modes of action are urgently needed to overcome some of these challenges. One of the most promising sources for novel targets is the apicoplast. The apicoplast is a unique chloroplast-like organelle and essential for parasite growth and pathogenesis (23-25). As humans lack chloroplasts targeting plastid function has great potential to yield interventions that specifically inhibit the parasite but not the host (26, 27). Driven by the availability of full genome sequence we have made solid progress in our understanding of the potential metabolic functions of the organelle, however, its true raison d'être remains to be defined (28, 29). The apicoplast has a fascinating evolutionary history. As schematically depicted in Fig. 1 this organelle is the product of two subsequent endosymbiotic events. Most remarkably the second step reflects the union of two eukaryotes (a red alga and a pre-alveolate) and has let to a complex sub cellular structure that is delineated by four membranes (30).

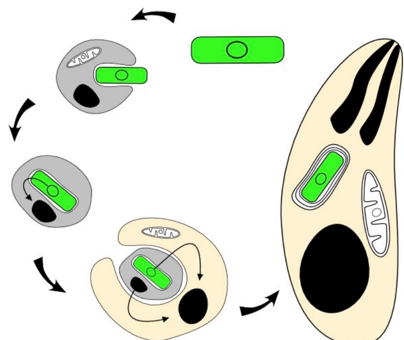


Figure 1: Schematic outline of apicoplast evolution. A cyanobacterium was endocytosed by a eukaryotic cell (primary endosymbiosis). Genes were transferred from the bacterial genome to the nucleus. Chloroplasts are bounded by two membranes. In a secondary endosymbiosis event a eukaryotic alga was phagocytosed by an ancestor of Apicomplexa. Gene transfer occurred from the endosymbiont to the host nuclear genome. The apicoplast is surrounded by four membranes.

The goal of this research project, which was first funded in December 2005, is to produce mechanistic insights into the biology of the apicomplexan plastid using *Toxoplasma gondii* as a genetic model. We anticipate that a detailed understanding of the function and cell biology of this organelle will help to prioritize the list of currently considered plastid drug targets and furthermore will lead to the discovery of new targets beyond a narrow focus on anabolic metabolism (e.g. interference with enzymes involved in apicoplast biogenesis rather than interference with the fatty acid synthesis pathway). We also expect this work to contribute in meaningful ways to our general understanding of the evolution of the eukaryotic cell and the biogenesis of organelles. Lastly, to achieve our specific experimental goals we have developed novel genetic approaches and reagents and will continue to do so in the future. We expect these advances to spur on molecular research on *T. gondii* and other Apicomplexa beyond our own research focus. We feel that our first four years of work on this project have delivered on some of this promise. We have discovered genes and proteins with important roles in apicoplast division (31-34), apicoplast metabolism (24, 35), and apicoplast protein import (36, 37) and the detailed characterization of mutants in these genes has let us to strong mechanistic models for key aspects of apicoplast and parasite biology (29, 31, 38, 39). We have developed an approach to forward genetic analysis through chemical mutagenesis and complementation cloning and a highly efficient strategy to engineer conditional gene deletions (35, 40-42). We have shared these new reagents with the community and numerous researchers have put our fluorescent organelle markers, parasite strains, cosmid clones and libraries and KO cassettes to great use (see e.g. our collaboration with the Robey lab (43-45)). Lastly, we note that there are a number of studies that we have conducted that are still in submission or preparation. These include a fully characterized knock out in the MORN1 protein demonstrating that this protein is essential for apicoplast fission

and segregation, a mutant in Tic22 showing that this protein is required for apicoplast import across the innermost membrane, and a mutant in the HU protein demonstrating that this protein is required for the maintenance of the apicoplast genome. We also have identified the genes for a suite of proteins with a likely role in apicoplast genome replication and we have tagged and successfully localized their protein products (PolA/Helicase/Primase, DNA-Helicase, Gyrase A and B). The space limitations of the new format preclude us from showing this extensive (yet still unpublished) data.

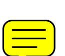
(B) Innovation
We would like to argue that our project has been highly innovative and we expect it to continue to be innovative. Innovation in this project is evident in the topic of the research, the concepts and hypotheses to be tested, and the approaches to be used. The apicoplast as a research topic has produced a truly new way to think about Apicomplexa that now permeates our view of their metabolism, development and cell biology. Studying the apicoplast has brought together biologist focused on different organisms that previously had little contact. This cross-fertilization has let parasitologists to consider pathways initially studied in plants and algae to explain parasite metabolism, drug sensitivity, gene expression control, and signaling and hormone action. Over the last year I have been invited to present our research on the apicoplast at Gordon Research Conferences in three different fields (parasite biology, chloroplast biology, and protein transport) and we view that as a testament to the fact that our specific questions and hypotheses have been innovative and are at the cutting edge of multiple fields. I am particularly excited about the potential role of ubiquitination in the apicoplast. This is a new concept that will lead us to discover new biology and might have strong potential for drug development. Lastly, as a research group and within this project we have invested considerable effort into the development of new experimental tools and approaches. This is particularly true for parasite genetics and cell biological analysis and the current proposal represents an extension of these efforts. We feel that overall this investment has paid off (at times in unexpected ways) and that taking the risk to develop new approaches in the future will keep our experiments fresh and will allow us to ask deeper and deeper mechanistic questions.

(C) Approach

Specific Aim1: Dissect the mechanism of apicoplast protein import.
The apicoplast maintains its own genome, however the bulk of the estimated 400-500 apicoplast proteins are nuclear encoded and imported into the organelle (46). Nuclear encoded apicoplast proteins are synthesized with a bipartite leader and initially routed through the secretory pathway (47, 48). Vesicle fusion with the outermost membrane is believed to deliver proteins to the organelle. How they then cross the remaining three membranes to reach the lumen has been the topic of many spirited discussions and countless review articles, yet until very recently little experimental evidence was available to test these various hypotheses. This has changed dramatically and a model is emerging that proposes three consecutive protein translocons that enable transport over consecutive membranes. Work performed under this proposal has made significant contributions to this progress. We have identified, cloned and localized several members of two candidate translocons in *T. gondii* *(*(29, 36, 37), van Dooren & Striepen unpubl., and this proposal). This work described a translocon related to the Tic (translocon of the inner chloroplast membrane) in the innermost apicoplast membrane and a translocon of the second or periplastid membrane derived from an endoplasmic reticulum associated mechanism of the algal endosymbiont. We have adapted a split GFP assay to define the various subcompartments of the apicoplast (37) that has subsequently also been used in other complex plastid systems (49, 50). We also have developed three biochemical assays to quantify apicoplast protein import and most importantly we have used these assays in combination with knock out studies to provide rigorous genetic support for a direct role of two translocons in protein import (36, 37). A recent publication from an algal model organism suggests that a Toc (tranlocon of the outer chloroplast membrane) derived mechanism might be responsible for the transport across the middle membrane (50).

The apicoplast ERAD system
The initial focus of our mechanistic studies will be the system that is now believed to be responsible for protein import across the second outermost apicoplast membrane. This membrane is of particular interest as it is thought to be homologous to the plasma membrane of the algal endosymbiont. Key to the discovery of the

mechanism was the sequencing of the nucleomorph genome of the cryptophyte alga *Guillardia theta* (the nucleomorph is the "fossil" remnant of the algal nucleus). Sommer and coworkers noted that this highly reduced genome encodes core elements of the endoplasmatic reticulum associated degradation (ERAD) system (51). ERAD usually acts in ER homeostasis by retrieving misfolded secretory proteins from the ER and funneling them for degradation to the proteasome in the cytosol. The core components of the ERAD transport machinery are Der-1, the ATPase Cdc48 and its co-factor Ufd-1. Der-1 is a favored candidate for the proteinaceous pore in the ER membrane and has been shown to be essential for retro translocation of misfolded luminal proteins (52). Protein substrates destined to be degraded are polyubiquitinated and subsequently extracted from the pore by the Cdc48-Ufd-1-Npl4 complex (53). Sommer and colleagues formulated the hypothesis that an ERAD translocon had been retooled to import proteins into complex plastids (51). The ERAD hypothesis has accumulated considerable support from a recent flurry of publications reporting the identification and plastid localization of ERAD components in cryptomonads, diatoms, and Apicomplexa (51, 54-56). In our own work supported by this grant we have demonstrated that the *T. gondii* genome encodes multiple homologs of Der1, Cdc48 and Ufd-1. Immunofluorescence analysis of parasite cell lines expressing epitope tagged forms of these proteins reveal that while one complete set of components is associated with the ER and likely performs their classical role in ERAD, at least one homolog of each of these components localizes to the outer membranes of the apicoplast. Furthermore phylogenetic analysis of the two *T. gondii* Cdc48 proteins demonstrates that they are of divergent evolutionary origins. The apicoplast localized Cdc48 forms a well-supported clade with its red algal lineage counterparts (including the protein encoded on the *G. theta* nucleomorph) while the cytoplasmic protein branches with proteins that reflect the current view of vertical evolution for Apicomplexa (56). Genetic ablation of Der1$_{Ap}$ in *T. gondii* results in swift and complete ablation of apicoplast protein import as measured using a variety of biochemical assays (56, 57) demonstrating a direct role of Der1$_{Ap}$ and the endosymbiont derived ERAD system in apicoplast protein import.



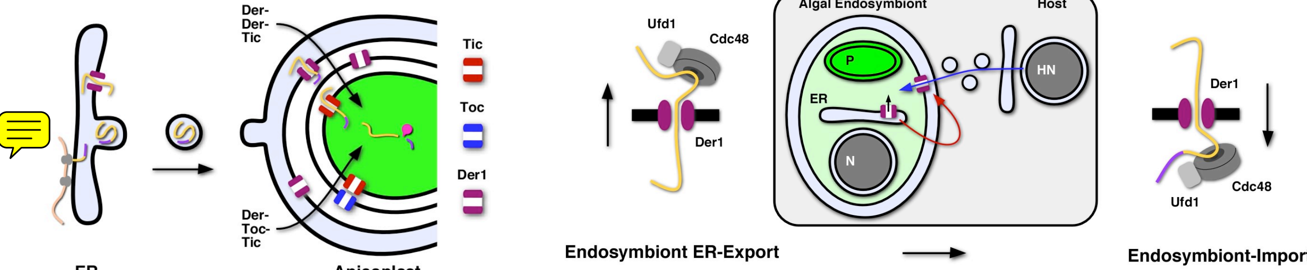


Figure 2: Schematic overview of protein trafficking in the apicoplast and the components and evolution of the apicoplast ERAD system (see our recent publications (29, 36) for further detail).

How is the apicoplast ERAD system powered?

In the next funding period we will further dissect the mechanism of protein import and initially focus on the apicoplast ERAD system. We hypothesize that Cdc48$_{AP}$ and more specifically its ATPase activity serves as the driving force of the translocation machine. This hypothesis predicts Cdc48$_{AP}$ to be essential for apicoplast protein import. To test this we will construct a conditional mutant. We already have engineered the coding sequence of Cdc48$_{AP}$ into a plasmid harboring a tetracycline regulatable promoter element and transfection with this construct results in regulated expression. Next we plan to disrupt the native copy of the gene. We have used modified cosmid clones with great success to delete genes (35). Unfortunately however, Cdc48$_{AP}$ is not covered in our cosmid collection and initial attempts with small plasmid-based constructs were unsuccessful. To be able to efficiently target the gene we have recently constructed a deletion mutant in the Ku80 gene in the tetracycline regulatable TATi strain (please refer to Specific Aim 3 for preliminary data and technical detail). This new strain combines the benefits of high homologous recombination (58, 59) to attain gene replacement with a robust regulatable expression system (60). We will use a plasmid-based targeting construct using 5' and 3' flanks of the Cdc48$_{AP}$ gene to drive gene replacement of the locus in this new strain using a chloramphenicol marker. We will isolate stable drug resistant clones and test them for gene replacement by PCR and Southern blot as we have done previously many times e.g. (61). While not likely it is possible that we might not be able to isolate a Cdc48$_{AP}$ knock out. This might not be due to lack of homologous recombination but due to the fact

that the regulated transgene does not sufficiently complement loss of the native copy. One possibility is that the epitope tag that we introduce interferes with function. Fortunately, this can now be conveniently tested by introducing an epitope tag directly into the genomic locus in a Ku80 KO strain (58). We have conducted such an experiment and found that the locus can be tagged. A second potential pitfall could be that the level of transgene expression driven by the regulated promoter is lower than the native promoter preventing effective complementation and mutant isolation. This is a significant problem that we (and others) have encountered in the past. We like to note that overall the tet7sag4 promoter we use appears to be well matched to the expression level of apicplast proteins as we were able to use this strategy to up to now disrupt seven different apicoplast proteins successfully (24, 35-37) and Nair & Striepen in preparation). However it is conceivable that some proteins require higher expression. We have engineered an alternative system using the tet-repressor model in the Ku80 background. This model is based on a stronger ribosomal promoter and drives higher expression (62) and we will explore this and other alternative models described in further detail in Aim 3 if needed. Having a mutant in hand we will test the impact of loss of $Cdc48_{AP}$ on apicoplast protein import. First we will establish the kinetics of regulation of a particular mutant clone. We will culture the parasites in the presence of anhydrotetracycline (ATc) and sample parasites over a four day time course. We will measure the concentration of $Cdc48_{AP}$ by Western blot using an antibody against the HA-epitope tag that we introduced at the 3' end of the coding region (29). Next we will evaluate apicoplast protein import by pulse chase assay. We have developed a number of independent markers for this assay that measure targeting peptide removal by signal peptidase, transfer of lipoic acid to apicoplast pyruvate dehydrogenase and biotinylation of acetyl CoA carboxylase. Note that we also have developed a number of control experiments to assure that effects are specific to the apicoplast and to protein import (see our recent publications (36, 37) for technical detail). Based on the critical importance of Cdc48 in the ER ERAD system (53) and its conservation among organisms with secondary plastids (36, 49, 55, 63) we expect to demonstrate that $Cdc48_{AP}$ is essential for import and parasite survival.

We hypothesize that $Cdc48_{AP}$ acts as the molecular motor of the translocon and more specifically that its ATPase activity is critically required to power protein translocation. Alternatively it is possible that $Cdc48_{AP}$ might have a more structural role or act as a cargo receptor. We will test this by complementation analysis. The biochemistry of AAA-ATPases including Cdc48 has been studied in considerable detail (64). Taking advantage of the detailed knowledge of yeast Cdc48 we have annotated a N-terminal domain (aa 162-248) and two ATPase domains containing Walker A and B motifs (410-594, 683-870) in *T. gondii* $Cdc48_{AP}$. Using site directed mutagenesis we will produce a series point mutations that disrupt conserved residues that are usually critical for ATP binding (e.g. a K/A exchange in the GxxGxGK motif of the first ATPase domain) or ATP hydrolysis (e.g. an E/Q exchange in the DExx motif of the Walker B motif of the second ATPase domain). We will transfect the conditional $Cdc48_{AP}$ mutant with plasmids that either express wild type $Cdc48_{AP}$ or one of the point mutants. We have used such complementation assays previously and we have a variety of vectors with suitable selectable markers, epitope tags and constitutive promoters available (e.g. a Ble marker after using DHFR and CAT to generate the conditional KO see our publication (35)). Next we will culture each clone in the absence or presence of ATc. Under ATc treatment the regulated gene will be suppressed and the parasite will depend on complementation from the constitutive transgene. We expect that the wild type gene will restore apicoplast protein import and parasite growth under ATc. Should point mutants in the ATPase motifs fail to complement the mutant we will conclude that this activity is essential to transport and that $Cdc48_{AP}$ likely acts as the motor of the translocon. A potential concern with these experiments could be that the mutants we engineer do not (or not only) affect ATPase activity but result in unstable or mislocalized proteins. The vectors we use also include a Ty-1 epitope tag. We will establish that mutant and wild-type proteins are targeted correctly to the apicoplast and are expressed at similar level by immunofluorescence and Western blot (see our recent publication for further detail (35)). Some point mutations in ATPase domains can result in dominant negative mutants which could complicate our analysis (we might not be able to isolate stable clones to test for complementation) and this is not always easily predicted. Should this become a concern we will engineer a range of different point mutants to identify recessive mutants. We would then also explore to introduce putative dominant negative transgenes using a conditional promoter element (please refer to Aim 3 for more detail and preliminary data on additional elements).

Is substrate recognition of $Cdc48_{AP}$ modulated by a cofactor complex?

Cytoplasmic Cdc48 acts as an unfolding and extracting enzyme in a tremendous variety of cellular processes. To lend specificity and thus control to this process Cdc48 has evolved to interact with numerous other proteins that modify substrate recognition and control activity (65). In the yeast and mammalian ERAD system Cdc48 forms a complex with Ufd1 and Npl4 (64, 66) and both cofactors are required for ERAD activity. Our studies indicate the presence of an apicoplast targeted Ufd1 but do not detect an Npl4 homolog. So far we have established that $Ufd1_{AP}$ is targeted to the apicoplast and based on immunofluorescence data we believe it to localize to the outer membrane compartment (36). The Cdc48 cofactors are of particular significance here as they tie ERAD to ubiquitination, more specifically the recognition of ubiquitin on proteins to be extracted (or imported in this case) by Cdc48. We will establish if Ufd1 is required for apicoplast protein import and what its interacting partners in the process are. We will construct a conditional mutant as detailed above for Cdc48. This should be straightforward as we have a cosmid clone in hand and we have shown that we can efficiently target the Ufd1 locus as we used a cosmid to introduce an epitope tag (36). Next we will measure apicoplast protein import in this mutant. Our hypothesis predicts loss of import. It is conceivable that deletion of $Ufd1_{Ap}$ does not impact protein import and we would then conclude that it acts in a different process (e.g. in a process linked to protein stability in the apicoplast). In other systems Cdc48 and Ufd1 form a stable complex. In mammalian cells Ufd1 and Npl4 from an initial binary complex that then recruits Cdc48 forming a tertiary complex. We will test if such interaction occurs between $Ufd1_{Ap}$ and $Cdc48_{Ap}$ in co-immunoprecipitation experiments. For these experiments we will construct a parasite strain in which Cdc48 is tagged with an HA epitope and $Ufd1_{Ap}$ tagged with a c-myc tag. We will lyse parasites in the presence of mild detergents and protease inhibitors and precipitate Ufd1 or $Cdc48_{Ap}$ using a polyclonal rabbit serum specific to cmyc or a rat anti-HA monoclonal antibody respectively. Next we will test for co-precipitation by reciprocal Western blot. The hypothesis of a Cdc48/Ufd1 complex in the apicoplast predicts that we detect $Ufd1_{Ap}$ upon $Cdc48_{Ap}$ pull down and vice versa. This result would provide critical mechanistic insight into the translocase, show that these components are truly in the same compartment (which is beyond the resolution of microscopy based approaches), and would be an important first link to ubiquitin biology as Ufd in other systems acts as a ubiquitin receptor.

There are some potential experimental challenges here. 1) Epitope tagging interferes with interaction. This is unlikely as we were able to tag both proteins in the genomic copy. As alternative we can also consider using the anti-$Cdc48_{Ap}$ serum that we raised against recombinant protein (36). 2) Only a subpopulation of both proteins interacts and/or native protein competes with tagged protein diminishing the sensitivity of the experiment. This is a significant concern. To gain maximum sensitivity we will use a strain in which both partners are tagged directly in the genomic locus (this is straightforward as we already have clones in hand in which one of the partners is tagged (in a Ku80 mutant background) and we have suitable vectors to introduce the second tag). 3) The complex is not sufficiently stable to remain intact through lysis and pull down. Again this is a valid concern. We will begin our experiments using detergent and ionic strength conditions that were suitable for the mammalian and yeast ERAD components (64, 67) and optimize conditions if necessary.

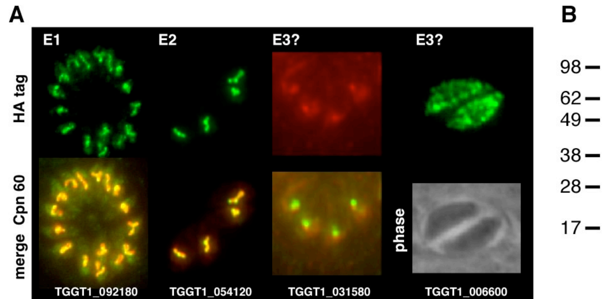
The pull down approach would be highly suitable to identify additional components of the translocation machinery. In particular to answer the question if there indeed not Npl4 component to the apicoplast translocon. In Aim 2 we will explore tap-tags and mass spectrometric analyses that might also provide future avenues for this Aim.

Overall we expect that Aim1 will define the core of the apicoplast ERAD system, show if all three identified components are equally important for translocation, demonstrate their interaction, and identify CDC48 ATPase activity as the motor of the system.

Specific Aim2: Understand the function of the apicoplast ubiquitination pathway.

Retro-translocation by the ER resident ERAD system is immediately followed by the degradation of extracted proteins in the cytosol. This occurs in the proteasome and proteins are routed this way by concurrent ubiquitination (68). Ubiquitin is a small protein that is transferred to lysine residues of the target protein is a complex process that in many cases requires three enzymes: a ubiquitin activating enzyme (Uba or E1), a ubiquitin conjugating enzyme (Ubc or E2) onto which E1 transfers a ubiquitin, and finally a ubiquitin transferease (Ubt or E3) that transfers ubiquitin units to the target protein. In some systems an E3 is dispensable.

The interaction between these three enzymes along with several cofactors and regulators provides an exquisite level of specificity and control to the process (69). Is the ubiquitination portion of the ERAD process conserved and active in the apicoplast? At first glance this would appear counter-productive as protein import and not protein degradation appears to be the function of the apicoplast $Der1_{Ap}$ translocon. Furthermore addition of ubiquitin might interfere with protein function and/or further transport by subsequent Toc and Tic translocons. However, there is emerging evidence for a set of tantalizing genes that may encode ERAD associated ubiquitination enzymes that localize to the plastid in certain chromalveolate algae and Apicomplexa (36, 55, 63). At the moment the best characterized is Ufd1, but there are also candidates for E1 and E2 enzymes as well as ubiquitin. Spork and colleagues cloned what they expected to be the plastid leaders of some of these genes from *Plasmodium falciparum* and introduced them into a GFP expression vector. Transfection of malaria parasites showed plastid labeling (55). While this is not proof that the native proteins are plastid-targeted it is certainly consistent with the notion of an apicoplast ubinquitination pathway. In this Aim we will dissect the identity and function of this pathway. As our first step we will annotate and localize the putative ERAD associated ubiquitination factors in *T. gondii*. As of today we have tested four candidate genes out of a larger pool that appeared more likely to be plastid targeted based on phylogenetic profile and/or cell cycle expression pattern (see Aim 3). Fig. 3 shows the localization pattern of HA epitope tagged clones constructed by either direct tagging of the genomic locus (58) or tagging of a cDNA clone. We identify one protein with cytoplasmatic localization (a putative E3), two proteins (an E1 and an E2) that show a pattern indistinguishable from the three apicoplast ERAD proteins that we identified earlier (36), and one E3 protein with weaker labeling in the apicoplast that requires further work to attribute localization with full confidence.



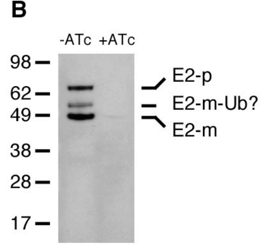


Figure 3: Putative apicoplast ubiquitination factors in *T. gondii*. (A) Immunoflorescence analysis of epitope tagged genes. Note robust apicoplast labeling for the putative E1 and E2 enzymes. **(B)** Westernblot of apicoplast E2. Note regulation by ATc and presence of intermediate size band.

We will complete this analysis tagging additional genes (and we will conduct further light and electron microscopic studies to confirm the apicoplast membrane localization of the factors identified), but it already appears that there are several strong candidates in hand. What could be the function of an apicoplast ubiquitination system? We consider a number of hypotheses. 1. Ubiquitination is an essential part of the translocation mechanism and each cargo protein is modified before or while passing through the Der1 pore. This idea would be consistent with the presence of the ubiquitin binding protein Ufd1 and the finding that polyubiquitination is required not only for degradation but also for translocation in the mammalian ERAD system (64, 67, 68). 2. Ubiquitination occurs only on a subset of cargo proteins and it acts as a specific sorting signal e.g. as a retention signal (70) for the periplastid space preventing further transport to the lumen. 3. Ubiquitination acts in a process independent of protein import (e.g. apicoplast division, protein folding and quality control etc.). We currently favor the translocation hypothesis but we will keep an open mind. We are planning a series of genetic and biochemical experiments to work our way to a mechanistic understanding of the process and to develop molecular reagents and assays.

Is apicoplast E2 required for apicoplast protein import?

Our first goal will be to demonstrate if ubiquitination is essential to apicoplast protein import. To do this we will construct a conditional mutant in this pathway and compare its effect to mutants in apicoplast Der1, Cdc48 or Ufd1. We have not concluded our annotation and localization project yet, however, at this time gene GT1_054120 the putative apicoplast E2 appears to be our best choice as an initial target. We have solid support for the plastid localization of its protein product, and we were able to construct a cDNA based expression vector that results in regulated expression (Fig. 3 A&B). We will engineer a conditional mutant in the TATi ΔKu80 strain using a plasmid construct. Once a mutant clone is established and validated we will test its ability to

import apicoplast proteins in the absence or presence of ATc. Strong inhibition of apicoplast import upon suppression of E2, comparable to the $Der1_{Ap}$ mutant (36), would argue for an essential mechanistic role of E2 in the translocation process. Alternatively, lack of a strong and direct effect on import would argue against such a role. We will interpret these genetic results with care. Multiple redundant E2 activities might be present in the apicoplast preventing a strong phenotype (further work might identify additional enzyme genes). It is also conceivable that loss of E2 might have a strong effect on transport even if proteins in transit to the lumen are not direct substrates of ubiquitination (e.g. if ubiquitination is required for the proper assembly and retention of the Der1 translocon). We will therefore complement our analysis of the E2 mutant phenotype with further genetic and biochemical experiments.

What is the molecular identity of the putative substrate of apicoplast E2?

The first goal of our biochemical studies might appear quite simple and basic: we wish to establish if ubiquitin is present in the apicoplast as suggested by the presence of E1 and E2 enzymes. So far there is no direct experimental evidence to support this notion. While Spork and colleagues show that a portion of the *P. falciparum* genome upstream of a putative ubiquitin gene can confer plastid targeting to GFP this same portion is not readily identified in other *Plasmodium* species (in their analysis (55)) or in *T. gondii* (our analysis) or in some of the algal systems with secondary plastids (49, 63). Blast searches identify a comparable number of ubiquitin (and ubiquitin-related) protein genes in all Apicomplexa including the plastid-less *Cryptosporidium* and no particular gene stands out due to its phylogentic or expression characteristics. This does not mean that an apicoplast-targeted ubiquitin does not exist, but it indicates that bioinformatics alone might not suffice to answer the question. Experimental detection of the protein in the apicoplast is not as trivial as one might think. Commercial antibodies against human ubiqitin are available (*T. gondii* ubiquitin is essentially identical to the human protein), but this is one of the most abundant cellular proteins, immunofluorescence shows labeling across the entire parasite and host cell and Western blot detects numerous bands due to the many functions of the protein (data not shown). Epitope tagging is complicated by the fact that the c-terminus of the protein is required for function (it is the site of peptide linkage to target proteins). We are working to epitope tag the two most conserved candidates, a polyubiquitin (the homolog of the gene studied by Spork) and an ubiquitin-ribosomal protein fusion directly in the genomic locus. Work on the first gene is still ongoing, transgenics for the second gene resulted in cytoplasmic staining. We will also introduce an N-terminal epitope tag as N-terminaly tagged ubiquitin is readily transferred by E1, E2 and E3 enzymes e.g. (71, 72). However, this might not be easily done without upsetting correct targeting which also depends on the N-terminus of the protein. Overall, while epitope tagging ubiquitin is a reasonable approach that very well might provide us with an apicoplast specific probe, and therefore is worth pursuing, there are some potential technical problems. To overcome this obstacle we will use the validated apicoplast E1 and E2 proteins in hand to track down apicoplast ubiquitin. We will take advantage of the fact that after activation through E1 ubiquitin is covalently linked to the E2 enzyme.

In preliminary Western blot analyses of the apicoplast E2 candidate and we detect three bands (Fig. 3B). The sizes of two of these bands are consistent with the predicted precursor and mature form of the protein (before and after removal of the apicoplast transit peptide). The third band is of intermediate size and consistent with an ubiquitinated form (+ 8 kDa). We will perform immunoprecipitation experiments using reagents specific to the E2 protein followed by Western blot using antibodies to ubiquitin. The hypothesis predicts that the intermediate band should react with the ubiquitin antibody while the other two will not. To independently evaluate the identity of these three bands we will replace the HA epitope tag in the expression construct with a tandem affinity purification (tap) tag ((73) we have used this tag successfully before to purify apicoplast acyl-carrier protein in unpublished work). We will establish a parasite line expressing this transgene and verify that localization and Western pattern are not perturbed by the tag. Next we will purify E2 from larger parasite preparations (we will start with 10e10 parasites a number easily in reach that was sufficient to identify ACP by MS in our previous studies). Our first experiment using purified protein will be a MALDI-TOFF analysis to measure the precise molecular weight of the three proteins. This can be done with very little material and to a resolution of +/- 20 Da, see our previous MALDI-TOFF experiments in (74). We expect three main peaks in our spectra, one consistent with the molecular weight of the entire protein minus the signal peptide and one

representing the mature peptide, subtraction of these masses will allow us to determine the site of transit peptide cleavage. More importantly, comparison of this mass with the mass of the intermediate band will allow us to test if the difference is consistent with the mass of *T. gondii* ubiquitin. In a second experiment we will separate the three protein species by SDS-PAGE and conduct LC-MS-MS experiments on trypsinized gel slices to generate peptide sequence information. Our hypothesis suggests that the internal band will yield peptides consistent with ubiquitin. It is not inconceivable that the apicoplast ERAD system might have evolved to use a specialized ubiquitin-like modifier rather than ubiquitin itself (numerous variations on this theme regulate a tremendous variety of cellular processes (75)). Pull down of apicoplast localized E2 and mass spectrometric analysis will allow us to identify such a modifier *de novo*. We have used the Harvard mass spectrometry and peptide sequencing facility in the past but we also have access to outstanding local proteomic facilities in UGA's Center for Complex Carbohydrate Research and many groups in our institute have extensive expertise in this area (e.g. Rick Tarleton (76, 77)). Overall we expect that these analyses will identify the nature of the apicoplast E2 modification and our favored hypothesis at this time is that this will be ubiquitin.

Where and how does ubiquination occur?

A concern one might raise is that the presence of ubiquitin on the apicoplast E2 might be non-specific. Many proteins are ubiquitinated e.g. upon misfolding or at the end of the cellular lifetime and we might potentially over-interpret our result. How can we test if ubiquitination only occurs in the apicoplast as part of a decicated activation process? We will explore three independent avenues of experimentation. The first set of simple experiments will test if apicoplast E2 is specifically ubiquitinated as a thiolester to its active site cysteine. One would expect this in *bona fide* activation (68, 69). In contrast if E2 is merely tagged for degradation ubiquitination should occur through a peptide linkage at another site (typically a lysine). Our alignments identify cys573 as the likely active site cysteine surrounded by a number of highly conserved residues. We will construct a cys/ala point mutation which in other systems abolishes thiolester formation and E2 activity while preserving the overall structure and stability of the enzyme e.g. (78). Engineering this mutant is straightforward and we usually use the Statagene quickchange site-directed mutatgenesis approach (our recent publication (35)). We will test the point mutant side by side with the wild type gene in transfection experiments. We expect that mutation of cys573 will abolish the intermediate band observed for E2 in Western blots (we will control for correct targeting and processing by IFA and Western blot). Lastly we will use the point mutants in complementation assays in the E2 mutant (assuming here that such a mutant has a detectable phenotype) to formally establish if enzymatic activity is required for biological function. In some cases catalytic cys mutants have been shown to have dominant negative effects (78, 79). If needed we could consider to use regulated expression systems, however we note that our primary experiment can be done by transient transfection and therefore does not depend on viable stable progeny.

Our second approach is based on the hypothesis that the apicoplast E1 enzyme that we have identified (GT1_092180) is the enzyme that transfers ubiquitin to the E2. To test this hypothesis we will construct a conditional mutant in the E1 gene. The hypothesis would predict that upon repression of E1 (by addition of ATc to the growth medium) E2 would no longer be modified. This should be straightforward to test by Western blot and pull down as outlined in the previous section. Furthermore, if ubiquitination is essential to apicoplast protein import we would also expect to detect loss of protein import providing a further test of the central hypothesis. In contrast, should we find no effect of loss of apicoplast E1 we would focus our attention on other potential E1 candidates. Demonstrating that E1 is required for E2 modification is consistent with the idea that his occurs in the apicoplast as both enzymes are localized there, however, it is not impossible that this might occur *en route* through the secretory pathway (see e.g. the ubiquitination process associated with endosomal sorting in the ESCRT pathway (80)). Having mutants in hand would permit to rigorously test the requirement for organellar targeting through complementation studies. We will introduce a copy of the E1 gene into the respective mutant and test for complementation. Alternatively we will construct strains in which we introduce mutant genes that maintain the signal peptide but lack the presumptive transit peptide required for apicoplast targeting, the particular sequence motifs have been dissected in considerable detail for several genes (46, 47, 81-83). Equivalent experiments can be conducted with the substrate E2. The hypothesis of apicoplast-localized ubiquitination would be strongly supported by the finding that only apicoplast targeted E1 and E2 restore

ubiquitination of E2 and plastid import. We are aware that such complementation assays are not without technical pitfalls and we will design control experiments to test expression, folding and localization of the various proteins.

Which apicoplast proteins are substrates of ubiquitination?

The import and retention hypotheses for the function of apicoplast ubiquitination make straightforward predictions with respect to the substrates of the process (all proteins inclucing those targeted to the lumen are modified or alternatively only periplastic proteins are modified). If all apicoplast proteins are ubiquitinated we expect a band (or a series of bands) of slightly higher molecular weight species for every protein. So far on Western blots of apicoplast proteins this is not readily appreciated, however, in pulse chase experiments higher molecular weight bands can be seen (see e.g. (36, 37)). Presence or absence of extra bands might be artifacutal for a variety of reasons (transient nature of modifications, bands are not related to ubiquitination but are degradation products, modification is labile and lost during preparation etc.). More robust experiments are needed to settle this issue. We will test for ubiquitination by immunoprecipation of apicoplast proteins followed by Western blot with ubiquitin antibodies as outline above for E2. We have reagents for a representative number of proteins either through epitope tags or antibodies raised against recombinant proteins. For luminal proteins we will use FNR-RFP, Cpn60, DOXP-RI and ACP and for proteins that remain in the apicoplast periphery $Cdc48_{Ap}$, $Der1_{Ap}$, $Ufd1_{Ap}$ and thioredoxin (36, 37, 47, 84, 85).

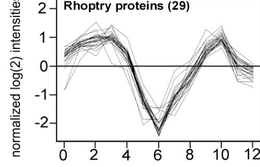
Low abundance of modified precursor due to a fast turnover could potentially make detection of such intermediates more difficult. There are several biochemical and genetic approaches that we will explore if this should be a problem. We will pre-treat samples with specific protease inhibitors used to preservere ubiquitination through sample preparation (see (86) for technical detail, we sought advice from a laboratory in our department with considerable experience in analysis of ubiquitination pathways). To increase sensitivity we can also use the tap-tagged ACP transgene to purifiy larger amounts of protein. We will follow the purification step by Western blot or characterization by mass spectrometry as outline earlier. Another option would be to increase the abundance of transport intermediates. We have isolated a number of apicoplast import mutants that we can use to arrest proteins in the pre-luminal compartments like our mutants in Tic20 and Tic22 ((37) and van Dooren & Striepen unpubl.). We could also consider to interfere with the presumptive subsequent deubiquination step akin to proteasome inhibition in studies of cytoplasmic ubiquitnation (most characterized deubiquitinases are cysteine proteases (87), the proteasome associated activity is a metallo-protease). It is important to note that we would not require tight specificity of inhibition as we have an apicoplast specific read out through out antibody reagents and therefore could explore a number of broader inhibitor classes. Lastly, a mutant in a putative apicoplast deubiquitinase, while beyond the initial scope of this proposal, would be an excellent tool here and we note that the *T. gondii* genome encodes two interesting yet uncharacterized candidates.

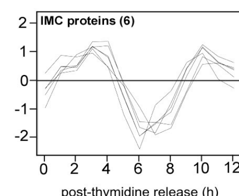
Overall we expect that Aim2 will provide us with robust information on the identity and composition of the core ubiquitination machinery of the apicoplast, its localization, mechanistic activity and substrates. Importantly this Aim should establish if these enzymes play an essential role in apicoplast protein import.

Specific Aim 3: Discover a comprehensive set of apicoplast proteins and characterize their function.

Our first two Aims test the mechanistic function of a smaller number of apicoplast proteins for which organelle localization is already established in considerable detail. In this third Aim we seek to complement these experiments by discovering new proteins with roles in apicoplast biology through a more comprehensive genome-wide effort. This approach will directly feed into the mechanistic studies outlined in Aim 1 and 2 as it will provide important candidates for functions that are not easily spotted by simple similarity search. It will also identify proteins with roles in other apicoplast functions that we are studying including plastid division, plastid genome maintenance and plastid metabolism. Our comprehensive genetic approach will not only provide insights towards the function of a particular protein but will also immediately highlight those that are essential to parasite survival and thus should receive particular attention as potential targets for intervention.

Identify the putative apicoplast proteome by comparative and functional genomic analyses
The proteome of the apicoplast is currently estimated to consist of roughly 500 proteins (46). Some of these proteins are readily identified because they carry a stereotypical apicoplast targeting peptide and/or have obvious chloroplast homologues. Good examples of this are the enzymes of the metabolic pathways that are of obvious chloroplast origin (25, 47, 82). However there are numerous proteins for which the leader is not as easily identified (e.g. due to an internal signal peptide (88)) or their function is not limited to chloroplasts (e.g. the ERAD translocon (29, 36, 54)). Many apicoplast proteins are likely specific to secondary plastids and the specific biology of these unique organelles. We have assembled several prioritized lists of candidate apicoplast proteins over the previous funding period and these have fed our mechanistic studies. One avenue that we used to prioritize genes is through phylogenetic analyses. We have used the OrthoMCL algorithm (89) to identify genes that are present in *T. gondii* and other plastid-bearing Apicomplexa but absent in the plastid-less *Cryptosporidium*. This is a very informative list with many validated plastid proteins that has served us well so far but likely provides and overestimate as *Cryptosporidium* shows considerable genome reduction not limited to plastid loss (90). To enhance specificity we used cross-referencing with plant and algal genomes in particular those of the red alga *Cyanidioschyzon merolae* (91) the nucleomorh genomes of cryptophytes and chlorachinophytes (92-94) and the nuclear genome the diatome *Phaedactylum* (95). Depending on the stringency and complexity applied the number of genes identified by these searches range from 99 to 550 (Sheiner & Striepen unpubl.). A second collection of genes was annotated independently of phylogenetic analyses. The underlying assumption for this list is that proteins targeted to the plastid are expressed in a similar window of time within the parasite cell cycle (96). There is accumulating evidence for a rigid transcriptional program unfolding over the cell cycle and this patterning appears to be regulated by a cascade of transcription factors with similarity to plant AP-2 proteins in *Plasmodium* (97-99) and *Toxoplasma* (still largely unpublished, see letter by Dr. White). Using extensive mircorarray analysis the White lab has deciphered numerous specific expression patterns that now can be applied to identify genes with simlar expression kinetics (see example figure kindly provided by our collaborator Michael White comparing rhoptry, inner membrane complex and apicoplast patterns). An initial bait set of 32 apicoplast proteins has identified a group of 459 genes. While this group clearly has some false-positives that are involved in mitochondrial and nuclear biology it contains a large number of recently identified apicoplast genes (importantly including essentially all of our recently validated apicoplast specific ERAD and ubiquitination components). We have produced several intersections between the phylogenetic and the expression lists and we will begin their experimental characterization by localizing the first 100 highest priority candidates (see below). We will use the results to iteratively improve the functional genomic annotation (see letter from Dr. White) and then move to the next set of 100 genes.





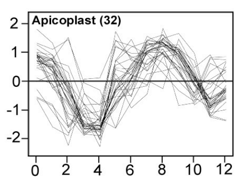


Figure 4: Cell cycle transcriptional pattern of organelle proteins (Microarray analyses kindly provided by Dr. Michael White, University of South Florida)

Test for apicoplast localization by epitope tagging of candidate genomic loci
Our first step is to localize the products of candidate genes. We will use high throughput HA epitope tagging of the c-terminus of predicted proteins by single crossover recombination directly into the genomic locus. This is straightforward using a ligation independent cloning protocol to generate a targeting vector using a single PCR product (58). We have tested this approach using a pilot set of 10 candidate genes from our list. We found four apicoplast tags, one mitochondrial tag, one nuclear tag, one protein that appears to localize to the secretory pathway very close to the apicoplast and for three genes we did not obtain any staining. Reasons for absence of labeling likely include incorrect gene models or the fact that an epitope tag interferes with protein function. Overall we found this to be a robust and scalable approach that delivered several interesting candidates with relatively modest effort. Based on the pilot we assume that testing 200 candidates will provide us with at least 50 validated apicoplast proteins for further analysis. There have been discussions in the *Toxoplasma* research community to potentially seek funding to test the localization of all *T. gondii* proteins. We will integrate our

experiments into this larger effort should that project come to pass and receive funding.

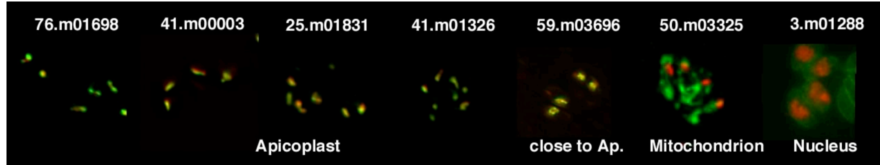


Figure 5: Immunofluorescence of genomic HA-epitope tags. Counterstained with anti-Cpn60 (apicoplast) antibody (first six panels) or myosin light chain (for parasite outline, last panel).

Establish the function of a prioritized set of apicoplast proteins by mutant analysis

In this part of the Aim we will evaluate the loss of gene function with respect to apicoplast function. The main goal here is not as much to validate suspected function but to survey a more comprehensive fraction of the apicoplast proteome to discover new function and new molecular players. We are particularly interested in targeting "conserved hypothetical" proteins that are shared among Apicomplexa and alga with secondary plastids (gene 76.m01698 shown in Fig. 5 is a good example as it is not only found in all Apciomplexa with the exception of *Cryptosporidium* but also has a homolog encoded in the nucleomorph of *Guillaridia theta*). We are also very interested in new apicoplast proteins with identifiable structural domains that may be helpful in subsequent mechanistic dissection. A number of our candidate genes e.g. have presumptive GTP binding domains (583.m05609, 20.m0376 or 55.m04878). How many genes we will be able to characterize by mutant analysis depends on the strategy that we will employ. As outlined in the next section we will make efforts to develop novel approaches with higher throughput to survey a large fraction of the validated gene set. However even with existing technology we should be able to target a sizable collection. Recombineering of cosmids is a highly efficient approach to gene targeting that we developed in the previous funding period (35) and we have constructed ten mutants using this approach in relatively short time (APT, 3 FA elongation enzymes, 2 DoxP pathway enzymes, MORN1, IMPDH, Ku80 and HU). Furthermore, in extensive preliminary work we have also generated strains that combine the high homologous recombination frequency of the Ku80 mutants (58, 59) with regulated transgene expression. We have introduced the YFP-TetR tetracycline-repressor gene (62) into a Ku80 mutant generously provided by the Carruthers lab (58). We have also used a targeting cosmid to delete the Ku80 gene in the orginal TATi transactivator strain ((60)). Lastly we have introduced Cre-recombinase (100) under the combined regulation of a DD destablilazation domain (101, 102) and the tet-repressor (62) into the Ku80 mutant strain (58). In all systems we observe robust regulation (Fig. 6). These models have their inherent strengths and weaknesses that we cannot discuss in full here due to the constraint of the new page limit (please see our recent publications (41, 42, 103) for further technical detail). However, we would like to point out that their differences (in particular with respect to strength and timing of the transcriptional activity and the mechanism and extent of control) should permit us to match the particular requirements of our gene set.

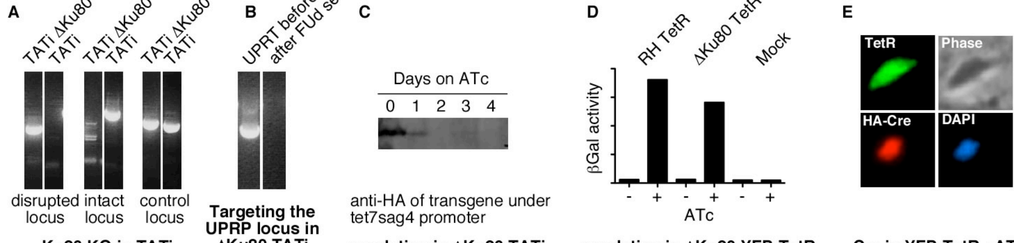


Figure 6: New parasite strains combining Ku80 deletion with regulated expression.

Our primary strategy will be to introduce a complementing cDNA-based clone of the target gene into the new TATi ΔKu80 strain (the Tet-R model will be our backup approach should we require higher expression levels). Stable transformation in ΔKu80 requires homologous recombination. To make this straightforward we have developed a plasmid into which a PCR product can be introduced by ligation independent cloning. This plasmid places the coding sequence under control of a tet-regulated promoter and introduces a c-terminal epitope tag. The entire expression cassette is flanked by sequences from the *T. gondii* uracil phosphoribosyltransferase gene and transfection of linearized plasmid into the TATi ΔKu80 results in targeting of the regulated expression cassette into the UPRT locus which can be conveniently selected for with 5-fluoro-uridine ((104, 105), Fig. 6B &C show targeting into the UPRT locus and regulated expression from this locus). Next we will target the

native locus using a modified cosmid (35) or if that is might not be available using a plasmid construct as detailed in our previous publications (24, 37).
After drug selection and cloning we will establish if loss of the gene interferes with parasite growth and plastid function. During the current funding period we have developed strong assays to measure apicoplast protein import (36, 37), apicoplast division and segregation (32), and apicoplast metabolism ((24, 35) and Nair & Striepen in preparation), and parasite growth (42, 106). Length restrictions prevent us from detailing these assays here. However, we would like to note that we now have numerous antibody reagents available that permit us to score apicoplast phenotypes in a simple Western blot (e.g. the absence of transit peptide maturation or the lipoic acid modification of apicoplast pyruvate dehydrogenase E2 subunit) or immunofluorescene assay (e.g. lack of plastid fission detected using antibody staining for plastid and inner membrane complex). This should be sufficient for an initial classification that could later be validated through more sophisticated pulse chase experiments or in vivo microscopy assays. In an independently funded project in collaboration with Dr. Malcolm McConville we are developing metabolomic approaches to dissect parasite metabolism. These assays would be marvelous tools to study the impact of mutants on the biochemistry of the apicoplast.

Develop a high-throughput strategy for the construction of conditional mutants in *T. gondii*
This part of the aim is not essential for the success of the remainder of the Aim. We can generate conditional mutants with high efficiency with existing technology (35-37, 61). However, a single step conditional mutagenesis protocol would considerably increase the throughput of our analyses, allow us to study “large” genes for which complementing minigenes might be challenging to construct, and might open the door for future efforts to use conditional insertional mutagenesis (42, 107) for forward genetic screens. The general idea of this approach is to separate the target gene from its native promoter by insertion of a targeting cassette by double crossover driven by homologous flanks. Along with a selectable marker and transcriptional termination elements this cassette will also contain a regulatable promoter element. The target gene can then be modulated using molecular features of the introduced promoter to produce a conditional phenotype. A number of recent advances now provide high homologous recombination frequencies and a choice of promoter elements and regulatory mechanisms. We will systematically evaluate three alternative models (tet-transactivator, tet-repressor and Cre-loxP) and we have already engineered “mutagenesis” strains that carry the Ku80 mutation and provide the required genetic background for regulated expression (Fig. 6). We have outlined tet-regulation in previous Aims, the Cre-loxP stragegy will employ a classic “floxed” promoter approach (a constitutive promoter element is flanked by short loxP recognition sites and will be excised upon activation of Cre-recombinase). While the tet-promoters have advantages when it comes to dissecting phenotypes, the Cre model might be superior as a discovery model. An obvious advantage is that we can use promoter elements from apicoplast proteins that precisely match the strength and timing of native transcription. We will evaluate the three approaches side by side by targeting the promoter regions of three apicoplast protein genes (ACP, $CDC48_{Ap}$, and apicoplast DNA polymerase). The rationale for this choice is that we have excellent reagents for these three proteins (antibodies to the native protein as well as epiotope tagged lines in the Ku80 background), they differ in their relative expression level, and they represent different apicoplast subcompartments. To make this comparison straightforward and to increase the throughput of our subsequent experiments we are using the gateway recombination system to construct the targeting vectors (108). This is a robust restriction enzyme independent approach that we have used extensively in the previous publications (109, 110). Should we find one of the three models to be suitable for the construction of conditional mutants we would make this our primary approach for genetic validation. We will establish a straightforward molecular pipeline from epitope tagging to mutant analysis by engineering a dedicated mutagenesis strain. This strain will carry 1) a Ku80 mutation, 2) the genes required for promoter regulation, and 3) fluorescent protein transgenes in the cytoplasm (YFP-YFP) and the apicoplast (FNR-RFP) to facilitate to measure mutant impact on parasite growth and plastid biology.

Overall we expect that this Aim will provide us with a broad collection of new apicoplast proteins. We will categorize proteins by localization and molecular function. We will highlight a subset of essential apicoplast proteins as potential targets and we will develop innovative approaches to the genetic analysis of parasite biology that will have a broaerd impact beyond our specific research goals.