Exploiting *Toxoplasma gondii MAF1* locus diversity to identify essential host proteins required for mitochondrial sequestration and manipulation

by

Matthew L. Blank

B.S. Biomedical Sciences, Texas A&M University, 2009

Submitted to the Graduate Faculty of

The Kenneth P. Dietrich School of Arts and Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2020

UNIVERSITY OF PITTSBURGH

DIETRICH SCHOOL OF ARTS AND SCIENCES

This dissertation was presented

by

Matthew L. Blank

It was defended on

March 3, 2020

and approved by

Mark Rebeiz, Associate Professor, Department of Biological Sciences

Allyson O'Donnell, Assistant Professor, Department of Biological Sciences

William Saunders, Associate Professor, Department of Biological Sciences

Martin Boulanger, Professor, Department of Biochemistry and Microbiology, University of Victoria

Dissertation Advisor: Jon P. Boyle, Associate Professor, Department of Biological Sciences

Exploiting *Toxoplasma gondii MAF1* locus diversity to identify essential host proteins required for mitochondrial sequestration and manipulation

Matthew L. Blank, Ph.D.

University of Pittsburgh, 2020

Recent genomic comparisons identified multiple expanded loci in *Toxoplasma gondii* that are unique compared to close Apicomplexan relatives. One of these loci, *mitochondrial association factor 1* (*MAF1*), encodes distinct paralogs of secreted dense granule effector proteins, some of which mediate the host mitochondrial association (HMA) phenotype. MAF1b drives HMA, MAF1a does not. Through sequence and functional analysis of multiple MAF1 paralogs, we have identified regions of the protein that have undergone paralog-specific selection-driven diversification. Using structure and alignment-guided site-directed mutagenesis of MAF1b and MAF1a, we identified three critical residues in the C-terminal helix that are required for HMA. Using this MAF1b mutant as a new tool to probe the function of MAF1b and HMA in *T. gondii* interactions, we performed an unbiased quantitative mass spectrometry screen comparing co-IP samples of MAF1b, MAF1a, and MAF1b mutant parasite infections. Of the 1,360 proteins identified in all samples, 13 candidate proteins were uniquely enriched in the MAF1b samples compared to both MAF1a and MAF1b mutant samples. Of these 13 candidate host binding proteins, nine play roles in mitochondrial biology. RNAi knockdown of each of the nine candidates followed by either Type II:MAF1b infection or GFP-MAF1b expression identified TOM70 and its mitochondrial chaperone, HSPA9, as partners required for MAF1b-mediated HMA. TOM70 enrichment at the *T. gondii* vacuolar membrane and parasite exclusion of the ER at this interface

illustrate an intricate manipulative strategy on the part of parasite. Both TOM70 and HSPA9 are implicated in specific ER-mitochondrial contact site functions like immune modulation, mitochondrial dynamics, autophagy, and calcium homeostasis. The requirement of both TOM70 and HSPA9 could explain key phenotypes previously described in an HMA+ *T. gondii* infection such as an *in vivo* growth advantage, modulated immune response, and mitochondrial fusion around the vacuole. Taken together, the neofunctionalization of the *MAF1* locus drove the evolution of an intermolecular network capable of mediating intimate interactions between host mitochondria and the *T. gondii*-containing vacuole.

Table of contents

List of tables

List of figures

Preface

I would like to thank many people who have mentored and guided me along this journey. First and foremost, my wife and partner, Emily, who took me up on the crazy idea of adventuring outside of Texas to Pittsburgh. Thank you for your continual support over the past five years of my graduate studies whether that was offering advice, listening to boring practice talks, or being alone at home on many weekends when I went into the lab. To my son, Ellis, who has taught me the importance of always being curious and exploring new adventures with confidence. Thank you for bringing so much joy to my life over the past year. Also, thank you to my parents, Steve and Debbie, for raising me to value education, my faith, and being a kind human being. Thank you to my sister, Erin, for engaging with me in scientific conversation and providing both academic and career advice. Thank you to my brother-in-law and nephew, Michael and James, for providing me with innovative desk puzzle toys to distract me from writing my dissertation.

I would also like to thank my advisor and mentor, Dr. Jon Boyle, who took me into his lab and did far more than simply support me as a scientist. Thank you for providing a safe environment to ask questions without feeling judged and thank you for treating me like a fellow colleague. You provided advice when needed, but also allowed me to figure out things on my own. Also, I will always cherish our open-door political conversations. You have not only made me a better scientist, but also made me a kinder and more open human being. None of this work would be possible without my fellow lab colleagues, particularly, my fellow graduate students, Dr. Elizabeth English, Rachel Coombs, and Sarah Sokol Borrelli. Thank you for scientific and emotional support over the past five years. I want to thank my undergraduate researcher, Mary Morcos, who executed

many experiments in the following pages. Thank you for your help in the lab as well as your humor.

Thank you to my committee members for your advice and guidance. Thank you to my cohort for coffee breaks, lunch appointments, and, of course, scientific discussions. I fell deeper in love with teaching over the course of my Ph.D. and that is only because of the incredible mentorship of Drs. Erica McGreevy, Debbie Chapman, Valerie Oke, Laura Zapanta, and Kate Gardner.

I would also like to thank all of the collaborators who have contributed to this work including Dr. Martin Boulanger and Cameron Powell at the University of Victoria. Additionally, thank you to the University of Pittsburgh Core Facilities including the Center for Biological Imaging, particularly, Mara Sullivan, and the Biomedical Mass Spectrometry Center. Good science is only made possible through collaboration and I am grateful for all of those who have contributed to the never-ending work of MAF1. Thank you.

1.0 Introduction: effector variation at tandem gene arrays in tissue-dwelling coccidia: who needs antigenic variation anyway?

The contents of this chapter are derived from Blank, M.L., and Boyle, J.P. (2018). Effector variation at tandem gene arrays in tissue-dwelling coccidia: who needs antigenic variation anyway? Curr Opin Microbiol *46,* 86-92. Additional introduction material is included.

1.1 Summary

Locus expansion and diversification is pervasive in apicomplexan genomes and is predominantly found in loci encoding secreted proteins that interact with factors outside of the parasite. In order to understand the impact of each of these loci on the host, it is critical to identify and functionally characterize their protein products, but multi-copy loci are often refractory to genome assembly. As a means to introduce the content in this volume, I will focus on *Toxoplasma gondii* and its nearest relatives to highlight the known impact of duplicated and diversified loci on the host-pathogen molecular arms race. I also describe current tools used for the identification and characterization of these loci, and discuss the most recent examples of how gene-expansion driven diversification can lead to novel gene function.

1.2 Background

Locus duplication events are an essential driver of genetic innovation, particularly with respect to the emergence of novel traits [1-5]. In the event that a given gene is perfectly duplicated, the most common fate is maintenance at low allele frequency or loss on the population level due to neutral impact, depending on the effective population size [6, 7]. However, in cases where gene dosage leads to improved fitness, the altered allele can be maintained in the population as either standing genetic variation or become fixed via drift or selective sweep [5, 8, 9]. If increases in gene dosage remain advantageous, the additional gene copy can be maintained under the same selective constraints as the original copy [8, 9]. However, selection-driven diversification can result in more dramatic changes in gene function, including sub- and neo-functionalization [3, 4], where functions are distributed across distinct copies or new functions arise. While the importance of Gene Expansion-Driven Diversification (GEDD) was known long before complete genome sequences were made available for any organism, in the post-genomic era structural variation analysis forms a consistent part of any genome-wide analysis, and variations in copy number have been extensively linked to a variety of phenotypes in humans and other organisms [3, 9-12].

GEDD is an appealing mechanism for virulence factor evolution in pathogens, given that duplicate gene copies can be used as fodder in the molecular arms race to counter host evolutionary adaptations. In Apicomplexans like *Plasmodium falciparum* GEDD has clearly played a role in the emergence of the expanded *var* family of genes which are the primary loci mediating antigenic variation in this organism [13, 14]. Outside of the Apicomplexa, the *T. brucei* genome contains a vast pool of extensively duplicated variant surface glycoprotein (VSG) genes [15] that have emerged via GEDD and expression of single of a small number of VSG genes underlie the canonical waves of parasitemia associated with *T. brucei* infection [16]. These classic examples

demonstrate how GEDD can provide the tools required to avoid parasite clearance by host adaptive immune responses, promoting parasite survival and ultimately parasite transmission.

We are very interested in the role that GEDD has played in the emergence of either refined or novel gene functions in apicomplexan parasites like *Toxoplasma gondii*, *Hammondia hammondi*, and *Neospora caninum*, tissue-dwelling coccidians responsible for billions of infections worldwide [17, 18]. In contrast to *T. brucei* and *Plasmodium spp.*, tissue-dwelling coccidians like *T. gondii* do not appear to have a sophisticated means of antigenic variation. However, our work and that of others [10, 12, 19-21] has shown that these parasite species do have a large collection of effectors, many of which have undergone extensive GEDD. Our work has also shown that the impact of these loci on infection outcome is most often NOT due to changes in gene dosage or temporal variation in expression, but rather on selection-driven changes in protein function across individual gene paralogs that are all expressed simultaneously. Others have thoroughly examined the extent of gene family expansion in *T. gondii* and other organisms and we refer the reader to these excellent papers if they are interested in gene duplicates that are distributed throughout the genome (e.g., [19, 22]). In this dissertation, I will discuss recent examples of the impact of GEDD on *T. gondii* interactions with its host, with an exclusive focus on *local* tandem gene expansions followed by diversification that have led to refined and/or completely distinct functions in individual paralogs.

1.3 Genome-wide characterization of expanded loci in *T. gondii* and its relatives

T. gondii genomics is very advanced. The genome for this parasite as well as many of its near and far relatives have been sequenced and extensively annotated, including the nearest extant

relative of *T. gondii*, *Hammondia hammondi* [19, 23], and the less-related cattle parasite *Neospora caninum* [22]. These annotated assemblies are available for download and query on ToxoDB.org [24]. We and others [10, 19] have used raw genome sequence data from these sequencing projects (mostly Sanger technology for *T. gondii*/*N. caninum* and a combination of 454 and Illumina short-read data for others) to identify genomic regions likely to contain tandem gene expansions, as evidenced by increased sequence coverage. After removal of low complexity reads with short repeats, we are then able to identify tandem gene arrays in multiple genomes, as we did for *T. gondii*, *N. caninum*, and *H. hammondi* [10], exploiting the high level of gene-by-gene synteny in these organisms. This pipeline was crucial for identification of loci with increased coverage in *T. gondii* compared to *N. caninum* and *H. hammondi*, and facilitated downstream studies aimed at understanding the functional impact of GEDD on the evolution of novel traits in this organism [10].

Our work and that of others thus far clearly shows that tandemly expanded loci are quite common in *T. gondii* and its near relatives and enriched for genes with the capacity to encode factors that work outside of the parasite rather than within it [10, 19]. During and after host cell invasion (Fig 1-1), *T. gondii* and its relatives secrete a suite of effector proteins from specialized secretory organelles (Fig 1-1) [25, 26]. Forty-two of the 53 identified expanded loci identified in *T. gondii* were predicted secretory proteins and 29 contained an N-terminal signal sequence [10]. Some of these loci had been previously annotated and studied because of their significant impact on pathogen virulence and host interactions. Members of the *ROP2* superfamily like *ROP5*, *ROP2* and *ROP4/7* are tandemly expanded [25, 27-29]. In most cases the impact of diversification in individual paralogs is unclear, with the exception of *ROP5* (discussed below), where there is clear evidence linking locus-specific diversification to gene function [29, 30]. *T. gondii* dense granule

proteins like Mitochondrial Association Factor 1a,b (MAF1a,b) mediate the bulk of their effects on the host-cytoplasmic side of the parasitophorous vacuole (Fig 1-1, green arrowheads) [12, 31]. The *SAG1*-related surface antigen family (*SRS* genes) are expressed at the parasite surface and have significant levels of local, tandem GEDD in both *T. gondii* and *N. caninum*. Select members of this family are known to play a critical role in host immune evasion [10, 32], but little else is known about their contribution to host interactions and/or parasite biology. Even though the majority of these expanded loci do not have a predicted function, they provide a curated foundation to identify potential and exciting host-pathogen genes for study [10].

In contrast to extensive tandem locus expansion at host-interacting loci, examples of duplication of genes encoding cytoskeletal constituents or metabolic enzymes are more rare, likely due to tighter constraints on gene dose [19]. One exception to this can be found in two multicopy gene cassettes encoding inner membrane complex (IMC) proteins which localize to the apical complex of *T. gondii* and maintain rigidity of the conoid structure [26, 33, 34]. Two of these proteins (IMC4 and IMC1; *TGME49_231630* and *TGME49_231640*, respectively) are derived from a gene duplication/inversion event on Chromosome VIII such that, unlike most tandem gene arrays they are found in a head-to-head orientation. Interestingly, the IMC8 and IMC5 genes are arranged in a perfectly analogous fashion on chromosome X (*TGME49_224520* and *TGME49_224530*) suggesting that there was an ancestral duplication and translocation event at this locus that predated *N. caninum*, *H. hammondi* and *T. gondii* [19]. Absent their role in cytoskeletal rigidity the significance of the diversification across these IMC gene paralogs is unclear, although given their deep ancestry it is highly likely that they have evolved refined or even novel functions within the IMC. The NTPase locus is also characterized by a tandem duplication (*TGME49_277240* and *TGME49_277270*) and the proteins encoded by these genes

(NTPases I and II) are 59% identical to one another [19]. This duplication is difficult to trace across other *T. gondii* strains and related species, likely due to issues with poor assembly. Interestingly, these NTPases are not cytoplasmic in the parasite but rather are secreted into the parasite-containing vacuole. Their role in *T. gondii* biology is unclear, in that they do not affect virulence in mice and do not subvert the murine immune response [35].

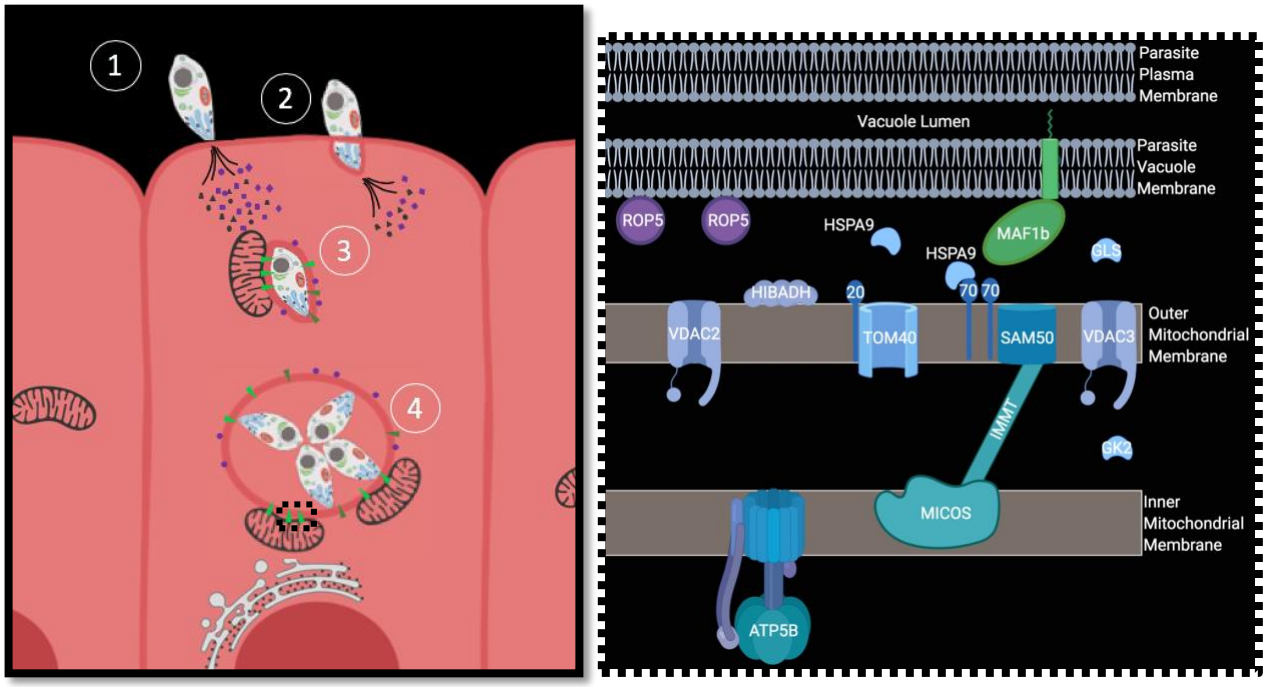


Figure 1-1 Events of *T. gondii* invasion and secretion of protein effectors.

(1) *T. gondii* attaches to plasma membrane and secretes a suite of effector proteins some of which originate in the *T. gondii* rhoptry organelles (i.e. ROP5). (2) *T. gondii* continues to secrete effectors from secretory organelles including rhoptries and dense granules. The parasite gains entry by establishing a moving junction which allows for the invagination of the plasma membrane and exclusion of host cell membrane proteins. (3) After entry, the parasite-containing vacuole (PV) associates with host mitochondria facilitated by the dense granule secretory protein, MAF1b (light green triangles), which is a predicted transmembrane protein. MAF1b binds to mitochondria via direct molecular interactions with mitochondrial proteins which were identified in the quantitative mass spectrometry screen (Chapter 3). MAF1a (dark green triangles) also localizes to the PV membrane but does not drive association with host

mitochondria. Upon formation of the PV, ROP5 (purple circles) localizes to the cytosolic face of the PV membrane. (4) *T. gondii* replicates within the PV until the parasite responds to environmental or growth stressors and signals for parasite egress. Dotted line box in left panel refers to zoomed in panel on right to illustrate predicted MAF1b orientation in membrane and relation to mitochondrial proteins.

1.4 Expanded loci in *T. gondii* and its close relatives are often found at the host-parasite interface

Secreted proteins from the rhoptry organelles of *T. gondii*, *N. caninum* and *H. hammondi* have been extensively studied because this class of proteins play key roles in pathogen virulence, particularly ROP18 and ROP5 (both members of the ROP2 superfamily) [25, 27, 36]. ROP5 is a pseudokinase that traffics to the PVM (Fig 1-1, purple circles) [28, 29] and is critical for *T. gondii* virulence. *T. gondii* harbors a diverse and expanded *ROP5* locus that has diverged both within and between strains [29], making it one of the more dramatic examples of tandem locus expansion followed by diversification. Within each strain, the *ROP5* locus encodes three distinct isoform clades based on amino acid sequences: A, B, and C [28, 29], suggesting that duplication was followed by selection-driven diversification at this locus [29]. In support of this, ectopic expression of individual ROP5 paralogs in *ROP5* null parasites and infection into mice show full recovery of the virulence phenotype only after complementation with both a ROP5A and B paralogs (but conversely not with two copies of paralog A) [29]. Based on quantitative genetic data the type I and III alleles are much more potent with respect to virulence in the mouse compared to the type II allele [29, 30, 37], although the root cause of this at the sequence level is unclear. Given the diversity present at this locus it is likely that multiple factors determine the phenotypic differences

between these loci, including the exact composition of the locus with respect to the A, B and C paralogs.

It is not just *T. gondii* that has a set of species-specific GEDD loci that are involved in mediating interactions with the host. Our analysis on the *N. caninum* genome identified uniquely expanded loci in this species as well [10], and we observed a remarkable lack of overlap in the putative biological function of expanded loci in this species compared to *T. gondii*. Specifically, nearly half of the expanded loci predicted to encode secreted proteins were clear members of the SRS superfamily of surface antigens [10, 22]. SRS proteins play a major role in parasite attachment and invasion [38], they are developmentally expressed, often the target of host antibody responses, and in some cases are required for establishing latent infections in the host [32, 39]. While SRS subfamilies are scattered throughout the genome (similar to members of the ROP2 superfamily), in *N. caninum* they are extensively duplicated in tandem [10, 21]. For example, the *NCLIV_002220* locus has five tandemly duplicated and diversified copies (ranging from 57-88% similar identities) as compared to *T. gondii* maintaining only one copy [10, 22]. While *N. caninum* harbors clusters of expanded SRS genes throughout the genome, it is rare that each copy is expressed in comparison to *T. gondii* [22]. The significance of certain copies being on or off remains elusive, but a proposed hypothesis describes that this differential expression leads to the restricted host range of *N. caninum* [22].

1.5 Neofunctionalization in the locus driving interactions between the parasite and host mitochondria

The expanded and diversified gene, *mitochondrial association factor 1* (*MAF1*), provides an example of the characterization of an expanded locus identified originally by QTL mapping and screening for expanded loci [10]. The *MAF1* locus (*TGME49_220950* and *TGME49_279100*) drives expression of unique classes of MAF1 paralogs ranging in 59% to 95% similar identity with one another.

MAF1 copy number varies across strain and species [10]. It is estimated that *N. caninum* harbors 1 or 2 identical copies of *MAF1* while *H. hammondi* has 2 diverse copies suggesting that a duplication and diversification event occurred during *Hammondia* speciation (Fig 1-2) [12]. *T. gondii* bears 6-10 copies of MAF1 depending on strain type (Fig 1-2) [12]. Some *MAF1* copies mediate the characteristic phenotype of host mitochondrial association (HMA) [12]. HMA has been recognized for quite some time as the intimate association of the host mitochondria with the parasite-containing parasitophorous vacuole membrane (PVM) [31, 40-42]. Attachment of the vacuole to the mitochondria is most readily quantified using transmission electron microscopy (TEM) because the distance between the vacuole and mitochondria is 12.04 nm ± 3.05 nm and starkly different than unattached mitochondria, which can be more difficult to distinguish using epifluorescence microscopy [41]. *T. gondii* types I and III are HMA-competent while Type II is HMA-incompetent [12]. The MAF1a and MAF1b paralogs were assembled into clades based on the presence (MAF1b) or absence (MAF1a) of a proline rich region between the putative transmembrane domain and C-terminus of the protein [12]. Interestingly, paralogs belonging to the MAF1b paralog clade are solely capable of complementing HMA in Type II *T. gondii* to near wildtype Type I levels while MAF1a paralogs do not [12]. These data suggest the requirement of

only the MAF1b paralogs in driving HMA [12]. Important to this finding is that *N. caninum* is HMA-incompetent and harbors only a MAF1a copy while *H. hammondi* is HMA-competent and has both MAF1a and MAF1b.

Crystallographic structures reveal that the MAF1a and MAF1b paralogs are highly similar to one another in their C-termini and adopt an α/β globular structure homologous to ADP-ribose (ADPr) binding macro-domains [11]. MAF1a and MAF1b bind ADPr with micromolar affinity, suggesting this site remains as a functional structure on both MAF1 paralogs (Fig 1-2) [11]. While the role of this macrodomain is unknown to date, there is no direct evidence linking HMA to ADPr coordination [11].

In contrast to the macrodomain, there are other regions in the solved MAF1 structure that are divergent. One such region is a charged basic patch in the last 16 C-terminal residues of MAF1a (Fig 1-2) [11, 12]. Upon mutation of the aligned uncharged residues in MAF1b to the charged residues in MAF1a, HMA is completely lost (Fig 1-2) [11]. The C-terminus of MAF1 encompassing these key residues has acquired significant diversity between the two paralog clades over time while the ADPr site has maintained somewhat similar binding pockets and is shared with the ancestral MAF1 paralog (Fig 1-2) [11]. This work suggests a modularity to MAF1 where only the components driving HMA were selected for over time, highlighting the significance of this neofunctionalized expanded locus and it evolutionary role in exploiting the host [11].

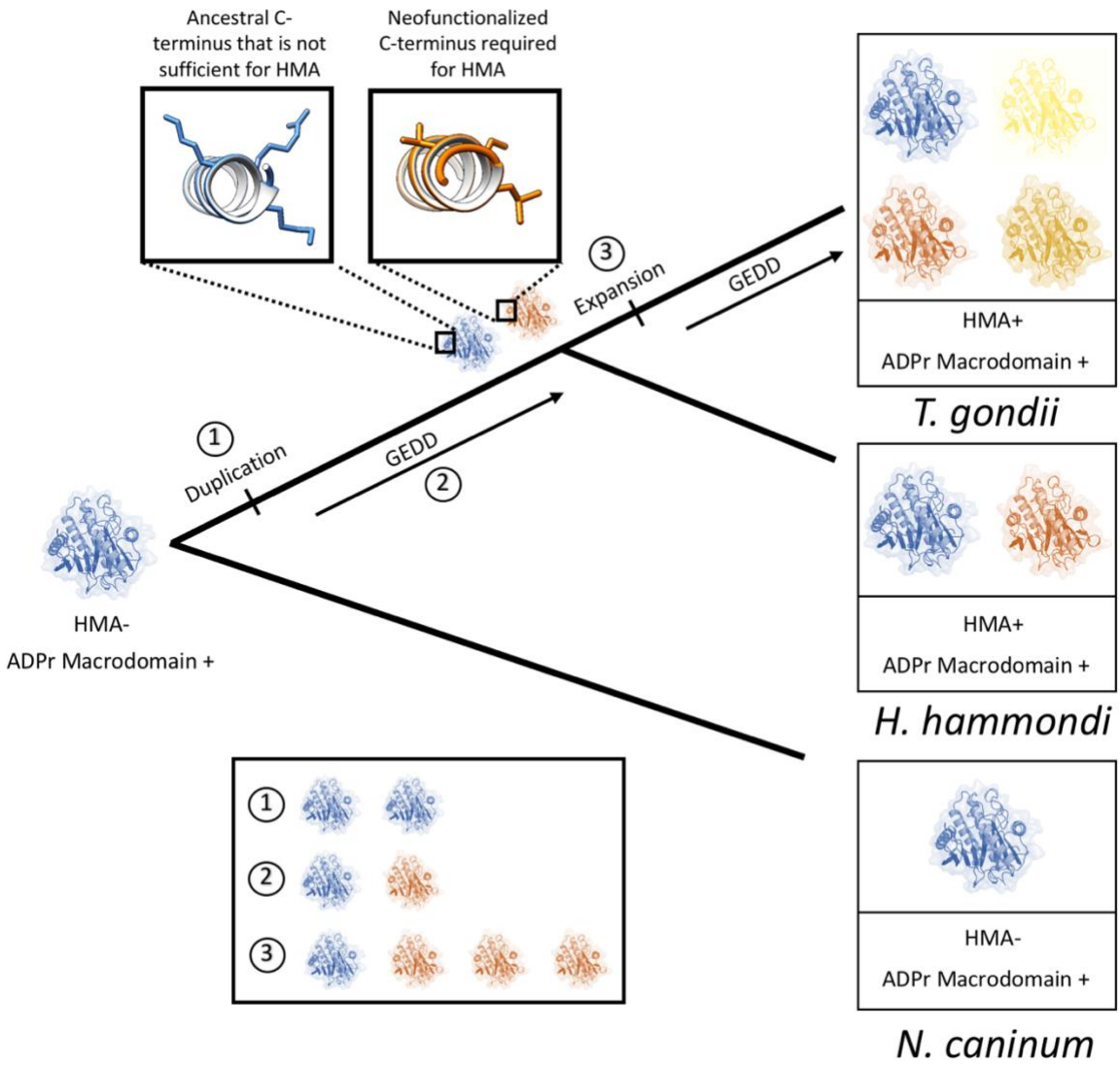


Figure 1-2 Stepwise evolution and neofunctionalization of *MAF1* locus in *T. gondii* and close relatives.

Our model for the stepwise evolution of the *MAF1* locus in *T. gondii* and its close relatives, *N. caninum* and *H. hammondi*. The MAF1a and MAF1b crystal structures [11] are in blue and orange, respectively. MAF1 maintains a vestigial ADPr binding macrodomain while duplication, expansion and GEDD events give rise to the neofunctionalized version of MAF1b driven at least in part by mutations of highly charged residues in the MAF1 C-terminus. Each significant change to the locus is marked and numbers correspond to MAF1 protein products in the key. MAF1b paralogs undergo subsequent diversification events following step 3 leading to the diverse shades of orange/yellow in the final *T. gondii MAF1* locus.

1.6 Implications of the neofunctionalized *MAF1* locus on host mitochondrial biology

While recent work highlights the intricacies of the neofunctionalized *MAF1* locus [11, 12], questions remain as to the physiological significance of this expanded and diversified gene. The selective pressure to maintain each of the *MAF1* copies suggests that each copy plays a role in *T. gondii* pathogenicity through its interactions with the host mitochondria. Pathogens associating with host organelles, particularly the mitochondria, is a long recognized phenotype in microorganisms like *Legionella pneumophila* and *Chlamydia psittaci* [43-45]. In the case of *L. pneumophila* the recently identified secreted pathogen effector, MitF, drives the association between the Legionella-containing vacuole and the host mitochondria [44, 46]. While MitF facilitates the fragmentation of attached mitochondria, the exact molecular mechanism leading to fragmentation is unknown. Similar shifts in mitochondrial dynamics exist in a *T. gondii* Type I (RH) infection where the attached mitochondria fuse around the vacuole during the first 12 hours of infection, followed by fragmentation at 24 hours [47]. In the cases of *L. pneumophila* and *T. gondii*, the exact pathways leading to mitochondrial manipulation remain elusive. The multi-copy nature of the *MAF1* locus provides an ideal system with built-in controls (MAF1b and MAF1a) to test hypotheses of what might be happening to elicit the phenotypic changes in mitochondrial biology leading to the observed alteration in parasite fitness.

A growing body of research in the field of mitochondrial biology and cellular bioenergetics continues to add to the mounting list of mitochondrial functions in the eukaryotic cell. Aside from the more traditional metabolic roles, mitochondrial proteins are initiators of cellular apoptosis pathways and release cytochrome c from Bak/Bax oligomerized pores leading to the formation of the apoptosome and caspase activation [48-51]. Mitochondria also function as essential initiators of the innate immune response triggered by dimerization of the outer mitochondrial membrane

protein, MAVS, and activation of downstream transcription factors, IRF3/7 and NF-κB [49]. Additionally, mitochondria have long been recognized for their role in calcium storage and trafficking via the endoplasmic reticulum, thus direct regulators of cell autophagy [48, 51-54]. Interestingly, recent work focuses on the interconnectedness of these mitochondrial-initiated pathways at physical locations in the cell where the mitochondria and ER come in contact with one another. These ER-mitochondrial contact sites (formerly known as MAMs) are hotspots for mitochondrial-initiated immune signaling, apoptosis, calcium flux, and autophagy [48, 50, 55, 56]. Key proteins at these sites include voltage dependent anion channel 1 (VDAC1), heat shock protein family A member 9 (HSPA9), inositol tri-phosphate receptor 3 (IP3R3), and translocase of the outer mitochondrial membrane receptor 70 (TOM70) [50, 51]. Each of these proteins is highly enriched at the ER-mitochondria contact sites and all are shown to interact with one another to maintain appropriate calcium flux [50]. Cells lacking the TOM70 receptor exhibit a significant decrease in division rate, decrease in mitochondrial activity, and increase in autophagy [50]. Additionally, MAVS clusters at these ER-mitochondrial contact sites in response to viral infection and triggers TOM70 recruitment and downstream cytokine production [54, 57, 58]. Needless to say, the ER-mitochondrial contact sites are hubs for mitochondria-initiated processes paramount for overall cellular health. Therefore, it is not surprising that parasites like *L. pneumophila* and *T. gondii* can interact with and manipulate host mitochondria. Specifically, gene expansion-driven diversification of the *MAF1* locus in *T. gondii* provides the perfect model to answer consequential questions surrounding intricate host-pathogen interactions through the lens of evolutionary biology.

1.7 Conclusion and perspective

The examples above highlight a large body of work describing GEDD-driven protein products involved in the host-pathogen molecular arms race. The *MAF1* locus is an excellent example of the importance of GEDD, in that it harbors paralogs that have maintained an ancestral macro domain while only certain paralogs are endowed with the ability to mediate host mitochondrial association (HMA) [11, 12]. In this case, GEDD led to the emergence of a novel function [12]. Similarly, the *ROP5* locus harbors 3 broad classes of paralogs which are synergistic in their impact on *T. gondii* virulence in mice [20]. Whether this is due to novel functions in each of the paralog families is unknown, but the impact of GEDD in the emergence of these paralog families within the tandem gene cluster is quite clear. The total number of loci that have undergone GEDD is difficult to determine since tandem gene arrays are almost always poorly assembled, but there are likely between 20-30 in the *T. gondii* genome alone [10]. The emergence of single molecule long read technologies such as PacBio SMRT cells and Oxford Nanopore DNA sequencing holds great promise to determine the exact sequences of these tandem gene arrays so that the precise amount of sequence variation between paralogs can be determined. As these methods are applied we anticipate refining our understanding of the extent of GEDD at loci like *MAF1* and *ROP5*, and also that GEDD is equally pervasive at other loci. This will greatly facilitate our efforts at better understanding the role of such loci in tissue dwelling coccidians like *T. gondii* and its near relatives, specifically with respect to how individual paralogs might differ in their interactions with the host.

2.0 A *Toxoplasma gondii* locus required for the direct manipulation of host mitochondria has maintained multiple ancestral functions

The contents of this chapter are taken from Blank ML, Parker ML, Ramaswamy R, Powell CJ, English ED, Adomako-Ankomah Y, et al. A *Toxoplasma gondii* locus required for the direct manipulation of host mitochondria has maintained multiple ancestral functions. Mol Microbiol. 2018. I performed all experiments and writing for sections 2.3.1, 2.3.4, 2.3.5, 2.3.6, 2.3.7, 2.4.1, 2.4.2, 2.4.3, 2.4.4, and 2.5.5-2.5.12. I am including sections 2.3.2, 2.3.3, and 2.5.1-2.5.4 to provide clarity and illustrate the complete narrative.

2.1 Summary

The *Toxoplasma gondii* locus *mitochondrial association factor 1* (*MAF1*) encodes multiple paralogs, some of which mediate host mitochondrial association (HMA). Previous work showed that HMA was a trait that arose in *T. gondii* through neofunctionalization of an ancestral MAF1 ortholog. Structural analysis of HMA-competent and incompetent MAF1 paralogs (MAF1b and MAF1a, respectively) revealed that both paralogs harbor an ADP ribose binding macro domain, with comparatively low (micromolar) affinity for ADP ribose. Replacing the 16 C-terminal residues of MAF1b with those of MAF1a abrogated HMA, and we also show that only three residues in the C-terminal helix are required for MAF1-mediated HMA. Importantly these same three residues are also required for the *in vivo* growth advantage conferred by MAF1b, providing a definitive link between *in vivo* proliferation and manipulation of host mitochondria. Co-

immunoprecipitation assays reveal that the ability to interact with the mitochondrial MICOS complex is shared by HMA-competent and incompetent MAF1 paralogs and mutants. The weak ADPr coordination and ability to interact with the MICOS complex shared between divergent paralogs may represent modular ancestral functions for this tandemly expanded and diversified *T. gondii* locus.

2.2 Introduction

Tight associations between pathogen-containing vacuoles and host organelles such as mitochondria have been described in a variety of intracellular pathogens, including *Chlamydia psittaci* [45], *Legionella pneumophila* [43], *Hammondia hammondi* [12] and *Toxoplasma gondii* [31, 40, 41]. While these phenotypes have been known for decades, the underlying molecular mechanisms are poorly understood. In most cases the pathogen molecules required for organellar association have not been identified, nor have their cognate binding partners in the host. This has hindered our ability to understand the relevance of this intimate association between the pathogen-containing vacuole and the host mitochondrion to infection outcome.

In *T. gondii*, we and others have identified the parasite locus that is required for HMA, *Mitochondrial Association Factor 1B* (*MAF1b*), and shown that expression of MAF1b increases cytokine signaling [31] during the acute phase of *in vivo* infections. In infected cell lysates and cells expressing MAF1b ectopically, MAF1b protein interacts with host mitochondrial outer membrane proteins belonging to the MICOS complex, which may be involved in the ability of MAF1b to mediate HMA [59]. During the acute phase of mouse infections, MAF1b-expressing

parasites outcompete their MAF1b-null counterparts [12], implying that MAF1b plays an important role in determining infection outcome.

The *T. gondii MAF1* locus encodes multiple tandemly duplicated paralogs that vary both in sequence and copy number across *T. gondii* strains [10, 12]. *Tg*MAF1 paralogs fall into two broad groups, which we have defined as 'A' and 'B' based on residue percent identity [12]. All non-pseudogenized MAF1 genes sequenced to date contain a signal peptide, a transmembrane (TM) domain and a large C-terminal region that lacks identifiable sequence homology to any known proteins ([12, 31] and Fig 2-1A). MAF1 paralogs harbor a repetitive, proline-rich region between the putative TM and C-terminal region that broadly distinguishes the 'A' and 'B' paralog groups. In genetic complementation experiments, only MAF1b paralogs are capable of complementing HMA- Type II strains, while complementation with MAF1a has no effect on HMA [12]. Numerous polymorphisms further distinguish *Tg*MAF1RHa1 and *Tg*MAF1RHb1, including in the C-terminus, which is more divergent across paralogs compared to the N-terminus [12]. Since MAF1b paralogs are present only in those species that are capable of mediating HMA and all other strains only harbor MAF1 paralogs most similar to the non-functional "A" copies, we hypothesized that the ability of MAF1b to intimately interact with host mitochondria evolved by neofunctionalization of an ancestral version of MAF1. While the ancestral role of MAF1a has yet to be determined, these non-functional (with respect to HMA) ancestral paralogs represent a robust comparative tool to further probe the mechanism, and ultimately the function, of HMA in *T. gondii*.

In the present study we first determined the overall structures of representative members of the A and B paralog groups using X-ray crystallography. These data allowed us to map the detailed structural differences between the A and B paralogs and identify specific residues within the C-terminal helix of MAF1a and MAF1b, that through mutagenesis studies, we showed are

required for MAF1b to mediate HMA. We further leveraged these mutants to provide the first direct evidence that the C-terminus of a MAF1b paralog is essential, not only to HMA itself, but to the *in vivo* selective advantage conferred by HMA+ *T. gondii*. We have also uncovered two putative ancestral functions for the *MAF1* gene family that are shared by MAF1a and MAF1b paralogs, providing further support for the theory that mitochondrial association arose in *T. gondii* via neofunctionalization of an ancestral version of *MAF1*.

2.3 Results

2.3.1 The *Tg*MAF1RHb1-specific proline-rich domain is dispensable for MAF1 function in HMA

Type II *T. gondii* strains (including *Tg*ME49 and PRU) are HMA-negative and express undetectable levels of *Tg*MAF1RHb1 [12]. Complementation of *Tg*ME49 and the near relative *Neospora caninum*, with *Tg*MAF1RHb1 but not *Tg*MAF1RHa1, confers the HMA phenotype [12]. To identify regions of *Tg*MAF1RHb1 that are required for HMA we compared all sequenced MAF1 paralogs [12]. Sequences were aligned using Clustal-Omega and visualized in JalView (Fig 2-1A). Consistent with previous observations, the sequences cluster into two main groups [12], distinguished by the absence ('A' group) or presence ('B' group) of a proline-rich stretch between the putative transmembrane and the C-terminal region. To investigate the significance of the prolines in this region of *Tg*MAF1RHb1, prolines 152-157 and 159-164 were mutated to alanine residues. However, these mutant paralogs were still capable of complementing HMA in *Tg*Me49 parasites (Fig 2-1B). In an effort to disrupt the proline-rich region using an alignment-guided

approach, separate mutations were made to the region flanking either side of the proline-rich region. Additionally, a third construct was made where the entirety of the proline-rich region, residues 131-174, were replaced with the aligned sequence in *Tg*MAF1RHb0 (HMA-) which is nearly identical to *Tg*MAF1RHa1 in this region (Fig 2-1A). Consistent with our point mutation analyses, this particular chimeric construct was still capable of mediating HMA in *T. gondii* strain *Tg*ME49 (Fig 2-1C). Additionally, work investigating the phosphoproteome following *T. gondii* infection identified significant MAF1 phosphorylation after secretion into the host cell [60]. To test the role of possible phosphorylation of the serines found within the MAF1b-specific proline-rich region, we mutated *Tg*MAF1RHb1 serine 158 to alanine. This mutation did not disrupt the ability of *Tg*MAF1RHb1 to confer HMA in *Tg*Me49 parasites (Fig 2-1B). Overall, these data demonstrate that the proline-rich region is dispensable for HMA, despite its presence in all functional MAF1b paralogs.

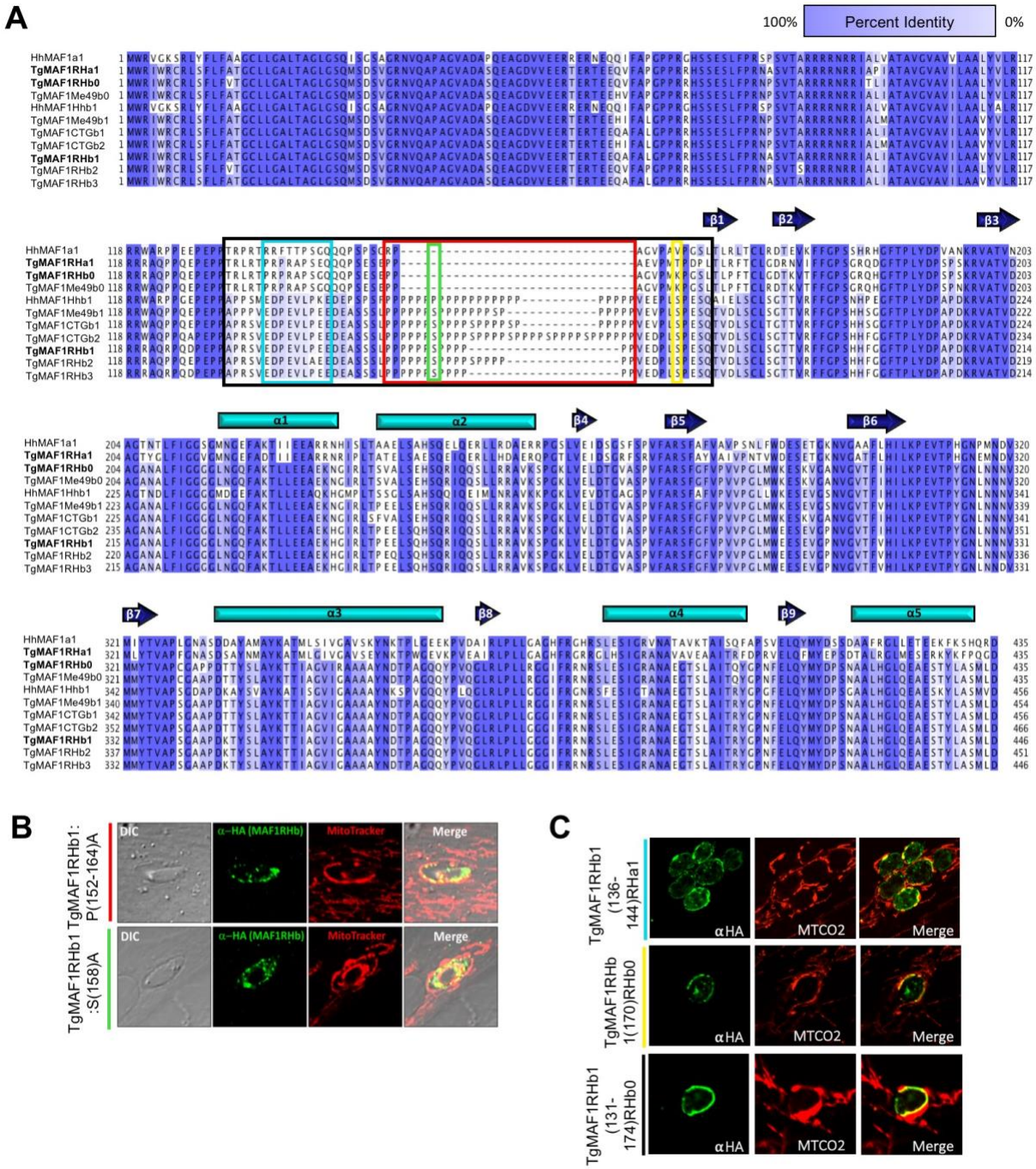

A
100%
Percent Identity
0%
HhMAF1a1
TgMAF1RHa1
TgMAF1RHb0
TgMAF1Me49b0
HhMAF1Hhb1
TgMAF1Me49b1
TgMAF1CTGb1
TgMAF1CTGb2
TgMAF1RHb1
TgMAF1RHb2
TgMAF1RHb3
β1
β2
β3
α1
α2
β4
β5
β6
β7
α3
β8
α4
β9
α5
B
TgMAF1RHb1 TgMAF1RHb1:P(152-164)A
TgMAF1RHb1 :S(158)A
DIC
α–HA (MAF1RHb)
MitoTracker
Merge
C
TgMAF1RHb1 (136-144)RHa1
TgMAF1RHb1(170)RHb0
TgMAF1RHb1 (131-174)RHb0
αHA
MTCO2
Merge

Figure 2-1 Proline-rich region of *Tg*MAF1RHb1 is not required for HMA.

(A) Alignment of MAF1 isoforms from T. gondii and H. hammondi based on percent identity. Dark blue and white indicate 100% and 0% percent identity, respectively. Alignments were visualized in JalView after alignment using ClustalOmega. Bold sequences refer to isoforms utilized for mutational analysis. Colored boxes correspond to the boundaries of the indicated mutation. Domain architecture is indicated above alignment and corresponds to Figure 2-2B. (B) HFFs were labeled with MitoTracker and infected with *Tg*Me49 parasites expressing an HA epitope tagged *Tg*MAF1RHb1 isoform with a disrupted proline-rich domain through replacement of prolines (P) or serines (S) with alanines (A). Cells were fixed at 18 hpi and visualized by confocal microscopy. (C) HFFs were infected with *Tg*Me49 parasites expressing an HA tagged mutated *Tg*MAF1RHb1 isoform. Site-directed mutations were made to the HMA(+), *Tg*MAF1RHb1 isoform to the aligned sequence in the HMA(-) paralog *Tg*MAF1RHa1 and *Tg*MAF1RHb0. Cells were fixed at 18 hpi and visualized with confocal microscopy. Immunofluorescence staining was performed with antibodies against both the HA epitope tag and the mitochondrial protein, cytochrome c oxidase II (MTCO2). All mutations to the proline-rich and flanking region were unable to disrupt *Tg*MAF1RHb1-driven HMA.

2.3.2 The C-terminal regions of both *Tg*MAF1Rha1 and b1 adopt an α/β globular structure with homology to ADP-ribose binding macro-domains

Aside from the proline-rich region characteristic of *Tg*MAF1RHb paralogs, there are no other N-terminal regions in the *Tg*MAF1 ectodomain with significant levels of polymorphism. We therefore turned our focus to the C-terminal region of the *Tg*MAF1 ectodomain, which harbors numerous amino acid polymorphisms between the 'A' and 'B' paralog groups (e.g. *Tg*MAF1RHb1 and *Tg*MAF1RHa1 are 57% identical in the C-terminal region) (Fig 2-1A). To best guide our approach of correlating sequence differences between *Tg*MAF1RHa1 and *Tg*MAF1RHb1 to their functional differences in genetic complementation, we determined the X-ray crystal structures of the C-terminal domains of *Tg*MAF1RHa1 and *Tg*MAF1RHb1 [12]. Constructs encoding the predicted C-terminal domain of *Tg*MAF1RHa1 (residues 159-435) and *Tg*MAF1RHb1 (residues

173-443: note, the C-terminal Met444, Leu445 and Asp446 residues are not included in the crystallization construct due to protein stability problems) (Genbank Accession numbers SCA78655 and AMN92247, respectively) were recombinantly produced in *E. coli*, purified to homogeneity and crystallized for X-ray diffraction experiments (Fig 2-2A). The structure of *Tg*MAF1RHb1 was phased by bromide single wavelength anomalous dispersion and refined to a resolution of 1.60 Å. Overall, the structure is well ordered with clear electron density extending from the Ser173 through Ser443 and including two C-terminal alanine residues derived from the expression vector. The 2.10 Å resolution structure of *Tg*MAF1RHa1 was solved by molecular replacement using *Tg*MAF1RHb1 as the search model with all three molecules in the asymmetric unit well-ordered. Despite possessing only 57% sequence identity within the C-terminal domain, structural overlays clearly showed that *Tg*MAF1RHa1 and *Tg*MAF1RHb1 adopt an overall similar architecture, with a rmsd of 0.6 Å over 240 Cα atoms (Fig 2-2B). Notably, these are the first structures of a MAF1 protein from any apicomplexan parasite.

Structural analysis of *Tg*MAF1RHa1 and *Tg*MAF1RHb1 revealed a compact, single domain with mixed α/β structure with a central, slightly curved 8-stranded β-sheet of mixed parallel and anti-parallel strands bound on one side by a 3 helical bundle and on the other side by a pair of helices (Fig 2-2B). Intriguingly, a DALI [61] structural homology search revealed significant similarity to macroH2A non-histone domains (also known as macro-domains) with the macro-domain from *Archaeoglobulus fulgidus* 1521 (Af1521) [62] identified as the most closely related structure (Z-score of 16; rmsd of 2.5 Å over 180 aligned Cα positions) (Fig 2-2C). A central feature of macro-domains is their ability to coordinate ADP-ribose (ADPr) and its derivatives through a surface cleft [63-65] that is conserved in both *Tg*MAF1RHa1 and *Tg*MAF1RHb1. In fact, sequences of all apicomplexan MAF1 homologs mapped onto the *Tg*MAF1RHb1 core

structure revealed the cleft as one of the most evolutionary conserved regions, supporting that this cleft is performing an important function (Fig 2-2D).

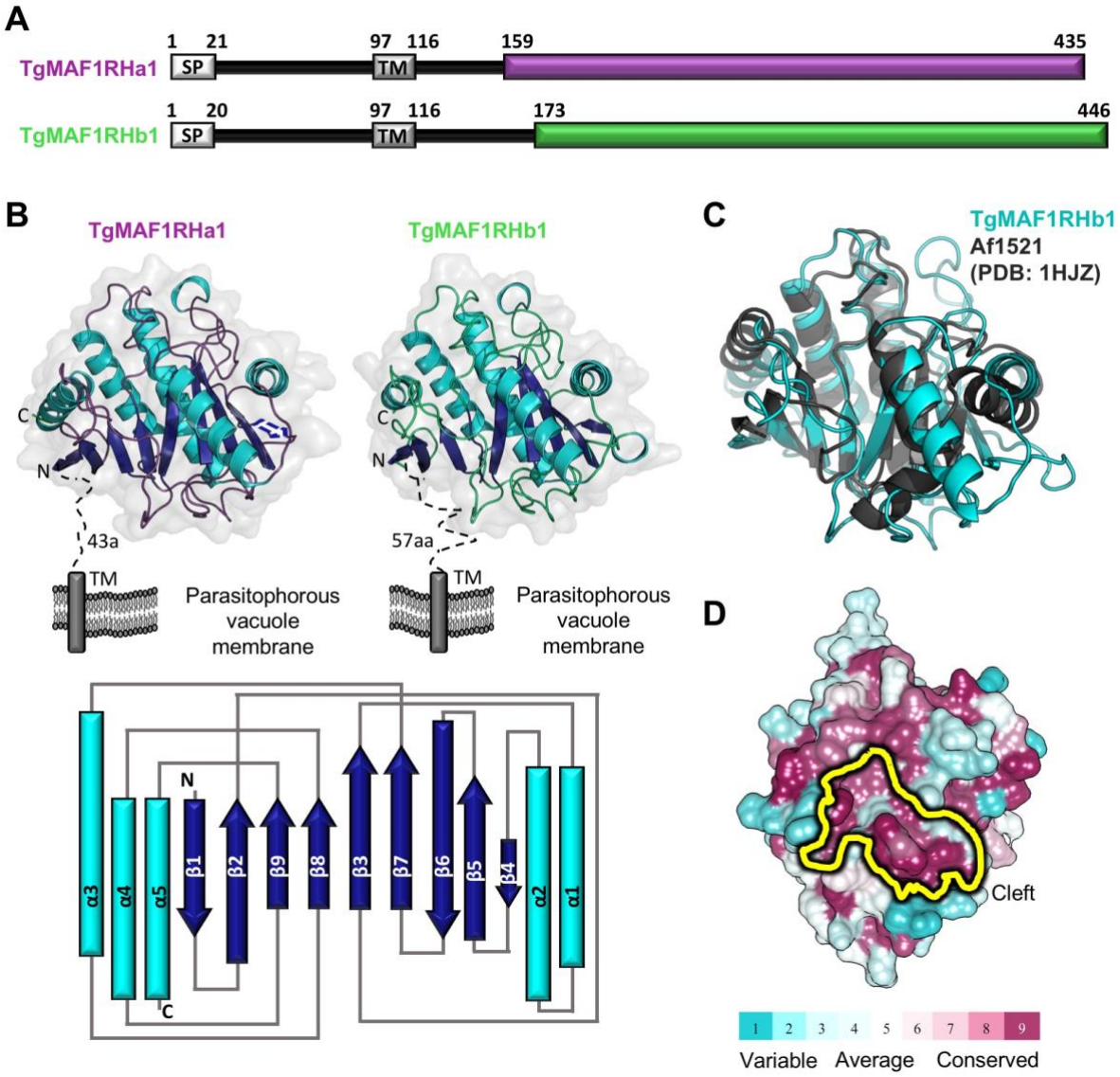


Figure 2-2 The C-terminal region of *Tg*MAF1 proteins adopts a conserved, well-ordered globular domain.

(A) Predicted domain architecture of *Tg*MAF1RHa1 and *Tg*MAF1RHb1. SP, signal peptide; TM, transmembrane. Colored boxes (*Tg*MAF1RHa1- deep purple; *Tg*MAF1RHb1- lime green) indicate C-terminal region with strongly

predicted secondary structure elements. Numbers correspond to amino acid positions. (B) Top: Tertiary structure of *Tg*MAF1RHa1 and *Tg*MAF1RHb1 colored based on secondary structure elements, with helices in cyan, strands in dark blue and loops colored as in (A). Dotted lines indicate unmolded regions and predicted features. The orientation of the proteins with respect to the parasitophorous vacuole membrane is shown. Bottom: Topology diagram of *Tg*MAF1RHa1 and *Tg*MAF1RHb1 colored as in Top. (C) Overlay of *Tg*MAF1RHb1 (cyan) with Af1521 (dark grey; PDB ID 1HJZ) showing conservation of the core macro-domain architecture. (D) Mapping of conserved (*burgundy*) and variable (cyan) residues of *Tg*MAF1RHa1 and *Tg*MAF1RHb1 homologs onto the *TgMAF1RHb1* surface using ConSurf [66]. The conserved cleft region is indicated in yellow.

2.3.3 ADP-ribose forms a low affinity complex with *Tg*MAF1RHa1 and *Tg*MAF1RHb1

The striking resemblance to a canonical macro-domain and the presence of a well conserved surface cleft led us to hypothesize that *Tg*MAF1s may be able to accommodate ADPr or a similar small molecule. Further support for this hypothesis was based on the observation that the structural homolog Af1521 was able to form a stable complex with ADPr (Figs 2-3A and B).

To investigate the ability of the *Tg*MAF1 macro-domain to bind adenine nucleotide derivatives, we first determined co-structures of ADPr bound to *Tg*MAF1RHa1 and *Tg*MAF1RHb1 to 2.7 Å and 1.65 Å resolution, respectively. Stabilizing ADPr in the surface cleft is a Phe198 (*Tg*MAF1RHb1) or Phe187 (*Tg*MAF1RHa1) that stacks onto one side of the ADPr adenine ring and the side chains of Leu227 (*Tg*MAF1RHb1) or Met216 (*Tg*MAF1RHa1) that pack against the opposite side (Fig 2-3B). In addition, the second phosphate fits into a specific pocket formed by Ile388 and Phe389 (*Tg*MAF1RHb1) or His377 and Phe378 (*Tg*MAF1RHa1) (Fig 2-3B). Structural overlays revealed that residues in Af1521 known to be crucial for ADPr binding are not conserved in the *Tg*MAF1 proteins; specifically, the strongly conserved Asp (labeled red

in Fig 2-3B) for selectivity of adenine-based nucleotides [63, 67], and the Asn (labeled red in Fig 2-3B) shown to be critical for phosphatase activity on ADP-ribose-1”-phosphate [67, 68] are not conserved (Fig 2-3B). Thus, it appears that both *Tg*MAF1RHa1 and *Tg*MAF1RHb1 have lost key residues involved in coordinating ADPr suggesting, at minimum, a weaker binding affinity, which we next measured by isothermal titration calorimetry (ITC). Binding data was only obtained for *Tg*MAF1RHa1, since the high salt required for *Tg*MAF1RHb1 stability obscured complex formation. A K_d of approximately 400 µM was measured for *Tg*MAF1RHa1 (Fig 2-3C), which is far weaker than the 130 nM K_d measured between ADPr and Af1521 [64]. It is worth noting that poly(ADPr) or poly-A binding does not necessarily require ADPr to bind with high affinity to form a functional complex [67, 69]. Thus, we cannot rule out the possibility that *Tg*MAF1RHa1 and/or *Tg*MAF1RHb1 are capable of binding these anionic polymers in a biologically relevant setting. Nor can we rule out that *Tg*MAF1RHb1 may be capable of binding other nucleotides or even oligonucleotides such as poly(A) or poly(ADP-ribose), as has been shown for other Macro-domain containing proteins [64, 69]. However, the low affinity complexes observed here in both MAF1a and MAF1b paralog classes are consistent with the hypothesis that ADPr binding is not relevant to HMA.

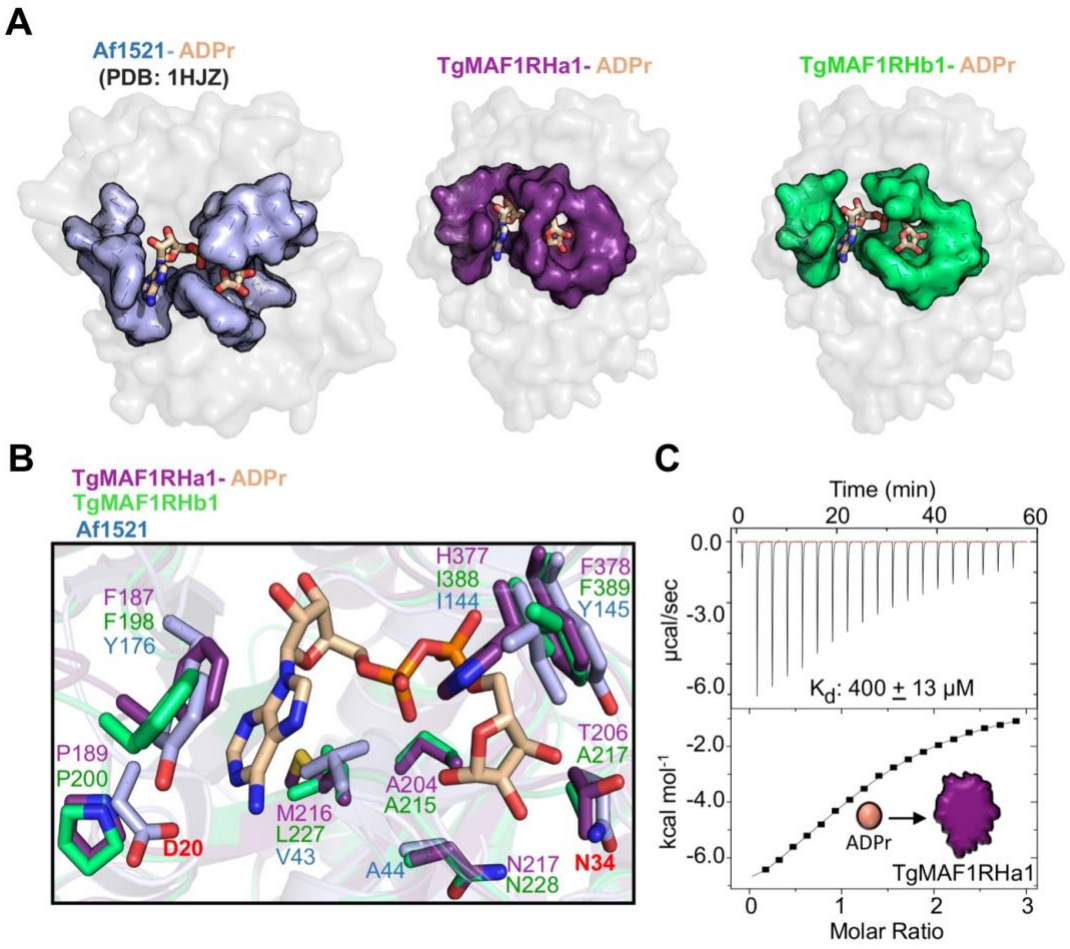


Figure 2-3 Structural characterization of *Tg*MAF1RHa1/b1 reveals that ADP-ribose coordination by the macro-domains may be vestigial.

(A) Co-structures of *Tg*MAF1RHa1, *Tg*MAF1RHb1 (grey surface) and ADP-ribose (ADPr) (wheat ball-and-stick colored by element), and Af1521 (PDB id- 1HJZ), highlighting the surface cleft of *Tg*MAF1RHa1 (Left- colored deep purple), *Tg*MAF1RHb1 (Right- colored lime green) and Af1521 (Left- colored light blue). (B) The residues coordinating ADPr in Af1521, *Tg*MAF1RHa1 and *Tg*MAF1RHb1 is shown based on the overlay of *Tg*MAF1RHa1 (deep purple), *Tg*MAF1RHb1 (lime green), and Af1521 (light blue) with ADPr (shown as wheat stick colored by element). (C) Representative ITC binding isotherm of ADPr titrating into *Tg*MAF1RHa1. The ITC clearly shows a low affinity binding with a K_d of ~400µM.

2.3.4 Unique *Tg*MAF1RHb1 C-terminal residues are required for HMA

The structures of *Tg*MAF1RHa1/b1 allowed us to identify additional residues for mutational studies by interrogating differences between the C- termini of both paralogs. We chose candidate residues by identifying residues that were 1) outward facing based on structural analysis and 2) consistently polymorphic among all members of the 'A' and 'B' groups. We first mutated peripheral *Tg*MAF1RHb1-specific β-sheet (residues 174-187) to their aligned sequence in *Tg*MAF1RHb0. Additionally, we mutated the outward-facing Ser339 to the aligned *Tg*MAF1RHa1 phenylalanine residue. Both mutant constructs sufficiently conferred HMA when expressed in *Tg*Me49 parasites (Fig 2-4A), indicating that these residues were not required for HMA.

We then focused on the 16 C-terminal residues of cloned MAF1b paralogs that are 94% similar to one another, but harbor multiple polymorphisms that distinguish the A and B lineages (Fig 2-4B). To investigate the significance of the C-terminus in MAF1-mediated HMA, we mutated the 16 C-terminal residues of *Tg*MAF1RHb1 to those found in *Tg*MAF1RHa1 to create a chimeric construct. Excitingly, the *Tg*MAF1RHb1 16 C-terminal chimera was incapable of conferring HMA in *Tg*ME49, indicating a central role for the C-terminus in HMA (Fig 2-4C). When we examined *Tg*ME49 expressing the *Tg*MAF1RHb1 16 C-terminal chimera, we found that, as for the WT *Tg*Me49, there was no significant HMA based on quantification of % vacuole coverage using EM (Fig 2-4D and E). The *Tg*MAF1RHb1 16 C-terminal chimera shows similar vacuole coverage to an *Tg*Me49 empty vector control further confirming the requirement of this region for HMA (Fig 2-4E). We then investigated the sufficiency of the C-terminus of *Tg*MAF1RHb1 in driving HMA. Using site-directed mutagenesis we created a *Tg*MAF1RHa1 chimeric construct with the 16 C-terminal residues of *Tg*MAF1RHb1. The *Tg*MAF1RHa1 16 C-

terminal chimera did not drive HMA upon expression in *Tg*Me49 parasites (Fig 2-4F), showing that while the 16 C-terminal residues of *Tg*MAF1RHb1 are necessary for HMA, when expressed in a Type II genetic background they are not sufficient.

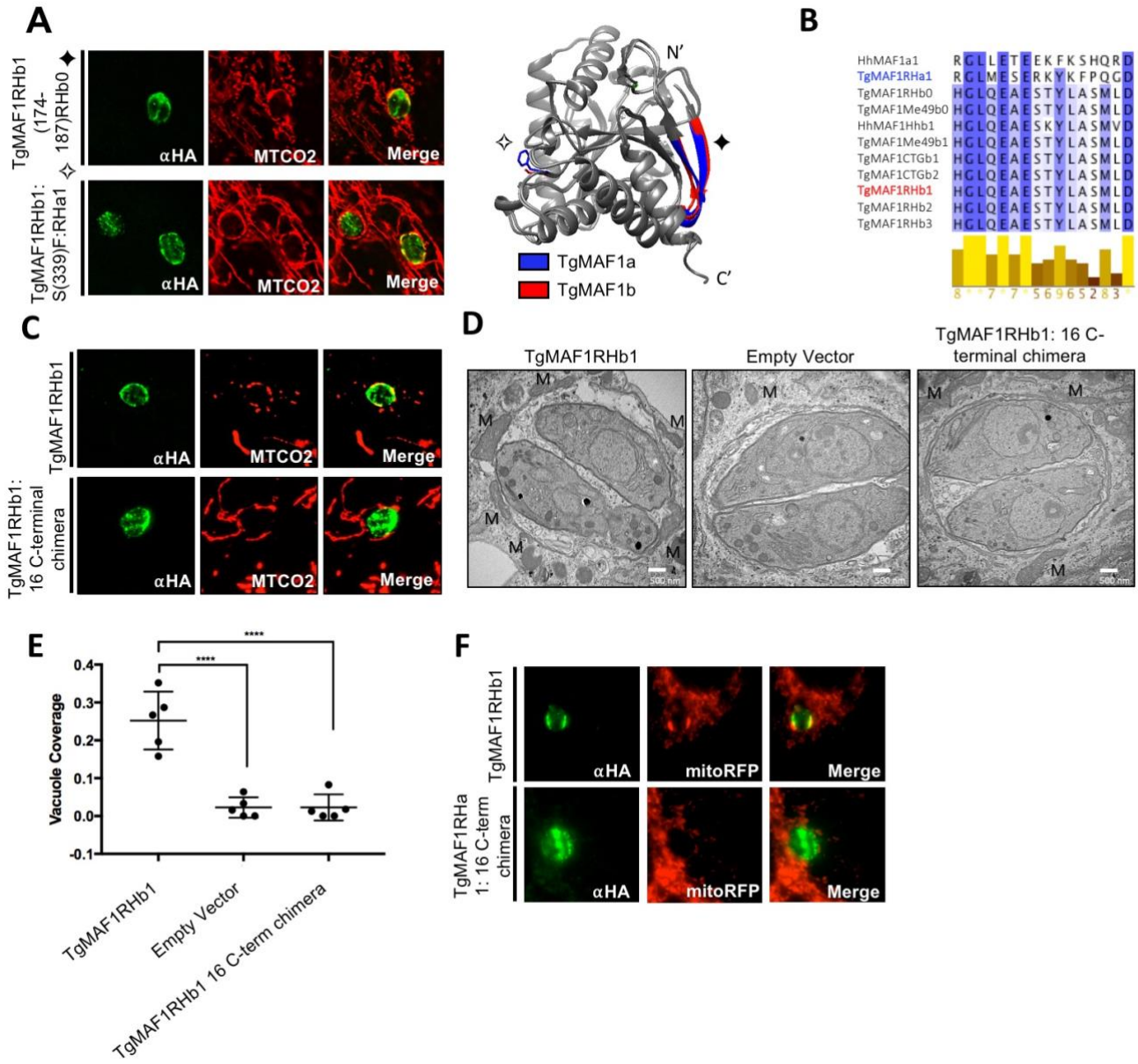


Figure 2-4 *Tg*MAF1RHb1 16 C-terminal residues are required for HMA.

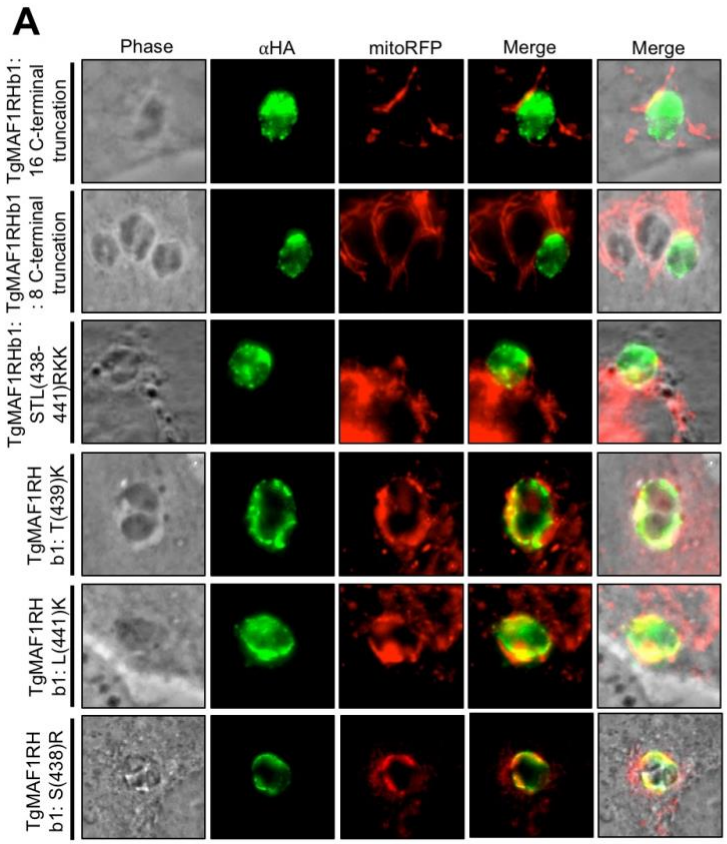
(A) HFFs were infected with *Tg*Me49 parasites expressing an HA tagged mutated *Tg*MAF1RHb1 isoform. Mutations were introduced using site-directed mutagenesis in an effort to disrupt key outward facing amino acids. Cells were

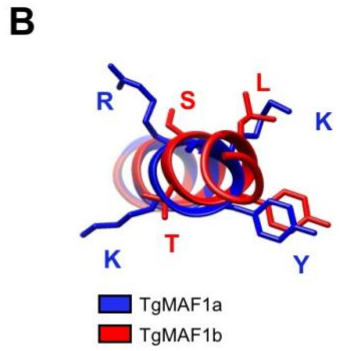
fixed at 18 hpi and visualized using confocal microscopy. Immunofluorescence staining was performed with antibodies against the HA epitope tag and the mitochondrial protein, cytochrome c oxidase II (MTCO2). Symbols on mutation titles correspond to mutations in the structural overlay of both MAF1a and MAF1b generated in UCSF Chimera. (B) Alignment of the 16 C-terminal residues of MAF1 isoforms. Dark blue and white indicate 100% and 0% percent identity, respectively. Alignments were visualized in JalView after alignment using Clustal-Omega. Yellow bars below alignment show residue conservation across MAF1 isoforms. (C) HFFs were infected with *Tg*Me49 parasites expressing either a WT HA-*Tg*MAF1RHb1 isoform or a *Tg*MAF1THb1 16 C-terminal chimeric form. The C-terminal chimera contained *Tg*MAF1RHb1 residues 1-430 and *Tg*MAF1RHa1 residues 420-435. Cells were fixed and treated with the same conditions outlined in 4A. (D) HFFs were infected with isolated *Tg*Me49 clones transfected with either HA-*Tg*MAF1RHb1, *Tg*MAF1RHb1-C-term chimera or empty vector. Cells were fixed 18 hpi and processed for transmission electron microscopy. Mitochondria (M) are labeled. (E) Quantification of percent vacuole coverage determined by electron microscopy. ****$p<0.0001$, one-way ANOVA. (F) Normal rat kidney epithelial cells expressing RFP in the mitochondrial matrix (NRK mitoRFPs) were infected with *Tg*Me49 parasites expressing either a WT HA-*Tg*MAF1RHb1 isoform or a *Tg*MAF1RHa1 16 C-terminal chimeric form. The C-terminal chimera is *Tg*MAF1RHa1 residues 1-419 and *Tg*MAF1RHb1 residues 431-446. Cells were fixed 18 hpi and visualized using epi-fluorescent microscopy. Immunofluorescence staining was performed with antibodies against the HA epitope tag.

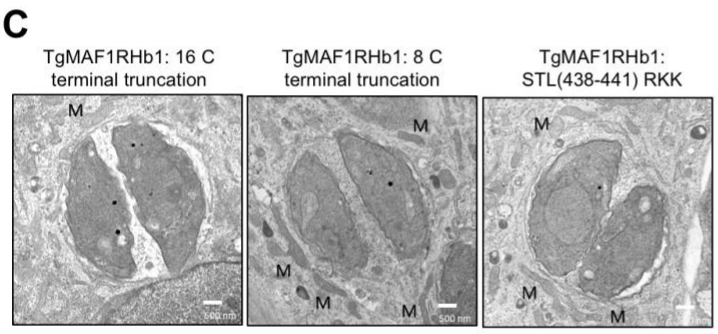
2.3.5 Three C-terminal residues in *Tg*MAF1RHb1 are required for HMA

As described above and shown in Figure 2-1A, the 16 C-terminal residues of *Tg*MAF1RHb1 and *Tg*MAF1RHa1 form the α5 helix that is unique to each *Tg*MAF1RHb1 and *Tg*MAF1RHa1 paralog. To further interrogate this region, truncation mutations were made to the C-terminus of *Tg*MAF1RHb1 in an effort to abrogate HMA without disrupting proper MAF1 localization. Truncation of both the 16 and 8 C-terminal residues resulted in *Tg*MAF1RHb1 products that did not confer HMA in *Tg*Me49 parasites, confirming the requirement of these residues (Fig 2-5A and C). The structural differences of both the *Tg*MAF1RHb1 and

*Tg*MAF1RHa1 helices and their opposing HMA phenotypes suggests residue-specific function within this region. Ten of the 16 C-terminal residues of this helix are polymorphic between the A and B paralogs, making it difficult to determine which residues to investigate based on primary sequence alone. However, when we examined the structure of the α5 helix, we identified a trio of basic residues unique to the C-terminus of *Tg*MAF1RHa1, specifically Arg427, Lys428 and Lys430 (Fig 2-5B). These are strikingly different than their uncharged counterparts in *Tg*MAF1RHb1 (Ser438, Thr439, and Leu441) (Fig 2-5B). To investigate the importance of these residues in HMA we mutated the STL residues in *Tg*MAF1RHb1 to RKK, and this mutant was completely unable to mediate HMA when expressed in Type II *T. gondii* (Fig 2-5A and C). Using electron microscopy, we found that the percentage of the vacuole with interacting mitochondria in parasites expressing the RKK mutant is nearly indistinguishable from wild type Type II *T. gondii* (Fig 2-5D). When we mutated each residue individually (Ser→Arg, Thr→Lys and Leu→Lys), all 3 constructs could still mediate HMA when expressed in Type II *T. gondii* (Fig 2-5A). Additionally, mutation of both the Ser438 and Leu441 to Arg427 and Lys430 which lie on the same face of the C-terminal alpha helix also mediated HMA when expressed in Type II parasites (Fig 2-5A), suggesting the collective requirement of all three uncharged MAF1b residues (Ser438, Thr439 and Leu441) for HMA to occur.


A
Phase
αHA
mitoRFP
Merge
Merge
TgMAF1RHb1: 16 C-terminal truncation
TgMAF1RHb1 : 8 C-terminal truncation
TgMAF1RHb1: STL(438-441)RKK
TgMAF1RHb1: T(439)K
TgMAF1RHb1: L(441)K
TgMAF1RHb1: S(438)R


B
R
S
L
K
K
T
Y
TgMAF1a
TgMAF1b


C
TgMAF1RHb1: 16 C terminal truncation
TgMAF1RHb1: 8 C terminal truncation
TgMAF1RHb1: STL(438-441) RKK
M
500 nm

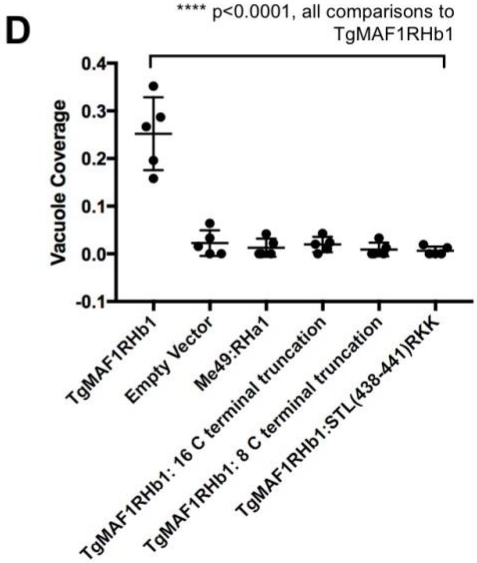

D
**** p<0.0001, all comparisons to TgMAF1RHb1
Vacuole Coverage
0.4
0.3
0.2
0.1
0.0
-0.1
TgMAF1RHb1
Empty Vector
Me49:RHa1
TgMAF1RHb1: 16 C terminal truncation
TgMAF1RHb1: 8 C terminal truncation
TgMAF1RHb1:STL(438-441)RKK

Figure 2-5 Three residues in the C-terminus of *Tg*MAF1RHb1 are required for HMA.

(A) NRK mitoRFP cells were infected with *Tg*Me49 parasites expressing HA-*Tg*MAF1RHb1 16 C-terminal mutant variants. Cells were fixed 18 hpi and visualized using epi-fluorescent microscopy. Immunofluorescence staining was performed with antibodies against the HA epitope tag. (B) *Tg*MAF1RHa1 and *Tg*MAF1RHb1 are highlighted in blue and red respectively which correspond to the color in the ribbon rendering of the C-terminal structures. Structure was visualized in UCSF chimera. (C) HFFs were infected with isolated *Tg*Me49 clones transfected with each of the HA-*Tg*MAF1RHb1 16 C-terminal mutants. Cells were fixed 18 hpi and processed for transmission electron microscopy. Mitochondria are labeled. (D) Quantification of percent vacuole coverage determined by electron microscopy. ****$p<0.0001$, one-way ANOVA ("all comparisons" means each comparison of *Tg*MAF1RHb1 C-terminal mutant and *Tg*MAF1RHa1 to *Tg*MAF1RHb1 is significant)

2.3.6 *Tg*MAF1RHb1 mutants retain binding to SAM50

*Tg*MAF1RHb1 has recently been shown to interact with the mitochondrial intermembrane space bridging (MIB) complex which consists of proteins spanning both the outer and inner mitochondrial membranes. One MIB component on the cytosolic face of the outer mitochondrial membrane interacting, directly or indirectly, with *Tg*MAF1RHb1 is SAM50 [70]. These data suggest SAM50 might function as the protein *Tg*MAF1RHb1 uses to anchor itself to the mitochondria. If interactions with SAM50 are the sole basis for mediating HMA, neither *Tg*MAF1RHa1 nor our *Tg*MAF1RHb1 with the chimeric C-terminus should interact with SAM50. To investigate the ability of each of the C-terminal mutants and *Tg*MAF1RHa1 to interact with SAM50, we conducted co-immunoprecipitations using HA conjugated beads. Consistent with published data we found that HA pulldowns from *Tg*ME49 clones expressing N-terminally HA-tagged *Tg*MAF1RHb1 co-precipitated SAM50, while cytosolic host protein (αTubulin) showed no evidence for association with *Tg*MAF1RHb1 (Fig 2-6A and B). Interestingly, however, we

found that immunoprecipitations of *Tg*MAF1RHa1, the *Tg*MAF1RHb1 C-terminal truncation mutant, and the *Tg*MAF1RHb1:STL→RKK mutant also specifically pulled down SAM50 (Fig 2-6A). Together, these data demonstrate that the specificity of the interaction between MAF1 and SAM50 is not dependent on the 16 C-terminal residues, and that MAF1-SAM50 interactions are not sufficient to mediate HMA.

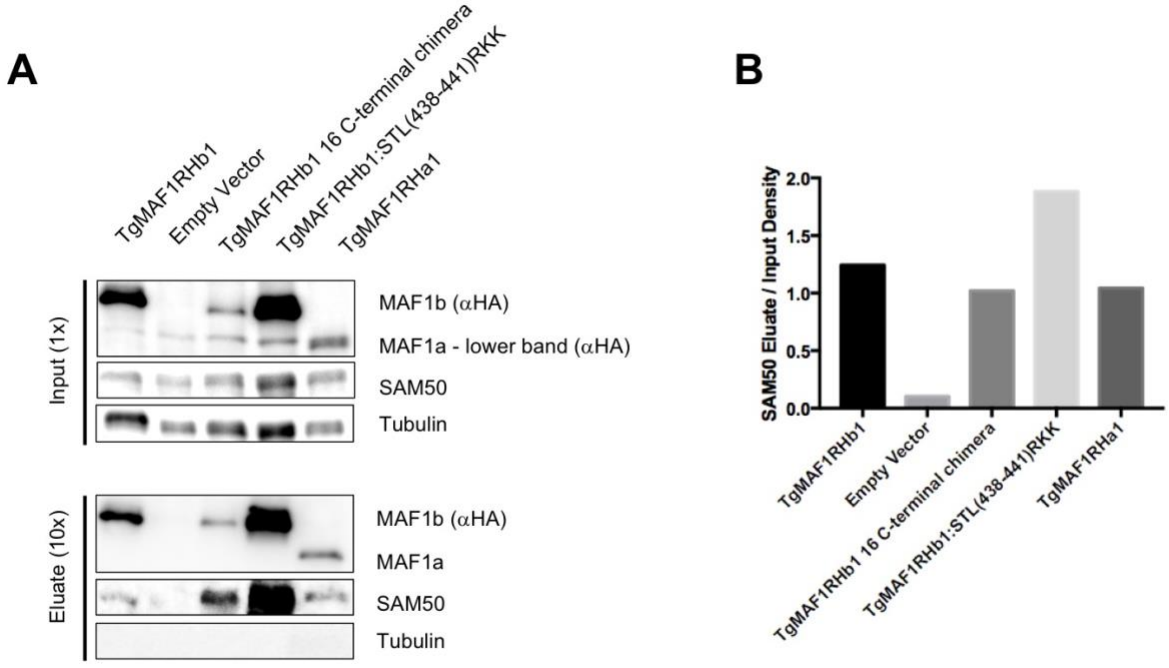


Figure 2-6 *Tg*MAF1RHb1 mutants bind to SAM50.

(A) Immunoprecipitation of lysed HFFs infected with cloned *Tg*Me49: *Tg*MAF1RHb1 16 C-terminal mutants, *Tg*Me49:*Tg*MAF1RHb1 and *Tg*Me49:EV. Lysates were incubated with HA-conjugated beads and eluted with LDS sample buffer. Western blotting analysis was performed with the listed primary antibodies and HRP-conjugated secondary antibodies. Each 16 C-terminal mutant is able to bind to SAM50 (B) Densitometric input/eluate quantification of SAM50 western blot in panel A using ImageJ (NIH). Pulldowns of *Tg*MAF1RHb1 C-terminal mutant infections have been repeated three times. Pulldowns of *Tg*MAF1RHa1 infection was only performed once.

2.3.7 HMA-functional paralogs confer *in vivo* competitive advantage, but *Tg*MAF1RHb1 C-terminal mutants do not

Previous work found that expression of *Tg*MAF1RHb1 and not *Tg*MAF1RHa1 in a Type II *T. gondii* strain provides a competitive advantage in an *in vivo* infection [12]. This growth advantage is not present during growth *in vitro* [3]. To test the *in vivo* selective advantage of expressing the C-terminal mutants, we performed *in vivo* competition assays by infecting mice with clonal lines expressing different HA-tagged WT and mutant *Tg*MAF1RHa1 and *Tg*MAF1RHb1 constructs with the same standard reference strain expressing only the empty vector (*Tg*ME49:EV). Mice were infected with total of 10_5 tachyzoites and the initial percent of parasites expressing HA-tagged MAF1 was quantified by immunofluorescence assay (IFA). Consistent with previous results [12], *Tg*MAF1RHb1-expressing parasites competed effectively with the reference strain (Fig 2-7A and B), while parasites expressing the *Tg*MAF1RHb1 16 amino acid truncation (Fig 2-7B), 16 amino acid C-terminal chimera (Figs 2-7A and B) or the 3 amino acid C-terminal chimera (STL→RKK; Fig 2-7A and B), competed less effectively with the reference strain. In contrast, one of the single amino acid mutants (Leu441Lys), that is HMA+ (see Fig 2-5A) had a competitive index similar to parasites expressing WT *Tg*MAF1RHb1 (Fig 2-7A, far right). It should be noted that *in vivo* competition assays are sensitive to parasite passage history, and we used a generic reference strain that was not passage matched to the strains expressing mutant MAF1b genes (although all of the HA-tagged lines were carefully passage matched). To further validate the competitive advantage of this mutant, a mixed, non-clonal population of Type II parasites expressing *Tg*MAF1RHb1:Leu441Lys (HMA+) was injected into three Balb/C mice and the Leu441Lys expressing parasites were found to have a significant *in vivo* growth rate as compared to their WT *Tg*Me49 counterparts (Fig 2-7C). Additionally, *Tg*Me49

parasites expressing *Tg*MAF1RHb1:SL(438/441)RK also outcompeted their passage-matched *Tg*Me49:EV clone (Fig 2-7D). These data show that the *in vivo* competitive advantage conferred by expressing *Tg*MAF1RHb1 in Type II *T. gondii* is almost certainly due to its impact on HMA, rather than the introduction of an additional copy of MAF1b.

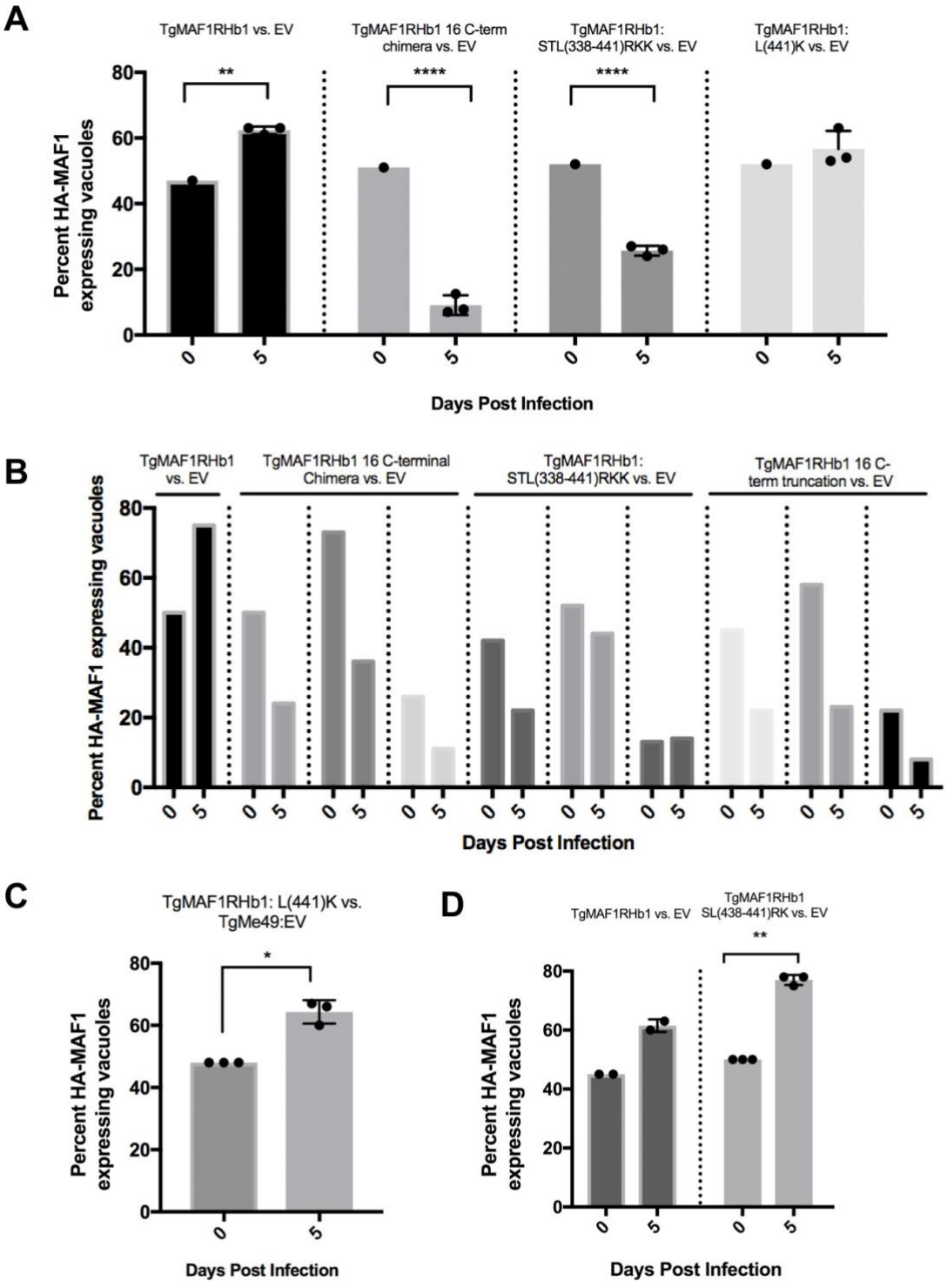

A
TgMAF1RHb1 vs. EV
TgMAF1RHb1 16 C-term chimera vs. EV
TgMAF1RHb1: STL(338-441)RKK vs. EV
TgMAF1RHb1: L(441)K vs. EV
**
****
****
Percent HA-MAF1 expressing vacuoles
80
60
40
20
0
0
5
Days Post Infection
B
TgMAF1RHb1 vs. EV
TgMAF1RHb1 16 C-terminal Chimera vs. EV
TgMAF1RHb1: STL(338-441)RKK vs. EV
TgMAF1RHb1 16 C-term truncation vs. EV
Percent HA-MAF1 expressing vacuoles
Days Post Infection
C
TgMAF1RHb1: L(441)K vs. TgMe49:EV
*
Percent HA-MAF1 expressing vacuoles
Days Post Infection
D
TgMAF1RHb1 vs. EV
TgMAF1RHb1 SL(438-441)RK vs. EV
**
Percent HA-MAF1 expressing vacuoles
Days Post Infection

Figure 2-7 HMA-functional paralog, *Tg*MAF1RHb1 confers *in vivo* competitive advantage but *Tg*MAF1RHb1 C-terminal mutants do not.

(A) Three mice per control or mutant were infected intraperitoneally (IP) with 50/50 mix of *Tg*Me49 parasites expressing either WT or C-terminal *Tg*MAF1RHb1 mutants and *Tg*Me49:EV parasites (i.e. 50% *Tg*Me49:*Tg*MAF1 16 C-terminal chimera vs. 50% *Tg*Me49 empty vector). HFFs were also infected with each input mixed population, fixed at 18 hpi and visualized utilizing epi-fluorescent microscopy. Immunofluorescence staining was performed with antibodies against the HA epitope tag. Following a five-day *in vivo* infection, extracted peritoneal content was used to infect a monolayer of HFFs. Cells were fixed at 18 hpi and visualized utilizing epi-fluorescent microscopy. Immunofluorescence staining was performed with antibodies against the HA epitope tag. Both input (0 dpi) and output (5 dpi) populations were quantified by measuring the percent of MAF1-HA expressing parasites. **p=0.0038, ****p<0.0001 two-way ANOVA (Sidak test). (B) Similar procedure as A, however input parasites populations were mixes of 25/75, 50/50, and 75/25 and only one mouse was infected for each of the mixes. (C) Similar procedure to A and B. Three mice were IP infected with a natural mixed population of *Tg*Me49:*Tg*MAF1RHb1:L(441)K parasites. HFFs were infected with both input populations and peritoneal population after a five day infection. Cells were fixed and probed for parasites expressing HA-epitope tag by IFA (Paired t-test, *p=0.0174). (D) Similar procedure to A and B. Three mice were IP infected with a 50/50 mixture of *Tg*Me49:EV and *Tg*Me49:*Tg*MAF1RHb1:SL(438/441)RK clones. HFFs were infected with both input populations and peritoneal population after a five-day infection. Cells were fixed and probed for parasites expressing HA-epitope tag by IFA (Paired t-test, **p=0.0014).

2.4 Discussion

The ability to closely associate with host cell mitochondria has evolved independently in multiple intracellular pathogens including *Legionella pneumophila, Chlamydia psittaci,* and *T. gondii* [31, 40, 41, 43, 45]. However, in most organisms the importance of this phenotype is unknown. In the case of *T. gondii,* the pathogen gene product required for this phenomenon has been identified as MAF1b, a gene present in multiple copies within a tandemly expanded locus

bearing extensive intra-strain and species sequence variation [10, 12]. In previous work we have shown that this variation has phenotypic consequences both with respect to HMA and infectivity *in vivo* [12]. In the present study we have exploited MAF1 locus diversity by using comparative structural biology and genetic complementation to identify regions of MAF1b required to mediate HMA.

2.4.1 Both MAF1 paralogs harbor vestigial ADP-ribose binding domains in their C-termini

Paramount to our interrogation of the MAF1 locus was firstly characterizing two distinct MAF1 paralogs with opposing HMA phenotypes through crystallographic and *in vitro* mutational studies: *Tg*MAF1RHa1 (*Tg*MAF1RHa1; "HMA-") and *Tg*MAF1RHb1 (*Tg*MAF1RHb1; "HMA+") [12]. Given the fact that this domain binds ADPr with comparatively low affinity [64], it is likely that this domain is a pseudo ADPr binding domain. ADP-ribosylation is an essential post translational modification (PTM) that regulates a wide range of cellular processes including host immune pathways [13]. Pathogens such as the hepatitis E virus (HEV) and SARS evolved an antagonistic ADPr binding macrodomain which reverses ADP-ribosylation events in order to subvert the host immune response [19]. Many poly-ADP-ribose polymerases (PARPs) members which catalyze this PTM have been shown to undergo recurrent positive diversifying selection in mammals which suggests a host-pathogen molecular arms race. These systems explain a possible ancestral role for the ADPr binding domain of MAF1 in *T. gondii* [71, 72].

In addition to an expanded family of Macro-domain containing proteins, the *T. gondii* genome also harbors a large number of rhoptry pseudokinases that have lost catalytic activity. Multiple rhoptry pseudokinases are found within tandemly expanded gene clusters (e.g., ROP2/3/4 locus [25, 73] and ROP5 [10, 20, 28, 30]). MAF1 may represent a new locus, which has been

modified from an existing protein domain for the purpose of manipulating the host cell. Interestingly the repurposing of this locus likely occurred in the most recent common ancestor of *T. gondii*, *H. hammondi* and *N. caninum* since all three parasite species harbor at least one MAF1 paralog, and based on sequence comparisons they all appear to have retained the pseudo-ADPr binding domain. To confirm whether the sequence comparison translated into structural conservation, we generated high confidence 3D models of *Hh*MAF1a/b and *Nc*MAF1b (Fig 2-8). Despite a similar charge distribution and high sequence conservation of the surface cleft forming residues, MAF1 paralogs appear to have lost the ability for high affinity coordination of ADPr and its derivatives, consistent with our hypothesis that the cleft region may be a vestigial domain. The polymorphic 16 residue C-terminal region, however, exhibits a remarkably divergent surface charge, suggesting that the evolutionary pressure on the C- terminal region is comparatively more than the cleft region, consistent with the importance of this C-terminal helix in gaining a central role in mediating HMA.

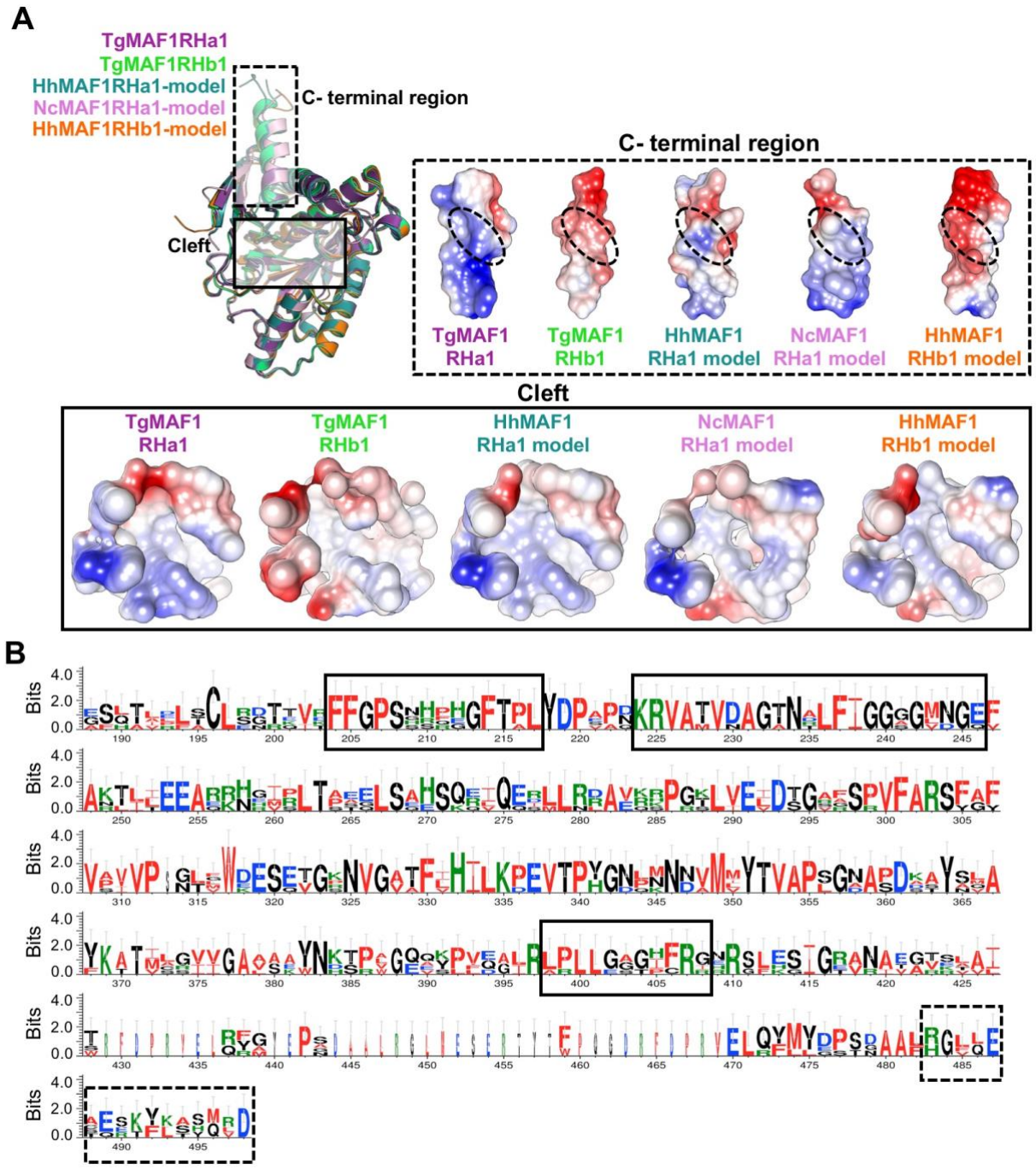

A
TgMAF1RHa1
TgMAF1RHb1
HhMAF1RHa1-model
NcMAF1RHa1-model
HhMAF1RHb1-model
C- terminal region
Cleft
C- terminal region
TgMAF1 RHa1
TgMAF1 RHb1
HhMAF1 RHa1 model
NcMAF1 RHa1 model
HhMAF1 RHb1 model
Cleft
TgMAF1 RHa1
TgMAF1 RHb1
HhMAF1 RHa1 model
NcMAF1 RHa1 model
HhMAF1 RHb1 model
B
Bits

Figure 2-8 The models of *Hammondia hammondi* and *Neospora caninum* MAF1RHa1/b1 paralogs reveal identical structural topology with *Tg*MAF1RHa1/b1 and exhibit major surface charge differences in the C-terminal region that may be responsible for HMA –ve and HMA +ve homologs.

(A) Structural overlap of *Tg*MAF1RHa1/b1 with HhMAF1RHa1/b1 and NcMAF1RHa1 models with C- terminal region and surface cleft indicated as dotted and solid black boxes. Inset 1. The surface charge distribution of the last 16 residues of the C terminal region in MAF1a/b homologs. The dotted circle indicates the major difference in charge distribution among the homologs. It is clear that the basic patch present in MAF1a homologs is absent in MAF1b. Inset 2. The surface charge distribution of the cleft region reveals all the homologs have a similar electrochemical cleft region. (B) Weblogo of the *T. gondii*, *H. hammondii*, *N. caninum* MAF1a/b sequences. The solid boxes indicate the cleft forming region whereas the dotted box depicts the last 16 residues of the C- terminal region. The C- terminal region is considerably more polymorphic than the cleft.

2.4.2 Structure-function analyses suggest that the C-terminus of *Tg*MAF1RHa1 has three residues that prevent HMA

We exploited sequence diversity across MAF1 paralog classes to ultimately determine that three residues in the C-terminus of MAF1b are necessary, but not sufficient, to mediate HMA. These data are consistent with previous studies of MAF1b, which found that C-terminally tagged MAF1b paralogs did not confer HMA in *Tg*Me49 parasites [31]. Interestingly, the 16 C-terminal residues of the A and B paralogs that comprise the α5 helix differ at 10 of 16 amino acid positions within a given *T. gondii* strain, and these differences are highly conserved across divergent *T. gondii* strains and between *T. gondii* and *H. hammondi*. Notably, the C-terminal α5 helices in *Tg*MAF1RHa1 and *Tg*MAF1RHb1 adopt slightly different positions with respect to the body of the macro-domain, which may be due to sequence polymorphisms. With respect to HMA, only three of the 10 polymorphic residues of *Tg*MAF1RHb1 are required. We hypothesize that Ser438,

Thr439 and Leu441 in MAF1b, which differ markedly from the structurally analogous Arg427, Lys428 and Lys430 residues in MAF1a, present an optimal surface enabling coordination of a yet unidentified protein partner that gives rise to the HMA phenotype. The basic patch in MAF1a may also support non-specific molecular interactions with anionic biomolecules that effectively disrupt recruitment of HMA proteins. In either scenario, the significant differences in size and charge between the STL versus RKK residues appear to have profound effects on the ability of a given MAF1 paralog to drive HMA.

2.4.3 *Tg*MAF1RHb1 mutants that do not mediate HMA still associate with members of the MICOS complex

*Tg*MAF1RHb1 exists primarily within the parasitophorous vacuole membrane (PVM) and has been found to interact with members of the MICOS complex in cellular lysates [70]. The MICOS complex spans the mitochondrial inner membrane and is responsible for the structural integrity of the mitochondria [74, 75]. Additionally, *Tg*MAF1RHb1 pulls down SAM50 on the outer mitochondrial membrane, which interacts with the MICOS complex, forming the membrane inner bridge complex (MIB) [70]. Upon RNAi knockdown of SAM50 and MIC60, the ability of *Tg*MAF1RHb1 to drive HMA was diminished [70], suggesting that associations between *Tg*MAF1RHb1 and the MICOS complex are required for HMA. In the present study we found that all tested MAF1 paralogs (e.g., A and B) and HMA-deficient MAF1b mutants were capable of pulling down SAM50, although we did not determine precise binding affinities between the a and b paralogs and therefore do not know if their binding to SAM50 is equally avid. An exciting potential outcome from these data is that the ability to interact with host SAM50 may represent an ancestral function of MAF1, but that other interactions are necessary for effective HMA. The

chimeric MAF1b/a constructs are perfect tools to identify the precise interactions necessary for this intriguing host cell manipulation phenotype [76, 77].

2.4.4 MAF1b confers *in vivo* competitive advantage and replication rate

Parasites expressing HMA-driving MAF1b paralogs do confer a selective advantage during mouse infections [12], but the direct link between HMA itself and this phenotype was lacking. In the present study our results provide the most convincing evidence to date that HMA itself (rather than other effects mediated by ectopic expression of MAF1b) confers increased parasite replication and/or survival in *vivo*. The competitive index of the Leu→Lys single mutant (which is $HMA_{(+)}$) was clearly superior to that of the STL→RKK triple mutant (which is $HMA_{(-)}$; Figs 2-7A and B). Arguably, loss of HMA function provides the parasites with a measurable disadvantage *in vivo* when competed with a WT *Tg*Me49 strain continuously cultured in the lab. These data directly link HMA itself to increased *in vivo* parasite survival and/or growth. The selective advantage conferred by HMA is consistent with the fact that MAF1b paralogs show signs of positive, diversifying selection (defined previously in [12]), while MAF1a paralogs are highly conserved between strains and across species (see Fig 2-1A). Somewhat paradoxically, *Tg*MAF1RHb1 expression in Type II parasites is also associated with an increased pro-inflammatory response *in vitro* (mouse embryonic fibroblasts) and *in vivo*, including the differential regulation of key Type 1 IFNs and pro-inflammatory cytokines in mice [31]. It is possible that increased cytokine production may either recruit more cells that are hospitable to *T. gondii* and/or block recruitment of cells that are more lethal to *T. gondii* (such as $GR1^{+}$ macrophages; [78, 79]). It is also possible that the differential cytokine response is a direct result of a gene dosage effect of Type II parasites retaining their endogenous MAF1 locus and exogenously expressing an additional MAF1b

paralog. This response is independent of the HMA phenotype and represents another ancestral function that is shared in both MAF1a and MAF1b paralogs.

In summary, using structure-function analyses we have identified three residues in *T. gondii* MAF1b that are required for MAF1b-mediated HMA, and have definitively linked HMA to increased parasite proliferation *in vivo*. In the process we have identified at least two functions associated with all MAF1 paralogs (ADP ribose coordination and association with the MICOS complex), which may represent the ancestral function of MAF1 prior to its neofunctionalization. This study has also enabled us to generate new MAF1 mutants that will serve as valuable reagents to probe the importance of HMA for *T. gondii* mediated host cell manipulation.

2.5 Materials and methods

2.5.1 *Tg*MAF1RHa1 and *Tg*MAF1RHb1 cloning, protein production and purification

Constructs encoding the predicted C-terminal region of *Tg*MAF1RHa1 (Genbank accession no. KU761333) (re-annotated *Tg*MAF1RHb1; Ser173 to Ser443) and *Tg*MAF1RHb1 (Genbank accession no. KU761342) (re-annotated *Tg*MAF1RHa1; Thr159 to Asp435) were cloned, produced and purified as previously described [80]. Each protein was in a final buffer of HBS (20 mM Hepes pH 7.5, 150-300 mM NaCl) with 1% glycerol and 1 mM dithiothreitol.

2.5.2 Crystallization and data collection

Crystals of *Tg*MAF1RHb1 were initially identified in the PEG/Ion Screen (Hampton Research) using sitting drops at 295 K. The final, refined drops consisted of 1.2 µL *Tg*MAF1RHb1 at 20 mg/mL with 1.2 µL of reservoir solution (0.2 M ammonium sulfate, 20% PEG3350) and were equilibrated against 120 µL of reservoir solution. For phase determination, *Tg*MAF1RHb1 crystals were soaked in a final cryoprotectant of reservoir solution with 12.5% glycerol and 1M NaBr for 3 min before flash cooling directly in liquid nitrogen. For co-crystallization with ADP-ribose (ADPr) (Sigma), *Tg*MAF1RHb1 was crystallized in the presence of 5 mM ADPr and the cryopreservation solution contained reservoir solution with 12.5% glycerol and 10 mM ADPr. Diffraction data were collected on beamline 08B1-1 at the Canadian Light Source (CLS) for bromide-derivatized crystals, and on beamline 12-2 at the Stanford Synchrotron Radiation Lightsource (SSRL) for ADPr bound *Tg*MAF1RHb1.

Crystals of *Tg*MAF1RHa1 were initially identified in the Index screen (Hampton Research) using sitting drops at 295 K. The final, refined drops consisted of 1.0 µL *Tg*MAF1RHa1 at 6.2 mg/mL with 1.0 µL of reservoir solution (0.9 M ammonium sulfate, 0.1 M Hepes pH 7.0, 0.5% PEG8000, 3% 2-methyl-2,4-pentanediol) and were equilibrated against 120 µL of reservoir solution. Crystals were cryopreserved in 80% saturated lithium sulfate and flash cooled in liquid nitrogen. A subset of crystals was soaked with 10 mM ADPr prior to cryopreservation. Diffraction data for *Tg*MAF1RHa1 were collected on beamline 08ID-1 at the CLS, and on beamline 11-1 at SSRL for ADPr bound *Tg*MAF1RHa1.

2.5.3 Data processing, structure solution and refinement

Diffraction data for *Tg*MAF1 crystals were collected and processed to 1.60 Å (*Tg*MAF1RHb1-Br), 1.65 Å (*Tg*MAF1RHb1-ADPr), 2.10 Å (*Tg*MAF1RHa1), and 2.70 Å (*Tg*MAF1RHa1-ADPr) resolution using Imosflm [81], Scala [82] and Aimless [83] in the CCP4 suite of programs [84]. The structure of *Tg*MAF1RHb1 was phased by bromide single wavelength anomalous dispersion. A total of 18 Br sites were identified and refined using the ShelxC/D/E pipeline [85]. High quality phases were obtained following density modification in dm [86] and enabled building and registering of approximately 85% of the backbone using buccaneer [87]. The *Tg*MAF1RHb1-ADPr structure was solved by molecular replacement using the final refined *Tg*MAF1RHb1 structure as the search model in PHASER [88], while the *Tg*MAF1RHa1 structures were solved using a chainsaw trimmed model of *Tg*MAF1RHb1 as the search model [89]. For each structure, COOT [90] was used for manual model building and selection of solvent atoms, and the models were refined in Phenix.Refine [91]. Complete structural validation was performed in Molprobity [92]. For each dataset, 5% of the reflections were set aside for calculation of R_{free}. Data collection and refinement statistics are presented in Table 1.

The atomic coordinates and structure factors have been deposited in the Protein Data Bank under the following codes: *Tg*MAF1RHb1-Br – 6BXR; *Tg*MAF1RHb1-ADPr – 6BXW; *Tg*MAF1RHa1 – 6BXS; *Tg*MAF1RHa1-ADPr – 6BXT.

2.5.4 In-silico homology modeling of *Hh*MAF1a1/b1 and *Nc*MAF1RHa1

The crystal structure of *Tg*MAF1RHa1 was used as a modeling template for *Hh*MAF1RHa1 and *Nc*MAF1RHa1 paralogs while the structure of *Tg*MAF1RHb1 was used as the

template for *Hh*MAF1RHb1. The sequence identity between target and template was 40% or greater for all the MAF1 paralogs. Using Modeller 9v18 [93], 10 models for each MAF1 paralog was generated and the best model chosen based on the low value of normalized discrete optimized protein energy (DOPE). The assessment of the final model was carried out with Ramachandran statistics [94], QMEAN [95], and ProSA[96] (Supplementary Table 1).

2.5.5 Cell maintenance and parasite infection

*Tg*Me49 in these experiments were regularly passed in human foreskin fibroblasts (HFFs) and incubated at 37$_{o}$C in 5% CO_2. NRK mito-RFP cells were a kind gift from Jennifer Lippincott-Schwartz (NIH, Bethesda, MD) [97]. Both cell types were grown in Dulbecco's modified Eagle's medium (DMEM) supplemented with 50 μg/ml of penicillin and streptomycin, 10% FBS, and 2mM glutamine (cDMEM).

2.5.6 Expression of MAF1 mutants and transgenic parasites

Parental plasmid used for cloning all *Tg*MAF1RHb1 constructs contains the *HXGPRT* gene. *Tg*MAF1RHb1 proline-alanine mutations were constructed using short overlap extension (SOE) PCR. *Tg*MAF1RHb1 and *Tg*MAF1RHa1 C-terminal mutants were generated using a site-directed mutagenesis kit (NEB Q5 site-directed mutagenesis). All constructs were confirmed by sanger sequencing methods using *Tg*MAF1RHb1 specific primers (Genewiz). The constructs contain the endogenous *Tg*MAF1RHb1 promoter followed by the start codon, signal peptide, N-terminal hemagglutinin (HA) epitope tag and end at the stop codon. Transgenic lines were generated using *Tg*Me49ΔHXGPRT:Luciferase parasites that were transfected with 55 μg of DNA

linearized with *Hind*III. A T25 flask of *Tg*Me49ΔHXGPRT:Luciferase parasites was scraped and passed through a 25 and 27 gauge needle followed by centrifugation for 10 minutes at 800xg. $2x10^7$ parasites were re-suspended in Cytomix (120 mM KCl; 0.15 mM $CaCl_2$; 10 mM KPO_4; 25 mM Hepes, 2mM EDTA, 5mM $MgCl_2$; pH to 7.6), GSH and ATP. Parasites were electroporated with 1.6Kv and 25μF. Following 24 hours of growth in cDMEM, parasites were selected with mycophenolic acid (MPA)/xanthine. Selected populations were then cloned via limited serial dilution in a 96-well plate. Cloned parasites were confirmed through immunofluorescence assays (IFA) by probing for HA epitope tag.

2.5.7 Immunofluorescence assays and microscopy

HFFs and NRK mito-RFP cells were grown to 100% and 60% confluency, respectively, on 0.7 cm_2 8-well glass chamber slide system (ThermoFisher Scientific) in cDMEM. Monolayers were infected at an MOI of 1 with transgenic parasites. Cell were fixed at 18 hpi with 4% paraformaldehyde for 15 min. and blocked/permeabilized with blocking buffer (5% BSA, 0.1% Triton X-100, PBS). HFFs and NRK mito-RFP cells were then probed with anti-HA rat monoclonal antibody (3F10 clone, Roche) diluted to 0.1 μg/mL in blocking buffer (see above) for 1 hour at room temperature while shaking. HFFs were also incubated in anti-MTCO2 mouse monoclonal antibody and cells were washed with PBS. HFFs were incubated in 488 goat anti-rat and 594 goat anti-mouse secondary antibody (Life Technologies Alexa Fluor H+L) for 1 hour at room temperature while shaking, followed by PBS washes. NRK mito-RFP cells were incubated in Alexa 488 goat anti-rat secondary antibody dilution for 1 hour followed by PBS washes. HFFs and NRK mito-RFP were then mounted in Vectashield mounting media (Vector laboratories) and sealed with cover glass. Slides were visualized using both confocal and epifluorescence

microscopy. Images were taken of the three channels: 488 (anti-HA), 594 (anti-MTCO2 and mito-RFP) and DIC/phase. Images were cropped and merged using ImageJ (NIH).

2.5.8 MAF1 paralog alignments and C-terminal structural views

All MAF1 paralog sequences were obtained from GenBank (NCBI accession numbers: SCA78755.1, ANN02899.1, AMN92255.1, AMN92254.1, AMN92252.1, AMN92247.1, AMN92246.1, AMN92253.1, AMN92250.1, AMN92249.1, AMN92248.1) and aligned by percent identity using Clustal Omega. Alignment was visualized using JalView [98]. *Tg*MAF1RHb1 and *Tg*MAF1RHa1 primary amino acid sequences were visualized using UCSF chimera software [99, 100].

2.5.9 TEM and quantification of vacuole coverage

HFFs were infected with *Tg*MAF1RHb1 C-terminal mutants. At 18 hpi cells were fixed with 2.5% glutaraldehyde in PBS for 1 hour at room temperature, washed three times with PBS for 10 minutes, post-fixed for 1 hour at 4C in 1% OsO_4 with 1% potassium ferricyanide, and washed three times with PBS. Samples were then dehydrated in a graded series of alcohol for ten minutes with three changes in 100% ethanol for 15 minutes and changed three times in epon for 1 hour each. Following the removal of epon, samples were covered with resin and polymerized at 37°C overnight and then 48 hours at 60°C (protocol: Center for Biological Imagining - CBI, University of Pittsburgh, Pittsburgh, PA, USA). Samples were cross sectioned and processed by the CBI. Five vacuoles containing two parasites were imaged for each of the infections. Vacuoles

were traced in ImageJ and the percent of the total distance around the vacuole in direct contact with the host mitochondria was quantified for each of the mutants and controls [101].

2.5.10 Immunoprecipitation and immunoblotting

HFFs were infected with cloned *Tg*Me49 parasites expressing either a *Tg*MAF1RHb1 C-terminal mutant, *Tg*Me49:*Tg*MAF1RHb1, *Tg*Me49:Empty Vector, or *Tg*Me49:*Tg*MAF1RHa1 at an MOI of 2. Cell were lysed in IP lysis buffer (50mM Tris - pH 8.0, 150 mM NaCl, 1% IGEPAL CA-630, 0.05% Tween 20) and treated with complete protease inhibitors (Roche) on ice. The insoluble fraction was pelleted at 700xg for 10 minutes at 4°C and the soluble fraction was incubated with Pierce anti-HA magnetic beads (Thermo scientific) for 2 hours at room temperature using a rotator. Beads were washed five times with IP lysis buffer and eluted by boiling in LDS sample buffer (Thermo scientific). Both input and eluate fractions were resolved on 10% SDS-PAGE gel and transferred to nitrocellulose membrane. Membranes were blocked in 5% BSA in PBST and probed with primary antibodies to HA, SAM50, and Tubulin followed by goat horse radish peroxidase (HRP) conjugated secondary antibodies. Bands were visualized with SuperSignal West Pico chemiluminescent substrates (Thermo Scientific). Antibodies used for these experiments: Anti-HA high affinity rat monoclonal antibody (clone 3F10) – Roche, Goat Anti-Rat IgG H&L HRP – Abcam, Anti-SAMM50 antibody – Abcam, Goat Anti-Rabbit IgG HRP – Southern Biotech and Goat Anti-mouse IgG HRP – Southern Biotech.

2.5.11 *In vivo* competition assay

Mice were 12-wk-old Balb/C female mice (Jackson Labs). Using the same mutant and control parasite clones previously listed, we created a 1:1 mix of *Tg*Me49:EV and *Tg*Me49:*Tg*MAF1RHb1, *Tg*Me49:EV and *Tg*MAF1RHb1 16 C-term chimera, *Tg*Me49:EV and *Tg*MAF1RHb1 STL(438-441)RKK, and *Tg*Me49:EV and *Tg*MAF1RHb1 L(441)K. *Tg*ME49:EV served as a baseline control and was not passage matched to the experimental strain. Three mice were intraperitoneally (IP) infected with 10^5 tachyzoites for each of the mixes allowing us to have three biological replicates for each treatment group. Coverslips with confluent HFFs were infected with the same mixed population preparations and probed for HA tag to quantify the input percentage of each mix. Parasite burden was measured daily over the course of the five day infection using in *vivo* bioluminescence imaging [23]. On day 6, each mouse was sacrificed and a peritoneal lavage was performed with PBS. A fraction of the peritoneal content was used to infect a confluent monolayer of HFFs, fixed at 18 hpi, and probed for HA tag for output percentage quantification.

2.5.12 Ethics statement

Animal experiments were conducted according to the guidelines of the American Veterinary Medical Association. Accordingly, all euthanasia of animals was carried out using controlled exposure to CO_2. All animal protocols were approved by the local Institutional Animal Care Committee at the University of Pittsburgh under IACUC protocol #12010130.

3.0 *Toxoplasma gondii* sequestration of host mitochondria requires key regulatory mitochondrial machinery

3.1 Summary

Host mitochondrial association (HMA) is a well-known phenomenon during *Toxoplasma gondii* infection of the host cell. The multicopy locus *mitochondrial association factor 1* (*MAF1*) is required for HMA and *MAF1* encodes distinct paralogs of secreted dense granule effector proteins, some of which mediate the HMA phenotype (MAF1b paralogs drive HMA, MAF1a paralogs do not). To identify host proteins required for MAF1b-mediated HMA, we performed quantitative proteomics on host cells infected with Type II parasites expressing MAF1b, MAF1a, and a C-terminal mutant of MAF1b that is also HMA incompetent. Across all three samples we identified ~1,360 MAF1-interacting proteins, but only 13 that were uniquely and significantly enriched in MAF1b pulldowns compared to controls. The gene products include multiple host mitochondria outer membrane proteins, including proteins critical for mitochondrial protein import and immune modulation. Through a focused esiRNA screen, we identified both the mitochondrial receptor protein, TOM70, and mitochondria-specific chaperone, HSPA9, as essential mediators of HMA. Additionally, the enrichment of TOM70 at the PVM interface suggests parasite-driven sequestration of TOM70 by the parasite. These results show that the interface between the *T. gondii* vacuole and the host mitochondria is characterized by direct and/or indirect interactions between a single parasite effector and multiple target host proteins, some of which may be critical for the HMA phenotype itself. The elucidation of the functional members of

this complex will permit us to explain the link between HMA and changes in the biology of the host cell and the inflammatory response.

3.2 Introduction

Direct interactions between intracellular microorganisms and their host mitochondria have been characterized in many pathogen species including *Legionella pneumophila* [43, 44], *Hammondia hammondi* [12], *Toxoplasma gondii* [31, 40, 41], and *Chlamydia psittaci* [45, 102]. In the case of *L. pneumophila*, the bacterium utilizes its type IV secretion system to inject the effector protein, MitF, which traffics to the Legionella-containing vacuole membrane (LVM) [44]. The presence of MitF at the LMV mediates recruitment of the GTPase dynamin 1-like protein (Drp1) to the mitochondrial membrane and induces mitochondrial fission in macrophages [44]. In the case of *T. gondii*, this eukaryotic parasite secretes the effector mitochondrial association factor 1b (MAF1b) to the parasitophorous vacuolar membrane (PVM), and this effector is required for the direct binding and manipulation of host mitochondria [11, 12, 31].

The *T. gondii MAF1* locus has a unique structure compared to its closest extant relatives, *Neospora caninum* and *H. hammondi* [12]. Based on extensive sequence analysis and molecular cloning, we found that the *MAF1* locus in *T. gondii*, which contains 6-10 copies dependent on the strain, originated from a single *MAF1* copy (MAF1a) in *N. caninum* which is incapable of driving HMA [12]. *H. hammondi* has two tandem copies of *MAF1* (*MAF1a* and *MAF1b*), and this organism is HMA-competent [12]. *H. hammondi* MAF1b facilitates HMA whereas its MAF1a paralog does not, suggesting that a neofunctionalization event occurred during *H. hammondi* speciation [10]. Here, we utilized both MAF1a and MAF1b paralogs, in addition to a mutated

HMA-incompetent version of MAF1b to identify the host proteins required for MAF1b binding at the PVM.

Through the use of quantitative mass spectrometry and a focused esiRNA knockdown [103] screen we were able to identify an outer mitochondrial membrane receptor, TOM70, as being critical for mediating HMA. TOM70 is a receptor protein and member of the translocase of the outer membrane (TOM) complex [50, 104, 105]. Collectively, the TOM complex is required for recognition and transport of mitochondrial-trafficked proteins [106, 107]. While decades of work focus on the TOM20 receptor, recent discoveries shed light on the requirement of TOM70 for maintaining mitochondrial health, dynamics, immunity, and calcium homeostasis [48, 49, 51, 55]. TOM70 binds pre-mitochondrial proteins containing internal mitochondrial targeting sequences through a binding groove [106-108]. Upon binding to TOM70, pre-proteins are thread through the TOM40 pore into the inner membrane space [104, 106]. Many of these proteins are destined for the sorting and assembly machinery component 50 (SAM50) which is known to interact with TOM70 [109]. Previous work identified SAM50 along with mitochondrial inner membrane bridging space complex (MIB) members as putative binding partners for MAF1b [59]. However, our data show that TOM70 is required for MAF1b-mediated HMA, while SAM50 is not.

Additionally, TOM70-specific substrates are targeted to TOM70 through the Hsp70-family of chaperone proteins, specifically HSPA9 (GRP75) [110]. We found that HSPA9 is also required to mediate HMA through our focused esiRNA screen. Both TOM70 and HSPA9 cluster at mitochondrial-ER contact sites formerly known as MAMs which are implicated in ER-mitochondrial calcium transfer, autophagy signaling, mitochondrial metabolism, and immune signaling [50, 51, 55, 58, 111]. Therefore, MAF1b binding to HSPA9 and/or TOM70 could disrupt these mitochondrial-ER contact sites and downstream signaling, leading to changes in host

responses to infection. Overall, these data identify both TOM70 and HSPA9 as essential host proteins for MAF1b binding and TOM70 enrichment at the PVM. These interactions may suggest manipulation of the mitochondrial-ER interface and contribute to parasite *in vivo* fitness.

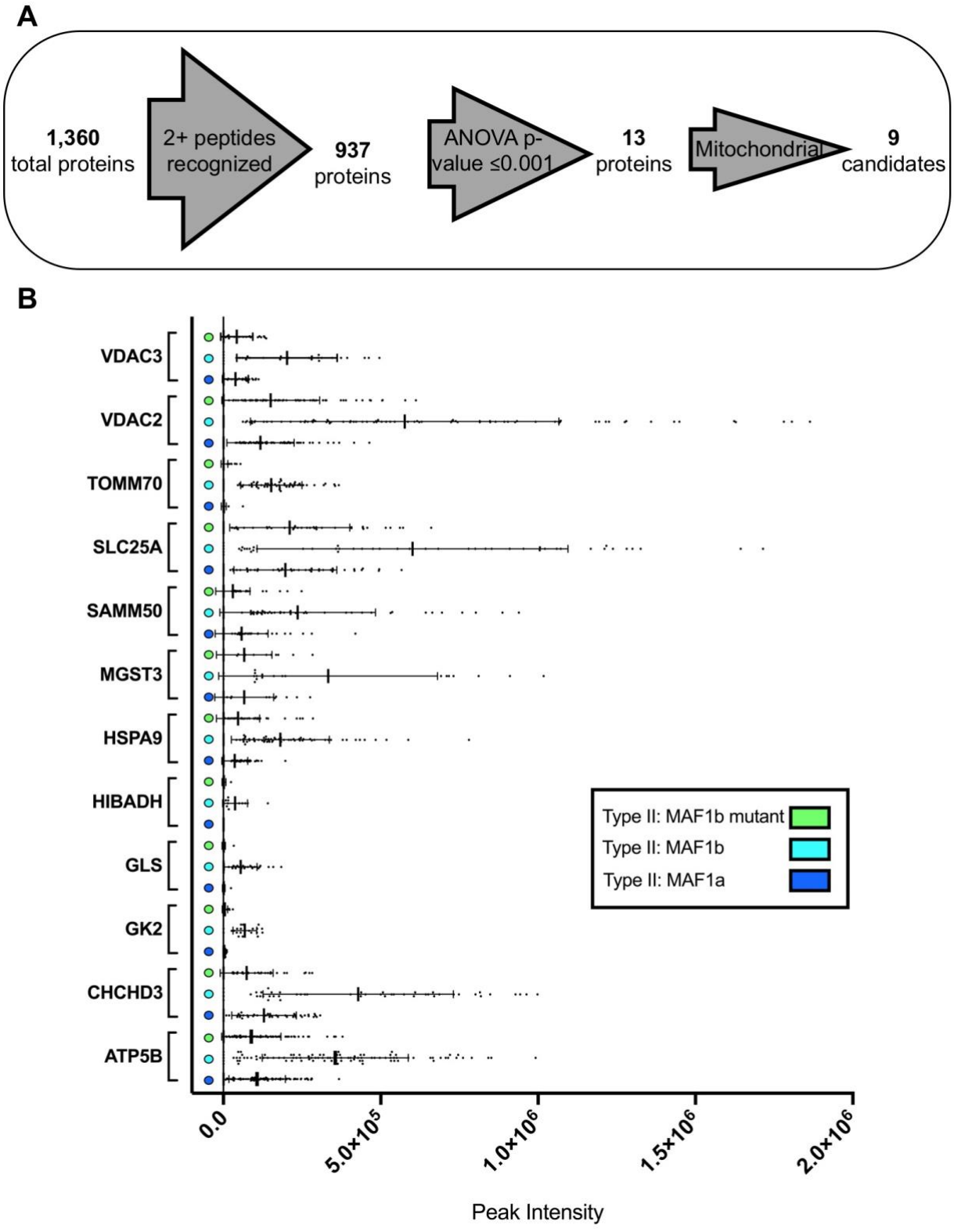
While the host mitochondrial association (HMA) phenomenon in the cases of *L. pneumophila* and *T. gondii* has been recognized for years in literature, the exact molecular mechanism and physiological impact of these interactions is unknown. Using quantitative mass spectrometry, we have now identified the host proteins required for *T. gondii* HMA.

3.3 Results

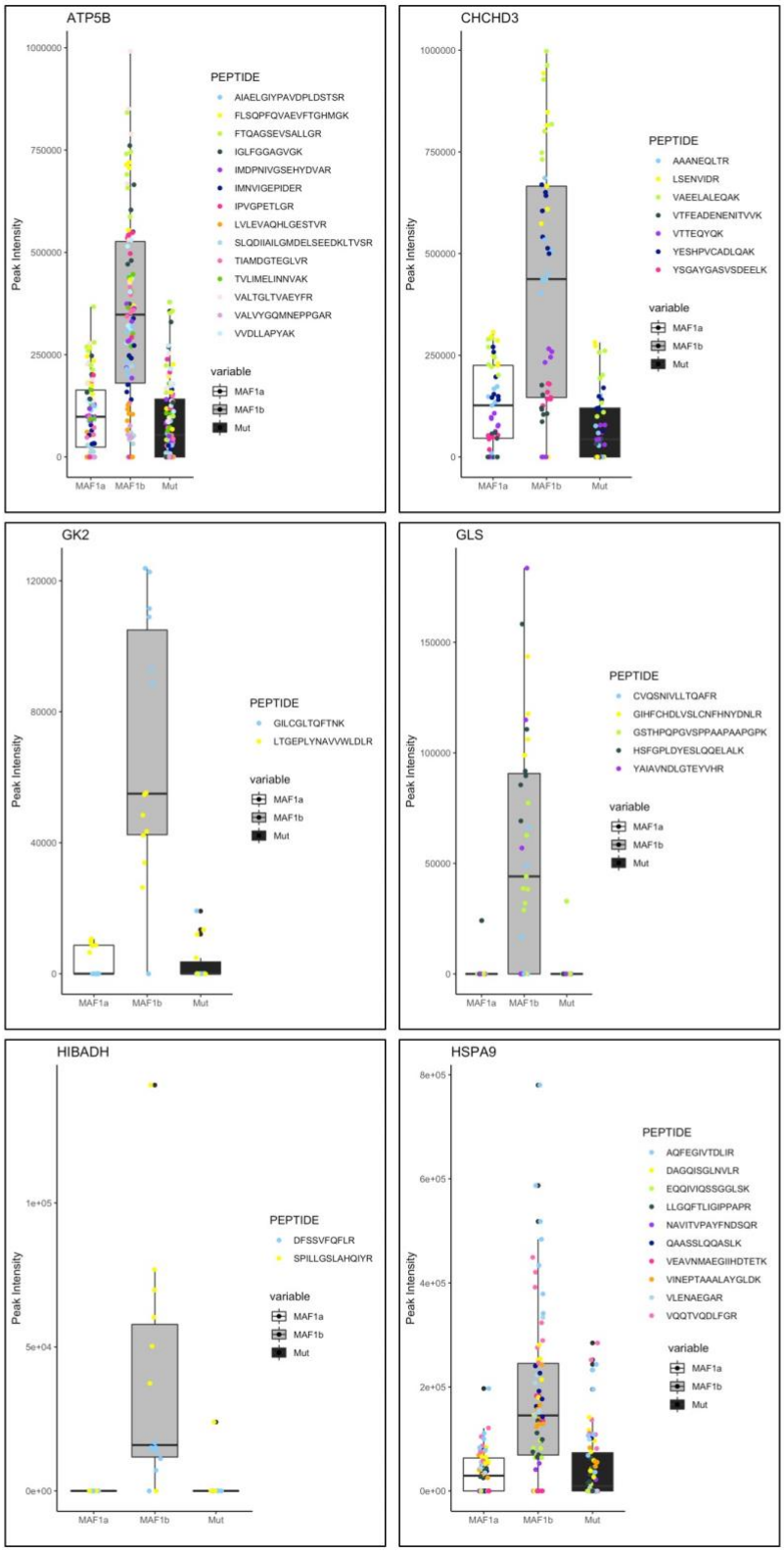
3.3.1 Quantitative mass spectrometry analysis identified a network of potential MAF1b host binding partners on the mitochondria

Type II *T. gondii* strains (including *Tg*Me49) are HMA negative and do not express *Tg*MAF1RHb1 (MAF1b) protein [12, 31]. We previously characterized three *Tg*Me49 parasite clones for their ability to mediate HMA: *Tg*Me49:*Tg*MAF1RHb1 (HMA+), *Tg*Me49:*Tg*MAF1RHa1 (HMA-), and *Tg*Me49:*Tg*MAF1RHb1 STYL(438-441)RKYK (HMA-)[11, 12]. While the HMA phenotype has long been recognized [11, 12, 31, 40, 59], the molecules modulating the interaction between the mitochondria and the parasite have remained elusive. To identify host proteins required for this interaction, we infected a monolayer of HFFs with each parasite clone described above and performed a co-immunoprecipitation assay followed by quantitative mass spectrometry. Prior to analysis and selection of candidate binding partners, we identified a total of 1,360 proteins in our IP samples (Fig 3-1A). In order to select candidate

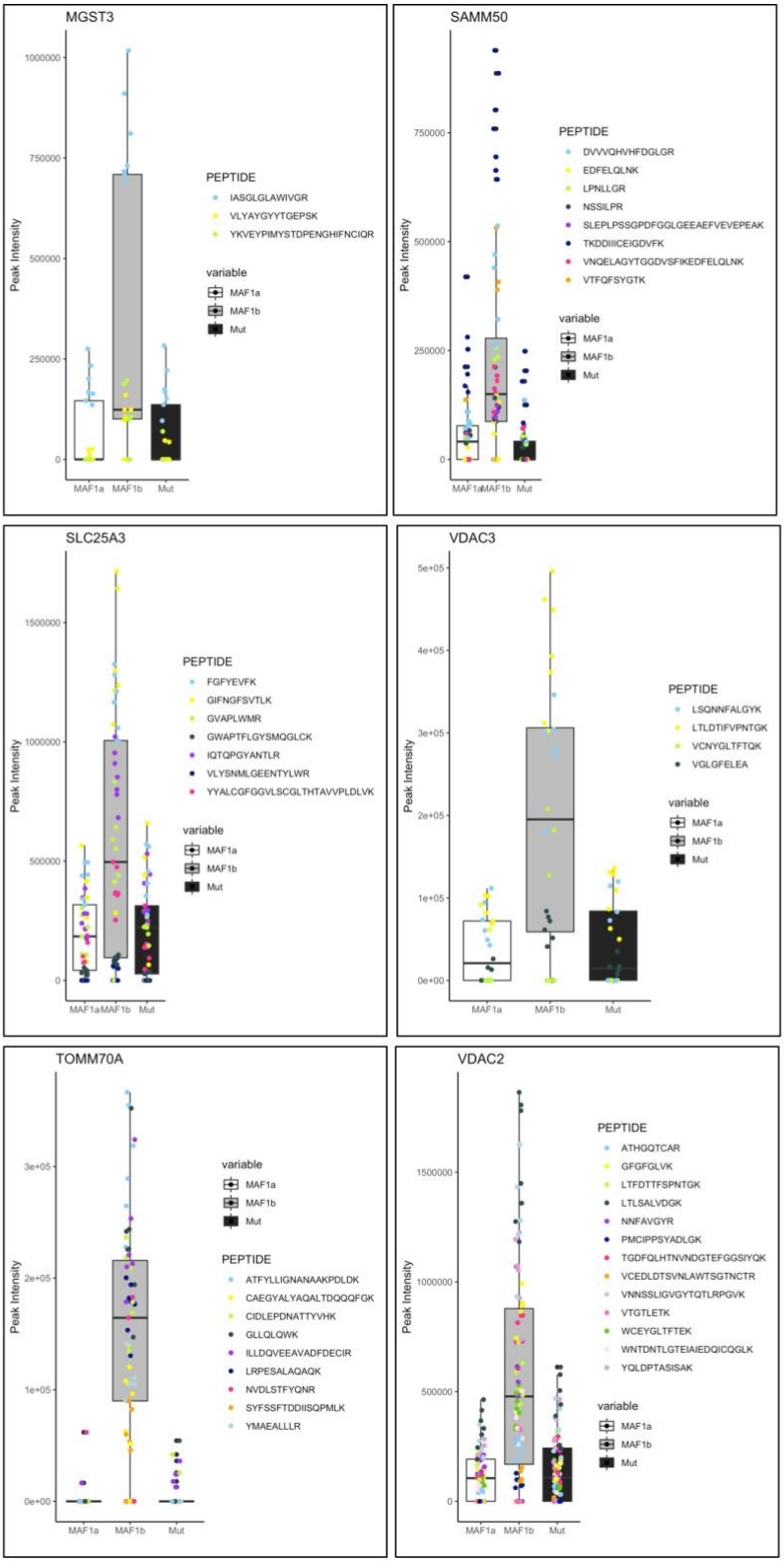
proteins we first excluded proteins that had a single peptide hit, leaving 937 proteins (Fig 3-1A). Each identified peptide yielded a peak intensity score directly related to its ability to ionize. We hypothesized that proteins required for binding MAF1b at the PVM would be enriched in the *Tg*Me49:*Tg*MAF1RHb1 (Type II: MAF1b) samples and less enriched in the *Tg*Me49:*Tg*MAF1RHa1 (Type II:MAF1a) and *Tg*Me49:*Tg*MAF1RHb1 STYL(438-441)RKYK (Type II:MAF1b mutant) IP samples. For this reason, we then selected candidate binding proteins that had a ≥2 fold change in the Type II:MAF1b samples compared to Type II:MAF1a and Type II:MAF1b mutant and an ANOVA p value of <0.001 (Fig 3-1A). These conditions revealed a final list of 13 candidate binding partners (Table 1). Candidate binding partners included proteins required for cellular metabolism including 3-Hydroxyisobutyrate dehydrogenase (HIBADH), glycerol kinase 2 (GK2), and ATP synthase F1 subunit beta (ATP5B) (Table 1). SAM50, which was previously identified as a potential binding partner of MAF1b was also enriched in the Type II:MAF1b samples [11, 59] (Table 1). Interestingly, proteins known to be enriched at mitochondrial-ER contact sites, including translocase of the outer mitochondrial membrane receptor (TOM70) and HSPA9, were also identified (Table 1). While ranking proteins based on their associated peptide ANOVA p-values is important for investigating candidates, we also prioritized proteins based on both their degree of variance in raw peak intensity (Fig 3-1B) and overall scatter of individual peptide ionizations (Fig 3-1C). High priority candidates had a narrow variance (Fig 3-1B and 3-1C). To assess the scatter of individual peptides, we normalized each of the peptide ionization intensity replicates for both SAM50 and TOM70. We found that upon normalization, individual peptides do not scatter as drastically compared to the raw data (Fig 3-1D).


A
1,360
total proteins
2+ peptides recognized
937
proteins
ANOVA p-value ≤0.001
13
proteins
Mitochondrial
9
candidates
B
VDAC3
VDAC2
TOMM70
SLC25A
SAMM50
MGST3
HSPA9
HIBADH
GLS
GK2
CHCHD3
ATP5B
Type II: MAF1b mutant
Type II: MAF1b
Type II: MAF1a
0.0
5.0×10^5
1.0×10^6
1.5×10^6
2.0×10^6
Peak Intensity

C


ATP5B
Peak Intensity
1000000
750000
500000
250000
0
MAF1a
MAF1b
Mut
PEPTIDE
AIAELGIYPAVDPLDSTSR
FLSQPFQVAEVFTGHMGK
FTQAGSEVSALLGR
IGLFGGAGVGK
IMDPNIVGSEHYDVAR
IMNVIGEPIDER
IPVGPETLGR
LVLEVAQHLGESTVR
SLQDIIAILGMDELSEEDKLTVSR
TIAMDGTEGLVR
TVLIMELINNVAK
VALTGLTVAEYFR
VALVYGQMNEPPGAR
VVDLLAPYAK
variable
MAF1a
MAF1b
Mut
CHCHD3
Peak Intensity
1000000
750000
500000
250000
0
MAF1a
MAF1b
Mut
PEPTIDE
AAANEQLTR
LSENVIDR
VAEELALEQAK
VTFEADENENITVVK
VTTEQYQK
YESHPVCADLQAK
YSGAYGASVSDEELK
variable
MAF1a
MAF1b
Mut
GK2
Peak Intensity
120000
80000
40000
0
MAF1a
MAF1b
Mut
PEPTIDE
GILCGLTQFTNK
LTGEPLYNAVVWLDLR
variable
MAF1a
MAF1b
Mut
GLS
Peak Intensity
150000
100000
50000
0
MAF1a
MAF1b
Mut
PEPTIDE
CVQSNIVLLTQAFR
GIHFCHDLVSLCNFHNYDNLR
GSTHPQPGVSPPAAPAAPGPK
HSFGPLDYESLQQELALK
YAIAVNDLGTEYVHR
variable
MAF1a
MAF1b
Mut
HIBADH
Peak Intensity
1e+05
5e+04
0e+00
MAF1a
MAF1b
Mut
PEPTIDE
DFSSVFQFLR
SPILLGSLAHQIYR
variable
MAF1a
MAF1b
Mut
HSPA9
Peak Intensity
8e+05
6e+05
4e+05
2e+05
0e+00
MAF1a
MAF1b
Mut
PEPTIDE
AQFEGIVTDLIR
DAGQISGLNVLR
EQQIVIQSSGGLSK
LLGQFTLIGIPPAPR
NAVITVPAYFNDSQR
QAASSLQQASLK
VEAVNMAEGIIHDTETK
VINEPTAAALAYGLDK
VLENAEGAR
VQQTVQDLFGR
variable
MAF1a
MAF1b
Mut

C


MGST3
Peak Intensity
1000000
750000
500000
250000
0
MAF1a
MAF1b
Mut
PEPTIDE
IASGLGLAWIVGR
VLYAYGYYTGEPSK
YKVEYPIMYSTDPENGHIFNCIQR
variable
MAF1a
MAF1b
Mut
SAMM50
Peak Intensity
750000
500000
250000
0
PEPTIDE
DVVVQHVHFDGLGR
EDFELQLNK
LPNLLGR
NSSILPR
SLEPLPSSGPDFGGLGEEAEFVEVEPEAK
TKDDIIICEIGDVFK
VNQELAGYTGGDVSFIKEDFELQLNK
VTFQFSYGTK
variable
MAF1a
MAF1b
Mut
SLC25A3
Peak Intensity
1500000
1000000
500000
0
PEPTIDE
FGFYEVFK
GIFNGFSVTLK
GVAPLWMR
GWAPTFLGYSMQGLCK
IQTQPGYANTLR
VLYSNMLGEENTYLWR
YYALCGFGGVLSCGLTHTAVVPLDLVK
variable
MAF1a
MAF1b
Mut
VDAC3
Peak Intensity
5e+05
4e+05
3e+05
2e+05
1e+05
0e+00
PEPTIDE
LSQNNFALGYK
LTLDTIFVPNTGK
VCNYGLTFTQK
VGLGFELEA
variable
MAF1a
MAF1b
Mut
TOMM70A
Peak Intensity
3e+05
2e+05
1e+05
0e+00
variable
MAF1a
MAF1b
Mut
PEPTIDE
ATFYLLIGNANAAKPDLDK
CAEGYALYAQALTDQQQFGK
CIDLEPDNATTYVHK
GLLQLQWK
ILLDQVEEAVADFDECIR
LRPESALAQAQK
NVDLSTFYQNR
SYFSSFTDDIISQPMLK
YMAEALLLR
VDAC2
Peak Intensity
1500000
1000000
500000
0
PEPTIDE
ATHGQTCAR
GFGFGLVK
LTFDTTFSPNTGK
LTLSALVDGK
NNFAVGYR
PMCIPPSYADLGK
TGDFQLHTNVNDGTEFGGSIYQK
VCEDLDTSVNLAWTSGTNCTR
VNNSSLIGVGYTQTLRPGVK
VTGTLETK
WCEYGLTFTEK
WNTDNTLGTEIAIEDQICQGLK
YQLDPTASISAK
variable
MAF1a
MAF1b
Mut

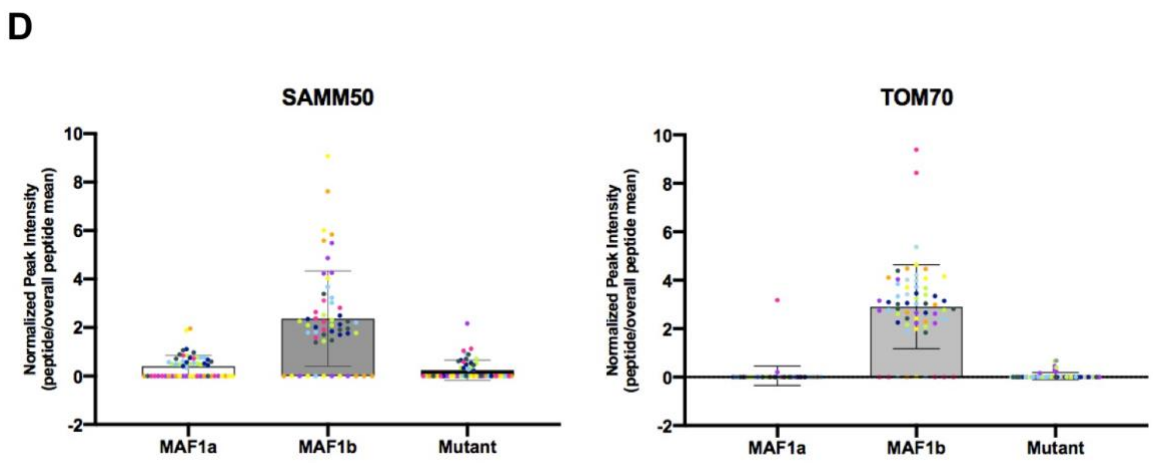


Figure 3-1 Quantitative mass spectrometry analysis identified 13 potential MAF1 host binding partners.

(A) Selection criteria for candidate proteins included identifying proteins with 2 or more peptide hits and proteins significantly higher in abundance (greater than 2-fold change) in Type II:MAF1b IP samples in comparison to the Type II:MAF1a and Type II:MAF1b mutant samples and must have an ANOVA p-value ≤0.001. Nine of the final 13 candidates have a mitochondrial function based on DAVID 6.8 analysis. (B) Peak intensity plots for the cumulative ionized peptides from the 13 candidate proteins. All identified peptides for each IP sample (n=7) are graphed against their peak ionization intensity. (C) Peak intensity plots for the cumulative ionized peptides for each mitochondrial candidate protein (9 total). Peptides are colored and labeled to illustrate individual recognized peptides for each infection condition (n=7). (D) Normalization of peak intensities for both SAM50 and TOM70 cumulative peptide ionizations. Ionization values for each peptides were normalized to overall mean peptide ionization. Peptides are colored as in (C).

Table 1. List of identified 13 candidate proteins sorted by their associated ANOVA p-value comparing mean of each peptide for each candidate protein described in Figure 3-1A. Gene names are provided with their relevance to mitochondrial biology.

Rank	Gene Name	ANOVA p-value (comparing mean of each peptide)	Mitochondrial Protein
1	IMMT	2.5E-11	✓
2	VDAC2	4E-10	✓
3	MGST3	1.5E-09	✗
4	CHCHD3	1.04E-08	✓
5	ATP5B	3.33E-07	✓
6	VDAC3	4.1E-07	✗
7	SLC25A3	4.28E-07	✓
8	GLS	6.03E-07	✗
9	TOMM70A	4.35E-06	✓
10	SAMM50	1.51E-05	✓
11	HIBADH	3.62E-05	✓
12	HSPA9	4.6E-05	✓
13	GK2	0.000443	✓

3.3.2 Focused RNAi screen reveals requirement for TOM70 and HSPA9 in MAF1b-driven HMA

To identify host proteins required for *T. gondii* recruitment of mitochondria, we generated endoribonuclease-prepared siRNAs (esiRNAs) [103] to knockdown candidate gene expression and assess co-localization of the mitochondria with the PVM. Human U2OS cells were transfected for 48 hours with either a target or non-target esiRNA followed by infection with Type II:MAF1b parasites. Knockdown was analyzed using RT-qPCR and candidate transcripts were successfully knocked down between 75-95% (Fig 3-2A). We performed an immunofluorescence assay (IFA), staining mitochondria and HA-tagged MAF1b to quantify the percent of HMA+ vacuoles. To

avoid inherent bias in this assay, we blinded each infection condition and counted three replicates (n=3 coverslips per treatment). In vehicle control and non-target (Cytophilin B esiRNA) treated cells, 76-84% of the Type II:MAF1b-containing vacuoles were HMA+ (Fig 3-2C). Knockdown of multiple candidates identified in our quantitative mass spectrometry screen, including VDAC2, GLS, GK2, HIBADH, SAM50, and ATP5B, did not significantly reduce HMA efficiency, where again between 71% and 81% of the MAF1b-containing vacuoles were HMA+ (Fig 3-2B and C). Interestingly, while many of these candidates are key proteins for the structural integrity and function of the mitochondria, only knockdown of ATP5B, a subunit of the multimeric ATPase F_1 subunit, resulted in strikingly unhealthy mitochondria. Importantly, even the fragmented bead-like mitochondria in these samples bound to the PVM (Fig 3-2B). TOM70 and HSPA9 esiRNA treatments were the only proteins to knockdown the HMA efficiency of the parasites by 30-40% (Fig 3-2B and C). We validated TOM70 knockdown by comparing TOM70 protein levels between vehicle and HIBADH esiRNA-treated cells (Fig 3-2D). The receptor protein, TOM70, recognizes key internal mitochondrial targeting sequences in pre-proteins destined for translocation through the TOM complex [106, 107, 112]. To confirm that the decrease of HMA was due to TOM70 and not an artifact of disrupting the TOM complex, we knocked down transcript of TOM40, which forms the pore of the complex, and this did not significantly alter HMA (Fig 3-2B and C). The degree of HMA is measured by quantifying the percent of the vacuole in contact with host mitochondria (percent vacuole coverage). Attachment of the vacuole to the mitochondria is most readily quantified using transmission electron microscopy (TEM) because the distance between the vacuole and mitochondria is 12.04 nm ± 3.05 nm and starkly different than unattached mitochondria, which can be more difficult to distinguish using epifluorescence microscopy [41]. To assess the attenuated HMA phenotype, we infected cells knocked down for either HIBADH or

TOM70 with Type II:MAF1b parasites and quantified the percent vacuole coverage using TEM. Coverage in the TOM70 treated cells was significantly lower than both the vehicle-treated and HIBADH knockdown cells (Fig 3-2E, ****$p<0.0001$) confirming the requirement of the mitochondrial TOM70 receptor in mediating HMA.

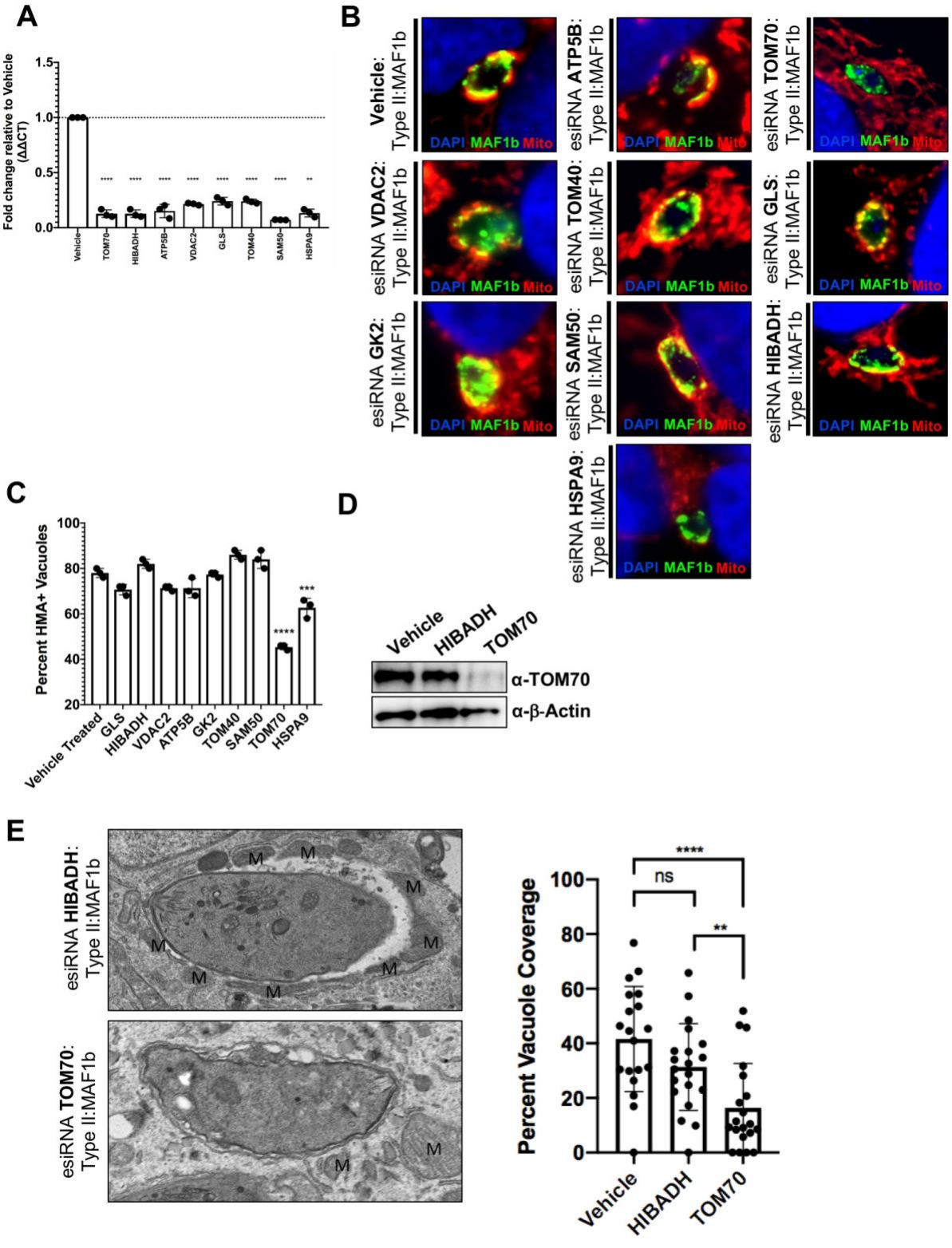

A
Fold change relative to Vehicle (ΔΔCT)
1.5
1.0
0.5
0.0
****
****
****
****
****
****
****
**
Vehicle
TOM70
HIBADH
ATP5B
VDAC2
GLS
TOM40
SAM50
HSPA9
B
Vehicle: Type II:MAF1b
esiRNA ATP5B: Type II:MAF1b
esiRNA TOM70: Type II:MAF1b
esiRNA VDAC2: Type II:MAF1b
esiRNA TOM40: Type II:MAF1b
esiRNA GLS: Type II:MAF1b
esiRNA GK2: Type II:MAF1b
esiRNA SAM50: Type II:MAF1b
esiRNA HIBADH: Type II:MAF1b
esiRNA HSPA9: Type II:MAF1b
DAPI MAF1b Mito
C
Percent HMA+ Vacuoles
100
80
60
40
20
****
***
Vehicle Treated
GLS
HIBADH
VDAC2
ATP5B
GK2
TOM40
SAM50
TOM70
HSPA9
D
Vehicle
HIBADH
TOM70
α-TOM70
α-β-Actin
E
esiRNA HIBADH: Type II:MAF1b
esiRNA TOM70: Type II:MAF1b
M
Percent Vacuole Coverage
100
80
60
40
20
0
****
ns
**
Vehicle
HIBADH
TOM70

Figure 3-2 TOM70 and HSPA0 are required for MAF1b-mediated HMA.

(A) U2OS cells were treated with 25nM esiRNA targeting the candidate gene of interest for 48 hours followed by Type II:MAF1b infection for 24 hours (MOI=2). RNA was collected at 48 hours and qPCR analysis was performed. **p=0.0021, ****p<0.0001 unpaired two-tailed t-test between cognate vehicle-treated control CT values. (B) U2OS cells were treated as in (A). Following 48 hours of esiRNA treatment, cells were infected for 24 hours with Type II: MAF1b parasites. Cells were fixed and visualized using fluorescence microscopy. Immunofluorescence staining was performed with antibodies against the HA epitope tag and mitochondria. (C) HMA+ and HMA- vacuoles were counted (n=3, 50 per treatment group) using fluorescence microscopy. ***p=0.0005, ****p<0.0001 one-way ANOVA multiple comparisons. (D) U2OS cells were treated as in (A) with 25nM esiRNA targeting TOM70, HIBADH, or vehicle control, and TOM70 protein levels were measured by Western blot analysis. β-Actin was used as a loading control. (E) U2OS cells were treated as in (A) with 25nM esiRNA targeting TOM70 or HIBADH. After 48 hours of treatment, cells were infected with Type II:MAF1b parasites for 24 hours followed by fixation in 2.5% glutaraldehyde and processed for transmission electron microscopy. For each esiRNA treatment, a minimum of 19 vacuoles were quantified. Mitochondria (M) are labelled. ****p<0.0001 **p=0.0055 unpaired two-tailed t-test.

3.3.3 MAF1 ectopic expression validates the requirement of TOM70 in mediating mitochondrial attachment at the vacuole

Our esiRNA investigation identified TOM70 as being a key mediator of Type II:MAF1b-driven interaction with the host mitochondria. To determine if the TOM70 requirement for MAF1b-mediated HMA required other parasite factors, we expressed GFP-tagged MAF1 in host cells directly. Overexpression of GFP-tagged *T. gondii* effectors is utilized to tease out effector localization, physiological response, and immune modulation on the host cell without the collective effect of the parasite's effector proteome. To first assess MAF1 localization and mitochondrial binding, we N-terminally GFP tagged both *Tg*MAF1RHb1 and *Tg*MAF1RHa1 (MAF1b and MAF1a respectively) using the pcDNA3.1/NT-GFP-TOPO expression plasmid

(ThermoFisher Scientific). Expression of both GFP-MAF1a and GFP-MAF1b was observed as early as 6 hours post transfection and up to 48 hours post transfection in U2OS cells (Fig 3-3A). Strikingly, GFP-MAF1b co-localized with the host cell mitochondria as soon as 6 hours post-transfection and mitochondrial morphology changed dramatically between the 6 and 48 hour time points (Fig 3-3A). Our observations of MAF1b-associated mitochondrial aggregation are consistent with previous work characterizing the changes to the mitochondrial cristae and overall increase in mitochondrial cross sectional area in vacuole-associated mitochondria [31]. Interestingly, the three MAF1b residues (STL→RKK) required for HMA during parasite infection (Chapter 2, [11]) are also required for MAF1b association with mitochondria during ectopic expression (Fig 3-3B).

Using this system we tested MAF1b binding partner candidates from our quantitative mass spectrometry screen for their ability to mediate mitochondrial localization of GFP-MAF1b. We knocked down each candidate in U2OS cells using esiRNAs as described above for 48 hours, followed by 24 hours of GFP-MAF1b expression by lipofection. Neither HIBADH or SAM50 knockdown disrupted GFP-MAF1b mitochondrial localization (Fig 3-3C). Additionally, the mitochondria morphology was severely altered and the mitochondria aggregated around the nucleus. Even in the aggregated mitochondria, GFP-MAF1b remained localized at the mitochondria in the knockdown cells (Fig 3-3C). In contrast, knockdown of TOM70 followed by GFP-MAF1b expression had markedly different results. TOM70 knockdown cells had severely impaired GFP-MAF1b co-localization with mitochondria (Fig 3-3C). These data further confirm the requirement of TOM70 in mediating MAF1b binding to the mitochondria.

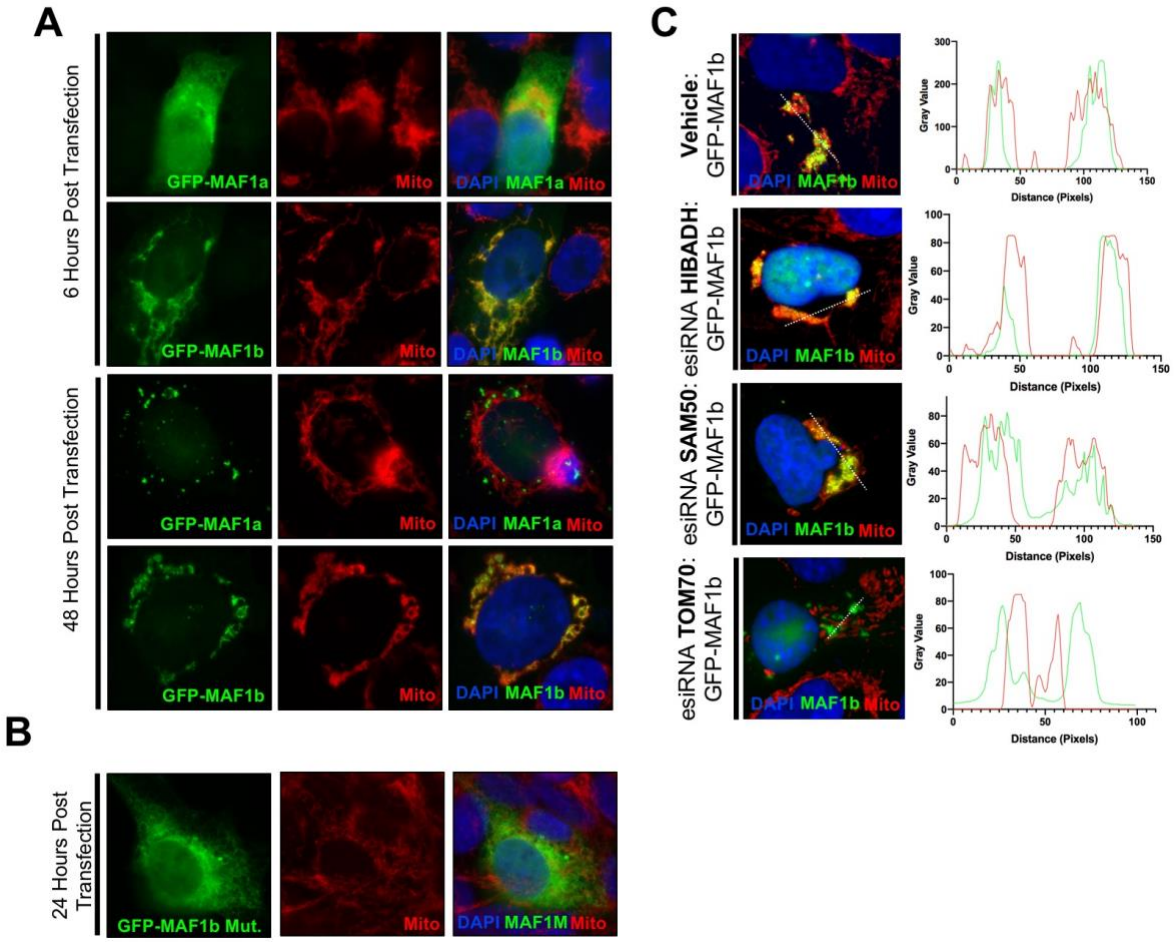


Figure 3-3 Ectopic expression of GFP-MAF1b/a confirms the requirement of TOM70 for mediating HMA.

(A) U2OS cells were transfected with either GFP-MAF1b or GFP-MAF1a. Cells were fixed at either 6 or 8 hours post transfection and probed with an antibody against mitochondria. Samples were visualized using epifluorescence microscopy. (B) U2OS cells were transfected with GFP-MAF1b mutant containing the C-terminal mutations STYL (438-441) RKYK for 24 hours followed by fixation and immunofluorescence analysis. (C) U2OS cells were treated with 25nM esiRNA against the indicated transcript for 48 hours following by transfection with GFP-MAF1b for 24 hours. Cells were fixed and immunofluorescence staining was performed with antibodies against the mitochondria. Profile plots of both the GFP-MAF1b and mitochondria correlate to dotted white line in the merged image (left to right).

3.3.4 MAF1 maintains multiple internal mitochondrial targeting sequences

Through a proteomic screen in yeast the mitochondrial ATPase, Atp25, was identified as a TOM70 binding partner [107]. Prior to mitochondrial import via the TOM complex, Atp25 binds to TOM70 through previously unclassified internal mitochondrial targeting sequences [107, 113]. Mitochondrial targeting sequence (MTS) identification algorithms were used to scan the sequence for alpha helices, amphipathic regions, and overall positive charge [113, 114]. These algorithms predicted MTSs predominantly at the N-terminus of the precursor protein. Through the use of the newer TargetP algorithm, mitochondrial sequences were identified throughout the protein with high confidence and these regions were termed internal mitochondrial targeting sequences (iMTSs) [115]. Atp25 is a prime example of a mitochondrial localized protein that contains multiple iMTSs, each of which is required for binding TOM70.

We hypothesized that MAF1b and MAF1a include distinct iMTSs, and that their differences would help to identify the region of MAF1b required for interacting with TOM70 and ultimately to drive HMA. To test this hypothesis, we used the TargetP algorithm to generate iMTS probabilities along the length of MAF1b/a. TargetP profiles for both MAF1b and MAF1a reveal five predicted targeting sequences along the length of both paralogs (Fig 3-4A). All but one of the predicted iMTS regions tightly overlap with each other between MAF1b and MAF1a (Fig 3-4A). The stretch of 40 residues in MAF1b iMT4 have a high TarpetP score as compared to their aligned residues in MAF1a (Fig 3-4A). To identify the specific region in MAF1b required for mitochondrial localization, we generated GFP-MAF1b/a chimeric constructs for ectopic expression in U2OS cells (Fig 3-4B). Each chimeric construct was designed in a manner so that the secondary structure would not be disrupted using both secondary structure prediction software (JPred 4) [116] and the solved structures for MAF1a (PDB:6BSX) and MAF1b (PDB: 6BSR)

(Chapter 2, [11]). For this reason, some chimeric regions extended beyond the boundaries of a predicted iMTS. MAF1 chimera 1 contained the N-terminus of MAF1b (including iMTSs 1 and 2) with the C-terminus of MAF1a (including iMTSs 3, 4, and 5). Each chimera with the C-terminus of MAF1a contained the three C-terminal residues of MAF1b (RKK→STL) required for HMA (Fig 3-4A and B) (Chapter 2, [11]).

Chimera 1 was incapable of mitochondrial localization revealing that iMTSs 1 and 2 were not sufficient to mediate HMA (Fig 3-4B). MAF1 chimera 2 contained the N-terminus of MAF1a (including iMTSs 1 and 2), MAF1b iMTS3, and the C-terminus of MAF1a (including iMTSs 4 and 5) (Fig 3-4B). Expression of MAF1 chimera 2 also failed to drive MAF1 association with the mitochondria (Fig 3-4B). Expression of chimeras 1 and 2 show that MAF1b iMTSs 1, 2, and 3 are not sufficient to mediate HMA. To probe the C-terminal iMTS regions, I generated a third MAF1 chimera which contained the N-terminus of MAF1a (including MAF1a iMTSs 1, 2, and 3) and the C-terminus of MAF1b (including MAF1b iMTSs 4 and 5) (Fig 3-4B). MAF1 chimera 3 was also unable to drive association with host mitochondria (Fig 3-4B). Finally, I expressed MAF1 chimera 4 which contained both the N-terminus and C-terminus of MAF1b (including MAF1b iMTSs 1-2 and iMTSs 4-5). This chimera drove HMA similar to WT GFP-MAF1b (Fig 3-4B). These data suggest the sequences at both the N-terminus and C-terminus of MAF1b are necessary to associate with mitochondria, while residues 102-211 of MAF1a are not.

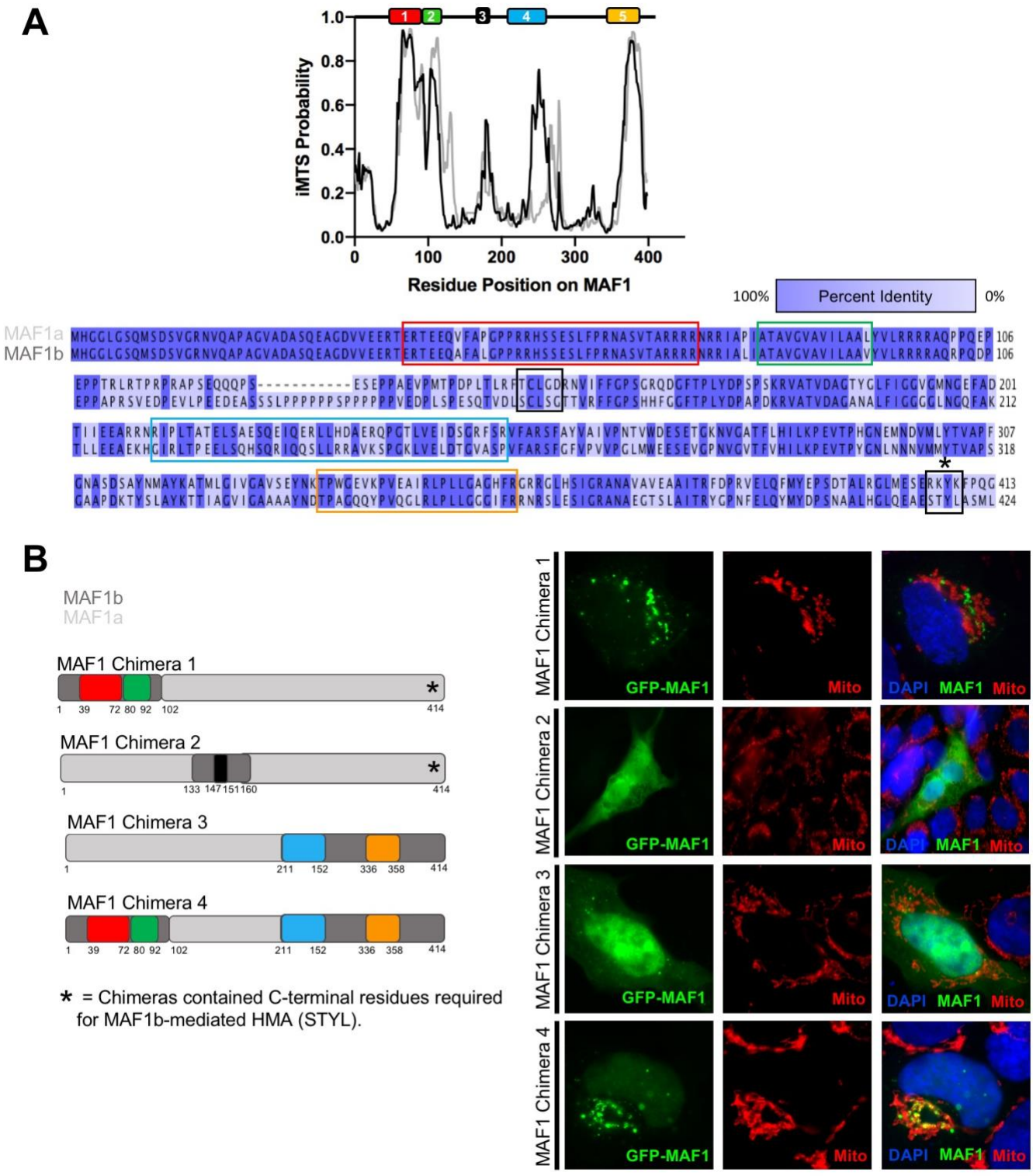

A
iMTS Probability
Residue Position on MAF1
Percent Identity
100%
0%
MAF1a
MAF1b
B
MAF1b
MAF1a
MAF1 Chimera 1
MAF1 Chimera 2
MAF1 Chimera 3
MAF1 Chimera 4
* = Chimeras contained C-terminal residues required for MAF1b-mediated HMA (STYL).
GFP-MAF1
Mito
DAPI MAF1 Mito

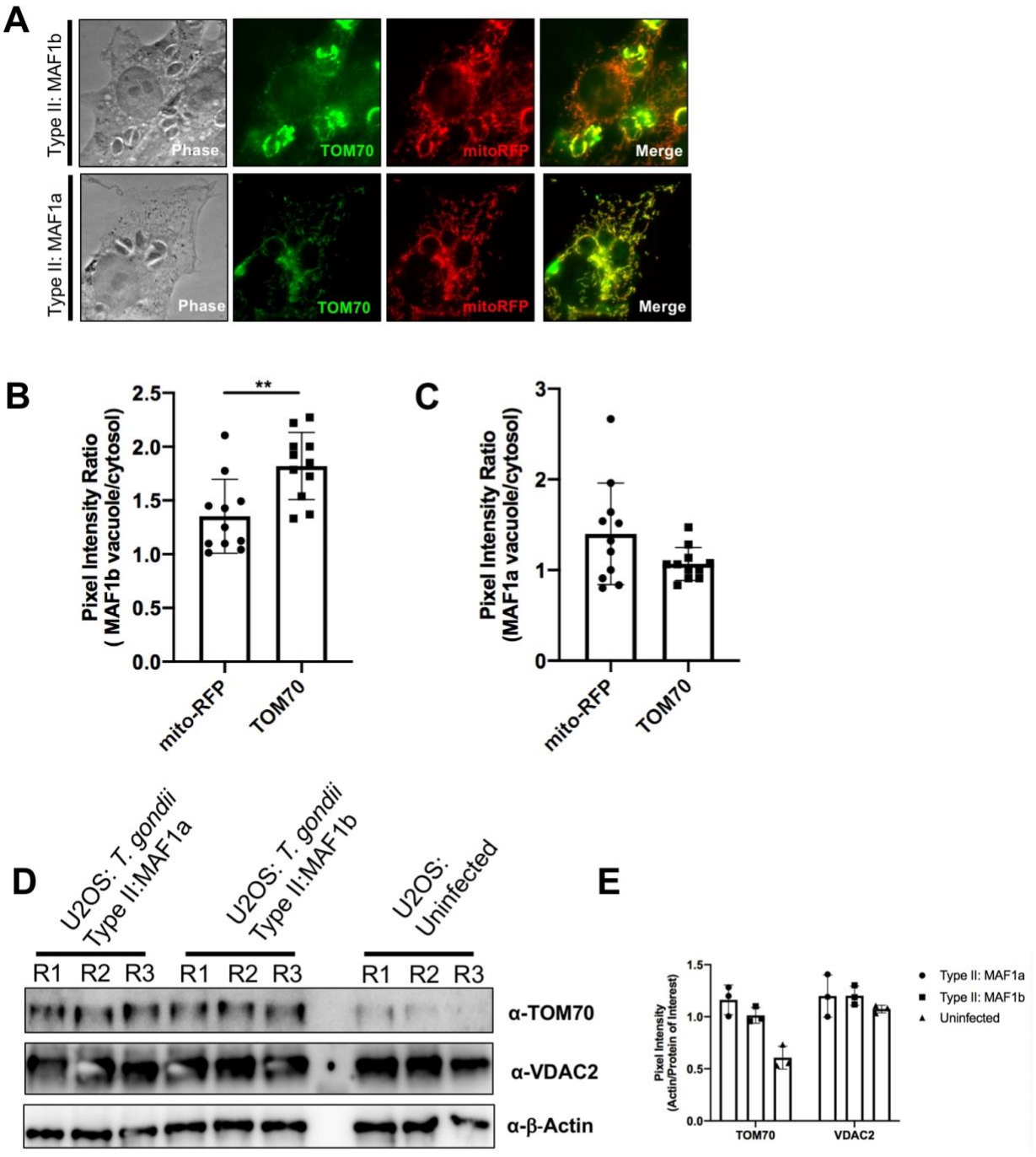
Figure 3-4 Internal mitochondrial targeting sequences (iMTSs) are present in both MAF1b/a paralogs.

(A) TargetP analysis predicts 5 IMTSs present in both MAF1a (light gray) and MAF1b (dark gray) sequences. Each iMTS is illustrated on the x-axis with different colors. Alignments were visualized in JalView after alignment using ClustalOmega. (B) U2OS were transfected with a MAF1 chimeric construct. Cells were fixed 24 hours post transfection and immunofluorescence staining was performed with antibodies against the mitochondria. Cartoon depictions of each generated GFP-MAF1a/b chimera are illustrated to highlight the boundaries and identity of each MAF1 iMTS. Associated iMTS colors are illustrated in (A). The asterisks in MAF1 Chimera 1 and MAF1 Chimera 2 refer to the C-terminal residues of MAF1b required for mediating HMA: STYL(338-441)RKYK.

3.3.5 TOM70 is enriched at the parasitophorous vacuolar membrane

Since most of the MAF1b binding candidates have clear mitochondrial functions, fluorescently labelling a potential binding partner only confirms MAF1b localization to the mitochondria but does not specifically delineate co-localization of that specific protein with MAF1b. We hypothesized that the requirement of TOM70 in mediating HMA would suggest localization of TOM70 in the MAF1b-rich regions on the PVM which could be detected using epifluorescence microscopy. To test this, we measured TOM70 raw intensity values following infection with either Type II:MAF1b or Type II:MAF1a parasites (Fig 3-5A). Following a 24-hour infection, we stained infected mito-RFP normal rat kidney (mito-RFP NRK) cells with a TOM70 antibody. TOM70 was significantly enriched at the MAF1b-PVM interface in comparison to unattached mitochondria. (Fig 3-5A and B). TOM70 was also significantly more enriched than the mitochondrial matrix-localized RFP indicator at the MAF1b-PVM interface (Fig 3-5A and B). This enrichment is lost in the Type II:MAF1a infected cells where we saw no enrichment of RFP-mitochondria or TOM70 at the PVM (Fig 3-5A and C). The enrichment of TOM70 at the vacuole

could be explained by upregulation or stabilization of the protein in cells infected with Type II:MAF1b. We wanted to address this question by infecting U2OS cells with either Type II:MAF1b or Type II:MAF1a parasites followed by Western analysis to assay overall TOM70 protein levels. There was no significant change in TOM70 abundance between infection conditions, suggesting no difference in overall MAF1b/a stabilization (Fig 3-5D and E). TOM70 enrichment at the vacuole surprised us, so we returned to previously VDAC2-stained cells that were ectopically expressing either GFP-MAF1b or GFP-MAF1a to see if another MAF1b candidate binding protein corroborated the enrichment phenotype. The voltage-gated dependent anion channel, VDAC2, is predominantly located in the outer mitochondrial membrane and is part of the VDAC family of proteins required for shuttling metabolites into the mitochondria [117, 118]. VDAC2 localized at the mitochondria as well as throughout the cytoplasm in both the GFP-MAF1b and GFP-MAF1a samples (Fig 3-5F). Because these U2OS were ectopically expressing GFP-MAF1a/b and were not infected with parasites, we quantified VDAC2 enrichment by first measuring the raw pixel intensity ratio of mitochondria (α-mitochondria Abcam) in the regions with MAF1 protein present to the regions where no MAF1 is present. We then compared this ratio to the raw pixel intensity ratio of VDAC2 at the MAF1-rich regions of the cell to VDAC2 regions where MAF1 was absent. These data indicate that VDAC2 is not enriched in the MAF1b regions of the cell. (Fig 3-5F and G). Additionally, there was no significant enrichment of either mitochondria or VDAC2 in the GFP-MAF1a samples (Fig 3-5F and H). Overall, we were able to identify TOM70-specific enrichment in the regions of the vacuole where MAF1b and not MAF1a localizes, suggesting a binding partner complex requiring TOM70.


A
Type II: MAF1b
Type II: MAF1a
Phase
TOM70
mitoRFP
Merge
B
Pixel Intensity Ratio
(MAF1b vacuole/cytosol)
**
mito-RFP
TOM70
C
Pixel Intensity Ratio
(MAF1a vacuole/cytosol)
D
U2OS: T. gondii Type II:MAF1a
U2OS: T. gondii Type II:MAF1b
U2OS: Uninfected
R1 R2 R3
α-TOM70
α-VDAC2
α-β-Actin
E
Pixel Intensity (Actin/Protein of Interest)
TOM70
VDAC2
Type II: MAF1a
Type II: MAF1b
Uninfected

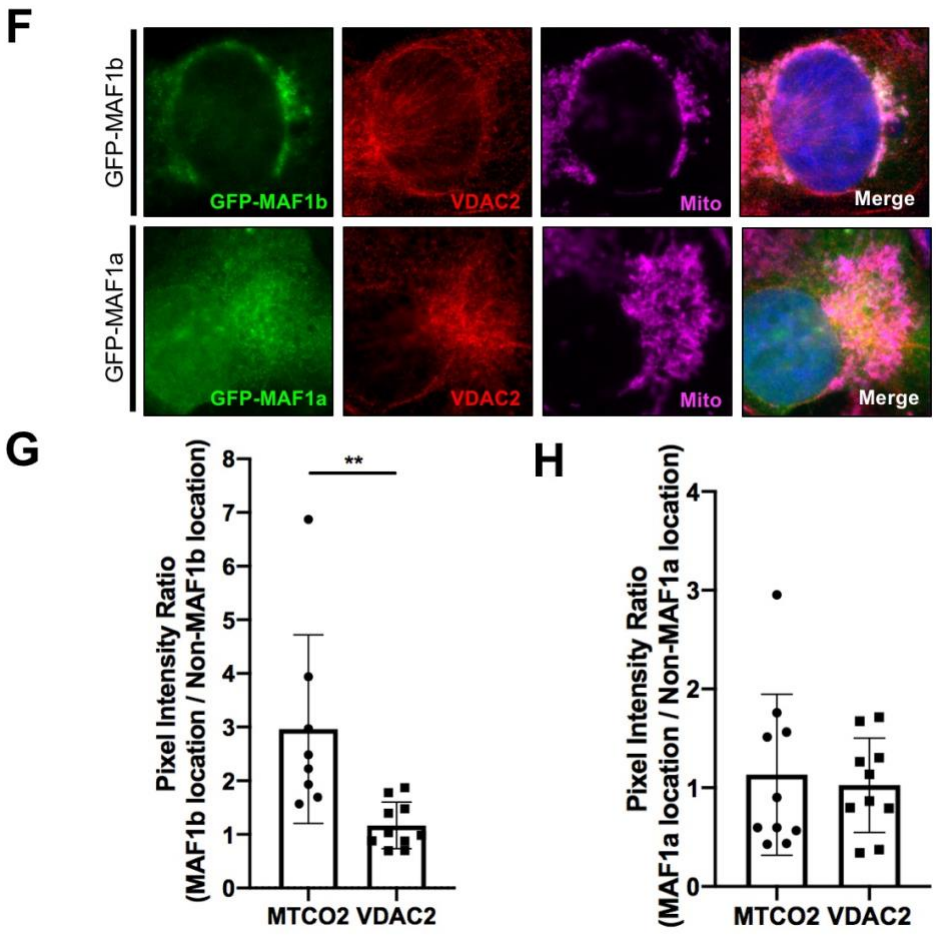


Figure 3-5 TOM70 is enriched at the parasitophorous vacuole.

(A) Normal rat kidney cells expressing RFP-labelled mitochondria were infected with Type II:MAF1b and Type II:MAF1a parasites. Cells were fixed at 18 hpi and visualized with epifluorescence microscopy. Immunofluorescence staining was performed with antibodies against TOM70. (B) Enrichment of either RFP-mitochondria or TOM70 at the parasitophorous vacuole was performed by quantifying pixel intensity of either the TOM70 or the mitochondria channel in a selected region on the vacuole membrane and a region off of the vacuole membrane. The ratio of these two areas was then calculated to measure enrichment of mito-RFP or TOM70 protein at the Type II:MAF1b vacuole. **p=0.0033 unpaired two-tailed t-test. (C) Similar to (B), enrichment ratios were calculated in cells infected with Type II:MAF1a parasites. (D) U2OS cells were infected with Type II:MAF1b or Type II:MAF1a parasites for 24 hours (MOI=2) in triplicate. Cells were lysed at 24 hours in IP lysis buffer and boiled in LDS sample buffer. Western blotting analysis was performed with the listed primary antibodies and HRP-conjugated secondary antibodies. (E)

Densitometric actin/TOM70 quantification of western blots depicted in (D) using FIJI ImageJ software (NIH). (F) U2OS cells were transfected with GFP-MAF1a and GFP-MAF1b. Cells were fixed at 24 hours post transfection and visualized with epifluorescence microscopy. Immunofluorescence staining was performed with antibodies against both VDAC2 and the mitochondria (MTCO2). (G) Enrichment of either mitochondria (MTCO2) or VDAC2 at the locations where MAF1b is present was performed by quantifying pixel intensity of either the VDAC2 or the mitochondria channel in a region overlapping with MAF1b and a region not-overlapping with MAF1b. The ratio of these two areas was then calculated to measure enrichment of that protein at MAF1b-localized sites. **p=0.0064 unpaired two-tailed t-test (H) Similar to (G), enrichment ratios were calculated in cells expressing GFP-MAF1a.

3.4 Discussion

The association of host cell organelles with their invading pathogens has long been observed, but the intricate molecular mechanisms underlying these interactions have remained elusive. For example, the striking cellular phenotype of some pathogens associating with host mitochondria has been phenotypically characterized by microbiologists for quite some time. This phenotype has been reported in *Legionella pneumophila* [43], *Chlamydia psittaci* [45, 102], *Hammondia hammondi* [12, 119], and *Toxoplasma gondii* [11, 12, 31]. While the secreted bacterium effector driving association of the infectious *C. psittaci* elementary body with its host mitochondria remains unknown, key mitochondrial functions like ATP synthesis and the regulation of glucose metabolism are enhanced upon pathogen infection [120]. Additionally, mitochondrial association is specific to *C. psittaci* and this association is absent in both *C. pneumonia* and *C. trachomatis* [121]. Recent discoveries in *L. pneumophila* answered a decades-long question about the impact of bacterial association with host mitochondria. *L. pneumophila*

utilizes the protein effector, MitF, to mediate downstream host GTPase recruitment leading to alterations in mitochondrial morphology, including fragmentation [44]. However, the mechanism behind MitF driving mitochondrial fragmentation is still unknown. Pinpointing the exact host proteins interacting with MitF would provide valuable insight into this mechanism. Similar work in the eukaryotic parasite, *T. gondii*, has resulted in the identification of the parasite effectors driving association with host mitochondria [31]. *T. gondii* secretes the effector protein, MAF1b, from its dense granules upon invasion, and after trafficking to the PVM, it is required for association with host mitochondria [12, 31]. This association is physiologically relevant, since HMA+ parasites outcompete their HMA- counterparts during acute infection in mice (Chapter 2, [11, 12]). This growth advantage was first recognized during acute infection, but there is also a MAF1b/HMA-dependent increase in cyst burden in the brain of mice during chronic infection [122]. The exact signaling pathways underlying these physiological effects of MAF1b and identification of host proteins required for HMA represent an important means to uncover these mechanisms. In this work, we use a non-biased quantitative mass spectrometry approach to show the requirement of the outer mitochondrial membrane receptor protein, TOM70, and mitochondrial-specific chaperone, HSPA9, in mediating MAF1b binding to the host mitochondria. While both proteins are directly required for mediating HMA, these data do not exclude the possibility of additional candidates being required to complex together and facilitate MAF1b-mitochondrial binding.

3.4.1 Quantitative mass spectrometry narrow MAF1b binding partners

The *MAF1* locus has undergone significant gene expansion-driven diversification and Type I *T. gondii* maintains 7 tandem *MAF1* copies in its genome (Chapter 1, [10, 12, 123]). Based on

the predicted primary sequence of each MAF1 protein product, these *MAF1* copies can be categorized as either MAF1a or MAF1b paralogs [12]. Both paralog types traffic to the PVM upon expression in Type II (HMA-) parasites, but only MAF1b drives association of the Type II PV with host mitochondria [12]. We leveraged the diversity of the *MAF1* locus to identify host proteins required for MAF1b-mediated HMA. We compared interactions between three parasite clones (MAF1b, MAF1a, and a non-functional MAF1b) with their host proteins. Through the use of quantitative mass spectrometry, we identified 13 candidate MAF1b host binding proteins. Some of these candidates including IMMT and SAM50 were previously identified as MAF1b binding partners [59].

3.4.2 *T. gondii* associates and manipulates host mitochondria through binding TOM70 and HSPA9

The esiRNA knockdown experiments are difficult to perform because we are measuring HMA competency following an extended time period of protein depletion. We were cautious of the effects of protein knockdown to overall mitochondrial and cellular health. For example, the candidate protein, ATP5B, is an essential component of the F_1/F_0 ATP synthase machinery. Previous literature shows that RNAi knockdown of ATP5B leads to a decrease in cellular respiratory rate, increased production of ROS, and initiation of cellular apoptosis [124, 125]. While fragmented mitochondria were visualized in these samples, Type II:MAF1b parasites were still capable of driving clear association with the fragmented mitochondria, indicating the robustness of the HMA phenotype and the tight interaction between MAF1b and the mitochondria. Additionally, in our experiments, knockdown of the previously identified MAF1b-interacting host protein, SAM50, did not disrupt HMA. Previous work reports MAF1b interacts with members of

the mitochondrial intermembrane-space bridging complex (MIB) including SAM50, IMMT, CHCHD3, and mitofilin [59]. They hypothesized that because SAM50 is the only member of the complex exposed to the cytosol, this is the reasonable point of contact with MAF1b [59]. Upon RNAi knockdown of the outermost member, SAM50, and internal candidate, Mitofilin, the percent of HMA+ vacuoles in both treatments decreased by 40-50% [59]. Parasite clones used in the previous experiments were an HMA+ Type I strain with an N-terminally HA-tagged MAF1b paralog and an HMA- Type I strain with a C-terminally HA-tagged MAF1b paralog [31, 59]. The presence of the C-terminal HA tag functions as a dominant negative and renders the parasites HMA incompetent. Our work utilized a more elegant approach by comparing evolutionary-relevant MAF1 paralogs (MAF1b:HMA+ and MAF1a:HMA-) and a single engineered MAF1b mutant (Chapter 2, [11]). While our data provides a more comprehensive view of the molecular mechanism driving MAF1b-mediated HMA, it is not in complete contrast to the previous results. Interestingly, SAM50 and IMMT, which were identified in our screen are known transient interactors with TOM70 and this might explain the previous identification of these proteins in pull down analyses [109].

The requirement of TOM70 in mediating HMA is also confirmed through our ectopic overexpression approach. The purpose of this method is to assay MAF1b interactions without the milieu of parasite effectors and subsequent pathway manipulations in a parasite infection. N-terminally GFP-tagged MAF1b traffics directly to the host mitochondria as early as six hours post transfection (Fig 3-3A). This is in direct contrast to the GFP-tagged MAF1a which remains cytosolic and occasionally forms puncta in both the cytosol and nucleus. Interestingly, overexpression of GFP-MAF1b in the host cell causes dramatic aggregation of the mitochondria which is not observed following GFP-MAF1a expression. Additionally, these morphological

changes to the mitochondria are not seen in the cells knocked down for TOM70 followed by GFP-MAF1b expression.

3.4.3 MAF1b binding to mitochondria may require multiple target sites

TOM70 is known to interact with its target proteins through internal mitochondrial targeting sequences (iMTSs) [107]. These regions function similarly to N-terminal signal sequences in that they maintain an amphipathic alpha helical structure and contain many hydroxylated and positively charged residues. Both MAF1b and MAF1a contain five predicted iMTSs with only the N-terminal iMTSs being conserved among the two paralogs, suggesting the MAF1b paralog interacts with TOM70 through its C-terminal iMTSs. However, almost all of the iMTSs including the N-terminal iMTSs are required for proper mitochondrial localization upon ectopic expression. The only region not required is the disordered region downstream of the MAF1b proline-rich stretch which was previously identified as being dispensable for MAF1b function (Chapter 2, [11]). This suggests that MAF1b and MAF1a are quite diverged from one another and MAF1b requires interactions on both its N-terminal and C-terminal ends to drive HMA.

3.4.4 MAF1b potentially disrupts ER-mitochondrial contact sites

While the physiological impact of the TOM70 and HSPA9-MAF1b interaction is not yet known, we did measure a significant enrichment of TOM70 at the MAF1b-mitochondria interface. Fluorescence microscopy does not allow us to visualize precise MAF1b-TOM70 interactions because TOM70 studs the outer mitochondrial membrane and serves as a good general

mitochondrial marker. However, through immunofluorescence we were able to quantify the significant enrichment of TOM70 at the vacuole in comparison to a mito-RFP mitochondrial marker, suggesting aggregation or stabilization of TOM70 at the regions on the vacuole in contact with MAF1b. In mammalian cells, TOM70 clusters at specific ER-mitochondria contact sites and recruits ER-bound inositol triphosphate receptor (IP3R3) and HSPA9 to transfer Ca^{2+} from the ER to the mitochondria [50]. This ER-mitochondrial contact site mechanism is essential for maintaining healthy mammalian cell homeostasis and mitochondrial health [50]. Upon knockdown of TOM70, the ER-mitochondrial transfer of Ca^{2+} is ineffective and the cells show a marked increase in autophagy and decreased proliferation [50]. Additionally, the attachment of the host mitochondria to the vacuole excludes the ER along the entire length of the interface which is observed on both an epifluorescence and TEM level (Fig 3-6), suggesting a disruption of ER-mitochondrial contact sites. Perhaps, *T. gondii* MAF1b binds the HSPA9-TOM70 complex and disrupts the ER-mitochondria contact sites along the length of the vacuole-mitochondria interface. These interactions might induce mitochondrial dysfunction and signal cellular autophagy leading to improved *T. gondii* growth conditions. The mitochondrial associated pathogen, *L. pneumophila*, utilizes a similar mechanism where bacterium induction of mitochondrial fragmentation leads to impairment of oxidative phosphorylation and Warburg-like metabolism promoting bacterium replication [44]. In the case of *T. gondii*, these conditions would describe the increase in parasite fitness in both acute and chronic *in vivo* infections.

In summary, by leveraging an evolved *T. gondii* locus, we identified the requirement of the host proteins, TOM70 and HSPA9, in mediating host mitochondrial association. These findings shed light on the intricate host-pathogen molecular arms race for pathogen manipulation of host organelles.

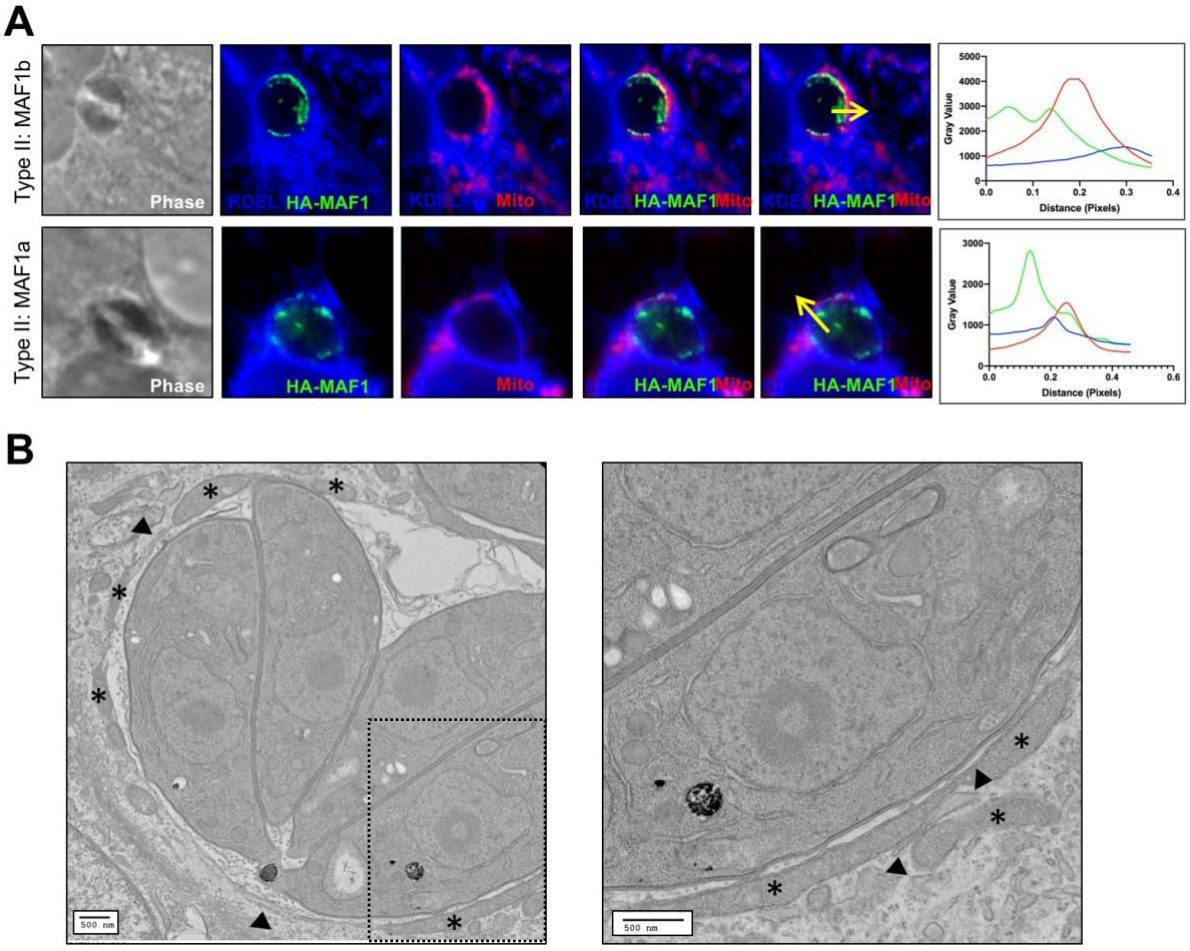


Figure 3-6 MAF1b preferentially binds mitochondria and excludes ER.

(A) mCHERRY-KDEL cells were infected with either Type II:MAF1b or Type II:MAF1a parasites. Cells were fixed 24 hpi and visualized using fluorescence microscopy. Immunofluorescence staining was performed with antibodies against the mitochondria and HA epitope tag. Profile plots of HA-MAF1b (green), KDEL(blue), and mitochondria (red) correlate to yellow arrow in the merged image. (B) TEM micrograph of HFFs infected with Type III (HMA+) parasites. Arrow heads indicate ER and asterisks indicate mitochondria. Right image is magnified region of left (dotted square).

3.5 Materials and methods

3.5.1 Parasite expression of MAF1 paralogs and transgenic parasites

Parental plasmid (pATTGra_HA_HPT) for expression of *Tg*MAF1RHb1 (MAF1b) and *Tg*MAF1RHa1 (MAF1a) paralogs contained the selectable *HXGPRT* gene marker. Transgenic clones were generated using Type II *Tg*Me49ΔHXGPRT:Luciferase parasites that were transfected by electroporation at 1.6Kv and 25μF with 50 μg of pATTGra *Tg*MAF1RHb1 (MAF1b) and *Tg*MAF1RHa1 (MAF1a) linearized with *Hind*III. *Tg*Me49ΔHXGPRT:Luciferase parasites were scraped and passed through a 25 and 27 gauge needle. Cell lysate was centrifuged for 10 minutes at 800xg and $2x10_7$ parasites were re-suspended in Cytomix (120 mM KCl; 0.15 mM $CaCl_2$; 10 mM KPO_4; 25 mM Hepes, 2mM EDTA, 5mM $MgCl_2$; pH to 7.6), ATP and GSH. Parasites were selected with mycophenolic acid (MPA)/Xanthine and cloned by limited series dilution in a 96-well plate. Positive and negative parasite clones were confirmed through immunofluorescence analysis by the presence or absence of an HA-tag. *Tg*Me49:*Tg*MAF1RHb1 (RHb1 clone 1 i.e. "Type II:MAF1b"), *Tg*Me49:*Tg*MAF1RHa1 (RHa1 clone 3 i.e. "Type II:MAF1a"), and *Tg*Me49:Empty Vector (EV clone 1) are passaged matched and used for the entirety of all experiments. Additionally, *Tg*Me49:*Tg*MAF1RHb1:STYL(438-441)RKYK clone 6 (STYL/RKYK clone 6 i.e. "Type II:MAF1b mutant") was similarly generated and used for all experimentation.

GFP-MAF1b/a constructs were generated using the pcDNA3.1/NT-GFP-TOPO GFP-fusion TOPO TA expression kit (ThermoFisher). The region of MAF1b/a cloned into the pcDNA3.1 expression plasmid excluded the first 23 residues containing the signal peptide. MAF1b

(residues 24-446) and MAF1a (residues 24-446) were PCR amplified and gel purified prior to GFP fusion.

MAF1 chimeras 1,3, and 4 were generated using overlapping PCR and Gibson assembly (NEB Gibson Assembly Master Mix) methods. MAF1 chimera 2 was generated using a site-directed mutagenesis kit (NEB Q5 Kit) to mutate both MAF1a iMTS3 and the C-terminal residues (STYL→RKYK). The same C-terminal residue changes (STYL→RKYK) were made in MAF1 chimera 1 using the same site-directed mutagenesis methods.

3.5.2 Cell maintenance and parasite infection

All *Tg*Me49 parasite clones were regularly passed in human foreskin fibroblasts (HFFs) and incubated at 37$_{o}$C in 5% CO_2. Following parasite selection and cloning, all HFFs were incubated in Dulbecco's modified Eagle's medium supplemented with 10% FBS, 2mM glutamine, and 50μg/ml of penicillin and streptomycin (cDMEM). Human bone osteosarcoma cells (U2OS) and NRK mito-RFP cells were also regularly passed in cDMEM media. All experimental parasite infections included scraping and passing cell lysate through a 25 and 27 gauge needle followed by counting on a hemocytometer and infection at the noted MOI.

3.5.3 cDNA synthesis and qPCR/TOM70 protein knockdown analysis

After 48 hours post 25nM esiRNA treatment, three wells were lysed using the Qiagen RNA Shredder and RNeasy RNA extraction kit. RNA samples were reverse transcribed using SuperScript III Reverse Transcriptase protocol (Thermofisher) and qPCR was performed using the SYBR Green Supermix (BioRad). Quantification and statistical analyses were performed in Prism

GraphPad. Replicate CT values were used to determine statistical significance in Prism GraphPad. Additionally, Vehicle, TOM70, and HIBADH esiRNA-treated wells were lysed in Pierce IP lysis buffer (ThermoFisher) and treated with complete protease inhibitors (Roche) on ice. Samples were boiled in SDS sample buffer for 10 min, resolved on 10% SDS-PAGE gel and transferred to nitrocellulose membrane. Membranes were blocked in 5% BSA in PBST and probed with primary antibodies to TOM70 and β-actin followed by HRP conjugated secondary antibodies. Membranes were visualized with SuperSignal West Femto chemiluminescent substrates. Antibodies used for these experiments: anti-TOM70 (sc-390545) – Santa Cruz Biotechnology, anti-β-actin (#4970) – Cell Signaling Technology, Goat Anti-Rabbit IgG HRP – Southern Biotech, Goat Anti-Mouse IgG HRP – Southern Biotech.

3.5.4 esiRNA generation and RNAi treatment

Silencing regions (400-600 bp) of a candidate gene were PCR amplified to include a 5' and 3' T7 RNA polymerase promoter sequence (FWD-TAATACGACTCACTATAGGGAGAG, REV-TAATACGACTCACTATAGGGAGAC) from human HEK cell cDNA. PCR products were gel purified and Sanger sequencing confirmed candidate sequences (Genewiz). 250 ng of verified product was *in vitro* transcribed using components of the MEGAscript T7 kit. Following DNase treatment, transcribed RNA product was digested with RNase III for 2 hours at 37$_{o}$C. Digested esiRNA products were purified by passing the solution in equilibrium buffer (20mM Tris-Cl, 1mM EDTA, 300mM NaCl) through Q-sepharose beads followed by wash buffer (20mM Tris-Cl, 1mM EDTA, 400mM NaCl). Purified product was eluted off the beads in elution buffer (20mM Tris-Cl, 1mM EDTA, 520mM NaCl). Eluted esiRNA precipitated overnight in isopropanol and resuspended 50uL of nuclease-free H_2O.

U2OS cells were seeded in a 24 well plate at 37$_{o}$C in 5% CO_2 for 24 hours to reach 80% confluency. Six wells for each candidate gene were treated with 25nM esiRNA using Dharfect-1 reagent for 48 hours. Three wells were infected with Type II:MAF1b parasites (MOI=2) and the remaining wells were lysed for RNA extraction and purification. Cells were fixed at 24 hpi with 4% formaldehyde in PBS followed by immunofluorescence analysis.

3.5.5 TEM and percent vacuole coverage quantification

U2OS were treated with 25nM Vehicle, TOM70, or HIBADH esiRNA for 48 hours followed by infection with Type II:MAF1b parasites (MOI=2). At 24 hpi cells were fixed with 2.5% glutaraldehyde in PBS overnight in 4°C and subjected to same TEM sample preparation previously described in Chapter 2 (protocol: Center for Biological Imagining - CBI, University of Pittsburgh, Pittsburgh, PA, USA). A minimum of 19 vacuoles for each esiRNA treatment were imaged. Vacuoles were traced in FIJI ImageJ and the percent of the total distance around the vacuole in contact with the host mitochondria was quantified for each esiRNA treatment group. Quantification and statistical analyses were performed in Prism GraphPad.

3.5.6 Immunofluorescence and immunoblotting

HFFs, U2OS, and NRK mito-RFP cells were grown in cDMEM on glass coverslips in a 24-well plate. Following either esiRNA knockdown, GFP-MAF1a/b expression, or Type II:MAF1a/b infection, cells were fixed with 4% paraformaldehyde for 15 min. and permeabilized with 0.1% Triton X-100 in PBS at room temperature. Cells were then probed with primary antibodies diluted in PBS for 1 hour at room temperature while shaking. Following PBS washes,

cells were incubated in secondary antibodies fused to a fluorescence indicator for 1 hour at room temperature while shaking (Life Technologies Alexa Fluor H+L) followed by PBS washes. Cells were mounted on glass slides using ProLong Diamond Antifade Mountant (ThermoFisher) and cured in the dark for 24 hours. Immunofluorescence primary antibodies used in this work include: anti-HA rat monoclonal antibody (3F10 clone, Roche), anti-mitochondria (ab92824, Abcam), anti-TOM70 (sc-390545, Santa Cruz Technologies), and anti-VDAC2 (PA5-28106, ThermoFisher). Images were taken of five total channels dependent on the experiment: 488, 594, 647, 350, and DIC/phase. Images were cropped and merged using FIJI ImageJ (NIH).

3.5.7 Immunoprecipitation and quantitative mass spectrometry

HFFs were seeded to confluency in a T175 (10^7 cells) and hyper-infected (MOI=10) with either Type II:MAF1b, Type II:MAF1a or Type II:MAF1b mutant parasites (n=7) for 24 hours. Cell were lysed in IP lysis buffer (50mM Tris - pH 8.0, 150 mM NaCl, 1% IGEPAL CA-630, 0.05% Tween 20) supplemented with complete ULTRA-tablet protease inhibitor (Roche) on ice. Following sonication using three 15mV 10s bursts with 25s intervals on ice, the insoluble fraction was pelleted at 800xg for 10 minutes at 4°C. The soluble fraction was incubated with Pierce anti-HA magnetic beads (Thermo scientific) overnight using a rotator. Beads were washed five times with IP lysis buffer and eluted by boiling in SDS sample buffer (Thermo scientific). Eluate fractions were resolved at 150V for 10 minutes on 10% SDS-PAGE gel to only allow the protein samples to enter the resolving portion of the gel. The eluted samples were loaded on to a 4-12% Bis-tris gradient gel in ~18µL aliquots at constant 150V 1X MES for 2 minutes or until the entire sample was loaded. The gel bands were digested with trypsin as described in [126] and tryptic peptides were desalted with Pierce C_{18} Spin Columns, speedvac dried and resuspended in 18µL

0.1% formic acid. A pooled instrument control (PIC) was made by taking 3µL from each of the 21 replicate samples. The injection order was randomized in a balanced block design.

All samples were analyzed by nano-flow liquid chromatography tandem mass spectrometry (nLC-MS/MS) using an injection volume of 1 µL on to a 25 cm C_{18} Picochip column. The LC Gradient ran at 300 nL/min with a 2-35% acetonitrile over a 66 minute gradient. Instrument parameters included a full scan with top13 ddMS_2 on the high resolution Thermo Orbitrap Velos Pro mass spectrometer.

Extracted ion chromatogram peak area for 32 selected peptides were monitored during the instrumental analysis to ensure proper instrument performance using Skyline software. MASCOT database search was performed using a modified UniProt human database with *Toxoplasma gondii* protein sequence database with the parameters: fixed modification carbamidomethylation (C), and variable modifications of acetyl (N-terminus), oxidation (M), and deamidation (NQ). Major identified proteins were reported in Scaffold v4.5.3. Proteomic features were aligned and quantitated using MaxQuant v1.6.2.1 software and Andromeda search engine with 1% FDR selected. Graphpad Prism was used to create bar charts. Matlab was used for the statistical analysis.

3.5.8 MAF1 ectopic expression

U2OS cells were plated to 80% confluency and transfected with 1µg of GFP-MAF1b/a plasmid by lipofection (Lipofectamine 3000). In some cases, U2OS were treated with 25nM target esiRNA for 48 hours followed by transfection with 1µg of GFP-MAF1b/a plasmid for 24 hours. Following transfection, cells were fixed and stained with anti-mitochondria antibody for immunofluorescence analysis. Additionally, VDAC2 localization was investigating by

transfecting both GFP-MAF1a/b for 24 hours followed by fixation and immunofluorescence analysis.

3.5.9 iMT targetP analysis

Both *Tg*MAF1RHa1 (MAF1a) and *Tg*MAF1RHb2 (MAF1b) sequences were subjected to TargetP prediction software as described in [107]. The algorithm N-terminally truncates the protein of interest by a single residue at a time and assigns a score to the truncated form. This allows us to identify regions of *Tg*MAF1b/a that are predicted iMTSs.

3.5.10 TOM70 and VDAC2 enrichment

Mito-RFP NRK cells were infected with either Type II:MAF1a or Type II:MAF1b parasites (MOI=2). Mito-RFP NRK cells were fixed at 24 hpi and stained for the HA-tag and TOM70 (antibodies previously described). Using FIJI ImageJ software, the raw intensity density of a 15x15 pixel area was measured in a region on the PVM and off of the PVM. An enrichment ratio was calculated (pixel intensity on the PVM/ pixel intensity off of the PVM) in both the RFP-mitochondria channel and the TOM70 channel (11 independent vacuoles were counted).

Additionally, U2OS cells were transfected with 1μg of GFP-MAF1b/a plasmid. At 24 hpi, cells were fixed and stained for the mitochondria and VDAC2 (as described above). Using FIJI ImageJ software, a 15x15 pixel area was measured in a MAF1-rich region and MAF1-absent region. An enrichment ratio was calculated (pixel intensity in MAF1-rich region/ pixel intensity MAF1-absent region) in both the mitochondria channel and the VDAC2 channel (9 independent cells were counted).

TOM70 stabilization experiment included seeding U2OS cells in a 24-well plate until they reached confluency. Wells were infected with Type II:MAF1b, Type II:MAF1a, or uninfected at an MOI of 5 (n=3 per treatment condition). Cells were lysed at 24 hours post infection in Pierce IP lysis buffer (ThermoFisher) and treated with complete protease inhibitors (Roche) on ice. Samples were boiled in SDS sample buffer for 10 min, resolved on 10% SDS-PAGE gel and transferred to nitrocellulose membrane. Membranes were blocked in 5% BSA in PBST and probed with primary antibodies to TOM70, β-actin, and VDAC2 followed by HRP conjugated secondary antibodies. Membranes were visualized with SuperSignal West Femto chemiluminescent substrates. Antibodies used for these experiments: anti-TOM70 (sc-390545) – Santa Cruz Biotechnology, anti-β-actin (#4970) – Cell Signaling Technology, anti-VDAC2 (PA5-28106) – ThermoFisher, Goat Anti-Rabbit IgG HRP – Southern Biotech, Goat Anti-Mouse IgG HRP – Southern Biotech.

4.0 Discussion and future directions

T. gondii and its closest extant relatives, *N. caninum* and *H. hammondi*, maintain many tandemly expanded loci, many of which contain N-terminal signal sequences required for secretion from the parasite [10]. Gene expansion-driven diversification (GEDD) in *T. gondii* has led to the evolution of novel mechanisms to interact with and manipulate its host [10, 123]. In Chapter 1, we show that a select group of characterized GEDD loci in tissue-dwelling coccidia maintain multiple tandemly expanded and diversified loci that play critical roles in host manipulation rather than providing a more classical function in antigenic variation. For example, in *P. falciparum* and *T. brucei*, the diversity of the *var* and variant surface glycoprotein (VSG) family genes, respectively, is a source for antigenic variation in the parasite allowing it to avoid parasite clearance by host antibody responses. In *T. gondii*, 53 of understudied GEDD loci drive expression of predicted secreted paralogs that each maintain an individualized role in their interaction with their associated host [13-15], and to date no such loci have been identified that perform any function akin to antigenic variation in *T. gondii* and *P. falciparum*. A prime example of a tandem GEDD locus in *T. gondii*, *ROP5*, illustrates the nuances of classes how gene diversification can impact parasite virulence. While the *ROP5* cassette maintains three paralog classes that each traffic to the PVM, parasite virulence is only driven by the combined expression of ROP5A and B. These finding suggest a neofunctionalized *ROP5* locus in *T. gondii* that maintains ROP5 paralog classes (ROP5C) which have no bearing on parasite virulence but are maintained in the genome [20, 29, 30]. While many of these tandem GEDD loci have been identified in *T. gondii* and its close relatives, including IMC8/5 [19], NTPases I and II [19], SRS paralogs [10, 22], and *MAF1* [11,

12], the multi-copy nature of these loci make them difficult to study with traditional sequencing techniques, and therefore difficult to clone and individually express in parasites.

The work in Chapters 2 and 3 along with previously published research [11, 12, 31], describe a comprehensive strategy to elucidate the mechanism and impact of GEDD on the host, with a focus on MAF1b-driven host mitochondrial association (HMA). Our work characterizes the functional significance of the expanded *MAF1* locus (Chapter 2) [12], tracks the expression of individual MAF1 paralogs (Chapter 2), and employs a structure-function approach to not only identify the functional regions of MAF1 (Chapter 2) but leverage the multi-copy locus to identify host proteins driving the interaction with the parasite (Chapter 3).

4.1 The shifting red herring: evolutionary significance of a diverse *MAF1* locus

T. gondii maintains a diverse and expanded *MAF1* locus in comparison to its closest extant relatives, *N. caninum* and *H. hammondi* [10, 12]. Interestingly, even within the *T. gondii* lineage, *MAF1* is diverse in copy number and sequence among the three major strains (Types I, II, and III) [12]. Previous literature characterized the neofunctionalized role of MAF1 in mediating HMA and highlighted the diverse protein paralogs expressed by each of the three strains [12]. Chapter 2 illustrates our structure-function approach to identify the regions of MAF1 critical for the HMA phenotype and reveal possible ancestral functions of MAF1 (prior to the emergence of HMA). The two paralogs most useful for our analyses were *Tg*MAF1RHb1 (MAF1b) and *Tg*MAF1RHa1 (MAF1a) cloned from RH (Type I) parasites. The HMA-incompetent, *N. caninum*, maintains two identical copies of MAF1a, while the more recently speciated *H. hammondi* maintains a single MAF1a and MAF1b copy, suggesting an ancestral role of MAF1a.

Structural analysis of both MAF1b and MAF1a in collaboration with the Boulanger lab, identifies similarities with the macroH2A (non-histone) macro-domain in *Archaeoglobulus fulgidus* 1521 (*Af*1521) [11, 62, 64]. This macro-domain is responsible for binding the nucleotide, ADPribose (ADPr) which functions as a histone post-translational modification involved in regulation of gene expression, however our work found minimal coordination of ADPr by either MAF1a or MAF1b [11], suggesting that ADPr binding may have been lost in MAF1a prior to its diversification in *N. caninum*. Interestingly, hepatitis E virus (HEV) expresses an antagonist ADPr-binding domain that subverts downstream host immune responses by reversing cellular ADP ribosylation [71]. While our work does not investigate further the pseudo-ADPr-binding domain on MAF1, it sheds light on the evolutionary history of host-interacting proteins secreted by *T. gondii*.

Our goal in identifying the MAF1b region required for mediating the HMA phenotype was two-fold. Firstly, identifying these regions would provide an evolutionary path for the neofunctionalization of the *MAF1* locus, pinpointing the residues maintained to provide increased parasite fitness in both acute and chronic infection. Secondly, identifying this region would create a powerful reagent to probe the molecular mechanism driving MAF1b-mediated HMA. Described in Chapter 2, significant differences were identified in a disordered repetitive proline-rich region near the N-terminus of only the MAF1b paralog. Proline-rich regions and motifs have been identified in budding viral proteins required for binding and manipulation of host ESCRT machinery [127]. While this region proved to have no direct role in mediating HMA, because MAF1 functions within the confines of a lipid bilayer originating from the host cell membrane, it might describe a yet unidentified role of MAF1b specifically in host immune evasion via ESCRT manipulation. Chapter 2 [11] describes the elucidation of three C-terminal residues in MAF1b

required for mediating binding to the host mitochondria and the *in vivo* growth advantage associated with HMA. These residues were pinpointed after ClustalO analysis revealed 100% conservation among the *T. gondii* Types I, II, and III MAF1b paralogs compared to 40% identity with MAF1a paralogs [11].

Crystallization and structural analysis of both the macrodomains of MAF1a and MAF1b reveal highly similar structures with an rmsd of 0.6 Å over 240 Cα atoms [11]. We focused our attention on differences in key phosphorylation sites, secondary structure, and charged residues between MAF1a/b structures. Finally after significant mutational analysis along the length of the MAF1b paralog, we identified a C-terminal MAF1b mutation that disrupts HMA (Chapter 2: STYL(338-441)RKYK) [11]. While expression of MAF1b in Type II parasites drives HMA to near WT Type I levels, expression of the C-terminal MAF1b mutant does not lead to HMA [11]. Additionally, Type II:MAF1b parasites outcompete their Type II HMA-incompetent counterparts in both *in vivo* acute and chronic infections, while Type II parasites expressing the MAF1b mutant do not [11, 122]. Interestingly, the MAF1b-driven *in vivo* growth advantage is only relevant *in vivo* and not during *in vitro* culture. These data suggest that MAF1b may manipulate host immune pathways not present in fibroblasts growth *in vitro*. However, additional work into these growth differences and fully characterizing immune modulation driven by MAF1b-mediated HMA in a more biologically relevant context (such as in primary immune cells or in the mouse model) is warranted.

4.2 The bargaining chip: leveraging the *MAF1* locus to probe molecular function

Previous work identified the mitochondrial membrane inner bridge (MIB) complex as a role in mediating HMA in *T. gondii* using a non-quantitative mass spectrometry approach [59]. Key players in this interaction were the outer membrane sorting assembly and machinery component 50 (SAM50) and inner membrane space protein, mitofilin (IMMT). However, our work in Chapter 2 reveals an interaction between SAM50 with *both* MAF1a and MAF1b [11], suggesting that the MIB complex may not be required for HMA. While this interaction could be explained by yet another ancestral role of MAF1, it is more likely that SAM50 was associated with Type I:HA-MAF1b based on belonging to a large, but functionally irrelevant complex of mitochondrial proteins. There is an inherent challenge in immunoprecipitation of membrane-bound proteins because they require optimized lysis conditions to solubilize the membrane-bound candidate without forcing it to aggregate with the lipid insoluble fraction. These lysis conditions could lead to membrane proteins remaining tethered together with their lipid constituents, and subsequent false positives in downstream proteomics.

Chapter 3 describes our unbiased approach to carefully identify MAF1b-binding candidates through co-IP of HFFs infected with either Type II:MAF1b (HMA+), Type II:MAF1a, or Type II:MAF1b (HMA-) parasites followed by quantitative mass spectrometry. The previous work which identified the binding of the MIB complex with MAF1b utilized an approach where they infected HFFs with Type I parasites or Type I parasites expressing a C-terminally tagged MAF1b which functioned as a dominant negative factor rendering the parasite HMA incompetent [31, 59]. Our strategy was different considering the mutated version of MAF1b consisted of only a three-residue change and we also compared Type II:MAF1b to an additional and naturally-occurring paralog, MAF1a (Type II:MAF1a). Interestingly, while SAM50 was significantly

enriched in our Type II:MAF1b samples in comparison to the Type II:MAF1a and Type II:MAF1b mutant samples, RNAi knockdown of SAM50 did not decrease the HMA efficiency of Type II:MAF1b parasites. These data confirmed our hypothesis that SAM50 affinity for MAF1b is unrelated to HMA.

The two host proteins we identified as being required for MAF1b-mediated HMA are the outer translocase of the mitochondria member 70 (TOM70) and the mitochondrial-specific heat shock protein, HSPA9. Both proteins are known to interact with one another and interact with SAM50 [109], illuminating a potential reason why SAM50 was previously identified as a MAF1b interactor. We found that upon esiRNA knockdown of TOM70 and HSPA9, the efficiency of HMA decreased. Knockdown of TOM70 had a more dramatic effect on HMA than HSPA9. Because of the role of HSPA9 in mediating the transport and docking of mitochondrial pre-proteins to TOM70 [110, 112], our data might suggest an intermediary interaction between HSPA9, TOM70, and MAF1b. Additionally, while *T. gondii* HMA drives a significant *in vivo* inflammatory response [31] as well as an acute growth advantage for the parasite in mice [11, 12], recent work described HMA as a potential host response to parasite infection [47]. This work reveals a decrease in parasite lipid scavenging upon *T. gondii* PVM association with mitochondria, suggesting mitochondrial binding and fusion around the vacuole is initially driven by the host and impairs parasite growth over the first 24 hours of infection [47]. After 24 hours of infection, the mitochondria fragment and ultimately expose the PV to the host cytosol. While our work described in Chapter 2 regarding the MAF1b-mediated acute *in vivo* growth advantage might seem contradictory to these most recent findings, we hypothesize that the timing of infection is critical and the intricacies of the host/parasite interaction is bound to vary over the course of infection and HMA may have divergent impacts *in vivo* and *in vitro*.

4.3 The ultimate tug of war: a model for MAF1b-mediated HMA

We now know that both TOM70 and HSPA9 are required for MAF1b-mediated HMA. A testable model for HMA could be described as initially host-derived where HSPA9 docks MAF1b and pulls the parasite-containing vacuole to the mitochondria via TOM70 (Fig 4-1). Over the first 24 hours of infection, parasite growth is restricted but alleviated upon mitochondrial fragmentation seen at the 24 hour timepoint, leading to the increase in MAF1b-mediated growth advantage we record in mice. Interestingly, both HSPA9 and TOM70 are implicated in regulating key cell signaling functions including calcium flux, autophagy, and oxidative phosphorylation [49-51, 55]. Recent work pinpoints their site of regulation as being ER-mitochondrial contact sites throughout the cell [50, 75]. Through interaction with one another in conjunction with ER-associated inositol tri-phosphate receptor 3 (IP3R3), calcium flux from the ER lumen into the mitochondria mediates calcium homeostasis, mitochondrial dynamics, and downstream immune signaling [50]. Highlighted in Chapter 3, many of the proteins we identified as candidate MAF1b-interactors are found in high abundance at ER-mitochondrial sites, including VDAC 1 and 2, TOM70, and HSPA9 [50, 51, 56]. Additionally, while the ER is known to loosely associate with the *T. gondii* PVM [128, 129], in all infections with MAF1b-expressing parasites, the mitochondria outcompete the ER for PVM binding (see Chapter 3 Fig 3-6), supporting a MAF1b-driven direct interaction with the mitochondria which is absent in ER association. Our data indicating the requirement of TOM70 to mediate HMA along with its significant enrichment at the vacuole and requirement of HSPA9 leads us to hypothesize that ER-mitochondrial contact sites are disrupted upon MAF1b binding to the mitochondria via TOM70.

ER-mitochondrial sites are thought the be an early form of the machinery required for recognizing pathogens (specifically pathogen PAMPs) as they drive intricate anti-microbial

immune responses such STING, RIG-I, and MAVS-associated pathways [49, 130]. For example, hepatitis C virus maintains a NS3/4a protease that cleaves the N-terminus of MAVS and disrupts the downstream immune response [131]. TOM70 plays a key role in MAVS-dependent immune signaling and cytokine production [49]. Type I parasites knocked out for the MAF1 locus show a significant reduction in MAVS-dependent cytokine transcript [31], suggesting a role for MAVS in MAF1b-specific immune modulation and further bolstering an argument for reduced ER-mitochondrial contact sites mediated through MAF1b binding.

Future work identifying the downstream effects of the requirement of TOM70 and HSPA9 in driving MAF1b-mediated HMA will require in depth analysis of ER-mitochondrial dynamics and identify mechanisms that are differentially disrupted between a Type II:MAF1b and Type II:MAF1a infection. It is quite possible that ER-mitochondrial effects can only be observed upon hyper infection and/or following multiple parasite replications when a large percentage of the host mitochondria are sequestered by the parasite (and thus disrupting a majority of the ER-mitochondrial contact sites). Overall, our work outlines an intricate host-parasite interaction at the PV-mitochondrial interface where parallel evolution of host and parasite factors shape the molecule arms race between the two.

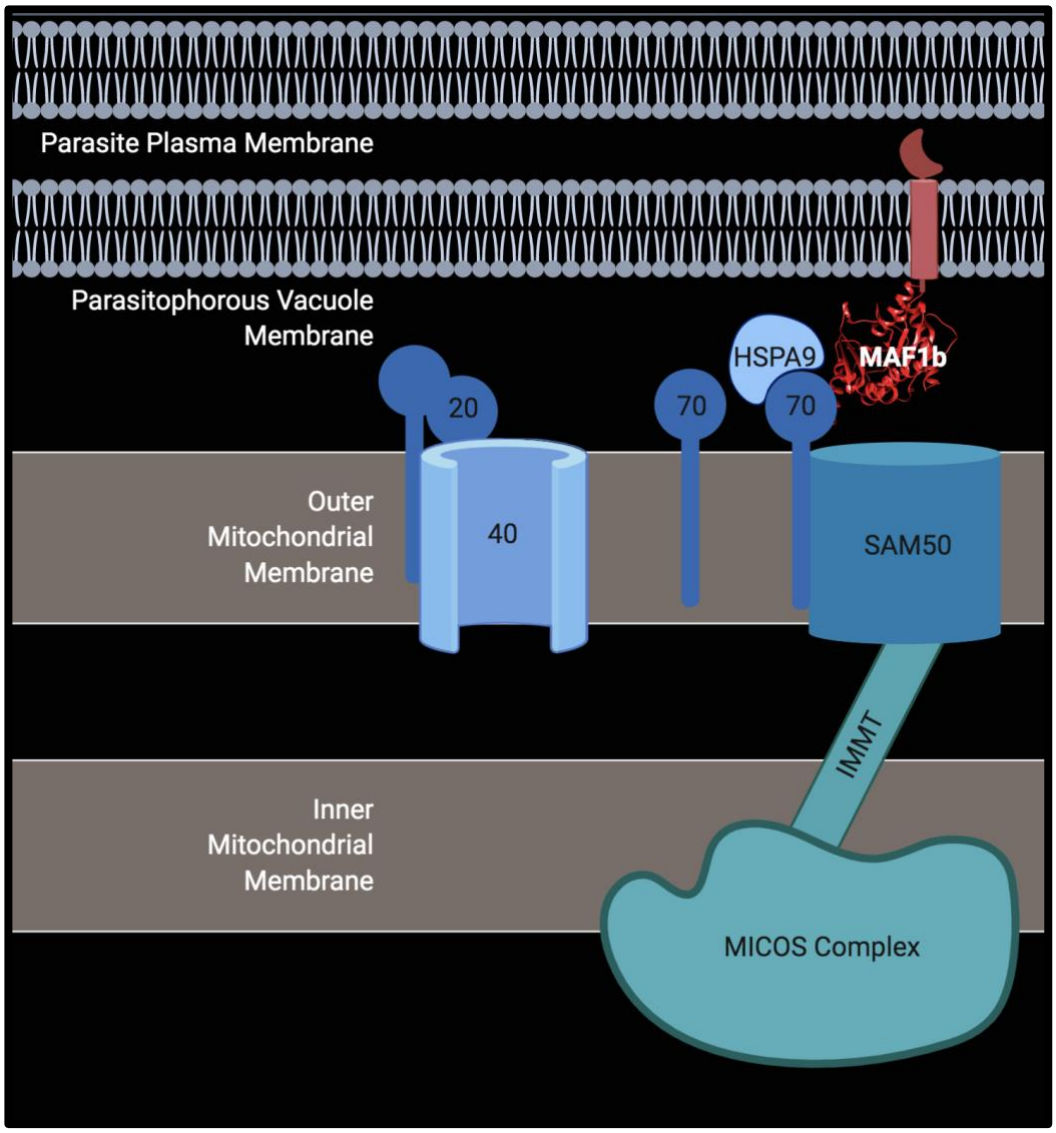


Figure 4-1 Model for MAF1b-mediated HMA.

T. gondii parasitophorous vacuole membrane (PVM) contains MAF1b which spans the PVM with the C-terminal portion in contact with the host cell cytosol. The requirement of HSPA9 and TOM70 for MAF1b-mediated HMA suggests a role of both proteins in tethering the PVM to the host mitochondria. Other members of the TOM complex, including the TOM40 pore (light blue) and TOM20 receptor (dark blue '20') are illustrated for reference but are not required for HMA. While MAF1b is not shown to directly interact with IMMT, members to the MIB complex (including IMMT, SAM50, and MICOS) are depicted in close proximity to the TOM complex because these are known transient interactors with members of the TOM complex (specifically TOM70).

Appendix A Investigating the role of the host protein RALGAPA1 in modulating HMA and its localization with MAF1b

Contents of this appendix were taken from work performed in collaboration with the Boulanger lab. I performed all experiments outlined in Fig A-3, A-4, and A-5. I am including Fig A-2 to provide clarity and illustrate the complete narrative.

Appendix A.1 Introduction

Currently unpublished work from our collaborators (Boulanger Lab, University of Victoria, BC) identified an interaction between the host GTPase-activating protein, RALGAPA1, and MAF1b. The specific interaction between MAF1b and the GAP domain of RALGAPA1 was identified through a Y2H screen of human host proteins and confirmed by isothermal titration calorimetry (ITC) using recombinant MAF1b (residues 173-443) and the RALGAPA1 GAP domain (T1807-P1988) protein. Interestingly, the HMA-competent MAF1b protein showed high affinity with RALGAPA1 (K_d of 334 ± 35nM), while MAF1a showed no affinity (Fig A-2). These data led us to hypothesize that the MAF1b-specific interaction with RALGAPA1 might be required for MAF1b-mediated HMA.

Appendix A.2 RNAi knockdown decreases HMA due to high dosage effects

As a GTPase activating protein, RALGAPA1 tightly regulates the dephosphorylation of Ras-like GTPases, RalA and RalB [132]. Dephosphorylation of RalA and RalB shifts the GTPases from their active (GTP-bound) to inactive (GDP-bound) states [132]. The oscillation of Ral proteins from their active to inactive forms drives a variety of complex signaling cascades in the cell [133]. For example, RALGAPA1 is known to induce mitochondrial fragmentation through its regulation of RalA at the site of the mitochondria [134].

To determine if RALGAPA1 was required for HMA, we used two separate siRNA reagents (siRNA "RG set 2" and "RG smart POOL") to knockdown RALGAPA1 transcript for 48 hours followed by infection with Type II:MAF1b parasites. In pilot studies, knockdown of RALGAPA1 proved challenging. Therefore we increased the concentration of siRNA to 40nM, which can be toxic to host cells and increases off-target effects [135]. Treatment with 40nM siRNA was successful at knocking down RALGAPA1 transcript and protein levels by 45-60% (Fig A-3 A-C). Following 48 hours of knockdown, we infected cells with Type II:MAF1b parasites. After 24 hours, the percent of HMA+ vacuoles was significantly less (**p=0.0024 and ***p=0.0008) in the siRNA-treated cells (Fig A-3D and E) compared to vehicle treatment. However, the number of HMA+ vacuoles also significantly decreased in the non-target siRNA (Cytophilin B) control-treated cells (Fig A-3E), suggesting that the impairment of HMA may have been an artifact of high siRNA dosage. Taken together, these data indicate that while RALGAPA1 knockdown decreases the HMA efficiency of Type II:MAF1b parasites, it is likely due to off target effects related to high siRNA toxicity.

Appendix A.3 MAF1b-RALGAPA1 binding loop is not required for mediating HMA in parasite infection

Through the use of hydrogen-deuterium exchange coupled with mass spectrometry (HDX-MS), the Boulanger lab discovered a loop on MAF1b driving the interaction with RALGAPA1 (boxed in Fig A-4A). We hypothesized that if the MAF1b-RALGAPA1 interaction drives MAF1b-mediated HMA, the MAF1b-RALGAPA1 binding loop would be required HMA. Using site-directed mutagenesis, we mutated the MAF1b-RALGAPA1 binding loop to its aligned sequence in MAF1a. Additionally, we mutated three loop residues shown to disrupt *in vitro* binding of MAF1b to RALGAPA1 (MAF1b:VAP→MAF1a:RFR). Expression of both constructs in Type II (Me49) parasites did not disrupt MAF1b trafficking to the PVM and HMA phenotype (Fig A-4B). We also transfected Type II parasites with a MAF1a construct where the aligned binding loop sequence was mutated to its MAF1b counterpart in an effort to complement the positive HMA phenotype. However, this chimeric construct was unable to drive association with the host mitochondria when expressed in Type II parasites (Fig A-4B).

Appendix A.4 RALGAPA1 localization is not dependent on MAF1b expression

To assess the significance of the MAF1b-RALGAPA1 interaction in the absence of the parasite proteome, we used ectopic GFP-MAF1b/a expression (see Chapter 3) by generating the same MAF1b/a-RALGAPA1 binding loop mutations fused to an NT-GFP tag (pcDNA3.1/NT-GFP-TOPO). Mutating the MAF1b-RALGAPA1 loop (residues 274-287) to their aligned sequence in MAF1a (residues 263-276) did not disrupt localization with host mitochondria 24

hours post transfection in U2OS cells, whereas mutating the MAF1a-RALGAPA1 loop sequence (residues 263-276) to their aligned sequence in MAF1b (residues 274-287) did not drive association with the host mitochondria (Fig A-5A) and the chimeric MAF1a construct remained largely cytosolic. In order to assess MAF1b-RALGAPA1 interaction in host cells, we quantified the differences between GFP-MAF1b and GFP-MAF1b RGLoop (274-287)MAF1a (i.e. "MAF1b mutant") localization with RALGAPA1 in the nucleus of the cell expressing GFP-MAF1, the mitochondria of the cell expressing GFP-MAF1, and the mitochondria in a neighboring cell not expressing GFP-MAF1. We hypothesized that if the MAF1b-RALGAPA1 *in vitro* interaction could be recapitulated in host cells, we would identify differences between MAF1b and MAF1b mutant co-localization with RALGAPA1. To first assess whether MAF1 expression re-localized and/or enriched RALGAPA1 at the mitochondria, we quantified the pixel intensity ratio of RALGAPA1 to mitochondria. There was no significant difference between RALGAPA1 localization in MAF1b-transfected and MAF1b mutant-transfected cells and no difference in RALGAPA1 localization when compared to a non-MAF1-expressing cell (Fig A-5B). We also quantified the pixel intensity ratio of MAF1 to mitochondria to measure MAF1 localization at the mitochondria. These data showed no significant difference between MAF1b-transfected and MAF1b mutant-transfected cells (Fig A-5C). Most importantly, there was no difference in RALGAPA1/MAF1 pixel intensity ratios between MAF1b-transfected and MAF1b mutant-transfected cells (Fig A-5D). These data indicate that MAF1b localization has no bearing on RALGAPA1 localization.

Overall, our data suggest that the MAF1b-specific interaction with the host GAP protein, RALGAPA1, has no bearing on the *T. gondii* HMA phenotype. It is possible that the observed *in vitro* MAF1b-RALGAPA1 interaction may point to an additional neofunctionalized role of

MAF1b, however, our data suggest that RALGAPA1 (and specifically the MAF1b-RALGAPA1 binding loop) is not required for *T. gondii* HMA.

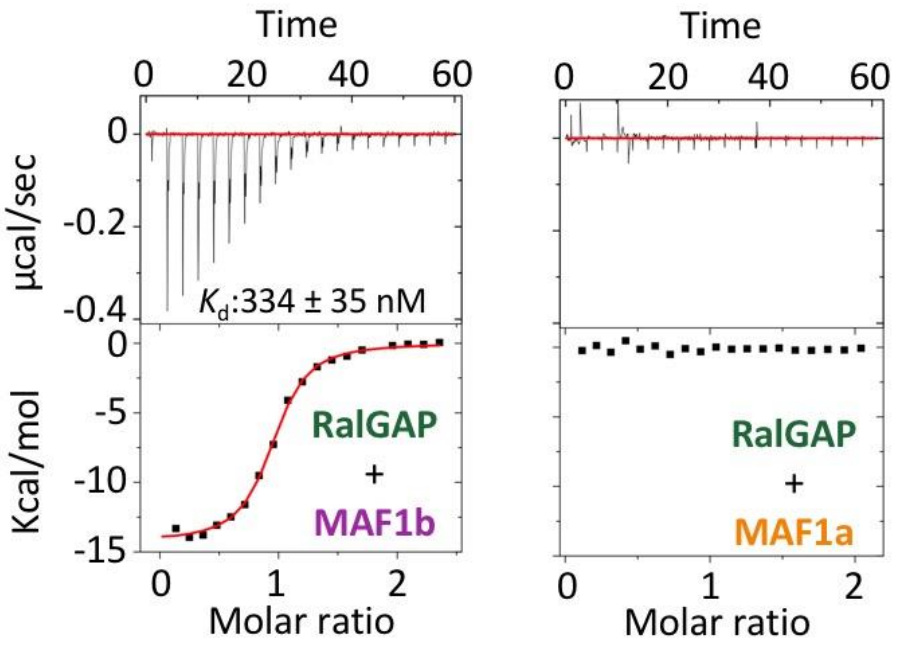


Figure A-1 *Tg*MAF1b interacts with the GAP domain of RalGAPa1.

Representative ITC binding isotherms following the titration of wild-type MAF1b (left) or MAF1a (right) into a cell containing RalGAP.

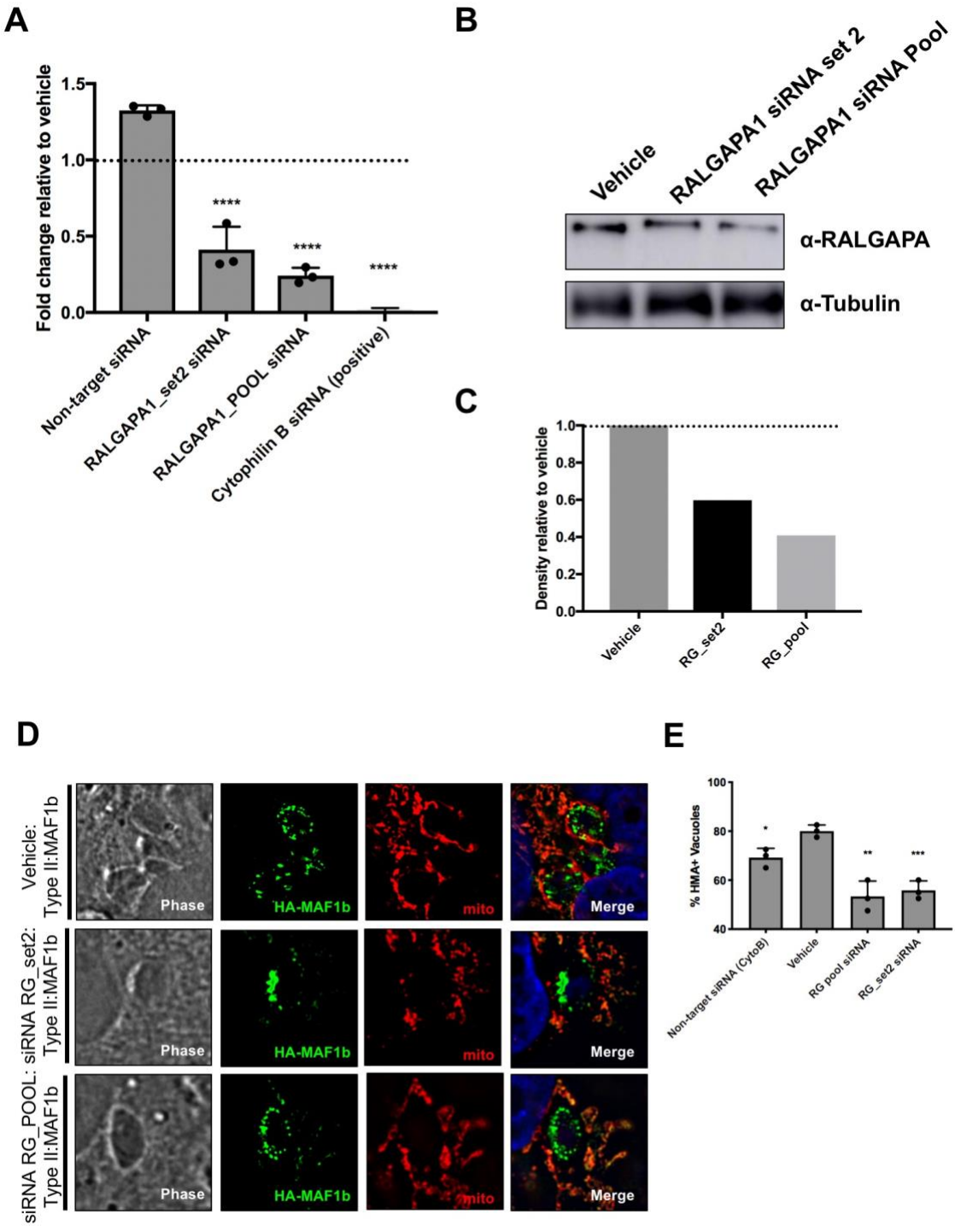


Figure A-2 Knockdown of RALGAPA1 protein decreases *T. gondii* association with host mitochondria similarly to siRNA off-target control.

(A) U2OS were treated with 40nM of RALGAPA1 (set 2) single target siRNA or RALGAPA1 siRNA smart pool (RG pool) in cDMEM. RNA was collected at 48 hours and qPCR analysis was performed (****p<0.0001, one-way ANOVA multiple comparisons). (B) U2OS cells treated with 40nM of RALGAPA1 siRNA set 2 or RALGAPA1 siRNA pool. At 48 hours, cells were lysed in Pierce IP lysis buffer and boiled in 1x SDS. Samples were resolved on 10% SDS-PAGE gel and transferred to nitrocellulose. Both RALGAPA1 (Sigma-Aldrich Atlas Antibodies #HPA000851) and αTubulin (Cell Signaling Technology #2144) antibodies were used for western analysis. (C) Quantification of pixel intensity density ratio of each treatment RALGAPA1/tubulin bands to vehicle control. (D) U2OS cells were treated as in (A) Following 48 hours of siRNA treatment, cells were infected for 24 hours with Type II: MAF1b parasites (MOI=2). Cells were fixed, permeabilized, and visualized using fluorescence microscopy. Immunofluorescence staining was performed with antibodies against the HA epitope tag and mitochondria. (E) HMA+ and HMA- vacuoles were counted (n=3, 50 per treatment group) using fluorescence microscopy. *p=0.0147, **p=0.0024, ***p=0.0008 unpaired t-test comparing treatment to vehicle.

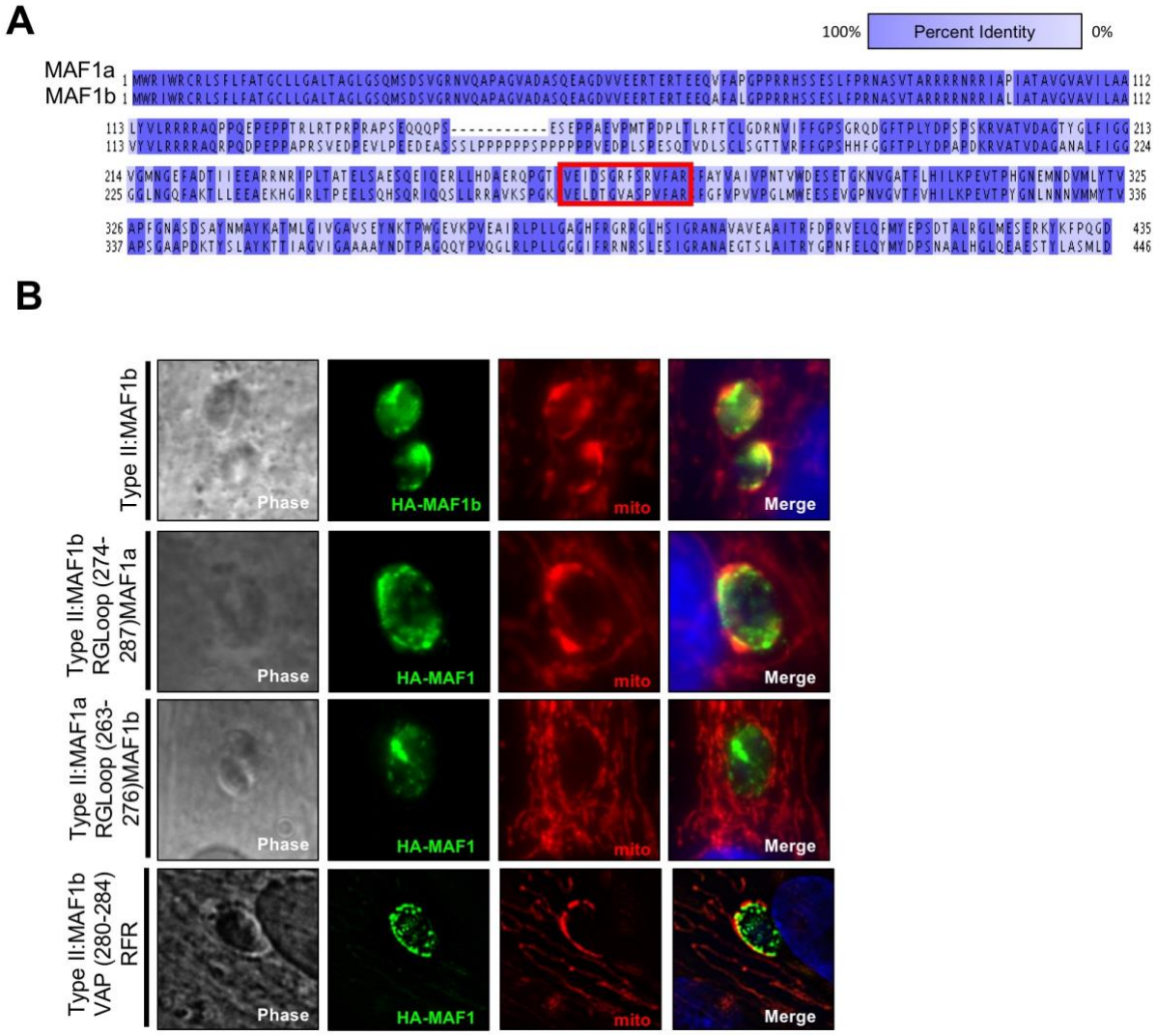


Figure A-3 MAF1b RALGAPA1-binding loop is not required to mediate HMA.

(A) Alignment of *Tg*MAF1RHb1 (MAF1b) and *Tg*MAF1RHa1 (MAF1a) visualized in JalView after alignment using ClustalOmega. MAF1-RALGAPA1 binding loop outlines in red box. (B) Site-directed mutagenesis of MAF1b-RALGAPA1 binding loop and the aligned region in MAF1a. HFFs were infected with stable *Tg*Me49 (Me49ΔHPT:Luc) clones of Type II:MAF1b-RGLoop(274-287)MAF1a and Type II:MAF1a-RGLoop(263-276)MAF1b. The three residue mutation (VAP→RFR) of MAF1b was transiently expressed in Type II (Me49ΔHPT:Luc) parasites and all infected cells were fixed 24 hpi and visualized using fluorescence microscopy. Immunofluorescence staining was performed with antibodies against the HA epitope tag and mitochondria.

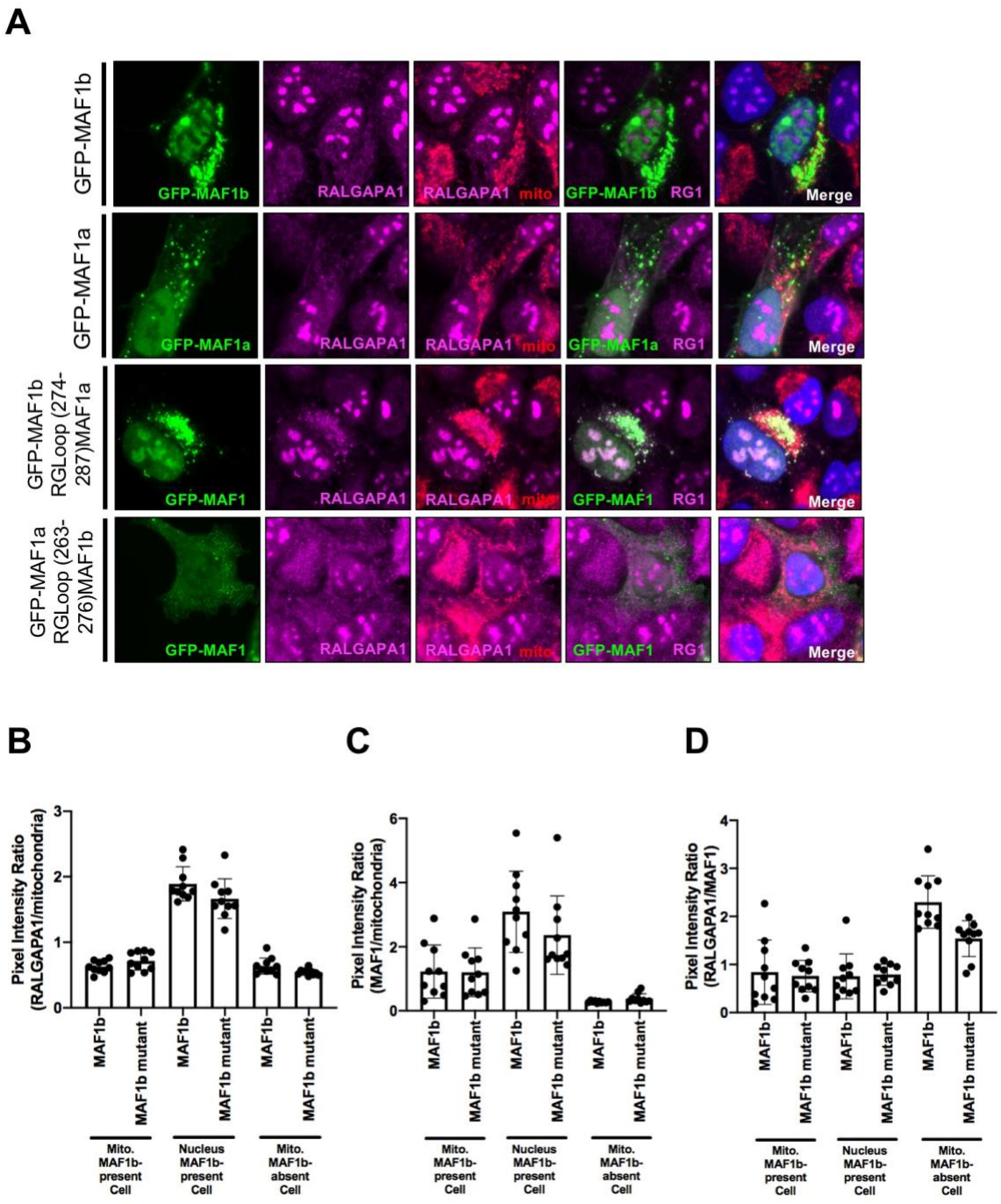


Figure A-4 RALGAPA1 localization does not depend on MAF1b localization.

(A) Mammalian GFP-MAF1 plasmids were generated using pcDNA3.1/NT-GFP TOPO plasmid cloning kit with the same mutations listed in Fig 4-4. U2OS cells were transfected with each construct by lipofection (Lipofectamine 3000, ThermoFisher). Cells were fixed and permeabilized at 24 hours post transfection and stained for RALGAPA1 (Sigma-Aldrich Atlas Antibodies #HPA000851) and mitochondria (Abcam #92824). (B) A selected 15x15 pixel area was

chosen for each measurement. Pixel intensity was measured in both the RALGAPA1 and mitochondria channel over the mitochondria of a MAF1-expressing cell, the nucleus of a MAF1-expressing cell, and the mitochondria of a neighboring non-MAF1-expressing cell. The ratio of RALGAPA1/mitochondria pixel intensity was measured for each of these three locations. This same procedure was performed for the MAF1b mutant (GFP-MAF1b-RGLoop(274-287)MAF1a) samples. (C) Similar to (B), the pixel intensity ratio of MAF1/mitochondria was measured. (D) Similar to (B), the pixel intensity ratio of RALGAPA1/MAF1 was measured.

Table 2. Primers used in Appendix A

Forward (5'-3') Primer	Reverse (3'-5') Primer	Purpose
CAGCTCCCTACCACCATTTAC	GGGAGTTTCACTGGGTAGAATC	RALGAPA1 qPCR primers
AGTGGGAGATTCTCCCGAGTG TTTGCAAGGTCGTTCGGCTTC	GTCTATCTCCACCAGTGTCCCCG GGCTTTTCACGGCACGGCG	MAF1 RALGAPA1 Loop Mutant TgMAF1RHb1 (274-287)RHa1
ACCGGGGTCGCTTCTCCAGTG TTCGCCAGGTCCTTTGCCTAC	GTCTAACTCCACCAGTTTCCCCGG TTGTCTCTCTGCATCGTGAAG	MAF1 RALGAPA1 Loop Mutant TgMAF1RHa1(263-276)TgMAF1RHb1
TCTAGAGTGTTCGCAAGGTCGTTC	GAACCGCCCGGTGTCTAACTCCAC	MAF1 RALGAPA1 Loop Mutant VASP→ RFSR TgMAF1RHb1 (280-284)TgMAF1RHa1
CACATCGCTCAGACACCATG	TGACGGTGCCATGGAATTTG	GAPDH qPCR primers

Appendix B

Table 3. Primers used for MAF1 mutational analysis in Chapter 2

Forward (5'-3') Primer	Reverse (3'-5') Primer	Purpose
GCCCTCCGGGCAAGACGAGGCTTCGTCTTCG	GCCCGGGGACGCGGCACAGAACGCGGTGCAGG	TgMAF1RHb1(136-144)RHa1
AGATCCGTTGAAGCCGGAGTCCCAAACAGTTGACCTCTCG	TCGACGGGCGGCGGCGGA	TgMAF1RHb1(170) RHb0
AGCCTTCTGAGTCAGAGCCGCCCGCGGGAGTTCCGATGAAGCCTGAGTCCCAAACAGTTGACCTCTC	GCTGTTGCCCGGAAGGAGCCCTGGGACGCGGCGTACGGAGTCGCGTAGGGGGTTCCGGATCCTG	TgMAF1RHb1(131-174)RHb0
TGTTTGGGAGACACGAAAGTGACATTCTTTGGGCCGTCACACCA	CGTGAAGGGAAGTGTTAGGGACTCCGGCGACAACGGATCTT	TgMAF1RHb1(174-187)RHb0
GTGGCCCCTTTTGGAGCTGCG	GGTGTACATCATGACGTTGTTGTTC	TgMAF1RHb1:S(339)F: RHa1
AAGTATAAGTTTCCCCAGGGAGATTGATTAATTAAGACTACGACGAAAGTGATGC	GCGTTCAGATTCCATTAAACCGCGAAGAGCCGCGTTGCTGGG	TgMAF1RHb1:16 C-terminal chimera
ACGTATCTGGCTAGCATGCTGGACTGATTAATTAAGACTACGACGAAAGTGATG	GCTTTCAGCCTCCTGTAAGCCGTGGAGGGCTGTGTCGCTAGG	TgMAF1RHa1:16 C-term chimera
TAACACGGCTTACAGGAGGCTG	AAGAGCCGCGTTGCTGGG	TgMAF1RHb1: 16 C-terminal truncation
TAAACGTATCTGGCTAGCATG	GCTTTCAGCCTCCTGTAAG	TgMAF1RHb1:8 C-terminal truncation
TACAAGGCTAGCATGCTGGACTGA	CTTCCTTTCAGCCTCCTGTAAGCC	TgMAF1RHb1: STL(438-441)RKK
GGCTGAAAGCAAATATCTGGCTAGCATG	TCCTGTAAGCCGTGAAGA	TgMAF1RHb1: T(439)K
AAGCACGTATAAAGCTAGCATGCTGGAC	TCAGCCTCCTGTAAGCCG	TgMAF1RHb1: L(441)K
GGAGGCTGAACGCACGTATCTGG	TGTAAGCCGTGAAGAGCC	TgMAF1RHb1: S(438)R
TATAAAGCTAGCATGCTGGACTGA	CGTTCTTTCAGCCTCCTGTAAGCC	TgMAF1RHb1: SL(438-441)RK

Table 4. Primers used in Chapter 3

Forward (5'-3') Primer	Reverse (5'-3') Primer	Purpose
TAATACGACTCACTATAGGGAGA GCCCAAGCAGCCAAGAATAAA	TAATACGACTCACTATAGGGAGA CGTGGAGATGGCATCAGAGGT	TOMM70 (set2) human primers for esiRNA generation
TAATACGACTCACTATAGGGAGA GCGGGTGATCTTTGGTCTCTT	TAATACGACTCACTATAGGGAGA CCTTCTCCACCTCGATCTTGC	Cytophilin B (set2) human primers for esiRNA generation
TAATACGACTCACTATAGGGAGA GTTCAGTCCCGATTTGTGGGG	TAATACGACTCACTATAGGGAGA CAAGGAATGCCTTTGATCACCA	GLS (set3) human primers for esiRNA generation
TAATACGACTCACTATAGGGAGA GAATGGGAGCAGTTTTCATGG	TAATACGACTCACTATAGGGAGA CGTCTTGTGCCAATCCCAGAT	HIBADH (set1) human primers for esiRNA generation
TAATACGACTCACTATAGGGAGA GCCCTGGAAGGTGTGCCAATA	TAATACGACTCACTATAGGGAGA CGGTTCAAGGCTCCAAACGC	GK2 (set6) human primers for esiRNA generation
TAATACGACTCACTATAGGGAGA GGGTGGCCTCAGGTGAAAATA	TAATACGACTCACTATAGGGAGA CACAAACAGACACACCCCACA	GLS (set2) human primers for esiRNA generation
TAATACGACTCACTATAGGGAGA GCGATTCTCGGAATTCTTCCA	TAATACGACTCACTATAGGGAGA CAGAGGTTTCCTGCGTTGAGA	SAMM50 (set2) human primers for esiRNA generations
TAATACGACTCACTATAGGGAGA GTGCTACCAAGCGTCTCATTG	TAATACGACTCACTATAGGGAGA CATCCCCATTTGTGGATTTCA	HSPA9 (set1) human primers for esiRNA generation
TAATACGACTCACTATAGGGAGA GTGGACAACAGTGGCAGTCTC	TAATACGACTCACTATAGGGAGA CCTCAAACTCCACACCCACCT	TOMM40 (set1) human primers for esiRNA generation
TAATACGACTCACTATAGGGAGA GTTGGCAAAGCTGCCAGAGAT	TAATACGACTCACTATAGGGAGA CCAAGCCAGCCCTCATAACCA	VDAC2 (set1) human primers for esiRNA generation
TCCCTGGAGAAGAGCTACGA	AGCACTGTGTTGGCGTACAG	ActinB qPCR human primers
GACAGACAGCCGGGATAAAC	CACAGACGGTCACTCAAAGAA	Cytophilin B qPCR human primers
CCCTGTCTCTGGGTATTGAAAC	CATCAGCGGCAGTAGAGAATAC	HSPA9 qPCR human primers
CCCAAGCAGCCAAGAATAAAG	CTCTGTAGGGCACAAGCTAATA	TOMM70 qPCR human primers
CGCACTTCCAAATGGGTTAGA	CTGCACGACCAAGAAGATTAGG	SAMM50 qPCR human primers
TCTCAACGCTCAGGTCATTC	CCTGCCAGTTCACAAACTTC	TOM40 qPCR human primers
CATGGGTACTATGCAGGAAAGA	AGTCATCAGCAGGCACATAG	ATP5B qPCR human primers
GGTTCAGCTGTCTTTGGTTATG	GTAGCCCACTGCAAAGTTATTC	VDAC2 qPCR human primers
CTGGAGGAAAGGTTGCAGATTA	GAATGCCTCTGTCCATCTACTG	GLS qPCR human primers
CGTTCCCTCGGCTAATAACTATC	CCAAGAAGGATTGGGCTCTT	HIBADH qPCR human primers
CCTGAAACAACTGCACTAGGA	TCCATCCTGAGAACTGACAAAG	GK2 qPCR human primers

ATGCATGGTGGTCTAGGCAG	TCAGTCCAGCATGCTAGCCA	GFP-MAF1b for TOPO NT-GFP fusion
ATGCATGGTGGTCTAGGCAG	TCAGTCGCCTTGCGGAAACT	GFP-MAF1a for TOPO NT-GFP fusion
ATGCATGGTGGTCTAGGCAG	TCAGTCCAGCATGCTAGCCT	GFP-MAF1b (STL→RKK) for TOPO NT-GFP fusion
agtgtggtggaattgcccttATGCATGGTGGTCTAGGC	tgtggtggctgtgcccggcgCCGCCTTAATACATACACCG	MAF1 chimera 1 MAF1b portion
CGCCGGGCACAGCCACCA	AAGGGCAATTCCACCACACTGGACTAGTG	MAF1 chimera 1 MAF1a portion for Gibson
tgacctctcgtgtttgtcaggcacgacagtgagaTTCTTTGGGCCGTCGGGC	actgtttgggactccggcgacaacggatcttcgacGGGCGGCTCTGACTCAGAAG	MAF1 chimera 2 for Q5 mutagenesis
tcgaggaggcaagacgaaacGGCATCCGGCTGACACCC	atatctgcagaattgcccttTCAGTCCAGCATGCTAGCCAG	MAF1 chimera 3 MAF1b portion for Gibson
AAGGGCAATTCTGCAGATATC	GTTTCGTCTTGCCTCCTC	MAF1 chimera 3 MAF1a portion for Gibson
ggcatccggctgacacccgaagag	ctgtgcccggcgccgcct	MAF1 chimera 4 MAF1b portion for Gibson
Taaggcggcgccgggcacagccaccacaggagccggaa	Tcgggtgtcagccggatgccgtttcgtcttgcctcctcg	MAF1 chimera 4 MAF1a portion for Gibson
tatctgTTTCCGCAAGGCGACTGA	cgtgctTTCAGACTCCATCAAGCCAC	MAF1 chimera 1&2 RKK→STL C-term Q5

Table 5. Data collection statistics for X-Ray crystallographic studies in Chapter 2

	*Tg*MAF1RHb1-Br	*Tg*MAF1RHb1-ADPr	*Tg*MAF1RHa1	*Tg*MAF1Rha1-ADPr
A. Data collection statistics				
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_1$	$P2_1$
a, *b*, *c* (Å)	45.09, 62.11, 89.21	45.47, 62.36, 89.69	79.89, 49.67, 114.75	78.23, 49.52, 115.71
α, *β*, *γ* (°)	90, 90, 90	90, 90, 90	90, 96.09, 90	90, 96.63, 90
Wavelength (Å)	0.9195	0.9795	0.9794	0.9795
Resolution (Å)	45.09-1.60 (1.69-1.60)	51.20-1.65 (1.68-1.65)	48.86-2.10 (2.16-2.10)	27.40-2.70 (2.86-2.70)
Measured reflections	280,473 (26,900)	114,544 (5,631)	197,078 (14,313)	117,630 (16,227)
Unique reflections	33,786 (4,786)	29,065 (1,489)	51,862 (3,990)	23,134 (3,436)
Redundancy	8.3 (5.6)	3.9 (3.8)	3.8 (3.6)	5.1 (4.7)
Completeness (%)	99.8 (98.9)	93.4 (97.4)	98.5 (92.4)	94.4 (87.9)
$I/\sigma(I)$	12.7 (3.5)	7.6 (2.2)	11.5 (3.0)	11.6 (2.2)
R_{merge} (%)	10.3 (47.5)	9.4 (55.3)	8.0 (44.4)	9.0 (61.1)
B. Refinement statistics				
Resolution (Å)	44.60-1.60	40.56-1.65	48.86-2.10	27.40-2.70
R_{work} / R_{free} (%)	17.0/21.0	20.5/22.8	21.2/24.4	23.3/28.3

No. of atoms				
Protein (A/B/C)	2057	2012	2114/2081/2099	2047/2005/1949
Solvent/Br/Na	257/11/3	105	159	16
Glycerol or Sulfate	18	N/A	70	N/A
ADPribose	N/A	36	N/A	72
Average B-values (Å$_2$)				
Protein (A/B/C)	14.2	22.3	30.5/36.8/33.4	55.1/63.9/63.6
Solvent/Br/Na	24.0/26.9/32.9	28.4	31.5	42.6
Glycerol or Sulfate	23.0	N/A	47.8	N/A
ADPribose	N/A	33.4	N/A	68.6
r.m.s. deviation from ideality				
Bond lengths (Å)	0.008	0.012	0.003	0.003
Bond angles (°)	1.17	1.02	0.67	0.67
Ramachandran statistics (%)				
Most favoured	99.3	98.9	98.5	96.5
Allowed	0.7	1.1	1.5	3.5
Disallowed	0.0	0.0	0.0	0.0

Values in parentheses are for the highest resolution shell

Bibliography

1. Sudmant, P.H., Kitzman, J.O., Antonacci, F., Alkan, C., Malig, M., Tsalenko, A., Sampas, N., Bruhn, L., Shendure, J., Genomes, P., et al. (2010). Diversity of human copy number variation and multicopy genes. Science *330*, 641-646.

2. Campbell, C.D., Sampas, N., Tsalenko, A., Sudmant, P.H., Kidd, J.M., Malig, M., Vu, T.H., Vives, L., Tsang, P., Bruhn, L., et al. (2011). Population-genetic properties of differentiated human copy-number polymorphisms. Am J Hum Genet *88*, 317-332.

3. Ohno, S. (1970). Evolution by gene duplication, (London, New York,:Allen & Unwin; Springer-Varlag).

4. Espinosa-Cantu, A., Ascencio, D., Barona-Gomez, F., and DeLuna, A. (2015). Gene duplication and the evolution of moonlighting proteins. Front Genet *6*, 227.

5. Magadum, S., Banerjee, U., Murugan, P., Gangapur, D., and Ravikesavan, R. (2013). Gene duplication as a major force in evolution. J Genet *92*, 155-161.

6. Koonin, E.V. (2016). Splendor and misery of adaptation, or the importance of neutral null for understanding evolution. BMC Biol *14*, 114.

7. Lynch, M., and Conery, J.S. (2000). The evolutionary fate and consequences of duplicate genes. Science *290*, 1151-1155.

8. Wagner, A. (2002). Selection and gene duplication: a view from the genome. Genome Biol *3*, reviews1012.

9. Reams, A.B., and Roth, J.R. (2015). Mechanisms of gene duplication and amplification. Cold Spring Harb Perspect Biol *7*, a016592.

10. Adomako-Ankomah, Y., Wier, G.M., Borges, A.L., Wand, H.E., and Boyle, J.P. (2014). Differential locus expansion distinguishes Toxoplasmatinae species and closely related strains of *Toxoplasma gondii*. MBio *5*, e01003-01013.

11. Blank, M.L., Parker, M.L., Ramaswamy, R., Powell, C.J., English, E.D., Adomako-Ankomah, Y., Pernas, L.F., Workman, S.D., Boothroyd, J.C., Boulanger, M.J., et al. (2018). A *Toxoplasma gondii* locus required for the direct manipulation of host mitochondria has maintained multiple ancestral functions. Mol Microbiol.

12. Adomako-Ankomah, Y., English, E.D., Danielson, J.J., Pernas, L.F., Parker, M.L., Boulanger, M.J., Dubey, J.P., and Boyle, J.P. (2016). Host Mitochondrial Association

Evolved in the Human Parasite *Toxoplasma gondii* via Neofunctionalization of a Gene Duplicate. Genetics *203*, 283-298.

13. Freitas-Junior, L.H., Bottius, E., Pirrit, L.A., Deitsch, K.W., Scheidig, C., Guinet, F., Nehrbass, U., Wellems, T.E., and Scherf, A. (2000). Frequent ectopic recombination of virulence factor genes in telomeric chromosome clusters of P. falciparum. Nature *407*, 1018-1022.

14. Pasternak, N.D., and Dzikowski, R. (2009). PfEMP1: an antigen that plays a key role in the pathogenicity and immune evasion of the malaria parasite Plasmodium falciparum. Int J Biochem Cell Biol *41*, 1463-1466.

15. Berriman, M., Ghedin, E., Hertz-Fowler, C., Blandin, G., Renauld, H., Bartholomeu, D.C., Lennard, N.J., Caler, E., Hamlin, N.E., Haas, B., et al. (2005). The genome of the African trypanosome Trypanosoma brucei. Science *309*, 416-422.

16. Mugnier, M.R., Cross, G.A., and Papavasiliou, F.N. (2015). The in vivo dynamics of antigenic variation in Trypanosoma brucei. Science *347*, 1470-1473.

17. Torgerson, P.R., and Mastroiacovo, P. (2013). The global burden of congenital toxoplasmosis: a systematic review. Bull World Health Organ *91*, 501-508.

18. Furtado, J.M., Smith, J.R., Belfort, R., Jr., Gattey, D., and Winthrop, K.L. (2011). Toxoplasmosis: a global threat. J Glob Infect Dis *3*, 281-284.

19. Lorenzi, H., Khan, A., Behnke, M.S., Namasivayam, S., Swapna, L.S., Hadjithomas, M., Karamycheva, S., Pinney, D., Brunk, B.P., Ajioka, J.W., et al. (2016). Local admixture of amplified and diversified secreted pathogenesis determinants shapes mosaic *Toxoplasma gondii* genomes. Nat Commun *7*, 10147.

20. Reese, M.L., and Boothroyd, J.C. (2011). A conserved non-canonical motif in the pseudoactive site of the ROP5 pseudokinase domain mediates its effect on Toxoplasma virulence. J Biol Chem *286*, 29366-29375.

21. Wasmuth, J.D., Pszenny, V., Haile, S., Jansen, E.M., Gast, A.T., Sher, A., Boyle, J.P., Boulanger, M.J., Parkinson, J., and Grigg, M.E. (2012). Integrated bioinformatic and targeted deletion analyses of the SRS gene superfamily identify SRS29C as a negative regulator of Toxoplasma virulence. MBio *3*.

22. Reid, A.J., Vermont, S.J., Cotton, J.A., Harris, D., Hill-Cawthorne, G.A., Konen-Waisman, S., Latham, S.M., Mourier, T., Norton, R., Quail, M.A., et al. (2012). Comparative genomics of the apicomplexan parasites *Toxoplasma gondii* and *Neospora caninum*: Coccidia differing in host range and transmission strategy. PLoS Pathog *8*, e1002567.

23. Walzer, K.A., Adomako-Ankomah, Y., Dam, R.A., Herrmann, D.C., Schares, G., Dubey, J.P., and Boyle, J.P. (2013). *Hammondia hammondi*, an avirulent relative of Toxoplasma

gondii, has functional orthologs of known T. gondii virulence genes. Proc Natl Acad Sci U S A *110*, 7446-7451.

24. Gajria, B., Bahl, A., Brestelli, J., Dommer, J., Fischer, S., Gao, X., Heiges, M., Iodice, J., Kissinger, J.C., Mackey, A.J., et al. (2008). ToxoDB: an integrated *Toxoplasma gondii* database resource. Nucleic Acids Res *36*, D553-556.

25. Boothroyd, J.C., and Dubremetz, J.F. (2008). Kiss and spit: the dual roles of Toxoplasma rhoptries. Nat Rev Microbiol *6*, 79-88.

26. Blader, I.J., Coleman, B.I., Chen, C.T., and Gubbels, M.J. (2015). Lytic Cycle of *Toxoplasma gondii*: 15 Years Later. Annu Rev Microbiol *69*, 463-485.

27. Saeij, J.P., Boyle, J.P., Coller, S., Taylor, S., Sibley, L.D., Brooke-Powell, E.T., Ajioka, J.W., and Boothroyd, J.C. (2006). Polymorphic secreted kinases are key virulence factors in toxoplasmosis. Science *314*, 1780-1783.

28. Behnke, M.S., Khan, A., Wootton, J.C., Dubey, J.P., Tang, K., and Sibley, L.D. (2011). Virulence differences in Toxoplasma mediated by amplification of a family of polymorphic pseudokinases. Proc Natl Acad Sci U S A *108*, 9631-9636.

29. Reese, M.L., Zeiner, G.M., Saeij, J.P., Boothroyd, J.C., and Boyle, J.P. (2011). Polymorphic family of injected pseudokinases is paramount in Toxoplasma virulence. Proc Natl Acad Sci U S A *108*, 9625-9630.

30. Behnke, M.S., Fentress, S.J., Mashayekhi, M., Li, L.X., Taylor, G.A., and Sibley, L.D. (2012). The polymorphic pseudokinase ROP5 controls virulence in *Toxoplasma gondii* by regulating the active kinase ROP18. PLoS Pathog *8*, e1002992.

31. Pernas, L., Adomako-Ankomah, Y., Shastri, A.J., Ewald, S.E., Treeck, M., Boyle, J.P., and Boothroyd, J.C. (2014). Toxoplasma effector MAF1 mediates recruitment of host mitochondria and impacts the host response. PLoS Biol *12*, e1001845.

32. Kim, S.K., Karasov, A., and Boothroyd, J.C. (2007). Bradyzoite-specific surface antigen SRS9 plays a role in maintaining *Toxoplasma gondii* persistence in the brain and in host control of parasite replication in the intestine. Infect Immun *75*, 1626-1634.

33. Hu, K., Johnson, J., Florens, L., Fraunholz, M., Suravajjala, S., DiLullo, C., Yates, J., Roos, D.S., and Murray, J.M. (2006). Cytoskeletal components of an invasion machine--the apical complex of *Toxoplasma gondii*. PLoS Pathog *2*, e13.

34. Gould, S.B., Tham, W.H., Cowman, A.F., McFadden, G.I., and Waller, R.F. (2008). Alveolins, a new family of cortical proteins that define the protist infrakingdom Alveolata. Mol Biol Evol *25*, 1219-1230.

35. Olias, P., and Sibley, L.D. (2016). Functional Analysis of the Role of Toxoplasma gondii Nucleoside Triphosphate Hydrolases I and II in Acute Mouse Virulence and Immune Suppression. Infect Immun *84*, 1994-2001.

36. Taylor, S., Barragan, A., Su, C., Fux, B., Fentress, S.J., Tang, K., Beatty, W.L., Hajj, H.E., Jerome, M., Behnke, M.S., et al. (2006). A secreted serine-threonine kinase determines virulence in the eukaryotic pathogen *Toxoplasma gondii*. Science *314*, 1776-1780.

37. Behnke, M.S., Khan, A., Lauron, E.J., Jimah, J.R., Wang, Q., Tolia, N.H., and Sibley, L.D. (2015). Rhoptry Proteins ROP5 and ROP18 Are Major Murine Virulence Factors in Genetically Divergent South American Strains of *Toxoplasma gondii*. PLoS Genet *11*, e1005434.

38. Jung, C., Lee, C.Y., and Grigg, M.E. (2004). The SRS superfamily of Toxoplasma surface proteins. Int J Parasitol *34*, 285-296.

39. Crawford, J., Lamb, E., Wasmuth, J., Grujic, O., Grigg, M.E., and Boulanger, M.J. (2010). Structural and functional characterization of SporoSAG: a SAG2-related surface antigen from *Toxoplasma gondii*. J Biol Chem *285*, 12063-12070.

40. Jones, T.C., Yeh, S., and Hirsch, J. (1971). The Interaction Between *Toxoplasma Gondii* and Mammalian Cells. J Exp Med *136*, 1157-1172.

41. Sinai, A.P., Webster, P., and Joiner, K.A. (1997). Association of host cell endoplasmic reticulum and mitochondria with the *Toxoplasma gondii* parasitophorous vacuole membrane: a high affinity interaction. J Cell Sci *110 (Pt 17)*, 2117-2128.

42. Pernas, L., and Boothroyd, J.C. (2010). Association of host mitochondria with the parasitophorous vacuole during Toxoplasma infection is not dependent on rhoptry proteins ROP2/8. Int J Parasitol *40*, 1367-1371.

43. Horwitz, M.A. (1983). Formation of a novel phagosome by the Legionnaires' disease bacterium (*Legionella pneumophila*) in human monocytes. J Exp Med *158*, 1319-1331.

44. Escoll, P., Song, O.R., Viana, F., Steiner, B., Lagache, T., Olivo-Marin, J.C., Impens, F., Brodin, P., Hilbi, H., and Buchrieser, C. (2017). *Legionella pneumophila* Modulates Mitochondrial Dynamics to Trigger Metabolic Repurposing of Infected Macrophages. Cell Host Microbe *22*, 302-316 e307.

45. Matsumoto, A., Bessho, H., Uehira, K., and Suda, T. (1991). Morphological studies of the association of mitochondria with chlamydial inclusions and the fusion of chlamydial inclusions. J Electron Microsc (Tokyo) *40*, 356-363.

46. Du Toit, A. (2017). Bacterial pathogenesis: Rewiring cellular dynamics and metabolism. Nat Rev Microbiol *15*, 576.

47. Pernas, L., Bean, C., Boothroyd, J.C., and Scorrano, L. (2018). Mitochondria Restrict Growth of the Intracellular Parasite *Toxoplasma gondii* by Limiting Its Uptake of Fatty Acids. Cell Metab *27*, 886-897 e884.

48. Anderson, A.J., Jackson, T.D., Stroud, D.A., and Stojanovski, D. (2019). Mitochondria-hubs for regulating cellular biochemistry: emerging concepts and networks. Open Biol *9*, 190126.

49. West, A.P., Shadel, G.S., and Ghosh, S. (2011). Mitochondria in innate immune responses. Nat Rev Immunol *11*, 389-402.

50. Filadi, R., Leal, N.S., Schreiner, B., Rossi, A., Dentoni, G., Pinho, C.M., Wiehager, B., Cieri, D., Cali, T., Pizzo, P., et al. (2018). TOM70 Sustains Cell Bioenergetics by Promoting IP3R3-Mediated ER to Mitochondria Ca(2+) Transfer. Curr Biol *28*, 369-382 e366.

51. Rieusset, J. (2018). The role of endoplasmic reticulum-mitochondria contact sites in the control of glucose homeostasis: an update. Cell Death Dis *9*, 388.

52. Bock, F.J., and Tait, S.W.G. (2019). Mitochondria as multifaceted regulators of cell death. Nat Rev Mol Cell Biol.

53. Parsons, M.J., and Green, D.R. (2010). Mitochondria in cell death. Essays Biochem *47*, 99-114.

54. Li, X.D., Chiu, Y.H., Ismail, A.S., Behrendt, C.L., Wight-Carter, M., Hooper, L.V., and Chen, Z.J. (2011). Mitochondrial antiviral signaling protein (MAVS) monitors commensal bacteria and induces an immune response that prevents experimental colitis. Proc Natl Acad Sci U S A *108*, 17390-17395.

55. Missiroli, S., Patergnani, S., Caroccia, N., Pedriali, G., Perrone, M., Previati, M., Wieckowski, M.R., and Giorgi, C. (2018). Mitochondria-associated membranes (MAMs) and inflammation. Cell Death Dis *9*, 329.

56. Lee, K.S., Huh, S., Lee, S., Wu, Z., Kim, A.K., Kang, H.Y., and Lu, B. (2018). Altered ER-mitochondria contact impacts mitochondria calcium homeostasis and contributes to neurodegeneration in vivo in disease models. Proc Natl Acad Sci U S A *115*, E8844-E8853.

57. Arnoult, D., Soares, F., Tattoli, I., and Girardin, S.E. (2011). Mitochondria in innate immunity. EMBO Rep *12*, 901-910.

58. Liu, X.Y., Wei, B., Shi, H.X., Shan, Y.F., and Wang, C. (2010). Tom70 mediates activation of interferon regulatory factor 3 on mitochondria. Cell Res *20*, 994-1011.

59. Kelly, F.D., Wei, B.M., Cygan, A.M., Parker, M.P., Boulanger, M.J., and Boothroyd, J.C. (2017). *Toxoplasma gondii* MAF1b Binds the Host Cell MIB Complex To Mediate Mitochondrial Association. mSphere *2*.

60. Treeck, M., Sanders, J.L., Elias, J.E., and Boothroyd, J.C. (2011). The phosphoproteomes of Plasmodium falciparum and *Toxoplasma gondii* reveal unusual adaptations within and beyond the parasites' boundaries. Cell Host Microbe *10*, 410-419.

61. Holm, L., and Rosenstrom, P. (2010). Dali server: conservation mapping in 3D. Nucleic Acids Res *38*, W545-549.

62. Allen, M.D., Buckle, A.M., Cordell, S.C., Lowe, J., and Bycroft, M. (2003). The crystal structure of AF1521 a protein from Archaeoglobus fulgidus with homology to the non-histone domain of macroH2A. J Mol Biol *330*, 503-511.

63. Han, W., Li, X., and Fu, X. (2011). The macro domain protein family: structure, functions, and their potential therapeutic implications. Mutat Res *727*, 86-103.

64. Karras, G.I., Kustatscher, G., Buhecha, H.R., Allen, M.D., Pugieux, C., Sait, F., Bycroft, M., and Ladurner, A.G. (2005). The macro domain is an ADP-ribose binding module. EMBO J *24*, 1911-1920.

65. Barkauskaite, E., Jankevicius, G., and Ahel, I. (2015). Structures and Mechanisms of Enzymes Employed in the Synthesis and Degradation of PARP-Dependent Protein ADP-Ribosylation. Mol Cell *58*, 935-946.

66. Ashkenazy, H., Erez, E., Martz, E., Pupko, T., and Ben-Tal, N. (2010). ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. Nucleic acids research *38*, W529-W533.

67. Malet, H., Coutard, B., Jamal, S., Dutartre, H., Papageorgiou, N., Neuvonen, M., Ahola, T., Forrester, N., Gould, E.A., Lafitte, D., et al. (2009). The crystal structures of Chikungunya and Venezuelan equine encephalitis virus nsP3 macro domains define a conserved adenosine binding pocket. J Virol *83*, 6534-6545.

68. Egloff, M.P., Malet, H., Putics, A., Heinonen, M., Dutartre, H., Frangeul, A., Gruez, A., Campanacci, V., Cambillau, C., Ziebuhr, J., et al. (2006). Structural and functional basis for ADP-ribose and poly(ADP-ribose) binding by viral macro domains. J Virol *80*, 8493-8502.

69. Neuvonen, M., and Ahola, T. (2009). Differential activities of cellular and viral macro domain proteins in binding of ADP-ribose metabolites. J Mol Biol *385*, 212-225.

70. Kelly, F.D., Wei, B.M., Cygan, A.M., Parker, M.L., Boulanger, M.J., and Boothroyd, J.C. (2017). *Toxoplasma gondii* MAF1b Binds the Host Cell MIB Complex To Mediate Mitochondrial Association. mSphere *2*.

71. Li, C., Debing, Y., Jankevicius, G., Neyts, J., Ahel, I., Coutard, B., and Canard, B. (2016). Viral Macro Domains Reverse Protein ADP-Ribosylation. J Virol *90*, 8478-8486.

72. Daugherty, M.D., Young, J.M., Kerns, J.A., and Malik, H.S. (2014). Rapid evolution of PARP genes suggests a broad role for ADP-ribosylation in host-virus conflicts. PLoS Genet *10*, e1004403.

73. Fox, B.A., Falla, A., Rommereim, L.M., Tomita, T., Gigley, J.P., Mercier, C., Cesbron-Delauw, M.F., Weiss, L.M., and Bzik, D.J. (2011). Type II *Toxoplasma gondii* KU80 knockout strains enable functional analysis of genes required for cyst development and latent infection. Eukaryot Cell *10*, 1193-1206.

74. Kozjak-Pavlovic, V. (2017). The MICOS complex of human mitochondria. Cell Tissue Res *367*, 83-93.

75. Harner, M., Korner, C., Walther, D., Mokranjac, D., Kaesmacher, J., Welsch, U., Griffith, J., Mann, M., Reggiori, F., and Neupert, W. (2011). The mitochondrial contact site complex, a determinant of mitochondrial architecture. EMBO J *30*, 4356-4370.

76. Forne, I., Ludwigsen, J., Imhof, A., Becker, P.B., and Mueller-Planitz, F. (2012). Probing the conformation of the ISWI ATPase domain with genetically encoded photoreactive crosslinkers and mass spectrometry. Mol Cell Proteomics *11*, M111 012088.

77. Tonkin, M.L., Beck, J.R., Bradley, P.J., and Boulanger, M.J. (2014). The inner membrane complex sub-compartment proteins critical for replication of the apicomplexan parasite *Toxoplasma gondii* adopt a pleckstrin homology fold. J Biol Chem *289*, 13962-13973.

78. Dunay, I.R., Damatta, R.A., Fux, B., Presti, R., Greco, S., Colonna, M., and Sibley, L.D. (2008). Gr1(+) inflammatory monocytes are required for mucosal resistance to the pathogen *Toxoplasma gondii*. Immunity *29*, 306-317.

79. Gay, G., Braun, L., Brenier-Pinchart, M.P., Vollaire, J., Josserand, V., Bertini, R.L., Varesano, A., Touquet, B., De Bock, P.J., Coute, Y., et al. (2016). *Toxoplasma gondii* TgIST co-opts host chromatin repressors dampening STAT1-dependent gene regulation and IFN-gamma-mediated host defenses. J Exp Med *213*, 1779-1798.

80. Adomako-Ankomah, Y., English, E.D., Danielson, J.J., Pernas, L.F., Parker, M.L., Boulanger, M.J., Dubey, J.P., and Boyle, J.P. (2016). Host mitochondrial association evolved in the human parasite *Toxoplasma gondii* via neofunctionalization of a gene duplicate. Genetics *203*, 283-298.

81. Battye, T.G., Kontogiannis, L., Johnson, O., Powell, H.R., and Leslie, A.G. (2011). iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. Acta Crystallogr D Biol Crystallogr *67*, 271-281.

82. Evans, P.R. (2005). Scaling and assessment of data quality. Acta Cryst *D*, 72-82.

83. Evans, P.R., and Murshudov, G.N. (2013). How good are my data and what is the resolution? Acta Crystallogr D Biol Crystallogr *69*, 1204-1214.

84. Winn, M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., Evans, P.R., Keegan, R.M., Krissinel, E.B., Leslie, A.G., McCoy, A., et al. (2011). Overview of the CCP4 suite and current developments. Acta Crystallogr D Biol Crystallogr *67*, 235-242.

85. Sheldrick, G.M. (2010). Experimental phasing with SHELXC/D/E: combining chain tracing with density modification. Acta Crystallogr D Biol Crystallogr *66*, 479-485.

86. Cowtan, K. (1994). dm: An automated procedure for phase improvement by density modification. Joint CCP4 and ESF-EACBM Newsletter on Protein Crystallography *31*, 34-38.

87. Cowtan, K. (2008). Fitting molecular fragments into electron density. Acta Crystallogr D Biol Crystallogr *64*, 83-89.

88. McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007). Phaser crystallographic software. J Appl Crystallogr *40*, 658-674.

89. Schwarzenbacher, R., Godzik, A., Grzechnik, S.K., and Jaroszewski, L. (2004). The importance of alignment accuracy for molecular replacement. Acta Cryst. *D60*, 1229-1236.

90. Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. Acta Crystallogr D Biol Crystallogr *66*, 486-501.

91. Afonine, P.V., Grosse-Kunstleve, R.W., Echols, N., Headd, J.J., Moriarty, N.W., Mustyakimov, M., Terwilliger, T.C., Urzhumtsev, A., Zwart, P.H., and Adams, P.D. (2012). Towards automated crystallographic structure refinement with phenix.refine. Acta Crystallogr D Biol Crystallogr *68*, 352-367.

92. Chen, V.B., Arendall, W.B., 3rd, Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S., and Richardson, D.C. (2010). MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr D Biol Crystallogr *66*, 12-21.

93. Webb, B., and Sali, A. (2014). Protein structure modeling with MODELLER. Protein Structure Prediction, 1-15.

94. RAMPAGE Server. p. http://ravenbioccam.ac.uk/rampage.php.

95. Benkert, P., Tosatto, S.C., and Schomburg, D. (2008). QMEAN: A comprehensive scoring function for model quality assessment. Proteins: Structure, Function, and Bioinformatics *71*, 261-277.

96. Wiederstein, M., and Sippl, M.J. (2007). ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic acids research *35*, W407-W410.

97. Mitra, K., and Lippincott-Schwartz, J. (2010). Analysis of mitochondrial dynamics and functions using imaging approaches. Curr Protoc Cell Biol *Chapter 4*, Unit 4 25 21-21.

98. Waterhouse, A.M., Procter, J.B., Martin, D.M., Clamp, M., and Barton, G.J. (2009). Jalview Version 2--a multiple sequence alignment editor and analysis workbench. Bioinformatics *25*, 1189-1191.

99. Meng, E.C., Pettersen, E.F., Couch, G.S., Huang, C.C., and Ferrin, T.E. (2006). Tools for integrated sequence-structure analysis with UCSF Chimera. BMC Bioinformatics *7*, 339.

100. Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera--a visualization system for exploratory research and analysis. J Comput Chem *25*, 1605-1612.

101. Schneider, C.A., Rasband, W.S., and Eliceiri, K.W. (2012). NIH Image to ImageJ: 25 years of image analysis. Nat Methods *9*, 671-675.

102. Matsumoto, A. (1981). Isolation and electron microscopic observations of intracytoplasmic inclusions containing *Chlamydia psittaci*. J Bacteriol *145*, 605-612.

103. Heninger, A.K., and Buchholz, F. (2007). Production of Endoribonuclease-Prepared Short Interfering RNAs (esiRNAs) for Specific and Effective Gene Silencing in Mammalian Cells. CSH Protoc *2007*, pdb prot4824.

104. Bausewein, T., Mills, D.J., Langer, J.D., Nitschke, B., Nussberger, S., and Kuhlbrandt, W. (2017). Cryo-EM Structure of the TOM Core Complex from Neurospora crassa. Cell *170*, 693-700 e697.

105. Rapaport, D. (2002). Biogenesis of the mitochondrial TOM complex. Trends Biochem Sci *27*, 191-197.

106. Yamamoto, H., Fukui, K., Takahashi, H., Kitamura, S., Shiota, T., Terao, K., Uchida, M., Esaki, M., Nishikawa, S., Yoshihisa, T., et al. (2009). Roles of Tom70 in import of presequence-containing mitochondrial proteins. J Biol Chem *284*, 31635-31646.

107. Backes, S., Hess, S., Boos, F., Woellhaf, M.W., Godel, S., Jung, M., Muhlhaus, T., and Herrmann, J.M. (2018). Tom70 enhances mitochondrial preprotein import efficiency by binding to internal targeting sequences. J Cell Biol *217*, 1369-1382.

108. Melin, J., Kilisch, M., Neumann, P., Lytovchenko, O., Gomkale, R., Schendzielorz, A., Schmidt, B., Liepold, T., Ficner, R., Jahn, O., et al. (2015). A presequence-binding groove

in Tom70 supports import of Mdl1 into mitochondria. Biochim Biophys Acta *1853*, 1850-1859.

109. Liu, X., Salokas, K., Tamene, F., Jiu, Y., Weldatsadik, R.G., Ohman, T., and Varjosalo, M. (2018). An AP-MS- and BioID-compatible MAC-tag enables comprehensive mapping of protein interactions and subcellular localizations. Nat Commun *9*, 1188.

110. Young, J.C., Hoogenraad, N.J., and Hartl, F.U. (2003). Molecular chaperones Hsp90 and Hsp70 deliver preproteins to the mitochondrial import receptor Tom70. Cell *112*, 41-50.

111. Liu, S., Cai, X., Wu, J., Cong, Q., Chen, X., Li, T., Du, F., Ren, J., Wu, Y.T., Grishin, N.V., et al. (2015). Phosphorylation of innate immune adaptor proteins MAVS, STING, and TRIF induces IRF3 activation. Science *347*, aaa2630.

112. Edmonson, A.M., Mayfield, D.K., Vervoort, V., DuPont, B.R., and Argyropoulos, G. (2002). Characterization of a human import component of the mitochondrial outer membrane, TOMM70A. Cell Commun Adhes *9*, 15-27.

113. Nielsen, H. (2017). Predicting Secretory Proteins with SignalP. Methods Mol Biol *1611*, 59-73.

114. Petersen, T.N., Brunak, S., von Heijne, G., and Nielsen, H. (2011). SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat Methods *8*, 785-786.

115. Almagro Armenteros, J.J., Tsirigos, K.D., Sonderby, C.K., Petersen, T.N., Winther, O., Brunak, S., von Heijne, G., and Nielsen, H. (2019). SignalP 5.0 improves signal peptide predictions using deep neural networks. Nat Biotechnol *37*, 420-423.

116. Drozdetskiy, A., Cole, C., Procter, J., and Barton, G.J. (2015). JPred4: a protein secondary structure prediction server. Nucleic Acids Res *43*, W389-394.

117. Colombini, M. (2016). The VDAC channel: Molecular basis for selectivity. Biochim Biophys Acta *1863*, 2498-2502.

118. Naghdi, S., and Hajnoczky, G. (2016). VDAC2-specific cellular functions and the underlying structure. Biochim Biophys Acta *1863*, 2503-2514.

119. Dubey, J.P., and Sreekumar, C. (2003). Redescription of *Hammondia hammondi* and its differentiation from *Toxoplasma gondii*. Int J Parasitol *33*, 1437-1453.

120. Ojcius, D.M., Degani, H., Mispelter, J., and Dautry-Varsat, A. (1998). Enhancement of ATP levels and glucose metabolism during an infection by Chlamydia. NMR studies of living cells. J Biol Chem *273*, 7052-7058.

121. Radomski, N., Einenkel, R., Muller, A., and Knittler, M.R. (2016). Chlamydia-host cell interaction not only from a bird's eye view: some lessons from *Chlamydia psittaci*. FEBS Lett *590*, 3920-3940.

122. English, E.D., and Boyle, J.P. (2018). Impact of Engineered Expression of Mitochondrial Association Factor 1b on *Toxoplasma gondii* Infection and the Host Response in a Mouse Model. mSphere *3*.

123. Blank, M.L., and Boyle, J.P. (2018). Effector variation at tandem gene arrays in tissue-dwelling coccidia: who needs antigenic variation anyway? Curr Opin Microbiol *46*, 86-92.

124. Guan, S.S., Sheu, M.L., Wu, C.T., Chiang, C.K., and Liu, S.H. (2015). ATP synthase subunit-beta down-regulation aggravates diabetic nephropathy. Sci Rep *5*, 14561.

125. Fu, X., Chin, R.M., Vergnes, L., Hwang, H., Deng, G., Xing, Y., Pai, M.Y., Li, S., Ta, L., Fazlollahi, F., et al. (2015). 2-Hydroxyglutarate Inhibits ATP Synthase and mTOR Signaling. Cell Metab *22*, 508-515.

126. Braganza, A., Li, J., Zeng, X., Yates, N.A., Dey, N.B., Andrews, J., Clark, J., Zamani, L., Wang, X.H., St Croix, C., et al. (2017). UBE3B Is a Calmodulin-regulated, Mitochondrion-associated E3 Ubiquitin Ligase. J Biol Chem *292*, 2470-2484.

127. Ren, X., and Hurley, J.H. (2011). Proline-rich regions and motifs in trafficking: from ESCRT interaction to viral exploitation. Traffic *12*, 1282-1290.

128. Coppens, I., and Romano, J.D. (2018). Hostile intruder: Toxoplasma holds host organelles captive. PLoS Pathog *14*, e1006893.

129. Romano, J.D., Nolan, S.J., Porter, C., Ehrenman, K., Hartman, E.J., Hsia, R.C., and Coppens, I. (2017). The parasite Toxoplasma sequesters diverse Rab host vesicles within an intravacuolar network. J Cell Biol *216*, 4235-4254.

130. Martinvalet, D. (2018). The role of the mitochondria and the endoplasmic reticulum contact sites in the development of the immune responses. Cell Death Dis *9*, 336.

131. Meylan, E., Curran, J., Hofmann, K., Moradpour, D., Binder, M., Bartenschlager, R., and Tschopp, J. (2005). Cardif is an adaptor protein in the RIG-I antiviral pathway and is targeted by hepatitis C virus. Nature *437*, 1167-1172.

132. Takai, Y., Sasaki, T., and Matozaki, T. (2001). Small GTP-binding proteins. Physiol Rev *81*, 153-208.

133. Shirakawa, R., and Horiuchi, H. (2015). Ral GTPases: crucial mediators of exocytosis and tumourigenesis. J Biochem *157*, 285-299.

134. Kashatus, D.F., Lim, K.H., Brady, D.C., Pershing, N.L., Cox, A.D., and Counter, C.M. (2011). RALA and RALBP1 regulate mitochondrial fission at mitosis. Nat Cell Biol *13*, 1108-1115.

135. Caffrey, D.R., Zhao, J., Song, Z., Schaffer, M.E., Haney, S.A., Subramanian, R.R., Seymour, A.B., and Hughes, J.D. (2011). siRNA off-target effects can be reduced at concentrations that match their individual potency. PLoS One *6*, e21503.